

German Longitudinal
Election Study



GLES

Langfrist-Online-Tracking

T39, 2018

ZA6823, Version 2.0.0

Studienbeschreibung

Inhaltsverzeichnis

Vorbemerkungen	3
1. Studieninformationen	4
2. Konzeption	5
3. Stichprobenziehung.....	8
4. Erhebung	11
5. Datenaufbereitung.....	16
6. Hinweise und Anmerkungen.....	33
Links	34
Literaturverzeichnis.....	35

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:

GLES (2022): Langfrist-Online-Tracking, T39 (GLES). GESIS, Köln: ZA6823 Datenfile Version 2.0.0, doi:10.4232/1.13989.

1. Studieninformationen

1.1. Studiennummer

ZA6823

1.2. Titel der Studie

German Longitudinal Election Study, Komponente 8, Langfrist-Online-Tracking, T39: Politische Einstellungen und Wahlverhalten

1.3. Version

2.0.0, 20.09.2022, doi:10.4232/1.13989

1.4. Erhebungszeitraum

20.04.2018 – 04.05.2018

1.5. Primärforscher/innen

Prof. Dr. Sigrid Roßteutscher - Goethe-Universität Frankfurt am Main, Deutschland

Prof. Dr. Rüdiger Schmitt-Beck - Universität Mannheim, Deutschland

Prof. Dr. Harald Schoen - Universität Mannheim, Deutschland

Prof. Dr. Bernhard Weßels - Wissenschaftszentrum Berlin für Sozialforschung, Deutschland

Prof. Dr. Christof Wolf - GESIS – Leibniz-Institut für Sozialwissenschaften, Deutschland

1.6. Projektmitarbeiter/innen

Bauer, Irina - GESIS – Leibniz-Institut für Sozialwissenschaften, Deutschland

Dietz, Melanie - Goethe-Universität Frankfurt am Main, Deutschland

Kratz, Sophia - GESIS – Leibniz-Institut für Sozialwissenschaften, Deutschland

Kühn, Marie - GESIS – Leibniz-Institut für Sozialwissenschaften, Deutschland

Roßmann, Joss - GESIS – Leibniz-Institut für Sozialwissenschaften, Deutschland

Scherer, Philipp - Goethe-Universität Frankfurt am Main, Deutschland

Stövsand, Lars-Christopher - Goethe-Universität Frankfurt am Main, Deutschland

1.7. Finanzierende Stelle

Prof. Sigrid Roßteutscher, PhD, Johann Wolfgang Goethe-Universität Frankfurt

GESIS – Leibniz-Institut für Sozialwissenschaften, Mannheim

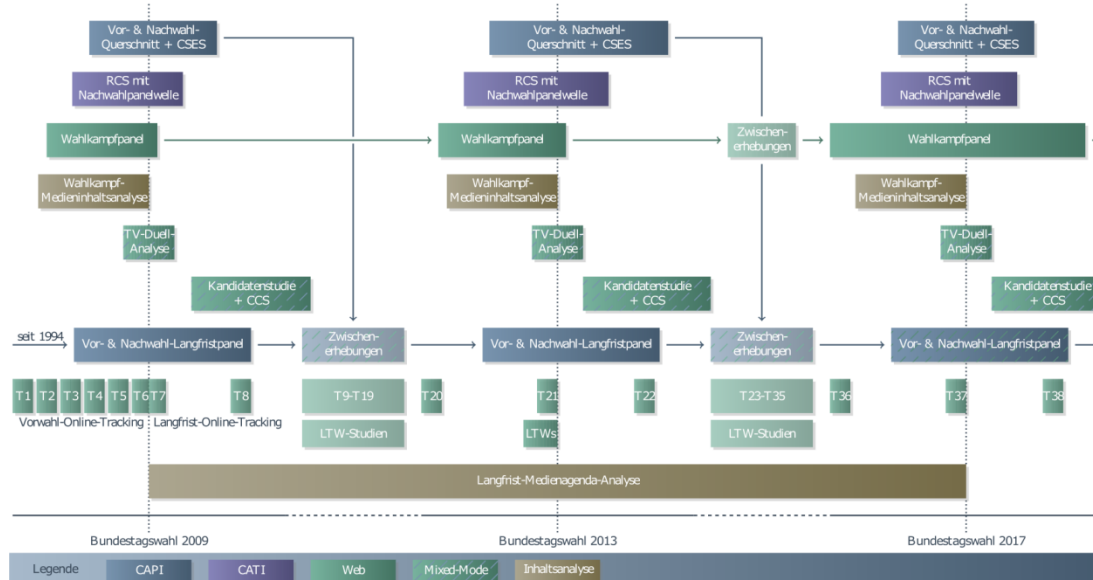
2. Konzeption

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 Wahlkampfes und auch langfristige soziale Wandlungsprozesse über den Wahlzyklus hinweg (Schmitt-Beck et al. 2010).

Abbildung 1: Das Design der German Longitudinal Election Study (GLES) zu den Bundestagswahlen 2009, 2013 und 2017



Das bis Frühjahr 2019 fortgeführte Langfrist-Online-Tracking ist dabei auf die Analyse langfristiger Prozesse der Formierung und des Wandels der öffentlichen Meinung ausgerichtet und zielt insbesondere darauf ab, die öffentliche Meinung und ihre Veränderungen zwischen den Bundestagswahlen zu untersuchen (Schmitt-Beck et al. 2010). Hierzu werden je Kalenderjahr in regelmäßigen Abständen bis zu vier webbasierte Querschnitte mit jeweils etwa 1.000 Befragten aus einem Online-Access-Panel erhoben. Mit Hilfe des Langfrist-Online-Trackings lassen sich die Einstellungen der Wähler zu den wichtigsten politischen und gesellschaftlichen Themen sowie zu den Parteien und ihrem Spitzenpersonal, aber auch zu den Leistungen der Bundesregierung und der Opposition analysieren (Schmitt-Beck et al. 2010).

2.2. Aufbau der Studie

Die Durchführung der Langfrist-Online-Trackings begann im Frühjahr vor der Bundestagswahl 2009. Die Grundlage der Studien bilden ein weitgehend konstanter Kernfragebogen zu zentralen Aspekten der Wahl- und politischen Einstellungsforschung und die Abfrage von soziodemographischen Angaben. Bis einschließlich dem Online-Tracking T38 wurde die

Befragung zusätzlich um rotierende thematische Module sowie situationsbezogene variierende Frageblöcke ergänzt. Eine Übersicht über die bisher durchgeführten Online-Trackings und die jeweiligen thematischen Schwerpunkte bietet die nachfolgende Darstellung.

Tabelle 1: Die Langfrist-Online-Trackings der GLES

Erhebung	Studien-nummer	Schwerpunkt	Feldbeginn	Feldende
T1	ZA5334	Grundlegende Tests	30.04.2009	05.05.2009
T2	ZA5335	Wählen auf mehreren Ebenen	27.05.2009	05.06.2009
T3	ZA5336	Wirtschaftliche Lage	03.07.2009	13.07.2009
T4	ZA5337	Koalitionen	31.07.2009	11.08.2009
T5	ZA5338	Skalen- und Reihenfolgeexperimente	24.08.2009	01.09.2009
T6	ZA5339	Wahlkampf	18.09.2009	27.09.2009
T7	ZA5340	Nachwahl	29.09.2009	08.10.2009
T7Exp	ZA5340	Experiment zur Bundestagswahl	08.10.2009	25.10.2009
T8	ZA5341	Nachwahl und Netzwerke	10.12.2009	20.12.2009
T9	ZA5342	Wirtschaftliche Lage	15.04.2010	23.04.2010
T10	ZA5343	Wählen auf mehreren Ebenen	24.06.2010	05.07.2010
T11	ZA5344	Psychologische Konstrukte	16.09.2010	26.09.2010
T12	ZA5345	Positionsissues	09.12.2010	19.12.2010
T12NB	ZA5345	Nachbefragung von Abbrechern	20.12.2010	30.12.2010
T13	ZA5346	Wirtschaftliche Lage	09.03.2011	19.03.2011
T13NB	ZA5346	Nachbefragung von Abbrechern	21.03.2011	30.03.2011
T14	ZA5347	Wählen auf mehreren Ebenen	23.05.2011	03.06.2011
T14NB	ZA5347	Nachbefragung von Abbrechern	03.06.2011	13.06.2011
T15	ZA5348	Psychologische Konstrukte	24.08.2011	03.09.2011
T15NB	ZA5348	Nachbefragung von Abbrechern	05.09.2011	14.09.2011
T16	ZA5349	Positionsissues	08.12.2011	18.12.2011
T17	ZA5350	Wirtschaftliche Lage und Wählen auf mehreren Ebenen	02.05.2012	15.05.2012
T18	ZA5351	Netzwerke, psychologische Konstrukte und Koalitionen	17.09.2012	01.10.2012
T19	ZA5719	EU-Krise und politische Partizipation	04.01.2013	19.01.2013
T20	ZA5720	Positionsissues und Europa	24.05.2013	08.06.2013
T21	ZA5721	Wahlkampf	06.09.2013	21.09.2013
T22	ZA5722	Koalitionen und psychologische Konstrukte	29.11.2013	13.12.2014
T23	ZA5723	Wirtschaftliche Lage und politisches Wissen	21.02.2014	07.03.2014
T24	ZA5724	Europa und Europawahl	09.05.2014	23.05.2014
T25	ZA5725	Psychologische Konstrukte	29.08.2014	13.09.2014
T26	ZA5726	Positionsissues, Netzwerke und Koalitionen	21.11.2014	05.12.2014
T27	ZA5727	Psychologische Konstrukte, Wirtschaftliche Lage, Politisches Wissen, Partnerschaft	27.02.2015	13.03.2015
T28	ZA5728	Europa, Institutionenvertrauen, Position-	05.06.2015	19.06.2015

		issues		
T32	ZA5732	Europa und Positionsissues	03.06.2016	17.06.2016
T33	ZA5733	Psychologische Konstrukte und Umfragen	19.08.2016	02.09.2016
T34	ZA5734	Positionsissues und Koalitionen	02.12.2016	16.12.2016
T35	ZA6815	Netzwerke, Psychologische Konstrukte, Wirtschaftliche Lage und Politisches Wissen	17.03.2017	31.03.2017
T36	ZA6816	Europa und Positionsissues	16.06.2017	30.06.2017
T37	ZA6817	Wahlkampf, politische Einstellungen und Verhaltensweisen	12.09.2017	23.09.2017
T38	ZA6818	Positionsissues, Koalitionen und psychologische Konstrukte	01.12.2017	15.12.2017
T39	ZA6823	Politische Einstellungen und Wahlverhalten	20.04.2018	04.05.2018

2.3. Inhalt der Studie

Diese Erhebung des Langfrist-Online-Trackings umfasst den Kernfragebogen zu zentralen Aspekten der Wahl- und politischen Einstellungsforschung sowie soziodemographische Angaben.

