



Arbeitspapier 01/2016

Unterschiedliche Ergebnisse beim deutschen Zensus 2011

Erklärung der Differenzen

Josef Schäfer



Impressum

Herausgegeben von
Information und Technik Nordrhein-Westfalen (IT.NRW),
Geschäftsbereich Statistik

Postfach 10 11 05, 40002 Düsseldorf
Mauerstraße 51, 40476 Düsseldorf

Telefon 0211 9449-01

Mail poststelle@it.nrw.de

Web www.it.nrw.de

Bestell-Nr. Z269 2016 51

Dieser Artikel wurde erstmals veröffentlicht in:

AStA Wirtschafts- und Sozialstatistisches Archiv; Dezember 2015, Volume 9, Issue 3, pp 173-190
(<http://link.springer.com/article/10.1007/s11943-015-0170-0>) und unterliegt den Bedingungen der
Creative Commons Attribution License. Dadurch sind die Nutzung, Verteilung und Reproduktion
erlaubt, sofern der/die Originalautor/in(en) und die Quelle angegeben sind.



Inhalt

- Zusammenfassung / Abstract 3
- 1. Einführung..... 4
- 2. Zwei Veröffentlichungstermine 6
- 3. Unterschiedliche Ergebnisse zum gleichen Sachverhalt..... 7
- 4. Einwohnerzahl 7
- 5. Demografie 7
- 6. Bildung und Migration 8
- 7. Erwerbstätigkeit 8
- 8. Religion 9
- 9. Haushalte und Familien 9
- 10. Gebäude und Wohnungen..... 9
- 11. Ergebnisdifferenzen nach ausgewählten Merkmalen..... 10
- 12. Differenzen innerhalb des Veröffentlichungstermins 2 16
- 13. Fazit 20
- Literatur 22

Zusammenfassung / Abstract

Der Zensus 2011 war zur Entlastung der Auskunftspflichtigen und zur Reduzierung der Kosten ein registergestützter Zensus, bei dem im Gegensatz zu früheren Volkszählungen nicht alle Personen und Haushalte befragt wurden. Der Zensus 2011 hat dabei unterschiedliche Datenquellen genutzt und unterschiedliche Erhebungsverfahren eingesetzt. Damit hat sich ein breites Spektrum an Auswertungsmöglichkeiten ergeben. Neben den Einwohnerzahlen konnten im Mai 2013 bereits weitere erste Ergebnisse vor dem Abschluss aller Arbeiten am Zensus 2011 erstellt und veröffentlicht werden. Zur Erstellung der endgültigen Ergebnisse wurden weitere Informationen genutzt, zum Teil beruhten die endgültigen Ergebnisse im Mai 2014 auf anderen Datenquellen und anderen Berechnungsverfahren.

Anhand von Boxplots über die Gemeinden Deutschlands ab 10 000 Einwohner werden die aus den genannten methodischen Unterschieden resultierenden Divergenzen in den Ergebnissen für ausgewählte Merkmale dargestellt.

Schlüsselwörter: Zensus 2011, zwei Veröffentlichungstermine, unterschiedliche Datenquellen und Erhebungsverfahren, unterschiedliche Ergebnisse

The German census 2011 has been a register based census. Deviant from former "Volkszählungen" not all persons and households has been surveyed to discharge respondents and to save costs. Therefore the Census 2011 has used different data sources and different methods of survey. This has opened a wide spectrum of potential analysis. In addition to the population size first results had already been compiled and published in May 2013 before the data processing was completed. For final results additional information had been included. In parts the final results in May 2014 were based on other data sources and other methods of calculation.

Boxplots on German municipalities with at least 10 000 inhabitants show the divergences in the results due to different methods of survey and analysis for selected variables.

