

## ANHANG

### Allgemeine und methodische Hinweise zur repräsentativen Wahlstatistik der Bundestagswahl 2017

#### 1 Allgemeines

Rechtsgrundlage für die Erstellung der repräsentativen Wahlstatistik ist das Gesetz über die allgemeine und die repräsentative Wahlstatistik bei der Wahl zum Deutschen Bundestag und bei der Wahl der Abgeordneten des Europäischen Parlaments aus der Bundesrepublik Deutschland (Wahlstatistikgesetz – WStatG) vom 21. Mai 1999 (BGBl. I S. 1023), zuletzt geändert durch Artikel 1a des Gesetzes vom 27. April 2013 (BGBl. I S. 962).

Das WStatG legt für die Durchführung der repräsentativen Wahlstatistik im Wesentlichen folgendes fest:

1. Aus dem Ergebnis der Bundestagswahlen sind unter Wahrung des Wahlgeheimnisses in ausgewählten Wahlbezirken repräsentative Wahlstatistiken über
  - die Wahlberechtigten, Wahlscheinvermerke und die Beteiligung an der Wahl nach Geschlecht und Geburtsjahresgruppen (Statistik der Wahlbeteiligung) sowie
  - die Wählerinnen und Wähler und ihre Stimmabgabe für die einzelnen Wahlvorschläge nach Geschlecht und Geburtsjahresgruppen sowie die Gründe für die Ungültigkeit von Stimmen (Statistik der Stimmabgabe)

als Bundesstatistik zu erstellen (§ 2 WStatG).

Für die Statistik der Wahlbeteiligung dürfen höchstens 10 Geburtsjahresgruppen gebildet werden, in denen jeweils mindestens drei Geburtsjahrgänge zusammengefasst sind. Für die Statistik der Stimmabgabe sind höchstens sechs Geburtsjahresgruppen zulässig, in denen jeweils mindestens sieben Geburtsjahrgänge zusammengefasst sind (§ 4 WStatG).

2. Die Auswahl der Stichprobenwahlbezirke trifft der Bundeswahlleiter im Einvernehmen mit den Landeswahlleitungen und den Statistischen Ämtern der Länder. Es dürfen nicht mehr als 5 % der Wahlbezirke und der Briefwahlbezirke des Bundesgebietes sowie nicht mehr als 10 % der Wahlbezirke und der Briefwahlbezirke eines Landes an der repräsentativen Wahlstatistik teilnehmen. Ein für die repräsentative Wahlstatistik ausgewählter Wahlbezirk

muss mindestens 400 Wahlberechtigte umfassen, ein ausgewählter Briefwahlbezirk mindestens 400 Wählerinnen und Wähler bei der voran gegangenen Bundestagswahl umfasst haben. Die Wahlberechtigten sind in geeigneter Weise darauf hinzuweisen, dass der Wahlbezirk in eine repräsentative Wahlstatistik einbezogen ist (§ 3 WStatG).

3. Die Statistik der Wahlbeteiligung wird von den Gemeinden, in denen ausgewählte Wahlbezirke liegen, unter Auszählung der Wählerverzeichnisse durchgeführt. Die Gemeinden teilen die Ergebnisse getrennt nach Wahlbezirken dem zuständigen Statistischen Amt des Landes mit (§ 5 Absatz 1 WStatG).

Die Statistik der Stimmabgabe wird unter Verwendung von amtlichen Stimmzetteln, die zudem Unterscheidungsmerkmale nach Geschlecht und Geburtsjahresgruppen enthalten, durchgeführt. Die Gemeindebehörden leiten die ihnen von den Wahlvorstehern übergebenen, verpackten und versiegelten Stimmzettel der für die Statistik ausgewählten Wahlbezirke ungeöffnet und getrennt nach Wahlbezirken zur Auswertung an das Statistische Amt des Landes weiter. Gemeinden mit einer abgeschotteten Statistikstelle (§ 16 Absatz 5 Satz 2 des Bundesstatistikgesetzes) können die Auswertung der Stimmzettel mit Zustimmung der Landeswahlleitung selbst in der Statistikstelle vornehmen; in diesem Fall teilen sie die Ergebnisse getrennt nach Wahlbezirken dem zuständigen Statistischen Amt des Landes mit. Wählerverzeichnisse und gekennzeichnete Stimmzettel oder Ergebnisaufzeichnungen von Wahlgeräten dürfen nicht zusammengeführt werden (§ 5 Absatz 2 WStatG).