3. Stichprobenziehung

3.1. Untersuchungsgebiet

Deutschland (DE)

3.2. Grund- und Auswahlgesamtheit

Die Grundgesamtheit der Studie bildet die zum Zeitpunkt der Erhebung für eine Wahl zum Deutschen Bundestag wahlberechtigte Bevölkerung der Bundesrepublik Deutschland, d.h. deutsche Staatsbürger, die zum Zeitpunkt der Studiendurchführung das 18. Lebensjahr erreicht hatten. Da die Studie webbasiert durchgeführt wird, haben nicht alle wahlberechtigten Bürger eine Chance, für die Befragung des Online-Trackings ausgewählt zu werden. Die dieser Studie zugrundeliegende Auswahlgesamtheit umfasst daher ausschließlich aktive Teilnehmer des von der respondi AG betriebenen Online-Access-Panels, die die deutsche Staatsbürgerschaft besitzen und zum Zeitpunkt der Studiendurchführung das 18. Lebensjahr erreicht hatten. Das respondi-Panel umfasste 2016 in Deutschland 65.000-70.000 aktive Personen. Als aktive Personen definiert respondi diejenigen Personen, die nach der (Double-opt-in-)Registrierung¹ die Stammdatenbefragung ausgefüllt haben und in den letzten drei Monaten an mindestens einer Umfrage teilgenommen haben.

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

	respondi
Geschlecht	
Weiblich	62
Männlich	38
Bildung	
Niedrig (d.h. kein Abschluss, Haupt- und Volksschule)	19
Mittel (d.h. Realschule, Mittlere Reife)	39
Hoch (d.h. Abitur, Fachhochschulreife, Studium)	42
Altersgruppen	
18-29 Jahre	30
30-39 Jahre	21
40-49 Jahre	18
50-59 Jahre	18
60 und älter	13

Fehlend zu 100%: keine Angabe.

Die Mitglieder des Online-Access-Panels werden von respondi auf unterschiedlichen Wegen rekrutiert. Zu diesem Zweck tritt respondi unter dem Markennamen „Mingle“ an die Personen

¹ Double-opt-in bedeutet, dass ein Nutzer nach seiner Registrierung eine E-Mail erhält, die ihn dazu auffordert, die Mitgliedschaft zu bestätigen. Erst danach wird der Nutzer zur Stammdatenbefragung weitergeleitet.

heran. Dies erfolgt überwiegend online, zu einem geringen Anteil aber auch offline. Die Mitglieder wirbt respondi mehrheitlich über Umfragen auf themenspezifischen Portalen, in Foren und Communities an.

Tabelle 3: Rekrutierungswege bei respondi

	Anteil in Prozent
MingleTrend Blog	2,1
Über Online-Werbung (davon beworbene Thementumfragen 50%; direkte Panelwerbung 50% inkl. Affiliate Marketing)	38
Selbsteinträge	11
Facebook Fanseite	35
Suchmaschinen	10
Empfehlungen (Freundschaftswerbung)	0,9
Kooperationen	2,9
Telefonakquise (CATI)	0,1

respondi gibt an, ein wirkungsvolles Qualitätsmanagement zu betreiben. Über ein internes Bewertungssystem wird das Antwortverhalten der Mitglieder des respondi-Panels kontinuierlich evaluiert. Gründe für die Deaktivierung von Mitgliedern aus der Datenbank sind beispielsweise fehlende Teilnahme über einen Zeitraum von drei Monaten (Voraussetzung bei respondi ist, dass die Mitglieder mindestens zehn Einladungen erhalten haben) sowie Doppelanmeldungen und bewusste Falschangaben bei mehreren Befragungen.

Das Qualitätsmanagement sowie eine maßvolle Einladungsfrequenz sollen dazu beitragen, unerwünschte Effekte wie Panel-Verzerrungen oder Professionalisierung der Mitglieder zu vermeiden. Ein durchschnittliches Mitglied verweilt elf Monate im respondi-Panel. Innerhalb eines Jahres werden ca. 40 Prozent der Panelmitglieder aufgrund von Qualitätskontrollen und Panel-Mortalität aus der Datenbank entfernt.

Die durchschnittliche Teilnahmequote der respondi-Mitglieder liegt bei einer Umfrage mit fünf Feldtagen bei etwa 50 Prozent. Die Teilnahmequote wird von respondi aus der Anzahl der begonnenen Interviews berechnet, also den beendeten Interviews, Screen-outs (Ausschluss bei Auswahl der zu befragenden Teilnehmer nach Thema), Quota-fulls (Ausschluss auf Grund erfüllter Quoten bei Quotenauswahl) und den abgebrochenen Interviews, die ins Verhältnis zur Gesamtzahl der Einladungen gesetzt werden.

Die Mitglieder des Access-Panels werden von respondi für die Teilnahme an Befragungen incentiviert. Sie erhalten normalerweise fünf sogenannte mingle-Punkte pro Minute, was 0,05 € entspricht. Sobald die Teilnehmer 1.500 mingle-Punkte gesammelt haben, kann dieser Betrag in Einkaufsgutscheine umgewandelt oder gespendet werden. Ab 2.000 mingle-Punkten können diese auch ausgezahlt werden. Zusätzlich werden zur Panelpflege regelmäßig Verlosungen unter den Mitgliedern durchgeführt.

3.3. Auswahlverfahren

Die Stichprobe für das Online-Tracking wird durch eine Quotenauswahl aus der Auswahlgesamtheit der bei respondi zur Verfügung stehenden aktiven Panelteilnehmer generiert. Dabei werden nur volljährige Personen mit deutscher Staatsangehörigkeit berücksichtigt.

Weiterhin wird darauf geachtet, dass die Panelteilnehmer nur einmal jährlich an einem Online-Tracking der GLES teilnehmen können.

Die Teilnehmer werden nach den Merkmalen Geschlecht, Alter und Bildung quotiert. Die Soll-Zahlen basieren auf den nach Internetnutzung und deutscher Staatsangehörigkeit gefilterten Verteilungen aus der „Best for Planning“-Studie, die von der Gesellschaft für integrierte Kommunikationsforschung gestellt wird.

Tabelle 3: Quotierung

Quotierungsmerkmale	Soll in Prozent	Ist in Prozent
Geschlecht		
Weiblich	47,6	48,2
Männlich	52,4	51,8
Bildung		
Niedrig (kein Abschluss, Haupt- oder Volksschule)	27,0	27,0
Mittel (Realschule, Mittlere Reife)	35,9	35,4
Hoch (Fachhochschulreife, Abitur)	37,1	37,6
Altersgruppe		
18-29 Jahre	21,6	20,7
30-39 Jahre	17,3	17,6
40-49 Jahre	23,7	24,1
50-59 Jahre	19,8	20,1
60 Jahre und älter	17,6	17,5

4. Erhebung

4.1. Erhebungsverfahren

Online-Befragung mit standardisiertem Fragebogen

4.2. Datenerhebung

GESIS übernahm die Programmierung des Fragebogens und die Durchführung der Datenerhebung.

4.3. Erhebungssoftware

Die Programmierung des Fragebogens und die Durchführung der Datenerhebung erfolgten mittels der Software Questback EFS Spring 2018.

4.4. Tests des programmierten Fragebogens

Der programmierte Fragebogen wurde intensiv geprüft. In den Tests gefundene Probleme oder Programmierfehler wurden bereinigt. Anschließend wurde der Fragebogen in der finalen Version von den Projektpartnern freigegeben.

4.5. Incentivierung

Um eine möglichst hohe Ausschöpfung zu erzielen, wurden Incentives eingesetzt. Für die Teilnahme an der ca. 20-minütigen Studie erhielten die Befragten 200 mingle-Punkte, was einem Gegenwert von 2,00 € entsprach.

4.6. Einladungen und Reminder

Die Einladung der Teilnehmer erfolgte durch respondi mit folgendem Standardtext:

Hallo Vorname Nachname,

heute haben wir ein neues Befragungsprojekt für Sie. Wir freuen uns auf Ihre Meinung!

Umfrage jetzt starten!

Umfragedauer	Vergütung
20 Minuten	200 mingle-Punkte

Ihre Teilnahme lohnt sich in jedem Fall!

Zur Bestimmung der genauen Zielgruppe werden in den meisten Projekten kurze Einstiegsfragen gestellt. Gehören Sie nicht zur Zielgruppe, werden Ihnen in jedem Fall 5 mingle-Punkte gutgeschrieben.

Viel Spaß bei der Befragung und bis bald auf mingle!



Herzlichst

Sarah Maiwald

Ihre Ansprechpartnerin bei mingle

Wenn ein Panelist die Einladung annahm, wurde er weitergeleitet und im Namen der für die Durchführung der Studie Verantwortlichen der GLES zur Befragung mit folgendem Text eingeladen:




Abbildung 2: Startbildschirm der Befragung

Herzlich willkommen!

Wir freuen uns, dass Sie an unserer Umfrage teilnehmen. Diese Umfrage ist Teil eines deutschlandweiten Projekts zur Untersuchung von Wahlen in Deutschland. Wir wünschen viel Freude beim Beantworten der Fragen und bedanken uns herzlich für Ihre Teilnahme.

Prof. Dr. Sigrid Roßteutscher und Melanie Dietz

Hinweis: Um zukünftige Befragungen einfacher und besser gestalten zu können, werden technische Informationen zum Befragungsverlauf erhoben (z.B. Datum, Zeiten, Bildschirmauflösung).

1%

WEITER

Zum Ende der Feldphase wurde folgender Reminder an Personen aus bisher noch nicht gefüllten Quoten versendet:

Hallo Vorname Nachname,

wir haben Sie vor kurzem zu einer neuen Befragung eingeladen und würden uns freuen, wenn Sie an diesem wichtigen Projekt teilnehmen:

Hier geht's zur Befragung: [Link](#)

Die Beantwortung des Fragebogens dauert ca. 20 Minuten

Gehören Sie zur Zielgruppe, erhalten Sie für die Beantwortung 200 mingle-Punkte.