Keywords: Census 2011, Two Dates of Publication, Different Data Sources and Methods of Survey, Different Results


1. Einführung

In Deutschland wurden 1983 (in den neuen Bundesländern) und 1987 (in den alten Bundesländern) die bislang letzten Volkszählungen als Totalerhebungen durchgeführt. In der Folge gab es für einen Zeitraum von 28 bzw. 24 Jahren keine aktuelle Bestandsaufnahme der Bevölkerung als Grundlage für deren Fortschreibung. Eine Aktualisierung des Bestands an Gebäuden und Wohnungen erfolgte 1995 nur in den neuen Bundesländern (vgl. Scholze 2006). Eine Inventur der statistischen Basis in Deutschland war damit im Vorfeld des Zensus dringend erforderlich. An der EU-weiten Zensusrunde 2001 hatte sich Deutschland als einziger Mitgliedsstaat nicht beteiligt. Eine herkömmliche Volkszählung war damals aus Kosten- und Akzeptanzgründen politisch nicht durchsetzbar, methodische Grundlagen für andere Modelle mussten erst noch geschaffen werden. Eppmann und Schäfer (2006) beschreiben die damalige Situation.

Im Rahmen des Zensustests 2001 (Statistische Ämter des Bundes und der Länder 2004) wurde ein neues registergestütztes Verfahren erprobt, das in wesentlichen Teilen das Modell des Zensus 2011 vorwegnahm und erfolgreich getestet hat. So haben sich unter anderem die Verwendung von Einwohnermelderegistern, die Gebäude- und Wohnungszählung als schriftliche Befragung der Gebäude- und Wohnungseigentümer und das Verfahren einer Haushaltegenerierung aus Melderegisterangaben und Angaben aus der Gebäude- und Wohnungszählung das für den Zensus 2011 weiter optimiert wurde (vgl. Kreuzmair, Reisch 2021), bereits beim Zensustest bewährt.

Nach dem Zensustest hat das Modell des Zensus 2011 noch Erweiterungen, Anpassungen und Optimierungen erfahren. So wurde im Zensustest deutlich, dass aufgrund der hohen Unter- und Übererfassungen der Melderegister und der starken Streuung dieser Fehler zwischen den Gemeinden eine objektive Kontrolle der Melderegister erforderlich war. Diese statistische Kontrolle der Melderegister erfolgte beim Zensus 2011 in den Gemeinden ab 10 000 Einwohner in Form einer Zufallsstichprobe (Haushaltsbefragung). Zur Genese dieser Idee vgl. Schäfer (2004). In den Gemeinden unter 10 000 Einwohnern wären allerdings zur Erreichung des dafür notwendigen Stichprobenumfangs hohe Auswahlsätze erforderlich gewesen, außerdem zeigte sich beim Zensustest in den kleineren Gemeinden eine tendenziell bessere Registerqualität als in den größeren. Als Ergebnis des Zensustests ließen sich in den kleinen Gemeinden mit einer gezielten Befragung von Adressen mit einer bewohnten Wohnung, bei denen sich Auffälligkeiten aus den ersten Schritten der Haushaltegenerierung gezeigt haben, mit erheblich geringerem Aufwand Einwohnerzahlen mit einer Genauigkeit ermitteln, die derjenigen der Zufallsstichprobe in den größeren Gemeinden vergleichbar war.

Zusätzliche Erweiterungen des Zensus 2011 gegenüber dem beim Zensustest 2001 geprüften Modell sind die vollständige Erhebung von Anstalten und Wohnheimen (Adressen mit Sonderbereichen) und die Erweiterung der Grundgesamtheit über die Meldeanschriften hinaus. Der Zensustest 2001



beschränkte sich noch auf Meldeanschriften und schloss Anstalten und einen Teil der Wohnheime aus der Untersuchung aus.

IT-gestützte Verfahren zur Erhebungsunterstützung (auch unter Einbeziehung kommunaler Erhebungsstellen), zur Aufbereitung der Daten, zur Datenbereitstellung (Auswertungsdatenbank) und zur statistischen Geheimhaltung mussten für den Zensus 2011 neu entwickelt und umgesetzt werden.