Nach Abschluss der Aufbereitung durch die Statistischen Ämter der Länder sind die Wahlunterlagen unverzüglich den Gemeindebehörden zurückzugeben und von diesen entsprechend den wahlrechtlichen Bestimmungen zu behandeln (§ 7 Absatz 3 WStatG).

4. Die Ergebnisse der repräsentativen Wahlstatistik dürfen nur für die Bundes- und Landesebene und ausschließlich durch das Statistische Bundesamt und die Statistischen Ämter der Länder veröffent-

licht werden. Die Bekanntgabe von Ergebnissen für einzelne Wahlbezirke ist unzulässig.

Bei Bundestagswahlen ist die repräsentative Wahlstatistik erstmals im Jahre 1953 (ohne Beteiligung der Länder Rheinland-Pfalz, Bayern und Saarland) und von 1957 bis 1990 sowie seit 2002 unter Beteiligung aller Bundesländer durchgeführt worden. Für die Bundestagswahlen 1994 und 1998 hat der Bundesgesetzgeber die Durchführung der repräsentativen Wahlstatistik ausgesetzt. Ergebnisse der repräsentativen Wahlstatistik liegen für jede der seit 1979 stattfindenden Europawahlen vor.

Im Gegensatz zur allgemeinen Wahlstatistik, bei der es sich um eine Dokumentation und Auswertung der bei den Wahlorganen angefallenen Wahlergebnisse handelt, wird in der repräsentativen Wahlstatistik das Wahlverhalten, das heißt die Wahlbeteiligung und die Stimmabgabe verschiedener Bevölkerungsgruppen, untersucht. Einfachere Untersuchungen dieser Art waren schon immer in der Weise möglich, dass man Wahlergebnisse für bestimmte regionale Einheiten mit der Zusammensetzung ihrer Bevölkerung verglich (zum Beispiel Gruppierung der Kreise oder Gemeinden nach dem Anteil der Religionszugehörigkeit, der sozial-ökonomischen Struktur und dergleichen). Ein solches Verfahren erfordert aber einen großen Aufwand und vermittelt nur gröbere Aussagen. Feststellungen über die Stimmabgabe nach dem Alter konnten überhaupt nicht getroffen werden.

Die repräsentative Bundestagswahlstatistik 2017 beruht auf den Wahlergebnissen in 2.723 ausgewählten Wahlbezirken für die Stimmabgabe (davon 2.254 allgemeine und 469 Briefwahlbezirke), die für die insgesamt 88.511 Wahlbezirke (einschließlich Briefwahlbezirke) als repräsentativ angesehen werden können. Sie erstreckt sich dadurch auf 2,2 Millionen der 61,7 Millionen Wahlberechtigten und 1,7 Millionen der 47,0 Millionen Wählerinnen und Wähler. Ihre Zahlenbasis ist danach weit breiter als bei entsprechenden Untersuchungen nichtamtlicher Stellen, die sich auf deutlich weniger Personen stützen. Außerdem handelt es sich bei der repräsentativen Bundestagswahlstatistik nicht um die Auswertung dessen, was Personen über ihr Wahlverhalten vor oder nach der Wahl aussagen, sondern um die Auswertung des tatsächlichen Wahlverhaltens nach den Wählerverzeichnissen und Stimmzetteln.