Zur Bestimmung der genauen Zielgruppe werden in den meisten Projekten kurze Einstiegsfragen gestellt. Gehören Sie nicht zur Zielgruppe, werden Ihnen in jedem Fall 5 mingle-Punkte gutgeschrieben.

Herzlichst

Sarah Maiwald

Ihre Ansprechpartnerin bei mingle

Tabelle 4: Übersicht über Einladungen und Reminder

Datum	Gruppe	N	Geschlecht		Alter					Bildung		
			Männlich	Weiblich	18-29	30-39	40-49	50-59	60+	Niedrig	Mittel	Hoch
Fr 20.04.2018	E	2.988	1.495	1.493	748	597	551	539	553	1.695	644	649
Mo 23.04.2018	E	792	328	286	210	139	109	83	73	614	0	0
Di 24.04.2018	E	1.971	983	988	705	394	615	198	59	0	986	985
Do 26.04.2018	E	1.969	819	1.150	1.077	394	395	103	0	0	1.281	688
Fr 27.04.2018	E	491	491	0	491	0	0	0	0	0	491	0
Mo 30.04.2018	R	460	460	0	460	0	0	0	0	0	460	0
Mi 02.05.2018	R	883	883	0	883	0	0	0	0	0	883	0
Fr 04.05.2018	R	864	864	0	864	0	0	0	0	0	864	0

Anmerkungen: E = Einladung, R = Reminder

4.7. Feldverlauf

Die nachfolgende Tabelle illustriert das Teilnahmeverhalten der zur Umfrage eingeladenen Teilnehmer des respondi-Panels über den Erhebungszeitraum hinweg. Partielle und vollständige Interviews umfassen alle Interviews, die bis zur letzten Frage des Fragebogens durchgeführt wurden, unabhängig davon, ob die Umfrage unterbrochen wurde oder nicht. Die Anzahl der begonnenen Interviews ergibt sich aus der Anzahl der partiellen und vollständigen Interviews, der abgebrochenen Interviews sowie der aufgrund der Quotierung nicht durchgeführten Interviews.

HINWEIS: Im Rahmen der Datenaufbereitung wurden 20 Teilnehmer*innen identifiziert, von denen insgesamt 40 Interviews vorlagen. Um diese Problematik der Mehrfachteilnahmen zu beheben, wurde folgendermaßen vorgegangen: Zunächst wurden Interviews, die keine Ausprägung auf der tic-Variable² vorweisen konnten, entfernt. Darunter fiel kein Interview, das beendet bzw. nach Unterbrechung beendet wurde. Dann wurde festgestellt, welches Interview den höchsten Anteil substanzieller Antworten an allen gesehenen Fragen aufweist. Substanzielle Antworten bezeichnen „echte“ inhaltliche Antworten der Befragten, also keine Antwortverweigerungen, Umfrageabbrüche, oder ähnliches. Interviews mit geringeren Anteilen wurden entfernt. Bei mehreren Interviews mit gleich hohem Anteil substanzieller Antworten, wurde in einem zweiten Schritt geprüft, welches Interview als erstes beendet wurde. Die übrigen Interviews wurden dann ebenfalls entfernt. Diese Vorgehensweise entfernte unter anderem auch 1 Interview, das beendet bzw. nach Unterbrechung beendet wurde. Die Fallzahl im finalen Datensatz reduziert sich damit um insgesamt 1 (0 ohne tic + 1 Mehrfachteilnahmen) Interviews, die beendet bzw. nach Unterbrechung beendet wurden.

² Die tic-Variable ermöglicht die Fallidentifikation innerhalb einer Umfrage.

Tabelle 5: Teilnahmestatistik nach Tagen

Datum		Begonnene Interviews			Partielle und vollständige Interviews		
		N	Prozent	kum. Prozent	N	Prozent	kum. Prozent
Fr	20.04.2018	120	6,4	6,4	92	8,3	8,3
Sa	21.04.2018	247	13,1	19,5	207	18,7	27,0
So	22.04.2018	93	4,9	24,4	80	7,2	34,2
Mo	23.04.2018	164	8,7	33,1	134	12,1	46,3
Di	24.04.2018	226	12,0	45,1	173	15,7	62,0
Mi	25.04.2018	239	12,7	57,8	174	15,7	77,8
Do	26.04.2018	252	13,4	71,2	192	17,3	95,1
Fr	27.04.2018	162	8,6	79,8	21	1,9	97,0
Sa	28.04.2018	58	3,1	82,9	7	0,6	97,7
So	29.04.2018	66	3,5	86,4	9	0,8	98,5
Mo	30.04.2018	69	3,7	90,1	7	0,6	99,1
Di	01.05.2018	55	2,9	93,0	3	0,3	99,4
Mi	02.05.2018	53	2,8	95,8	4	0,4	99,7
Do	03.05.2018	47	2,5	98,3	1	0,1	99,8
Fr	04.05.2018	32	1,7	100,0	2	0,2	100,0
Gesamt		1.883	100,0		1.106	100,0	

4.8. Ausschöpfung

Systematischer Unit Nonresponse ist eine Fehlerquelle in Umfragestudien, die die Qualität der erhobenen Daten erheblich mindern kann. 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 Indikatoren des Teilnahmeverhaltens der Personen in der Stichprobe sichert die Vergleichbarkeit über Studien hinweg. Für diese Studie werden zwei Indikatoren des Teilnahmeverhaltens nach den Empfehlungen zur Berechnung der Ausschöpfung von Callegaro und DiSogra (2008) angegeben.

Die Completion Rate (Callegaro und DiSogra 2008) berechnet sich als

$$\text{Completion Rate (COMR)} = \frac{(I + P)}{(I + P) + (R + NC + O)}$$

wobei I die Anzahl der vollständigen Interviews, P die Anzahl der partiellen Interviews, R die Anzahl der Verweigerungen und Befragungsabbrüche, NC die Anzahl der nicht hergestellten Kontakte und O die Anzahl der Ausfälle aus anderen Gründen ist.

Als weiterer informativer Indikator wird der Anteil der Befragungsabbrüche in der Online-Befragung berichtet. Die Breakoff Rate (Callegaro und DiSogra 2008) berechnet sich 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 umfasst. Die Breakoff Rate gibt somit den Anteil der Befragungsabbrüche an allen begonnen Interviews wieder.

Die Brutto-Stichprobe für die Online-Befragung wurde aus dem nicht-probabilistischen Online-Access-Panel der respondi AG gezogen. Die ausgewählten Panelmitglieder wurden von respondi zur Teilnahme eingeladen. Panelteilnehmer, die das Interview bis zur abschließenden Seite der Befragung absolvierten, wurden der Gruppe der vollständigen Interviews (I) zugeordnet. Wurde das Interview begonnen, während der Teilnahme unterbrochen und bis zum Ende der Feldzeit nicht vollständig absolviert, so wird der Panelteilnehmer der Gruppe der Befragungsabbrüche (R_{Breakoff}) zugeordnet. Um den Quotenvorgaben zu genügen, wurden Panelteilnehmer mit bestimmten Ausprägungen bei den Variablen Geschlecht, Alter und höchstem allgemein bildendem Schulabschluss nach Erreichen der einzelnen Vorgaben von der Teilnahme ausgeschlossen. Die betroffenen Teilnehmer wurden abgewiesen. Erfolgte keine Reaktion auf eine Befragungseinladung, so wurde dies als nicht hergestellter Kontakt klassifiziert (NC). Teilnehmer, die nach ihren eigenen Angaben nicht die deutsche Staatsangehörigkeit besitzen, wurden ausgescreent.

Tabelle 6: Ausschöpfung

Kategorie/Indikator	Anzahl/Rate in Prozent
Anzahl der von respondi versandten Befragungseinladungen	8.211
I = Vollständige Interviews	1.106
R = Verweigerungen und Befragungsabbrüche	101
Quote erfüllt/abgewiesen	513
Ausgescreent	72
NC = Nicht hergestellter Kontakt	6.417
Completion Rate	in Prozent 13,5
Breakoff Rate	in Prozent 8,4

4.9. Interviewdauer

Die Interviewdauer ist ein Indikator für den Umfang einer Befragung. Sie lässt Rückschlüsse darauf zu, wie groß der Aufwand war, den Befragte auf sich nehmen mussten, um den Fragebogen zu beantworten. Die nachfolgende Tabelle präsentiert deskriptive Statistiken zur Interviewdauer.

Tabelle 7: Interviewdauer

	N	Min.	Max.	Arithmetisches Mittel	Median
Interviewdauer	1.105	0:03:52	49:28:05	0:35:35	0:23:15

Anmerkung: Hier wird die über das ECSP-Skript nutzerseitig gemessene Interviewdauer (Variable `durati-on_client`) ausgewiesen. Bei einem Fall ist es im Rahmen der Datenerhebung zu einer fehlerhaften Aufzeichnung der Interviewdauer gekommen. Für diesen Fall liegt die minimale Interviewdauer bei 0 Sekunden, weshalb er hier nicht mitgezählt wird.

5. Datenaufbereitung

5.1. Datenkontrollen

Die Daten dieser Studie wurden nach der Erhebung sorgfältig geprüft. Hierunter fiel die Überprüfung auf Vollständigkeit der Variablen im Datensatz, die Überprüfung auf Richtigkeit von Variablennamen und -labels, die Prüfung auf Vollständigkeit von Wertelabels, die Überprüfung der Codierung von Missing Values sowie eine Kontrolle der Randverteilung. Darüber hinaus erfolgte eine technische Kontrolle der Filterführungen. Alle gefundenen Fehler werden während der Datenaufbereitung dokumentiert und korrigiert. Trotz gründlicher Überprüfung der Daten kann es vorkommen, dass Fehler in den Datensätzen unentdeckt bleiben. Fallen diese zu einem späteren Zeitpunkt auf, werden sie in der Errataliste im GESIS Datenbestandskatalog (www.gesis.org/dbk) dokumentiert und zeitnah behoben.

Die Datensatzkontrolle wurde unterstützt durch die Verwendung des frei verfügbaren Stata-Moduls scandata (Kaukal 2016), das die Variablen in einem Datensatz auf Grundlage von vorgegebenen Spezifikationen (Verwendung von Großbuchstaben oder Umlauten, Länge von Variablennamen und Labels, ungerade Verteilungen) untersucht, Abweichungen von diesen anzeigt und – sofern möglich und vom Nutzer gewünscht – automatisch korrigiert.