Im Resultat ist der Zensus 2011 in mehrfacher Hinsicht komplexer als vorangegangene Volkszählungen. Wesentliche Elemente dieser Komplexität sind unter anderem:

- Aus methodischen wie aus Effizienzgründen gab es in Gemeinden unter 10 000 Einwohnern und in Gemeinden mit 10 000 und mehr Einwohnern unterschiedliche Verfahren zur Ermittlung der Einwohnerzahlen (vgl. Begründung zu § 7 Abs. 1 ZensG 2011).
- Datenquellen waren sowohl Verwaltungsregister (kommunale Einwohnermelderegister, erwerbsstatistische Register der Bundesagentur für Arbeit und der durch die nach dem Finanz- und Personalstandsstatistikgesetz auskunftspflichtigen Stellen) als auch Primärerhebungen, die miteinander verknüpft wurden.
- Primärerhebungen erfolgten teils als Totalerhebung (Gebäude- und Wohnungszählung, Erhebungen in Sonderbereichen), teils auf Basis einer Zufallsstichprobe (Haushaltsbefragung), teils durch gezielte Nachfragen (Personen nur mit Nebenwohnsitz in Deutschland, Personen mit mehr als einem Hauptwohnsitz - soweit zumindest ein Hauptwohnsitz in einer Gemeinde unter 10 000 Einwohnern lag, Befragung zur Klärung von Unstimmigkeiten in Gemeinden unter 10 000 Einwohnern bei auffälligen Anschriften mit nur einer bewohnten Wohnung).
- Für die Haushaltgenerierung wurden unterschiedliche Datenquellen zusammengeführt. Unmittelbar wurden die Angaben aus den Einwohnermelderegistern und der Gebäude- und Wohnungszählung, mittelbar wurden bestimmte Eckwerte zur Korrektur von Über- und Untererfassungen der Melderegister sowie zur statistischen Verteilung bestimmter Haushaltstypen auch aus der Haushaltsbefragung gewonnen.
- Der Umfang von Korrekturen der Über- und Untererfassungen der Melderegister auf Einzeldatenbasis in Gemeinden ab 10 000 Einwohnern wurde zunächst aus der Haushaltsbefragung hochgerechnet. Die Korrekturen wurden anschließend durch ein iteratives Verfahren im Rahmen der Haushaltgenerierung auf Einzelfallbasis umgesetzt (vgl. Hofmeister, Fürnrohr 2014).
- Die unterschiedlichen Datenquellen und Erhebungsmethoden führten bei der Ergebniserstellung zu einer Kombination von Auszählungen und Hochrechnungen (Multiple Source Mixed Mode Design).
- Ein neuentwickeltes, die Daten prätabular veränderndes Verfahren zur Sicherung der Statistischen Geheimhaltung gemäß § 16 BstatG (SAFE – Sichere Anonymisierung für Einzeldaten, zur Beschreibung des Verfahrens vgl. Höhne 2011, zur Anwendung beim Zensus 2011 Gießing et al. 2014) ermöglicht eine sichere Darstellung von Ergebnissen in mehrdimensionalen Datenquadranten – hier würden Zellsperrverfahren an die Grenzen ihrer Möglichkeiten gelangen – verändert aber die Daten.

2. Zwei Veröffentlichungstermine

Der Wunsch nach einer frühzeitigen Bereitstellung der Einwohnerzahlen vor Abschluss aller Arbeiten am Zensus 2011 führte zu einer Veröffentlichung von Ergebnissen des Zensus 2011 zu zwei unterschiedlichen Zeitpunkten, dem Veröffentlichungstermin 1 im Mai 2013 und dem Veröffentlichungstermin 2 im Mai 2014. Der Veröffentlichungstermin 1 umfasste dabei neben den Einwohnerzahlen

- (vorläufige) Ergebnisse aus der Gebäude- und Wohnungszählung,
- (vorläufige) aus der Haushaltsbefragung hochgerechnete Angaben zur Person für alle Gemeinden ab 10 000 Einwohner sowie für alle Kreise, Länder und Deutschland und
- Auszählungen demografischer Merkmale für Gemeinden unter 10 000 Einwohnern.

Die Ergebnisse zum Veröffentlichungstermin 1 wurden ohne die Einbeziehung der Angaben aus den erwerbsstatistischen Registern sowie der Ergebnisse der Haushaltegenerierung einschließlich des darin eingebundenen Verfahrens zur Korrektur der Melderegisterdaten um Über- und Untererfassungen auf Einzelfallbasis erstellt. Ferner waren Korrekturen im Gebäude- und Wohnungsbestand, die sich im Rahmen der Haushaltegenerierung beim Vergleich von Melderegistern und Ergebnissen der Gebäude- und Wohnungszählung ergaben, noch nicht berücksichtigt.