Für die Feststellung der Wahlberechtigten wurden die Wählerverzeichnisse der Auswahlbezirke herangezogen, die auch Angaben über den Vornamen (Geschlecht) und den Geburtstag enthalten. Die Feststel-

lungen über die Stimmabgabe wurden durch Ausgabe von Stimmzetteln mit Unterscheidungsaufdruck für Frauen und Männer nach jeweils sechs Geburtsjahresgruppen in den Stichprobenbezirken ermöglicht. Eine Verletzung des Wahlgeheimnisses war hierbei nicht zu befürchten. In die zur Feststellung der Wahlbeteiligung herangezogenen Wählerverzeichnisse können die Gemeindebehörden, die Wahlvorstände und die Öffentlichkeit (während der Auslegungsfrist der Wählerverzeichnisse) ohnehin Einblick nehmen. Auch die Methode zur Feststellung der Stimmabgabe der Frauen und Männer nach dem Geburtsjahr lässt keine Verletzung des Wahlgeheimnisses zu. Bei der Auswertung der Stimmzettel mit Unterscheidungsbezeichnungen wird beispielsweise festgestellt, wie viele Frauen einer jeden der sechs Geburtsjahresgruppen eine bestimmte Partei gewählt haben. Da aber zu jeder Geburtsjahresgruppe der Frauen und Männer zahlreiche Personen gehören, können daraus keine Anhaltspunkte für die Stimmabgabe einer bestimmten Person gewonnen werden. Die Wahlberechtigten werden durch amtliche Bekanntmachungen der Gemeindebehörden, in denen ausgewählte Wahlbezirke liegen, sowie in den betroffenen Wahllokalen durch Aushänge und Faltblätter über die Durchführung der repräsentativen Wahlstatistik informiert. In den ausgewählten Briefwahlbezirken wurden Faltblätter mit den Briefwahlunterlagen versandt.

## 2 Grundlagen des Stichprobenplans

Die Auswahl der Stichprobenwahlbezirke erfolgte durch eine Ziehung nach dem Zufallsprinzip auf Basis mathematischer Grundsätze. Dadurch sind zum einen die Objektivität des Auswahlvorgangs und die daraus resultierende Akzeptanz der Stichprobe sichergestellt. Zum anderen kann nur so die Präzision der Ergebnisse zuverlässig abgeschätzt werden.

Als Auswahlgrundlage für die Ziehung der Stichprobenwahlbezirke diente die allgemeine Wahlbezirksstatistik zur Bundestagswahl 2013.

Die Auswahlsätze beziehungsweise der auf die einzelnen Länder entfallende Stichprobenumfang wurden weitgehend proportional vorgenommen. Im Einzelnen setzt sich die Stichprobe wie folgt zusammen:

Land	Auswahlsatz
Deutschland	3,6 %
Schleswig-Holstein	3,9 %
Mecklenburg-Vorpommern	4,4 %
Hamburg	3,2 %
Niedersachsen	3,7 %
Bremen	4,1 %
Brandenburg	5,0 %
Sachsen-Anhalt	3,6 %
Berlin	3,4 %
Nordrhein-Westfalen	3,4 %
Sachsen	3,7 %
Hessen	3,9 %
Thüringen	4,6 %
Rheinland-Pfalz	4,3 %
Bayern	3,5 %
Baden-Württemberg	2,7 %
Saarland	3,9 %

Die dazugehörigen Stichprobenumfänge sind in der Übersicht 2.1 dargestellt.

Stichprobenmethodisch gesehen liegt eine geschichtete einfache Zufallsauswahl vor. Dies bedeutet, dass die Auswahlgrundlage vor der Ziehung der Stichprobe in Schichten unterteilt wird, wobei das Minimum von drei Wahlbezirken für jede Schicht festgelegt wurde. Dadurch und durch die Tatsache, dass die Anteile der Wahlbezirke, welche in die Auswahlgrundlage fließen (mindestens 400 Wahlberechtigte bei Urnen- beziehungsweise 400 Wählerinnen und Wähler bei Briefwahlbezirken), ergeben sich die Unterschiede in den Auswahlätzen zwischen den Bundesländern. Jede dieser Schichten bildet für sich genommen eine eigene Auswahlgrundlage aus der unabhängig von den übrigen Schichten eine Teilstichprobe gezogen wird. Die Verei-

nigung aller so erhaltenen Teilstichproben bildet dann die Stichprobe der Wahlbezirke für die repräsentative Wahlstatistik. Eine solche Schichtung dient hauptsächlich der Präzisionssteigerung der Ergebnisse gegenüber einer einfachen Zufallsauswahl.