5.2. Einheiten im Datensatz

Der Datensatz umfasst 1.106 realisierte Interviews.

5.3. Variablen im Datensatz

Der Datensatz setzt sich aus verschiedenen Typen von Variablen bzw. Daten zusammen. Diese umfassen Meta-, Para- und substantielle Daten sowie Kontext- und Gewichtungsvariablen (siehe auch die Abschnitte 5.4 bis 5.8). Die Tabelle 8 gibt einen Überblick über die Anzahl der Variablen im Datensatz nach Variablentyp.

Tabelle 8: Anzahl der Variablen im Datensatz nach Variablentyp

Variablentyp	Anzahl
Metadaten	7
Paradaten	130
Substantielle Daten	254
Kontextvariablen	6
Gewichtungsvariablen	4
Gesamt	401

5.4. Metadaten

Metadaten beschreiben den Datensatz und seine Erstellung (Callegaro 2013). Sie umfassen Variablen, die für die Archivierung und Distribution der Studie benötigt werden. In der GLES umfasst dies die Studiennummer des Datensatzes im vierstelligen Format, unter welcher die Studie in den Archivsystemen von GESIS abgelegt und zu finden ist, die Versionskennung des Datensatzes, Angaben zum Erhebungsjahr und –zeitraum, die Zuordnung der Studie zu den Komponenten der GLES sowie die Kennung der Erhebung innerhalb der Komponente der GLES (siehe Tabelle 9).

Tabelle 9: Metadaten

Variable	Label
study	Studiennummer
version	GESIS Archiv Version
doi	Digital Object Identifier
glescomp	GLÉS-Komponente
field_start	Beginn Erhebungszeitraum
field_end	Ende Erhebungszeitraum
sample	Stichprobe

5.5. Paradata

Paradata sind Informationen, die aus dem Erhebungsprozess resultieren (Couper 2000), z.B. das Datum und die Uhrzeit des Beginns und des Endes der Interviewteilnahme, die Interviewdauer, Antwortzeitmessungen oder Informationen zum verwendeten Gerät und dem verwendeten Webbrowser (Callegaro 2013). Auch administrative Informationen wie Identifikationsnummern, Informationen zum Teilnahmestatus von Befragungsteilnehmern und Informationen zur Zuweisung von Teilnehmern zu Gruppen auf Grund von inhaltlichen Angaben oder zufallsbasierten Verfahren werden hier als Paradata behandelt. Die Tabelle 10 gibt einen Überblick über Paradata im Datensatz.

Tabelle 10: Paradata

Variable	Label
lfdn	Laufende Nummer
browser	Genutzter Browser
device	Genutztes Geraet
smartphone	Nutzung Smartphone
tablet	Nutzung Tablet
javascript	javascript
datetime	datetime of survey start
date_of_last_access	date of last access
dispcode	disposition code
intstatus	Interviewstatus: Anteil beantworteter Fragen, kategorial
intstatus_prc	Interviewstatus: Anteil beantworteter Fragen, prozentual
duration_client	Interviewdauer (nutzer:innenseitig gemessen)
duration_server	Interviewdauer (serverseitig gemessen)
speederindex	Zeitunterschreiter-Index
speederflag	Zeitunterschreiter (10% der Befragten mit den niedrigsten Indexwerten)
lastpage	last page
zstart-zende	Zeitvariablen

Darüber hinaus gibt es einen erweiterten Paradatenatz, der zusätzliche Informationen zum Erhebungsprozess bereithält. Dieser Datensatz kann bei Interesse unter gles@gesis.org angefordert werden.

Diese Paradata wurden mittels des frei verfügbaren Systems „Embedded Client Side Paradata ECSP“ (Schlosser & Höhne 2018) erfasst. ECSP ist lizenziert unter der Creative Commons Attribution 4.0 International License. Es basiert auf unterschiedlichen Programmiersprachen, wie JavaScript und HTML. ECSP kann in onlinebasierte Umfragesoftware implementiert werden, die Zugang zum Programmcode ermöglicht, und erfasst eine Vielzahl an unterschiedlichen Paradatentypen, wie Antwortzeiten oder Mausklicks, unabhängig vom Internetbrowser und Betriebssystem. Die Paradata werden auf Seitenebene erfasst und gemeinsam mit den Antworten der Befragungspersonen in demselben Datensatz gespeichert.

5.5.1 Laufende Nummer und Teilnahmestatus

Die Variable *lfdn* „Laufende Nummer“ ist eine eindeutige Identifikationsnummer für die Teilnehmer an der Befragung. Eine Identifikation mittels der laufenden Nummer von Personen, die an mehreren Erhebungen des Online-Trackings teilgenommen haben, ist nicht möglich, da diese für jede Erhebung neu generiert wird.

5.5.2 Unterbrechung der Befragung

Die Teilnehmer der Studie konnten die Beantwortung des Fragebogens jederzeit unterbrechen und innerhalb der Feldzeit zu einem von ihnen selbst gewählten Zeitpunkt fortsetzen. Hat ein Befragter die Beantwortung unterbrochen und den Browser oder das Browserfenster auf seinem Computer geschlossen, so wurde diese Information in der Variable *dispcode* „disposition code“ gespeichert. Die Unterbrechung der Beantwortung der Umfrage ohne Schließen des Fragebogens im Browser kann nicht direkt festgestellt werden.

5.5.3 Zeitvariablen

Die Antwortzeiten wurden mithilfe des Embedded Client Side Paradata (ECSP) Skriptes (Schlosser & Höhne 2018) gemessen. Die Antwortzeitmessungen werden dabei nutzerseitig („client side“) erhoben, also direkt im Browser der befragten Person. Dies ermöglicht die Dokumentation der Befragtenaktivität im Zeitraum zwischen Empfang und Absenden einer Befragungsseite. Allerdings muss die befragte Person dazu JavaScript aktiviert haben.

Die Antwortzeit umfasst dabei die Zeit des vollständigen Seitenaufbaus bis zum Versenden der Umfrageseite mit Hilfe der "Weiter"-Schaltfläche (Schlosser 2016, 29). Das ECSP Skript misst diesen Zeitraum in Millisekunden, allerdings wurden die Antwortzeiten im Datensatz (Variablen *zstart-zend*) in Sekunden umgewandelt und auf ganze Zahlen gerundet. Für die Interviewdauer (Variable *duration_client*) wurden diese seitenbezogenen Antwortzeiten summiert.

Zusätzlich zu der nutzerseitigen Messung der Interviewdauer veröffentlichen wir außerdem die serverseitig gemessene Interviewdauer in der Variable *duration_server*. Das ist die Zeit, die eine befragte Person zwischen dem Betrachten und Bearbeiten der Begrüßungsseite bis zum Schließen der letzten Seite der Befragung benötigt hat – inklusive der Zeit, die das Laden der einzelnen Seiten beansprucht hat. Bei der Analyse der serverseitigen Interviewdauer ist zu beachten, dass Befragte den Fragebogen verlassen und die Befragung zu einem späteren Zeitpunkt fortsetzen konnten. Wenn die Befragten die Umfrage unterbrechen und zu einem späteren Zeitpunkt weiterführen, wird ihnen der Wert -99 zugewiesen, da dadurch keine sinnvolle Berechnung der Interviewdauer möglich ist. Die Information zu Unterbrechungen der Befragung ist in der Variable *dispcode* enthalten.

5.5.4 Zeitunterschreiter

Aufgrund der fehlenden Kontrolle der Befragten durch einen Interviewer in Online-Befragungen und der besonderen Anreiz- und Belohnungsstruktur durch die Incentivierung der Befragten ist diese Studie mit dem Problem außergewöhnlich schneller Antwortzeiten („Zeitunterschreitung“) konfrontiert. Zeitunterschreitung bedeutet, dass einige Befragte einzelne Fragen oder auch die gesamte Umfrage erheblich schneller beantworten als der Großteil der Teilnehmer. Eine schnelle Beantwortung der Fragen in der Umfrage ist noch kein Problem an sich, da es aufgrund bestimmter sozialstruktureller und persönlicher Merkmale (z.B. Bildung, Alter, Intelligenz, Reaktionsgeschwindigkeit) deutliche Unterschiede in der Beantwortungsgeschwindigkeit zwischen Befragten geben kann. Dennoch ist davon auszugehen, dass bei einer erheblichen Unterschreitung der mittleren Antwortzeit die Antwortqualität leidet, da sich Befragte mutmaßlich wenig bei der Beantwortung der Fragen bemühen und daher oberflächliche oder willkürliche Antworten abgeben, keine Angabe machen oder „weiß nicht“ antworten, obwohl sie eine substantielle Antworten hätten abgeben können (vgl. Krosnick 1991).

In der Fachliteratur gibt es keine etablierten Standards für die Identifikation von Zeitunterschreitern. In der Regel beziehen die Maße für deren Identifikation den Median bzw. Mittelwert der Verteilung und die Streuung mit ein und wählen auf dieser Basis ein Abschnidekriterium, das nicht unterschritten werden darf (vgl. Mayerl und Urban 2008). Diese Befragten werden dann entweder aus dem Datensatz ausgeschlossen oder durch Markervariablen gekennzeichnet.

In dieser Studie werden Zeitunterschreiter anhand einer modifizierten Routine des von Roßmann (2010) weiterentwickelten Algorithmus identifiziert. Hierfür wird mittels des frei verfügbaren Stata-Moduls *speedergles* (Roßmann 2015) ein Index der mittleren Antwortgeschwindigkeit gebildet (Variable *speederindex*). Der Index bezieht sowohl die Antwortzeit der Befragten auf allen Bildschirmseiten der Befragung als auch die Interviewdauer je Befragtem ein. Die Indexwerte können als Maß für die mittlere Antwortgeschwindigkeit eines Befragten interpretiert werden. Ein Indexwert von 1 entspricht dem Mittelwert der mittleren Antwortgeschwindigkeit der Befragten in der untersuchten Stichprobe. Indexwerte gegen 0 bedeuten eine sehr schnelle mittlere Antwortgeschwindigkeit, während Werte gegen 2 eine sehr langsame mittlere Antwortgeschwindigkeit bedeuten.

