

GERMAN LONGITUDINAL ELECTION STUDY (GLES)

ZA5301 (Version 4.0.0)

Komponente 1: Nachwahl-Querschnitt

Änderungen und Korrekturen von Version 3.0.0 zu Version 4.0.0

Von Version 3.0.0 zu Version 4.0.0 des Datensatzes von Komponente 1: Nachwahl-Querschnitt (ZA5301) wurden die folgenden Änderungen/Korrekturen vorgenommen:

(1) Im Datensatz wurden in allen Variablen- und Wertelabels sämtliche Umlaute (ä, ö, ü) sowie sämtliche "ß" durch "ae", "oe", "ue" bzw. "ss" ersetzt. Dies soll die problemlose internationale Nutzbarkeit des Datensatzes sicherstellen.

(2) Im Bestreben die Datensätze der GLES hinsichtlich der Codierungen zu harmonisieren, wurden die fehlenden Werte im vorliegenden Datensatz teilweise recodiert. Die Antwort "weiß nicht" wird im Allgemeinen mit dem Code (98), (998) oder (9998) codiert. Die Antwortverweigerung oder eine fehlende Angabe ("keine Angabe") wird mit dem Code (99), (999) bzw. (9999) codiert. Wenn eine Frage auf Grund einer Filterführung nicht auf einen Befragten zutraf ("trifft nicht zu"), so wird dies mit dem Code (100), (1000) oder (10000) codiert.

(3) Zur Sicherstellung der Nutzbarkeit der Datensätze der GLES mit möglichst vielen Analyseprogrammen wurde die maximale Länge von Variablenlabels auf 80 Zeichen und die maximale Länge von Wertelabels auf 120 Zeichen beschränkt. Bei einzelnen Variablen kam es vor, dass diese Beschränkungen übertroffen wurden. In diesen Fällen wurden die Variablen- und Wertelabels so gekürzt, dass die inhaltliche Bedeutung bestehen bleibt.

(4) Im Sinne der Harmonisierung der Datensätze der GLES wurden einige Variablen, bei denen nach Parteien gefragt wurde, mittels eines einheitlichen Codierschemas codiert. Dies betrifft etwa Fragen nach der Wahlabsicht, nach der Wahlentscheidung, nach der Identifikation mit einer Partei, der Nähe zu bestimmten Parteien usw. Das Codierschema "Parteien" kann über die Internet-Seite der GLES bei GESIS unter www.gesis.org/gles bezogen werden.

(5) Für die vorliegende Version des Datensatzes wurden mehrere Variablen mit offenen Nennungen codiert, die bei den vorherigen Versionen noch nicht berücksichtigt werden konnten. Oft betrifft dies Fragen nach "anderen" Parteien. Wenn möglich, wurden diese Codierungen nicht separat in Variablen ausgewiesen, sondern mit den gestützt abgefragten Nennungen der zugehörigen Frage in einer Variable zusammengefasst.

(6) Bei einigen Fragen der vorliegenden Studie wurden die Antworten der Befragten in mehreren Schritten erfasst. So wurden etwa Nennungen der im Bundestag vertretenen Parteien sowie Nennungen von "anderen" Parteien zumeist in separaten Variablen erfasst. Um den Nutzern der Daten die Arbeit mit dem Datensatz zu erleichtern, wurden die Antworten auf solche Fragen in der Regel in zwei Variablen zusammengefasst. Die erste Variable, die zumeist mit "Version A" im Variablenlabel gekennzeichnet ist, umfasst z.B. die Nennungen der im Bundestag vertretenen Parteien sowie eine "andere"-Kategorie, in der sämtliche Nennungen weiterer Parteien zusammengefasst sind. Die zweite Variable, die als "Version B" gekennzeichnet ist, schlüsselt zudem auf, welche anderen Parteien ge-

nannt wurden. Dieses Vorgehen wurde auch bei anderen Fragen gewählt, bei denen nicht nach Parteien gefragt wurde. Die erste Variable, "Version A", enthält dann in der Regel die vorgegebenen Antwortmöglichkeiten zuzüglich einer "andere"-Kategorie, während die zweite Variable, "Version B", die weiteren Nennungen aufgeschlüsselt wiedergibt. Lagen gute Gründe dafür vor, auf die Erstellung von zwei zusammenfassenden Variablen zu verzichten, so wurden alle vorgegebenen sowie alle weiteren Nennungen in einer Variable zusammengefasst. Auf eine spezielle Kennzeichnung wurde in diesem Falle verzichtet.