Die Konstruktion der Schichtung erfolgte in einem hierarchischen Verfahren. Die Auswahlgrundlage wurde zunächst nach Bundesland und nach der Art des Wahlbezirkes (Urnen- und Briefwahlbezirk) unterteilt. Innerhalb der so erhaltenen Kreuzkombinationen aus Land und Wahlbezirksart wurden dann durch den Einsatz von Verfahren aus dem Bereich der Clusteranalyse weitere Schichtstufen gebildet, indem solche Wahlbezirke zusammengefasst wurden, die sich bezüglich der (geeignet standardisierten) Zahl an gültigen Zweitstimmen für die Parteien CDU, SPD, DIE LINKE, GRÜNE, CSU, FDP und AfD möglichst ähnlich sind.

Die Anzahl der gebildeten Schichten wurde dabei darüber bestimmt, inwieweit die Hinzunahme von einer weiteren Schicht die Stichprobenvarianz verkleinert.

Für Änderungen in den Abgrenzungen der Wahlbezirke zwischen dem Gebietsstand der Auswahlgrundlage (Bundestagswahl 2013) und der Bundestagswahl 2017 wurden eindeutige Nachfolgeregelungen gegeben. Eine nachträgliche Ergänzungsstichprobe von neuen Wahlbezirken aus Gebieten, die in der Auswahlgrundlage noch zu keinem Wahlbezirk gehörten, war nicht erforderlich.

### 3 Auszählung und Aufbereitung der Ergebnisse

Für genauere Feststellungen über die Wahlbeteiligung wurden bei der Bundestagswahl 2017 10 Geburtsjahresgruppen gebildet, die ungefähr folgenden Altersgruppen entsprechen:

Geburtsjahresgruppe	entspricht etwa Altersgruppe
1997 – 1999	unter 21 Jahre
1993 – 1996	21 – 24 Jahre
1988 – 1992	25 – 29 Jahre
1983 – 1987	30 – 34 Jahre
1978 – 1982	35 – 39 Jahre
1973 – 1977	40 – 44 Jahre
1968 – 1972	45 – 49 Jahre
1958 – 1967	50 – 59 Jahre
1948 – 1957	60 – 69 Jahre
1947 und früher	70 Jahre und älter

Für die Feststellungen über die Stimmabgabe wurden sechs Altersgruppen gebildet:

Geburtsjahresgruppe	entspricht etwa Altersgruppe
1993 – 1999	unter 25 Jahre
1983 – 1992	25 – 34 Jahre
1973 – 1982	35 – 44 Jahre
1958 – 1972	45 – 59 Jahre
1948 – 1957	60 – 69 Jahre
1947 und früher	70 Jahre und älter

Die Angaben über die Wahlberechtigten, Wählerinnen und Wähler beziehungsweise Nichtwählerinnen und -wähler nach Geschlecht und Altersgruppen wurden nach der Wahl anhand der Wählerverzeichnisse gewonnen. In den Wählerverzeichnissen sind die Wahlberechtigten eingetragen und es ist bei jedem Wahlberechtigten ohne Sperrvermerk W für ausgegebene Wahlscheine vermerkt, wenn er seine Stimme abgegeben hat.

Dem Inhalt der Wählerverzeichnisse entsprechend wurden folgende Zahlen unterschieden:

- A. Wahlberechtigte insgesamt, davon
  - Wahlberechtigte ohne Sperrvermerk W (A1)
  - Wahlberechtigte mit Sperrvermerk W für ausgegebene Wahlscheine (A2)
- B. Wählerinnen und Wähler ohne Wahlschein, die in dem jeweiligen Urnenwahlbezirk ihre Stimme abgegeben haben
- C. Nichtwählerinnen und -wähler ohne Wahlschein, das heißt Wahlberechtigte ohne Sperrvermerk W,

die in dem jeweiligen Urnenwahlbezirk ihre Stimme nicht abgegeben haben

Die Anzahl der Wählerinnen und Wähler beziehungsweise Nichtwählerinnen und -wähler unter den Wahlscheininhabern (im Allgemeinen sind dies Briefwählerinnen und -wähler) konnten nicht festgestellt werden, da die Personen mit Wahlschein ohne Unterlagen für die Briefwahl in einem beliebigen Wahlbezirk des Wahlkreises wählen können und die Personen mit Wahlschein und Unterlagen für die Briefwahl ihren Wahlbrief an die dafür zuständigen Stellen senden.