Als Zeitunterschreiter werden alle Befragten im unteren 10%-Perzentil der Verteilung des Index gekennzeichnet. Die Kennzeichnung von Zeitunterschreitern ist im Datensatz in der Variable *speederflag* enthalten. Diese Markervariable kann zum Ausschluss der Fälle aus Analysen als auch zur Durchführung von Robustheitsanalysen verwendet werden.

In dieser Studie wurden 111 Fälle (10,04%) als Zeitunterschreiter markiert.

Tabelle 11: Statistiken zur Stichprobe mit und ohne Zeitunterschreiter

Merkmal	Mit Zeitunterschreitern	Ohne Zeitunterschreiter
Interviewdauer (hh:mm:ss)		
Min.	0:03:52	0:12:42
Mittelwert	0:35:35	0:38:26
Median	0:23:15	0:24:38
Geschlecht (in %)		
Männlich	51,72	51,11
Weiblich	48,28	48,89
Altersgruppe (in %)		
18 bis unter 30 Jahre	21,43	19,62
30 bis unter 45 Jahre	28,12	26,66

45 bis unter 60 Jahre	34,36	35,92
60 Jahre und älter	16,09	17,81
Bildung (in %)		
Niedrig	27,40	28,47
Mittel	34,99	35,51
Hoch	37,61	36,02
N	1.106	994

Anmerkung: Bei einem Fall ist es im Rahmen der Datenerhebung zu einer fehlerhaften Aufzeichnung der Interviewdauer gekommen. Für diesen Fall liegt die minimale Interviewdauer bei 0 Sekunden, weshalb er hier nicht mitgezählt wird.

5.5.5 Browser, Smartphone und Java

Die verwendete Erhebungssoftware liest den sogenannten „user agent header“ aus. Der „user agent“ ist eine Software, die in der Kommunikation im Internet als Interface zwischen den Servern, auf denen die Information gespeichert sind, und dem Nutzer fungiert, z.B. der auf dem Computer des Nutzers verwendete Browser. Es ist somit möglich zu identifizieren, welcher Browser verwendet wurde, ob die Befragung mittels eines Smartphones ausgefüllt wurde und ob der Browser Javascript unterstützt. Für die Codierung der entsprechenden Paradata-Variablen wurde das frei verfügbare Stata-Modul `parseuas` eingesetzt (Roßmann/Gummer 2020).

5.6. Substantielle Daten

Substantielle Daten bilden den Hauptbestandteil des Datensatzes. Sie umfassen inhaltliche und soziodemographische Variablen.

Tabelle 12: Inhaltliche und soziodemographische Variablen

VariableLabel	
t1026	Deutsche Staatsangehörigkeit
t1	Geschlecht
t2	Alter
t3	Schulabschluss
t4	Bundesland
t5	Politisches Interesse
t6	Demokratiezufriedenheit
t7	Wahlbeteiligungsabsicht
t8aa	Beabsichtigte Stimmabgabe (Erststimme, Version A)
t8ab	Beabsichtigte Stimmabgabe (Erststimme, Version B)
t8ba	Beabsichtigte Stimmabgabe (Zweitstimme, Version A)
t8bb	Beabsichtigte Stimmabgabe (Zweitstimme, Version B)
t9aa	Hypothetische Stimmabgabe (Erststimme, Version A)
t9ab	Hypothetische Stimmabgabe (Erststimme, Version B)
t9ba	Hypothetische Stimmabgabe (Zweitstimme, Version A)
t9bb	Hypothetische Stimmabgabe (Zweitstimme, Version B)

t10s	Wichtigstes Problem
t11a	Wichtigstes Problem, Lösungskompetenz (Version A)
t11b	Wichtigstes Problem, Lösungskompetenz (Version B)
t12s	Zweitwichtigstes Problem
t13a	Zweitwichtigstes Problem, Lösungskompetenz (Version A)
t13b	Zweitwichtigstes Problem, Lösungskompetenz (Version B)
t14a	Skalometer Parteien (CDU)
t14b	Skalometer Parteien (CSU)
t14c	Skalometer Parteien (SPD)
t14d	Skalometer Parteien (AfD)
t14e	Skalometer Parteien (FDP)
t14f	Skalometer Parteien (DIE LINKE)
t14g	Skalometer Parteien (GRÜNE)
t15a	Skalometer Politiker (Angela Merkel)
t15b	Skalometer Politiker (Andrea Nahles)
t15c	Skalometer Politiker (Horst Seehofer)
t15d	Skalometer Politiker (Alexander Gauland)
t15e	Skalometer Politiker (Christian Lindner)
t15f	Skalometer Politiker (Sahra Wagenknecht)
t15g	Skalometer Politiker (Katrín Göring-Eckardt)
t16	Leistung Bundesregierung
t17a	Leistung Regierungsparteien (CDU)
t17b	Leistung Regierungsparteien (CSU)
t17c	Leistung Regierungsparteien (SPD)
t18a	Leistung Oppositionsparteien (AfD)
t18b	Leistung Oppositionsparteien (FDP)
t18c	Leistung Oppositionsparteien (DIE LINKE)
t18d	Leistung Oppositionsparteien (GRÜNE)
t21	Eigene wirtschaftliche Lage, aktuell
t19	Eigene wirtschaftliche Lage, retrospektiv
t20	Verantwortlichkeit eigene wirtschaftliche Lage
t22	Eigene wirtschaftliche Lage, prospektiv
t23a	Links-Rechts-Einstufung Parteien (CDU)
t23b	Links-Rechts-Einstufung Parteien (CSU)
t23c	Links-Rechts-Einstufung Parteien (SPD)
t23d	Links-Rechts-Einstufung Parteien (AfD)
t23e	Links-Rechts-Einstufung Parteien (FDP)
t23f	Links-Rechts-Einstufung Parteien (DIE LINKE)
t23g	Links-Rechts-Einstufung Parteien (GRÜNE)
t24	Links-Rechts-Selbsteinstufung

t25	Unterschied Regierung
t26	Unterschiede Parteien
t29	Allgemeine wirtschaftliche Lage, aktuell
t27	Allgemeine wirtschaftliche Lage, retrospektiv
t28	Verantwortlichkeit allg. wirt. Lage Bundesregierung
t30	Allgemeine wirtschaftliche Lage, prospektiv
t145	Regionale wirtschaftliche Lage, aktuell
t31	Recall, Wahlberechtigung BTW 2017
t32	Recall, Wahlteilnahme BTW 2017
t33aa	Recall, Wahlentscheidung, BTW 2017 (Erststimme, Version A)
t33ab	Recall, Wahlentscheidung, BTW 2017 (Erststimme, Version B)
t33ba	Recall, Wahlentscheidung, BTW 2017 (Zweitstimme, Version A)
t33bb	Recall, Wahlentscheidung, BTW 2017 (Zweitstimme, Version B)
t225f	Emotionen: Angst, Intensität (Zuzug von Ausländern nach D)
t225a	Emotionen: Angst, Intensität (die derzeitige Wirtschaftslage)
t225c	Emotionen: Angst, Intensität (die globale Klimärwärmung)
t225d	Emotionen: Angst, Intensität (der internationale Terrorismus)
t225g	Emotionen: Angst, Intensität (der Zustand der Europäischen Union)
t225j	Emotionen: Angst, Intensität (die Globalisierung)
t225k	Emotionen: Angst, Intensität (Opfer eines Verbrechens zu werden)
t308a	Issuebatterie (Einwanderer verpflichten, sich der deutschen Kultur anzupassen)
t308b	Issuebatterie (Für Aufsichtsräte gr. Unternehmen gesetzl. Frauenquote)
t308c	Issuebatterie (Der Staat sollte sich aus der Wirtschaft heraushalten.)
t308d	Issuebatterie (Regierg: Maßn. ergreifen, um Einkommensuntersch. z verringern)
t308e	Issuebatterie (In Schuldenkrise soll D EU-MGstaaten finanziell unterstützen)
t308f	Issuebatterie (Bürger sollen Volksentscheid auf Bundesebene herbeiführen)
t308g	Issuebatterie (Energieversorgung sollte auch durch Atomkraft gesichert werden)
t308h	Issuebatterie (Ehe sollte für gleichgeschlechtliche Paare möglich sein)
t308i	Issuebatterie (Besserverdienende Bürger sollen mehr Steuern bezahlen)
t308j	Issuebatterie (Arbeitslose: gemeinnützige Arbeit für staatl. Unterstützung)
t308k	Issuebatterie (Menschen sollen sich stärker an Geboten Gottes orientieren)
t308l	Issuebatterie (Das Internet sollte stärker kontrolliert werden)
t308m	Issuebatterie (Die D-Mark sollte wieder eingeführt werden)
t308n	Issuebatterie (In D: wicht. pol. Themen v Medien u Politik einseitig diskutiert)
t45	Gespräche über Politik
t304	Netzwerkgröße
t444	Wahlbeteiligung, Nachbarschaft
t443	Wahlbeteiligung, Bekanntenkreis
t153a	Positionsissue: Libertär-autoritäre Dimension, Parteien (CDU)
t153b	Positionsissue: Libertär-autoritäre Dimension, Parteien (CSU)