Zum Veröffentlichungstermin 2 lagen für alle Gemeinden ab 10 000 Einwohnern sowie alle Kreise, Länder und Deutschland Ergebnisse zu folgenden Bereichen vor:

- Einwohnerzahl (unverändert gegenüber dem Veröffentlichungstermin 1)
- Demografie (feiner gegliedert als zum Veröffentlichungstermin 1 und nach anderer Methodik ermittelt)
- Erwerbstätigkeit (erweitert und verändert unter Verwendung erwerbsstatistischer Register, sogenannter „Baukasten“ (vgl. Sedmihradsky et al. 2012))
- Weitere Merkmale zur Person, z. B. Bildung und Migration (veränderte Hochrechnung)
- Haushalte und Familien (neu)
- Gebäude und Wohnungen (korrigiert unter Berücksichtigung von Erkenntnissen aus der Haushaltegenerierung)

3. Unterschiedliche Ergebnisse zum gleichen Sachverhalt

Anders als bei früheren Volkszählungen gab es beim Zensus 2011 unterschiedliche Ergebnisse zum jeweils gleichen Sachverhalt:

- Die Auswertungen zu den Veröffentlichungsterminen 1 und 2 erfolgten aus unterschiedlichen Quellen und nach unterschiedlichen Verfahren. Demografische Angaben für die Gemeinden ab 10 000 Einwohner wurden zum Beispiel zum Veröffentlichungstermin 1 aus der Haushaltsbefragung hochgerechnet zum Veröffentlichungstermin 2 aus statistisch korrigierten Melderegistern ausgezählt. Die statistischen Korrekturen der Melderegister zu Über- und Untererfassungen berücksichtigen aus der Haushaltsbefragung hochgerechnete Randverteilungen u.a. demografischer Merkmale. Dies verringert zwar Ergebnisdifferenzen löst diese aber nicht vollständig auf.
- Gleiche Merkmale wurden sowohl aus Registern als auch aus den Angaben der Befragten ermittelt. Die Hochrechnungen aus den Haushaltebefragungen basierten auf den Angaben der Befragten, die Auszählungen auf Registerangaben.
- Die Hochrechnungen unterlagen dabei einer Fehlerrechnung mit einer Unterdrückung von Angaben bei einem zu hohen Standardfehler.
- Die Auszählungen durchliefen einen Daten verändernden Geheimhaltungsprozess, der den Persönlichkeitsschutz sicherstellte, flexible Auswertungen ermöglichte, aber Abweichungen vom Originalwert zur Folge hatte. Bei zu großen Abweichungen wurden Angaben unterdrückt, um die Darstellung zu stark verzerrter Ergebnisse zu vermeiden.
- Der Bestand an Gebäuden und Wohnungen wurde nach Erkenntnissen aus der Haushaltegenerierung korrigiert.


4. Einwohnerzahl

Die Ermittlung der Einwohnerzahl war das zentrale Ziel des Zensus 2011. Sowohl in Bezug auf das Design als auch auf das Hochrechnungsverfahren wurde die Haushaltsbefragung im Rahmen eines Forschungsprojekts vor allem im Hinblick auf die Ermittlung der Einwohnerzahl optimiert (vgl. Münnich et al. 2012). Die Einwohnerzahl war zum Veröffentlichungstermin 1 bereits festgelegt.

5. Demografie

Demografische Angaben zur Person (Alter, Geschlecht, Familienstand und Staatsangehörigkeit) lagen sowohl im Melderegister als auch aus der Haushaltebefragung vor. Bei einer Auswertung nur demografischer Merkmale (in den Gemeinden ab 10 000 Einwohnern sowie auf höheren regionalen Ebenen) erfolgte zum Veröffentlichungstermin 1 eine Hochrechnung aus der Haushaltsbefragung, zum Veröffentlichungstermin 2 eine Auszählung aus statistisch korrigierten Melderegistern. Der Beitrag der Gemeinden unter 10 000 Einwohnern auf die Kreisebene und auf höhere regionale Aggregate wurde sowohl zum Veröffentlichungstermin 1 als auch zum Veröffentlichungstermin 2 ausgezählt.