Bei der Anzahl der **Wahlberechtigten** wird unterschieden zwischen den Wahlberechtigten ohne und mit Sperrvermerk für ausgegebene Wahlscheine. In der Tabelle 1 werden folgende Werte ausgewiesen:

- a) Die Wahlberechtigten insgesamt (A), das heißt Wahlberechtigte mit und ohne Sperrvermerk W für ausgegebene Wahlscheine.
- b) Nur Wahlberechtigte mit Sperrvermerk für ausgegebene Wahlscheine (A2).

Die **Wahlbeteiligung** wurde anhand der Anzahl der Wählerinnen und Wähler ohne Wahlschein zuzüglich der Anzahl der Wahlberechtigten mit Wahlschein gemessen an der Anzahl der Wahlberechtigten insgesamt berechnet.

Bei den Wahlberechtigten mit Wahlschein wurde berücksichtigt, dass nicht alle Wahlscheininhaberinnen und -inhaber an der Wahl teilnahmen. Dazu wurde je Land der aus der allgemeinen Wahlstatistik bekannte Anteil der tatsächlichen Wahlscheinwählerinnen und -wähler (siehe folgende Tabelle) mit der Anzahl der Wahlberechtigten mit Sperrvermerk über alle Bevölkerungsgruppen gleichermaßen multipliziert. Dadurch wurde eine Annäherung der Wahlbeteiligung an die amtlich festgestellte Wahlbeteiligung erreicht.

Land	Wahlbeteiligung unter Wahlscheininhaber/-innen
Deutschland	95,9 %
Schleswig-Holstein	96,0 %
Mecklenburg-Vorpommern	96,5 %
Hamburg	95,3 %
Niedersachsen	95,6 %
Bremen	94,9 %
Brandenburg	95,7 %
Sachsen-Anhalt	95,4 %
Berlin	93,3 %
Nordrhein-Westfalen	96,0 %
Sachsen	96,3 %

Land	Wahlbeteiligung unter Wahrscheinhaber/-innen
Hessen	96,0 %
Thüringen	96,0 %
Rheinland-Pfalz	96,1 %
Bayern	96,3 %
Baden-Württemberg	95,7 %
Saarland	96,8 %

Für die Ermittlung der nach Geschlecht und Geburtsjahresgruppen differenzierten **Stimmabgabe** wurden alle in den ausgewählten Wahlbezirken abgegebenen Stimmzettel herangezogen. Das heißt sowohl die Stimmzettel der im Wählerverzeichnis eingetragenen Wahlberechtigten ohne Sperrvermerk W als auch die Stimmzettel der Personen mit Wahrschein, die nicht per Briefwahl, sondern in dem jeweiligen Urnenwahlbezirken gewählt haben. Im Allgemeinen ist daher die Gesamtanzahl der auf Grund der Stimmzettel ermittelten abgegebenen Stimmen höher als die Anzahl der Wählerinnen und Wähler ohne Wahrschein (Wählerinnen und Wähler mit Stimmabgabevermerk im Wählerverzeichnis).

Briefwahlbezirke wurden erstmals bei der repräsentativen Bundestagswahlstatistik 2002 berücksichtigt. Briefwählerinnen und -wähler werden in die Statistik der Stimmabgabe einbezogen, um systematische Verzerrungen durch das Fehlen der Briefwahlstimmen zu vermeiden. Seit Einführung der Briefwahl zur Bundestagswahl 1957 ist der Anteil der Briefwähler an den Wählern von 4,9 % auf 28,6 % bei der Bundestagswahl 2017 angestiegen.

Das Ergebnis der repräsentativen Bundestagswahlstatistik 2017 wurde in den Ländern aus den Urnen- und Briefwahlbezirken zusammengefasst. Bedingt durch den jeweiligen begrenzten Auswahlsatz in den Ländern sind hochgerechnete Ergebnisse der Stimmabgabe nach Geburtsjahresgruppen und Geschlecht beschränkt auf die Briefwählerinnen und -wähler nicht möglich. Für das Bundesgebiet liegen jedoch Ergebnisse nach den Urnen- und Briefwahlbezirken vor.