t153c	Positionsissue: Libertär-autoritäre Dimension, Parteien (SPD)
t153d	Positionsissue: Libertär-autoritäre Dimension, Parteien (AfD)
t153e	Positionsissue: Libertär-autoritäre Dimension, Parteien (FDP)
t153f	Positionsissue: Libertär-autoritäre Dimension, Parteien (DIE LINKE)
t153g	Positionsissue: Libertär-autoritäre Dimension, Parteien (GRÜNE)
t154	Positionsissue: Libertär-autoritäre Dimension, Ego
t155	Positionsissue: Libertär-autoritäre Dimension, Salienz
t113a	Positionsissue: Sozioökonomische Dimension, Parteien (CDU)
t113b	Positionsissue: Sozioökonomische Dimension, Parteien (CSU)
t113c	Positionsissue: Sozioökonomische Dimension, Parteien (SPD)
t113d	Positionsissue: Sozioökonomische Dimension, Parteien (AfD)
t113e	Positionsissue: Sozioökonomische Dimension, Parteien (FDP)
t113f	Positionsissue: Sozioökonomische Dimension, Parteien (DIE LINKE)
t113g	Positionsissue: Sozioökonomische Dimension, Parteien (GRÜNE)
t114	Positionsissue: Sozioökonomische Dimension, Ego
t115	Positionsissue: Sozioökonomische Dimension, Salienz
t317a	Positionsissue: Klimawandel, Parteien (CDU)
t317b	Positionsissue: Klimawandel, Parteien (CSU)
t317c	Positionsissue: Klimawandel, Parteien (SPD)
t317d	Positionsissue: Klimawandel, Parteien (AfD)
t317e	Positionsissue: Klimawandel, Parteien (FDP)
t317f	Positionsissue: Klimawandel, Parteien (DIE LINKE)
t317g	Positionsissue: Klimawandel, Parteien (GRÜNE)
t318	Positionsissue: Klimawandel, Ego
t319	Positionsissue: Klimawandel, Salienz
t486	Positionsissue: Globalisierung
t36	Wichtigste Informationsquelle
t37	Nutzung Tageszeitung
t38	Printmedien, Nutzung, Wochenzeitschriften
t39a	Nutzung TV-Nachrichten (Tagesschau oder Tagesthemen)
t39b	Nutzung TV-Nachrichten (Heute oder das Heute Journal)
t39c	Nutzung TV-Nachrichten (RTL Aktuell)
t39d	Nutzung TV-Nachrichten (Sat.1 Nachrichten)
t39e	Nutzung TV-Nachrichten (Pro 7 Newstime)
t80	Internetnutzung, allgemein
t40	Internetnutzung, politische Information
t1022	Internetnutzung, News
t1023a	Allgemeine Nutzung von sozialen Medien (Facebook)
t1023b	Allgemeine Nutzung von sozialen Medien (Twitter)
t1023c	Allgemeine Nutzung von sozialen Medien (Instagram)

t1023d	Allgemeine Nutzung von sozialen Medien (Snapchat)
t1023e	Allgemeine Nutzung von sozialen Medien (WhatsApp)
t1023f	Allgemeine Nutzung von sozialen Medien (XING)
t1023g	Allgemeine Nutzung von sozialen Medien (Google+)
t1023h	Allgemeine Nutzung von sozialen Medien (LinkedIn)
t1023i	Allgemeine Nutzung von sozialen Medien (YouTube)
t1023j	Allgemeine Nutzung von sozialen Medien (Sonstige, und zwar)
t1023k	Allgemeine Nutzung von sozialen Medien (nutze keine sozialen Medien)
t1024a	Politische Nutzung von sozialen Medien (Facebook)
t1024b	Politische Nutzung von sozialen Medien (Twitter)
t1024c	Politische Nutzung von sozialen Medien (Instagram)
t1024d	Politische Nutzung von sozialen Medien (Snapchat)
t1024e	Politische Nutzung von sozialen Medien (WhatsApp)
t1024f	Politische Nutzung von sozialen Medien (XING)
t1024g	Politische Nutzung von sozialen Medien (Google+)
t1024h	Politische Nutzung von sozialen Medien (LinkedIn)
t1024i	Politische Nutzung von sozialen Medien (YouTube)
t1024j	Politische Nutzung von sozialen Medien (Sonstige, und zwar)
t1025a	Soziale Medien, politische Partizipation (Facebook)
t1025b	Soziale Medien, politische Partizipation (Twitter)
t1025c	Soziale Medien, politische Partizipation (Instagram)
t1025d	Soziale Medien, politische Partizipation (Snapchat)
t1025e	Soziale Medien, politische Partizipation (WhatsApp)
t1025f	Soziale Medien, politische Partizipation (XING)
t1025g	Soziale Medien, politische Partizipation (Google+)
t1025h	Soziale Medien, politische Partizipation (LinkedIn)
t1025i	Soziale Medien, politische Partizipation (YouTube)
t1025j	Soziale Medien, politische Partizipation (Sonstige, und zwar)
t1025k	Soziale Medien, politische Partizipation (gar keine Beiträge)
t160	Vertrauen
t156a	Einstellungen, Politik/Gesellschaft (Politische Fragen sind schwer zu verstehen)
t156b	Einstellungen, Politik/Gesellschaft (Parteien wollen nur Wählerstimmen)
t156c	Einstellungen, Politik/Gesellschaft (Politik kann Probleme nicht lösen)
t156d	Einstellungen, Politik/Gesellschaft (Sozialismus ist eine gute Idee)
t156e	Einstellungen, Politik/Gesellschaft (Welt ist zu unübersichtlich geworden)
t156f	Einstellungen, Politik/Gesellschaft (Wählen ist Bürgerpflicht)
t46a	Parteiidentifikation (Version A)
t46b	Parteiidentifikation (Version B)
t47	Parteiidentifikation, Stärke
t48	Parteiidentifikation, Dauer

t474a	Einstellung zu politischen Gruppen (Gewerkschaften)
t474b	Einstellung zu politischen Gruppen (Unternehmer-/Arbeitgeberverband)
t474c	Einstellung zu politischen Gruppen (Berufsvereinigung/-verband)
t474d	Einstellung zu politischen Gruppen (Bauern- bzw. Landwirtschaftsverband)
t474e	Einstellung zu politischen Gruppen (Religiöse/kirchliche Gruppen)
t474g	Einstellung zu politischen Gruppen (Umweltschutzgruppen)
t474h	Einstellung zu politischen Gruppen (Globalisierungskritische Gruppen,z.B.ATTAC)
t474k	Einstellung zu politischen Gruppen (Islamkritische Gruppen,z.B.PEGIDA)
t474i	Einstellung zu politischen Gruppen (Frauenorganisationen)
t473a	Organisationsmitgliedschaft (Gewerkschaft)
t473b	Organisationsmitgliedschaft (Unternehmer-/Arbeitgeberverband)
t473c	Organisationsmitgliedschaft (Berufsvereinigung/-verband)
t473d	Organisationsmitgliedschaft (Bauern- bzw. Landwirtschaftsverband)
t473e	Organisationsmitgliedschaft (Religiöse/kirchliche Gruppen)
t473f	Organisationsmitgliedschaft (Sport-/Hobbyverein)
t473g	Organisationsmitgliedschaft (Umweltschutzgruppe)
t473h	Organisationsmitgliedschaft (Globalisierungskritische Gruppe,z.B.ATTAC)
t473k	Organisationsmitgliedschaft (Islamkritische Gruppe,z.B.PEGIDA)
t473i	Organisationsmitgliedschaft (Frauenorganisation)
t473j	Organisationsmitgliedschaft (Partei, und zwar)
t473jc1	Organisationsmitgliedschaft (Partei, Codierung)
t57	Familienstand
t50	Haushaltsgröße
t348a	Haushaltsmitglieder, Alter (2. Person)
t348b	Haushaltsmitglieder, Alter (3. Person)
t348c	Haushaltsmitglieder, Alter (4. Person)
t348d	Haushaltsmitglieder, Alter (5. Person)
t348e	Haushaltsmitglieder, Alter (6. Person)
t348f	Haushaltsmitglieder, Alter (7. Person)
t348g	Haushaltsmitglieder, Alter (8. Person)
t303a	Berufliche Bildung (berufl.-betriebl. Anlernzeit mit Zeugnis/keine Lehre)
t303b	Berufliche Bildung (Teilfacharbeiterabschluss)
t303c	Berufliche Bildung (abgeschlossene gewerbliche oder landwirtschaftliche Lehre)
t303d	Berufliche Bildung (abgeschlossene kaufmännische Lehre)
t303e	Berufliche Bildung (berufliches Praktikum, Volontariat)
t303f	Berufliche Bildung (Berufsfachschulabschluss)
t303g	Berufliche Bildung (Fachschulabschluss)
t303h	Berufliche Bildung (Meister, Technikerabschluss)
t303i	Berufliche Bildung (Fachhochschulabschluss)
t303j	Berufliche Bildung (Hochschulabschluss)

t303k	Berufliche Bildung (andere beruflichen Ausbildungsabschluss)
t303ks	Berufliche Bildung (andere beruflichen Ausbildungsabschluss, Nennung)
t303l	Berufliche Bildung (noch in beruflicher Ausbildung)
t303m	Berufliche Bildung (keine abgeschlossene Ausbildung)
t52	Erwerbstätigkeit
t53	Angst Stellenverlust
t54	Erwerbstätigkeit, früher
t55	Beruf
t56	Beruf, früher
t351	Beruf, Nachfrage I
t352	Beruf, Nachfrage II
t353	Beruf, Nachfrage III
t354	Beruf, Nachfrage IV
t1012	Berufsfeld
t58	Beschäftigungssektor
t59	Beschäftigungssektor, früher
t60	Religionszugehörigkeit
t320	Häufigkeit Gottesdienst
t64	Deutsche Staatsbürgerschaft Geburt
t65	Deutsche Staatsbürgerschaft, Dauer
t66	Geburtsland (V2)
t72	Geburtsland, Bundesland
t73	Geburtsland, Ausland
t68	Geburtsland, Mutter (V2)
t67	Geburtsland, Vater (V2)
t454	Erfassung Migration in 3. Generation
t458a	Geburtsland, Großeltern (Die Mutter Ihrer Mutter)
t458b	Geburtsland, Großeltern (Der Vater Ihrer Mutter)
t458c	Geburtsland, Großeltern (Die Mutter Ihres Vaters)
t458d	Geburtsland, Großeltern (Der Vater Ihres Vaters)
t459	Geburtsland, Großmutter mütterlicherseits
t460	Geburtsland, Großvater mütterlicherseits
t461	Geburtsland, Großmutter väterlicherseits
t462	Geburtsland, Großvater väterlicherseits
t69	Schichtzugehörigkeit, subjektiv
t70	Nettoeinkommen HH
v_1	Gerätetyp
v_3s	Hersteller des Geräts

5.6.1 Codierung von offenen und halboffenen Fragen

Die offenen und halboffenen Fragen werden als String-Variablen bereitgestellt. Hierfür werden die offen erhobenen Antworten intensiv auf datenschutzrechtlich bedenkliche Nennungen geprüft und nötigenfalls zum Schutz der Befragten zensiert.

Für die Arbeit mit den offen erhobenen Antworten empfehlen wir die Nutzung des SPSS-Datensatzes, da aufgrund von Programmrestriktionen im Stata-Datensatz alle offenen Angaben bei 244 Zeichen abgeschnitten werden.

5.6.2 Codierung der Parteien

Die Codierung der Parteien erfolgte bei allen geschlossenen, halboffenen und offenen Fragen nach dem einheitlichen Codierschema der GLES. Das Codierschema kann auf den Seiten der GLES bei GESIS abgerufen werden (<http://www.gesis.org/gles>).