Eine Kreuztabellierung demografischer Merkmale mit Merkmalen, die ausschließlich aus der Haushaltsbefragung ermittelt wurden (zum Beispiel zur Bildung oder zur Migration) wurde sowohl zum



Veröffentlichungstermin 1 als auch zum Veröffentlichungstermin 2 hochgerechnet. Im Vergleich zur Hochrechnung (rein) demografischer Merkmale zum Veröffentlichungstermin 1 sind bei dieser Hochrechnung demografischer Merkmale folgende Abweichungen zu beachten:

- Bestimmte Formen von Sonderanschriften, bei denen die Bewohnerinnen und Bewohner nicht selbst befragt wurden (sog. sensible Sonderbereiche wie z. B. Justizvollzugsanstalten, psychiatrische Kliniken oder Meldeanschriften für Obdachlose sowie Kasernen) blieben bei den mit weiteren Merkmalen kombinierten Hochrechnungen unberücksichtigt.
- Angaben aus Gemeinden unter 10 000 Einwohnern wurden auf der Kreisebene und darüber hinaus in die Hochrechnung einbezogen – nicht ausgezählt wie bei der Auswertung ausschließlich demografischer Merkmale.

In Bezug auf Gliederung nach demografischen Merkmalen erfolgte zum Veröffentlichungstermin 2 eine Anpassung (Kalibrierung mittels Raking) an die Auszählung der statistisch korrigierten Melderegister. Dadurch erklären sich die in der Regel kleinen Unterschiede zwischen ansonsten gleichen Auswertungen zu den Veröffentlichungsterminen 1 und 2.

6. Bildung und Migration

Angaben zur Bildung und zur Migration wurden sowohl zum Veröffentlichungstermin 1 als auch zum Veröffentlichungstermin 2 aus der Haushaltsbefragung hochgerechnet. Unterschiedliche Ergebnisse sind hier auf Unterschiede im Hochrechnungsverfahren (Anpassung an die Auszählung der statistisch korrigierten Melderegister bezüglich demografischer Merkmale) zurückzuführen.

7. Erwerbstätigkeit

Zur Erwerbstätigkeit lagen zum Veröffentlichungstermin 1 nur die Angaben aus der Haushaltsbefragung vor, die hochgerechnet wurden. Zum Veröffentlichungstermin 2 kamen gemäß § 4 ZensG 2011 die Register der sozialversicherungspflichtig beschäftigten Personen und der als arbeitslos gemeldeten oder Arbeit suchenden oder nicht zu aktivierenden Personen der Bundesagentur für Arbeit sowie gemäß § 5 ZensG 2011 Daten der durch die nach dem Finanz- und Personalstandsstatistikgesetz auskunftspflichtigen Stellen zu Beamten, Richtern, Soldaten und Dienstordnungsangestellten als erwerbsstatistische Register hinzu (vgl. Sedmihradsky et al. 2012).

Die beim Zensus 2011 verwendeten erwerbsstatistischen Register deckten die Menge der Erwerbstätigen nicht vollständig ab, es fehlten unter anderem die Selbstständigen, die mithelfenden Familienangehörigen und die geringfügig Beschäftigten. Im sogenannten „Baukasten“ wurden daher ausgezählte Registerangaben um Hochrechnungen aus der Haushaltsbefragung ergänzt. Dabei war für den Auszählungsteil die statistische Geheimhaltung zu beachten, für den Hochrechnungsteil die Fehlerrechnung. Beides konnte im „Baukasten“ zum Veröffentlichungstermin 2 zur Unterdrückung von Tabellenfeldern führen.