#### 4 Hochrechnung der Ergebnisse

Für die Erstellung der Ergebnisse zur repräsentativen Wahlstatistik kommt ein Kalibrierungsverfahren auf Basis einer Regressionsschätzung zum Einsatz. Die Regressionsschätzung hat den Charakter einer Hochrechnung mit Anpassung an „Eckwerte“ aus externer Quelle. Im Falle der repräsentativen Wahlstatistik ist das das amtliche Endergebnis der jeweiligen Wahl.

Der verallgemeinerte Regressionsschätzer<sup>1</sup> für den Totalwert  $t_Y$  eines interessierenden Merkmals Y lässt sich darstellen als:

$$\hat{t}_{Y,GREG} = \sum_{k=1}^n w_k y_k,$$

wobei  $y_k$  die Merkmalsausprägung im k-ten Stichprobenwahlbezirk ist und das Hochrechnungsgewicht für den k-ten Stichprobenwahlbezirk ( $w_k$ ) bestimmt wird durch:

$$w_k = \left[ 1 + (t_X - \hat{t}_{X,HT})' \left( \sum_{k=1}^n d_k \mathbf{x}_k \mathbf{x}_k' \right)^{-1} \mathbf{x}_k \right] d_k.$$

In obiger Formel bezeichnet  $\mathbf{x}_k$  den Vektor der Ausprägungen der sog. Hilfsvariablen im k-ten Stichprobenwahlbezirk,  $t_X$  sind die aus dem amtlichen Endergebnis bekannten Totalwerte der Hilfsvariablen (Eckwerte) und  $\hat{t}_{X,HT}$  ein durch „freie Hochrechnung“ ermittelter Schätzwert für den Totalwert  $t_X$ :

$$\hat{t}_{X,HT} = \sum_{k=1}^n d_k \mathbf{x}_k.$$

Ferner bezeichnet  $d_k$  die reziproke (schichtspezifische) Auswahlwahrscheinlichkeit des k-ten Stichprobenwahlbezirkes.

Es kann gezeigt werden, dass der Schätzer  $\hat{t}_{Y,GREG}$  die Eigenschaft

$$\hat{t}_{X,GREG} = \sum_{k=1}^n w_k \mathbf{x}_k = t_X$$

besitzt, das heißt bei Schätzung der Totalwerte der Hilfsvariablen durch Hochrechnung der Stichprobe mittels der Gewichte  $w_k$  werden diese exakt getroffen. In anderen Worten,  $\hat{t}_{Y,GREG}$  liefert eine an die Eckwerte  $t_X$  gebundene Schätzung.

Die Bestimmung der Hochrechnungsgewichte  $w_k$  wird mit Hilfe einer statistischen Analysesoftware durchgeführt.

Wie bereits erwähnt, werden im Fall der repräsentativen Wahlstatistik die Eckwerte, bei denen es sich um

<sup>1</sup> „Generalized Regression Estimator“ (GREG) in der englischsprachigen Literatur.

Stimmzahlen auf Landesebene handelt (Details siehe unten), aus der allgemeinen Wahlstatistik herangezogen. Dies hat den angenehmen Effekt, dass die Randsummen aus der repräsentativen Wahlstatistik mit den entsprechenden amtlichen Endergebnissen übereinstimmen, was bei einer freien Hochrechnung nicht gegeben wäre. Ferner wirkt sich das Schätzverfahren positiv auf die Präzision der Schätzungen aus. Insbesondere wirkt das Hochrechnungsverfahren dem Mangel entgegen, dass Wahlbezirke mit weniger als 400 Wahlberechtigten (Urne) bzw. Wählerinnen und Wählern (Brief) nicht in die Stichprobe gelangen dürfen.