In der Regel werden die Angaben der Befragten zu Parteien in zwei Variablen veröffentlicht, die mit Version A und Version B bezeichnet werden. Version A enthält für Parteienfragen stets die Parteien, die auf dem verwendeten Musterstimmzettel aufgeführt sind, sowie eine Kategorie für „andere“ Parteien. Version B unterscheidet sich von Version A ausschließlich durch eine detaillierte Ausweisung der von den Befragten genannten „anderen“ Parteien.

5.7. Kontextvariablen

Kontextvariablen geben Auskünfte über den regionalen Kontext der Befragten, z.B. den Wahlkreis und die Gemeinde, in welchen Befragte leben.

Tabelle 13: Kontextvariablen

Variable	Label
ostwest	Ost/West (Zuordnung von Berlinern zu Ostdeutschland)
elecdist17_	Wahlkreis
elecdist17_1	Nicht-eindeutige Zuordnung: 1. Möglichkeit
elecdist17_2	Nicht-eindeutige Zuordnung: 2. Möglichkeit
elecdist17_3	Nicht-eindeutige Zuordnung: 3. Möglichkeit
elecdist17_4	Nicht-eindeutige Zuordnung: 4. Möglichkeit

5.7.1 Ost/West

Die Zuordnung zu Ost- und Westdeutschland erfolgte auf Grundlage der Angabe zum Bundesland, in dem die Befragten ihren Hauptwohnsitz haben. Da bei der Abfrage des Bundeslands keine Differenzierung zwischen Ost- und Westberlin vorgenommen wurde, werden Berliner zu Ostdeutschland zugeordnet.

5.7.2 Wahlkreise

Die Zuordnung zu einem Wahlkreis beruht auf der Wahlkreiseinteilung zur Bundestagswahl 2017 und wurde unter Verwendung der von den Befragten genannten Postleitzahl an ihrem Hauptwohnsitz vorgenommen. Die Zuordnung der Postleitzahlen zu Wahlkreisen wurde mittels des frei verfügbaren Stata-Ados `plztowknr` (Glinitzer et al. 2018) vorgenommen.

Mit Hilfe des Stata-Moduls `plztowknr` kann eine Zuordnung von Postleitzahlen zu Wahlkreisen gemäß den Wahlkreiseinteilungen bei den Bundestagswahlen 1998, 2002, 2005, 2009, 2013 oder 2017 vorgenommen werden. Da in der Bundesrepublik Deutschland keine offiziellen Listen für die Zuordnung von Postleitzahlen zu Wahlkreisen verfügbar sind, beru-

hen die Zuordnungstabellen auf verschiedenen administrativen Datenbeständen. Die Zuordnungstabellen für die Wahlkreiseinteilung zu den Bundestagswahlen in den Jahren 2009, 2013 und 2017 wurden im Rahmen der Datenaufbereitung der GLES erstellt. Sie beruhen auf dem vom Statistischen Bundesamt (www.destatis.de) herausgegebenen Gemeindeverzeichnis sowie der Wahlkreissuche des Deutschen Bundestages (www.bundestag.de/abgeordnete/wahlkreise/).

Bei der Verwendung der Wahlkreisinformationen ist zu beachten, dass einzelne Postleitzahlen in großstädtischen Gebieten zwischen zwei und vier Wahlkreisen zugeordnet werden können, d.h. eine eindeutige Zuordnung der Postleitzahl zu einem Wahlkreis ist auf Grundlage der vorhandenen Informationen nicht immer möglich. Die Variable *elecldist17_* „Wahlkreis“ enthält die Angabe zum Wahlkreis des Befragten, wenn eine eindeutige Zuordnung möglich ist. Ist dies nicht der Fall, so sind in den Variablen *elecldist17_1-elecldist17_4* „Wahlkreis nicht eindeutig zuordenbar: 1.-4. Möglichkeit“ die Nummern der Wahlkreise erfasst, denen der Befragte anhand der von ihm genannten Postleitzahl potenziell zugeordnet werden kann. In einzelnen Fällen ist eine Zuordnung zu einem oder mehreren Wahlkreisen nicht möglich, da eine Angaben zur Postleitzahl fehlt oder die vom Befragten genannte Postleitzahl nicht existiert.

Tabelle 14: Zuordnung von Postleitzahlen zu Wahlkreisen

Art der Zuordnung	N	Prozent
Eindeutig einem Wahlkreis zuordenbar	1.035	93,58
Nicht eindeutig einem Wahlkreis zuordenbar	53	4,79
Fehlende Angabe oder nicht zuordenbare Postleitzahl	18	1,63
Gesamt	1.106	100,0

5.8. GewichtungsvARIABLEN

Mit Hilfe von Anpassungsgewichten kann die Verteilung von 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. Die für die GLES-Daten erstellten Gewichte werden mittels einer stets gleichen Vorgehensweise berechnet, so dass sie über die verschiedenen GLES-Komponenten hinweg vergleichbar sind.

Zur Berechnung der Anpassungsgewichte der GLES wurden sowohl sozial- als auch regionalstrukturelle Merkmale ausgewählt. Bei der gleichzeitigen Anpassung an mehrere Merkmale ergeben sich häufig Nullzellen oder sehr kleine Zellbesetzungen, was bei einer einfachen Soll/Ist-Gewichtung zu Problemen führt. Daher wurde bei der Berechnung der Anpassungsgewichte für die GLES auf das iterative Anpassungsverfahren („iterative proportional fitting“, IPF)³ zurückgegriffen. 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 unter-

³ Siehe zum IPF-Verfahren Deming und Stephan (1940).

schreitet.⁴ Um sehr große Gewichtungsfaktoren zu vermeiden wurden die Faktoren (nach jedem Iterationsschritt) auf den fünffachen Mittelwert der Gewichtungsvariable (also fünf) getrimmt.⁵

Bei der Erstellung der Anpassungsgewichte für das Tracking wurde auf die aktuellsten verfügbaren Randverteilungen des Mikrozensus⁶ und des (N)Onliner Atlas⁷ zurückgegriffen. Bei der Erstellung der Sollverteilungen des Mikrozensus ging nur die Bevölkerung über 18 Jahren mit deutscher Staatsbürgerschaft in Privathaushalten am Ort der Hauptwohnung ein. Die Sollverteilungen des (N)Onliner Atlas beruhen auf den Angaben zu Personen mit deutscher Staatsbürgerschaft ab einem Alter von 18 Jahren.

Für diese Studie wurden insgesamt vier Anpassungsgewichte erstellt, wobei die Gewichte einmal mit und einmal ohne Zeitunterschreiter berechnet wurden.

Tabelle 16: Übersicht über die sozial- und regionalstrukturellen Gewichtungsvariablen

Gewichtungsvariable	Anpassung an Randverteilungen des...	Berücksichtigung von Zeitunterschreibern
wei_mzz	Mikrozensus 2017	ja
wei_mzoz	Mikrozensus 2017	nein
wei_onz	(N)Onliner Atlas 2016	ja
wei_onoz	(N)Onliner Atlas 2016	nein

Angepasst wurde an soziodemographische und regionalstrukturelle Merkmale: Geschlecht, Alter, Bildung und Zugehörigkeit zu West- bzw. Ostdeutschland (inkl. Berlin).

Die Variable Alter wurde kategorisiert und umfasst nun vier Gruppen: „18 bis unter 30 Jahre“, „30 bis unter 45 Jahre“, „45 bis unter 60 Jahre“ und „60 Jahre und älter“.

Bei der Kategorisierung der Variable Bildung wurden die drei folgenden Gruppen gebildet:

- niedrige Bildung: Schule beendet ohne Abschluss, Hauptschulabschluss, Volksschulabschluss, bin noch Schüler/in;
- 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).

Fehlende Werte wurden für die Erstellung der Gewichtungsvariablen auf den Modalwert der Verteilung gesetzt.

⁴ Die Anpassungsgewichte wurden mit dem Statistikprogramm Stata berechnet, wobei auf das frei verfügbare Ado „ipfweight“ (Bergmann 2011) zurückgegriffen wurde. Zur Installation muss lediglich „ssc install ipfweight“ in die Kommandozeile von Stata eingegeben werden.

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

⁶ Bei den vom Mikrozensus ausgewiesenen Werten handelt es sich um absolute Personenzahlen nach gebundener Hochrechnung.

⁷ Die Randverteilungen (für Personen ab 18 Jahren) aus dem (N)Onliner Atlas wurden für die Gewichtung freundlicherweise von TNS Infratest zur Verfügung gestellt.

Tabelle 17: Soll- und Ist-Verteilungen der zur Erstellung der Gewichte verwendeten Variablen

Merkmal	Ist in Prozent		Soll in Prozent	
	Mit Zeitunter- schreibern (N=1.106)	Ohne Zeitun- terschreiber (N=994)	Mikrozensus	(N)Onliner Atlas
Geschlecht				
Männlich	51,72	51,11	48,65	51,6
Weiblich	48,28	48,89	51,35	48,4
Altersgruppe, 4 Kat.				
18 bis unter 30 Jahre	21,43	19,62	15,68	21,0
30 bis unter 45 Jahre	28,12	26,66	20,65	25,2
45 bis unter 60 Jahre	34,36	35,92	28,88	33,8
60 Jahre und älter	16,09	17,81	34,79	20,0
Bildung				
Niedrig	27,40	28,47	36,67	35,5
Mittel	34,99	35,51	30,67	32,1
Hoch	37,61	36,02	32,65	32,4
Bundesland				
Alte Bundesländer	78,66	78,67	79,21	80,7
Neue Bundesländer (inkl. Berlin)	21,34	21,33	20,79	19,3

Die Tabelle 18 gibt die Anzahl der zur Berechnung notwendigen Iterationen sowie deskriptive Statistiken zu den resultierenden Gewichtungsfaktoren wieder.