(7) Dem Datensatz wurden mit der Version 4.0.0 insgesamt neun neue Gewichte hinzugefügt. Neben einem Ost/West-Gewicht, einem Transformationsgewicht und einem kombinierten Transformations- und Ost/West-Gewicht stehen dem Nutzer insgesamt sechs verschiedene sozial- und regionalstrukturelle Anpassungsgewichte für Ost-, West- sowie gesamtdeutsche Analysen mit und ohne Berücksichtigung des Transformationsgewichts zur Verfügung. Eine detaillierte Beschreibung zu allen verfügbaren Gewichten sowie deren Erstellung findet sich in der Studienbeschreibung zu dieser Studie.

(8) Es wurde eine neue Variable *ostwest* ("Ost/West (Zuordnung von Berlinern zu Ostdeutschland)") erstellt, die angibt, ob Befragte in Ost- oder Westdeutschland wohnen, wobei Berlin vollständig zu Berlin gezählt wird. Diese Variable wurde auch als Grundlage für die Erstellung des Ost/West-Gewichts verwendet. Weiterhin wurde die Variable *quelle* in *ostwest2* ("Ost/West (genaue Zuordnung von Berlinern zu Ost- und Westdeutschland)") umbenannt. Letztere weist ost- und westdeutsche Befragte genau aus, d.h. innerhalb Berlins wurde nach (ehemaliger) Zugehörigkeit zur Deutschen Demokratischen Republik (DDR) und der Bundesrepublik Deutschland (BRD) unterschieden.

(9) In einigen, wenigen Fällen kam es in unterschiedlichen Kombinationen vor, dass der Befragte keine valide Angabe zur Frage nach dem wichtigsten, zweitwichtigsten oder drittwichtigsten Problem in Deutschland (*Q5*, *Q6*, *Q7*) oder die Antwort verweigerte und dennoch die Nachfragen zu dem zweitwichtigsten bzw. drittwichtigsten Problem und/oder der Lösungskompetenz für diese Probleme (*Q8*, *Q9*, *Q10*) bekam. Diese fehlerhaft gefilterten Befragten wurden bei den entsprechenden Nachfragen auf den Code "trifft nicht zu" gesetzt.

(10) Die Nennungen der Befragten zu den Fragen *D251* ("Geburtsland"), *D253* ("Geburtsland Partner/in"), *D255* ("Geburtsland Mutter") sowie *D256* ("Geburtsland Vater") wurden entsprechend des einheitlichen Codierschemas "Staaten" recodiert. Weiterhin wurden alle offenen Nennungen der Befragten codiert und mit den vorgegebenen Nennungen in einer Variable zusammengefasst. Das Codierschema "Staaten" kann über die Internet-Seite der GLES bei GESIS unter www.gesis.org/gles bezogen werden.

(11) Auf Grund datenschutzrechtlicher Erwägungen wurde die Angabe des Sample Points (point), in dem ein Interview realisiert wurde, anonymisiert und als Variable *vpoint* ("Virtueller Sample Point") dem Datensatz hinzugefügt.

(12) Die Variablenlabels der Fragen *Q106*, *Q109*, *Q112* und *Q115* waren missverständlich. Die Variablen wurde umgelabelt in "Nutzungshäufigkeit TV-Nachrichten Sendername".

(13) Die Variablen *kand1* bis *kand5* enthielten zum Teil die Namen der falschen Wahlkreiskandidaten. Es wurden daher auf Basis der offiziellen Dokumente des Bundeswahlleiters die Variablen *wkkunion*, *wkkspd*, *wkkgruen*, *wkkfdp* und *wkklinke* generiert, die die Namen der Wahlkreiskandidaten in den Wahlkreisen bei der Bundestagswahl 2009 nach Parteien untergliedert enthalten.