Die Ergebnisse der repräsentativen Wahlstatistik lassen sich drei thematischen Gruppen zuordnen: Wahlbeteiligung, Wahlverhalten und Analyse der Ungültigkeit von Wahlzetteln. Die repräsentative Wahlstatistik vertieft die Ergebnisse der allgemeinen Wahlstatistik zu diesen Themen, indem sie darstellt, wie die Ausprägungen von den Merkmalen Alter und Geschlecht der Wähler abhängen. Die Ziele lassen sich genauer formulieren als:

1. Schätzung der Anzahl der Wahlberechtigten, der Wahrscheinlichkeiten und der Beteiligung an der Wahl in Abhängigkeit von Geschlecht und 10 Geburtsjahresgruppen.
2. Ermittlung des Wahlverhaltens in Erst- und Zweitstimme in Abhängigkeit von Geschlecht und sechs Geburtsjahresgruppen.
3. Aufschlüsselung der Häufigkeit ungültiger Stimmen nach Art und Geschlecht.

Für jede der drei thematischen Gruppen werden, für jedes Bundesland (bei der Stimmabgabe auch separat für das Bundesgebiet) getrennt, die Hochrechnungsgewichte  $w_k$  bestimmt, wobei stets der oben beschriebene Regressionsschätzer benutzt wird. Neben den Ergebnissen aus der Stichprobe gehen hier die Eckwerte  $t_x$  ein, die in den drei Fällen jeweils unterschiedlich sind:

1. Als Eckwerte dienen hier die Gesamtzahlen der Wahlberechtigten mit und ohne Wahrschein und der Wählerinnen und Wähler mit und ohne Wahrschein auf Landesebene.
2. In diesem Fall wird an die Gesamtzahlen der Stimmen nach Partei bzw. Gültigkeit auf Landes- und Bundesebene angepasst. Bei der Bundestagswahl 2017 werden hier sowohl für die Erst- als auch für die Zweitstimmen die zehn Kategorien „ungültig“,

CDU, SPD, DIE LINKE, GRÜNE, CSU, FDP, AfD, FREIE WÄHLER, NPD und „Sonstige“ betrachtet.<sup>2</sup>

3. Hier werden die Gesamtzahlen ungültiger Erst- und Zweitstimmen auf Landesebene als Eckwerte verwendet.

## 5 Weitere Hinweise

Wie bereits beschrieben, wurde die Hochrechnung für jedes Bundesland separat durchgeführt. Da nicht alle Ergebnisse bei der Stimmabgabe in den Ländern kalibriert werden konnten, wurde hier ein separates kalibriertes Bundesergebnis berechnet. Dies führt dazu, dass die Aufsummierung einzelner Merkmalsgruppen in den Länderergebnissen leicht zum entsprechenden Bundesergebnis abweichen können. Bei der Wahlbeteiligung und der Art der ungültigen Stimmabgabe wurde das Bundesergebnis durch Kumulation der Länderergebnisse gebildet.

Da es sich um Ergebnisse einer Stichprobe handelt, weisen die Ergebnisse mit einer bestimmten Wahrscheinlichkeit einen Zufallsfehler auf. Dieser Zufallsfehler wird umso größer, desto kleiner die Zahl der in der Stichprobe erfassten Merkmalsträger ist. Die Größe des Fehlerbereiches jedes einzelnen Stichprobenergebnisses hängt dabei nicht nur von der Zahl der Erhebungseinheiten und dem Auswahlatz ab, sondern auch von der Merkmalshäufigkeit. Deshalb hat jedes einzelne Merkmal seinen eigenen Fehlerbereich.

In einigen Übersichten und Tabellen, die nach Geburtsjahresgruppen differenzieren, sind zur besseren Verständlichkeit die Ergebnisse nach Altersgruppen dargestellt. Da bei der repräsentativen Wahlstatistik jedoch Geburtsjahresgruppen erhoben werden, können die Altersgruppen nicht exakt abgegrenzt werden. Übersicht 1.2 zeigt eine Gegenüberstellung der Alters- und Geburtsjahresgruppen für die Bundestagswahl 2017.

---

<sup>2</sup> In einigen Ländern erfolgte eine Reduzierung des Modells, sodass dort nicht alle oben genannten Eckwerte erreicht wurden. Die daraus resultierenden Abweichungen in den Randsummen können der folgenden Anhangtabelle entnommen werden.