Tabelle 18: Deskriptive Statistiken zu den Gewichtungsfaktoren

Statistik	wei_mzz	wei_mzoz	wei_onz	wei_onoz
Anzahl Iterationen ⁸	4	4	4	4
Mittelwert	1,00	1,00	1,00	1,00
Standardabweichung	0,53	0,46	0,20	0,18
Minimum	0,60	0,63	0,75	0,76
Maximum	2,63	2,33	1,50	1,44
1. Quartil	0,71	0,74	0,84	0,86
Median	0,79	0,82	0,92	0,94
3. Quartil	0,89	0,95	1,18	1,17
Max./Min. ⁹	4,38	3,70	3,64	1,90
N	1.106	994	1.106	994

5.9. Unveröffentlichte Variablen

Aus datenschutzrechtlichen Gründen können nicht alle erhobenen Variablen zum freien Download zur Verfügung gestellt werden. Generell müssen frei verfügbare Datensätze so aufgebaut sein, dass ein unverhältnismäßiger Aufwand zur De-Anonymisierung betrieben werden müsste. Das heißt, die Datensätze müssen faktisch anonym sein. Um dies zu gewährleisten, wurden einige wenige Variablen aus den Datensätzen entfernt bzw. Ausprägungen zusammengefasst. Die Tabelle 19 gibt eine Übersicht über die unveröffentlichten Variablen in dieser Studie.

Tabelle 19: Übersicht über unveröffentlichte Variablen

Variable	Label
browserversion	Browser version
operatingsystem	Operating system
useragent	Useragent
t1023js	Allgemeine Nutzung von sozialen Medien (Sonstige, Nennung)
t73s	Geburtsland, Ausland (anderes Land)
t68s	Geburtsland, Mutter (V2) (anderes Land)
t67s	Geburtsland, Vater (V2) (anderes Land)
t459s	Geburtsland, Großmutter mütterlicherseits (anderes Land)
t460s	Geburtsland, Großvater mütterlicherseits (anderes Land)
t461s	Geburtsland, Großmutter väterlicherseits (anderes Land)
t462s	Geburtsland, Großvater väterlicherseits (anderes Land)
t348h_1-t348h_7	Haushaltsmitglieder, Alter (9.-15. Person)
t71	Postleitzahl
v_3	Hersteller und Bezeichnung des Geräts

⁸ Die Anzahl der Iterationen gibt an, wie viele Anpassungsschritte bei der Gewichtung vorgenommen wurden bis die Abweichung zwischen den gewichteten Istwerten in der Stichprobe und den Sollwerten aus der Referenzverteilung weniger als 0,05 betrug. Bei Erreichen von 20 Iterationsschritten wird die Anpassung ebenfalls beendet.

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.

Diese Variablen stehen interessierten Nutzern zur Verfügung, je nach Variable in einem Secure Data Center (SDC) bei GESIS (Köln, Mannheim) oder nach Abschluss eines Nutzungsvertrags. Bei Interesse melden Sie sich bitte unter gles@gesis.org.

Weiterhin werden Systemvariablen aus dem veröffentlichten Datensatz entfernt, die lediglich zur Durchführung der Online-Befragung relevant sind, aber keine inhaltliche Aussagekraft haben. Darüber hinaus werden offen erhobene Antworten im Normalfall nur als vercodete Variablen veröffentlicht. Bei Interesse an den entfernten Systemvariablen oder den nicht vercodeten offenen Antworten, melden Sie sich bitte ebenfalls unter gles@gesis.org.

5.10. Fehlende Werte

Fehlende Werte wurden bei allen Variablen nach dem einheitlichen Codierschema der GLES vergeben, welches negative Werte im Bereich -71 bis -99 und in Stata Codes von .a bis .p vorsieht. In den SPSS-Datensätzen sind diese Werte standardmäßig als fehlende Werte definiert. Den Stata-Datensätzen hingegen liegt ein Do-File bei, anhand dessen die fehlenden Werte in die vorgesehenen Missing-Value-Codes recodiert werden können.

Tabelle 20: Übersicht über die Missing-Value-Codes der GLES

Code	Stata Missing-Value-Code	Label
-99	.a	keine Angabe
-98	.b	weiß nicht
-97	.c	trifft nicht zu
-96	.d	Split
-95	.e	nicht teilgenommen
-94	.f	nicht in Auswahlgesamtheit
-93	.g	Interview abgebrochen
-92	.h	Fehler in Daten
-86	.i	nicht wahlberechtigt
-85	.j	nicht wählen
-84	.k	keine Erst-/Zweitstimme abgeben
-83	.l	ungültig wählen
-82	.m	keine andere Partei wählen
-81	.n	noch nicht entschieden
-72	.o	nicht einzuschätzen
-71	.p	nicht bekannt

5.11. Analyse-Systeme

Der Datensatz ist für die Analyse-Systeme SPSS (IBM) und Stata (StataCorp LP) verfügbar.

Tabelle 21: Analysesysteme

Analyse-System	Dateiformat	Versionsbeschränkungen
SPSS	.sav	Version 18 oder höher
Stata	.dta	Version 12 oder höher

6. Hinweise und Anmerkungen

6.1. Errata

- t18*: Aufgrund eines Programmierfehlers wurde Code 8 „+2“ nicht als Code 8 abgespeichert sondern als -99 „keine Angabe“. Aus diesem Grund kann man bei diesen Variablen nicht mehr unterscheiden, ob die jeweiligen Befragten „+2“ oder „keine Angabe“ ausgewählt haben und wurden auf -92 „Fehler in Daten“ recodiert. In der gesamten Batterie fehlen deswegen die Codes 8 und -99.

6.2. Versionshistorie

Änderungen zwischen Version 1.0.0 und Version 2.0.0:

- Mehrfachteilnahmen wurden entfernt (siehe 4.7 Feldverlauf)
- Die Zeitvariablen wurden neu berechnet (siehe 5.5.3 Zeitvariablen), die Variable `duration_client` wurde hinzugefügt und die Variable `duration` wurde in `duration_server` umbenannt. Infolge der Neuberechnung der Zeitvariablen hat sich auch die Verteilung der Zeitunterschreiter:innen verändert (siehe 5.5.4 Zeitunterschreiter).
- Die Variablen `intstatus` und `intstatus_prc` mit Informationen zum Interviewstatus der Befragten wurden hinzugefügt.
- Die Variablen `field`, `year` und `survey` wurden durch die Variablen `field_start`, `field_end` und `sample` ersetzt.
- Korrektur der Wertelabels der Wahlkreisvariablen (`elecdist17_*`)
- Das Stata-Modul „`parseuas`“ zur Auslesung des „user agent header“ wurde aktualisiert. Dadurch verändert sich auch die Kodierung des genutzten Browsers und des genutzten Gerätes in den Variablen `browser` und `device`.

Eine aktuelle Errataliste kann über die GESIS-Suche (https://search.gesis.org/research_data/ZA6823) abgerufen werden. Dort werden auch alle Änderungen zwischen den einzelnen Versionen der Datensätze dokumentiert.

Links

Projektseite der GLES: www.gles.eu

Deutsche Gesellschaft für Wahlforschung e. V. (DGfW): www.dgfw.info

GESIS – Leibniz-Institut für Sozialwissenschaften: www.gesis.org/gles

Universität Frankfurt: www.uni-frankfurt.de

Universität Mannheim: www.uni-mannheim.de

Wissenschaftszentrum Berlin für Sozialforschung gGmbH (WZB): www.wzb.eu

Statistisches Bundesamt: www.destatis.de

Wahlkreissuche des Deutschen Bundestages: www.bundestag.de/abgeordnete/wahlkreise/

Literaturverzeichnis

- AAPOR (2011): Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 7. Aufl. The American Association for Public Opinion Research.
- Bergmann, Michael (2011): IPFWEIGHT: Stata module to create adjustment weights for surveys. Statistical Software Components S457353, Boston: Boston College Department of Economics.
- BIK Aschpurwis+Behrens (2001): BIK Regionen: Ballungsräume, Stadtregionen, Mittel-/Unterzentrengebiete. Methodenbeschreibung zur Aktualisierung 2000, www.bik-gmbh.de/texte/BIK-Regionen2000.pdf [05.11.2011].
- Callegaro, Mario (2013): Paradata in Web Surveys, in: Kreuter, Frauke (Hrsg.), Improving Surveys with Paradata: Analytic Uses of Process Information. (Aufl.) Hoboken, NJ: Wiley, 261-280.
- Callegaro, Mario und Charles Disogra (2008): Computing Response Metrics for Online Panels, in: Public Opinion Quarterly 72, 1008-1032.
- Couper, Mick P. (2000): Usability Evaluation of Computer-Assisted Survey Instruments, in: Social Science Computer Review 18, 384-396.
- 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 und 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: 427-444.
- Glinitzer, Konstantin, Tobias Gummer, Malte Kaukal und Joss Roßmann (2018): plztowknr: Stata module to translate German zip codes into electoral districts (Version: 1.0) [Computer Software]. Chestnut Hill, MA: Boston College.
- Kaukal, Malte (2016): SCANDATA: Stata module to scan a dataset for specified characteristics. Available from <http://EconPapers.repec.org/RePEc:boc:bocode:s458186> [27.11.2018].
- Krosnick, Jon A. (1991): Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys, in: Applied Cognitive Psychology 5, 213-236.
- Mayerl, Jochen und Dieter Urban (2008): Antwortreaktionszeiten in Survey-Analysen. Messung, Auswertung und Anwendung. Aufl. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Roßmann, Joss (2010): Data Quality in Web Surveys of the German Longitudinal Election Study 2009. 3rd ECPR Graduate Conference. Dublin.
- Roßmann, Joss (2015): SPEEDERGLES: Stata module for the computation of the GLES response speed index (Version: 1.0) [Computer Software]. Chestnut Hill, MA: Boston College.
- Roßmann, Joss und Tobias Gummer (2020): PARSEUAS: Stata module to extract detailed information from user agent strings (Version: 1.4) [Computer Software]. Chestnut Hill, MA: Boston College.
- Schmitt-Beck, Rüdiger, Hans Rattinger, Sigrid Roßteutscher und Bernhard Weißels (2010): Die deutsche Wahlforschung und die German Longitudinal Election Study (GLES), in: Faulbaum, Frank & Wolf, Christof (Hrsg.), Gesellschaftliche Entwicklungen im Spiegel der empirischen Sozialforschung. Wiesbaden: VS Verlag für Sozialwissenschaften, 141-172.
- Schlosser, S. (2016). ECSP - Embedded Client Side Paradata. Zenodo. Retrieved from <https://doi.org/10.5281/zenodo.55169>

Schlosser, Stephan und Jan Höhne (2018): Embedded Client Side Paradata.
https://zenodo.org/record/1218941#.WuNKNpdCSM_ [27.11.2018].