(14) Bei der Frage *D244* ("Religionszugehörigkeit") wurden auf Grund datenschutzrechtlicher Erwägungen die Antwortmöglichkeiten "dem Judentum", "dem Islam" sowie "einer anderen Glaubensge-

meinschaft" zu "andere Glaubensgemeinschaft" zusammengefasst. Die weiteren Nennungen wurden entsprechend des Codierschemas "Religionszugehörigkeit" recodiert. Dieses kann über die Internet-Seite der GLES bei GESIS unter www.gesis.org/gles abgerufen werden.

(15) Auf Grund datenschutzrechtlicher Erwägungen wurden die Variablen *D245* "Kirchgangshäufigkeit", *D246* "Häufigkeit Synagoge" sowie *D247* "Häufigkeit Moschee" zu der Variable *dkirchg* "Häufigkeit Besuch Kirche/Moschee" zusammengefasst.

(16) Das Variablenlabel der Variable *D249* ("Deutsche Staatsbürgerschaft") war missverständlich, da in dieser Frage nicht danach gefragt wird, ob der Befragte die deutsche Staatsbürgerschaft besitzt, sondern ob er sie schon seit der Geburt besitzt. Das Variablenlabel wurde folglich in "Deutsche Staatsbürgerschaft seit Geburt" geändert.

(17) Auf Grundlage des Geburtsdatums (*Q2b*) wurde das Alter in Jahren der Befragten zum Tag des Interviews berechnet (*Q2ad*).

(18) Auf Grundlage des Alters der Befragten zum Zeitpunkt des Interviews wurde die Variable *ragelt18* ("Befragte, die jünger als 18 Jahre sind (Quelle: *Q2ad*)") erzeugt, die Befragte unter 18 Jahre im Datensatz markiert.

(19) Basierend auf den Kategorien der Variable *D241b* ("Nettoeinkommen HH, mit Kategorien") wurde die Variable *D241* ("Nettoeinkommen HH, kategorisiert") erstellt, die die Angaben aus den Variablen *D241a* und *D241b* zusammenfasst.

(20) Die Labelung der Variablen *Q48m1-19* "nicht wählbare Partei" wurde geändert in "Welche Partei nicht wählbar".

(21) Bei den Variablen *Q118i* und *Q118j* ("Nutzung Tageszeitung") waren die Variablenlabels im Datensatz vertauscht. Sie wurden in *Q118i* "Nutzung Tageszeitung: Tageszeitung, die nur im Internet erscheint" sowie *Q118j* "Nutzung Tageszeitung: eine oder mehrere andere Tageszeitungen" geändert.

(22) Bei der Variable *Q51a* ("Recall vorangegangene BTW Erststimme") war das Wertelabel für den Wert 4 fehlerhaft. Dieser Fehler wurde im Zuge der Vereinheitlichung der Codierungen korrigiert.

(23) Bei der Frage *D257* ("Im HH gesprochene Sprache") hat in einem Fall ein Interviewer eine offene Nennung notiert, die jedoch als Precode vorgegeben war. Die entsprechende Nennung wurde im Nachhinein so behandelt, als hätte der Interviewer den entsprechenden Precode ausgewählt.

(24) Bei der Frage *D255* ("Geburtsland Mutter") gibt es einen Fall, bei dem sowohl die Nennung "Deutschland" als auch die offene Nennung eines anderen Landes bestehen. Vermutlich entstand dies durch das Zurückgehen eines Interviewers im CAPI-Instrument. In diesem Fall wurde die Nennung "Deutschland" als das Geburtsland der Mutter belassen. Die Variable *problem1* markiert den betreffenden Fall.

(25) In einem Fall hat der Interviewer bei der Frage *Q18* ("Repraesentation Spitzenkandidaten") zwei Nennungen von Kandidaten notiert, die nicht auf der vorgegebenen Liste vertreten waren. Die erste Nennung wurde in der Variable *q18c1* und die zweite Nennung in der Variable *q18c2* erfasst. Für die Erstellung der Variable *q18rc* ("Repraesentation Spitzenkandidaten (*q18+q18c1*)") wurde nur die erste Nennung berücksichtigt.