8. Religion

Die Frage einer Zugehörigkeit zu einer (öffentlich-rechtlichen) Religionsgesellschaft nahm eine Sonderrolle ein. Die Ausprägungen „Römisch-katholische Kirche“ und „Evangelische Kirche“ waren aus (kirchen-)steuerlichen Gründen auch in den kommunalen Melderegistern enthalten. In der groben Gliederung „katholisch“, „evangelisch“ und „sonstige, keine, ohne Angabe“ gab es Unterschiede aufgrund der Hochrechnung aus der Haushaltsbefragung zum Veröffentlichungstermin 1 und der Auszählung aus korrigierten Melderegistern zum Veröffentlichungstermin 2. Wurden sämtliche Ausprägungen dieser Frage (Römisch-katholische Kirche, Evangelische Kirche, Evangelische Freikirchen, Orthodoxe Kirchen, Jüdische Gemeinschaften, Sonstige öffentlich-rechtliche Religionsgesellschaften, Keiner öffentlich-rechtlichen Religionsgesellschaft) ausgewertet, erfolgte dies zu den Veröffentlichungsterminen 1 und 2 über eine Hochrechnung aus der Haushaltsbefragung.

Die freiwillige Beantwortung der Frage nach dem Bekenntnis zu Religionen, Glaubensrichtungen oder Weltanschauungen war nur an die Personen gerichtet, die keiner (öffentlich-rechtlichen) Religionsgesellschaft angehörten. Diese selektive Filterführung erschwerte, verbunden mit der Freiwilligkeit und der Sensibilität dieser Frage, eine Interpretation der Ergebnisse. Dies hat die Statistischen Ämter des Bundes und der Länder veranlasst, die Auswertung der Frage nach dem Bekenntnis zu Religionen, Glaubensrichtungen oder Weltanschauungen nicht in das Standardauswertungsprogramm aufzunehmen.


9. Haushalte und Familien

Angaben zu Haushalten und Familien sind Ergebnisse der Haushaltgenerierung und lagen erst zum Veröffentlichungstermin 2 vor. Damit ergibt sich hierfür ein einheitliches Ergebnis im Zensus 2011. Da die Haushaltgenerierung flächendeckend erfolgte, konnten die Angaben zu Haushalten und Familien ausgezählt werden.

Erhoben wurden Wohnhaushalte, das heißt die Bewohnerinnen und Bewohner je einer Wohnung bilden einen gemeinsamen Haushalt. Hiermit ergeben sich systematische Unterschiede zu früheren Volkszählungen und anderen Haushaltserhebungen wie zum Beispiel dem Mikrozensus, bei denen Wohn- und Wirtschaftshaushalte (einen Haushalt bilden dabei die Personen, die gemeinsam wohnen und gemeinsam wirtschaften) erfasst wurden bzw. werden.

10. Gebäude und Wohnungen

Angaben zu Gebäuden und Wohnungen wurden flächendeckend im Rahmen der Gebäude- und Wohnungszählung erhoben. Zum Veröffentlichungstermin 1 wurden die Ergebnisse dieser Primärerhebung ausgezählt. Daneben waren die Ergebnisse der Gebäude- und Wohnungszählung neben den Melderegistern Grundlage für die Haushaltgenerierung. Beim Vergleich zwischen den



statistisch korrigierten Melderegistern und dem ermittelten Bestand an Wohngebäuden und bewohnten Wohnungen mussten im Rahmen der Haushaltegenerierung Widersprüche aufgelöst werden, hierzu einige Beispiele:

- Von einem Haushalt genutzte Wohnungen wurden zusammengelegt.
- Bei Wohngebäuden, zu denen keine Angaben ermittelt werden konnten, wurden fehlende Angaben imputiert, das heißt, es wurden Angaben anderer Wohngebäude übernommen (vgl. Grunwald, Krause 2014). Bei dieser unsicheren Feststellung durch Imputation konnte die Haushaltegenerierung aus den Melderegisterangaben zur gleichen Anschrift weitere Erkenntnisse gewinnen. Imputierte Gebäude und Wohnungen, für die keine gemeldeten Bewohner vorhanden waren, wurden gelöscht.
- Für Haushalte, die keiner gemäß Gebäude- und Wohnungszählung vorhandenen Wohnungen zugeordnet werden konnten, wurden Wohnungen angelegt.

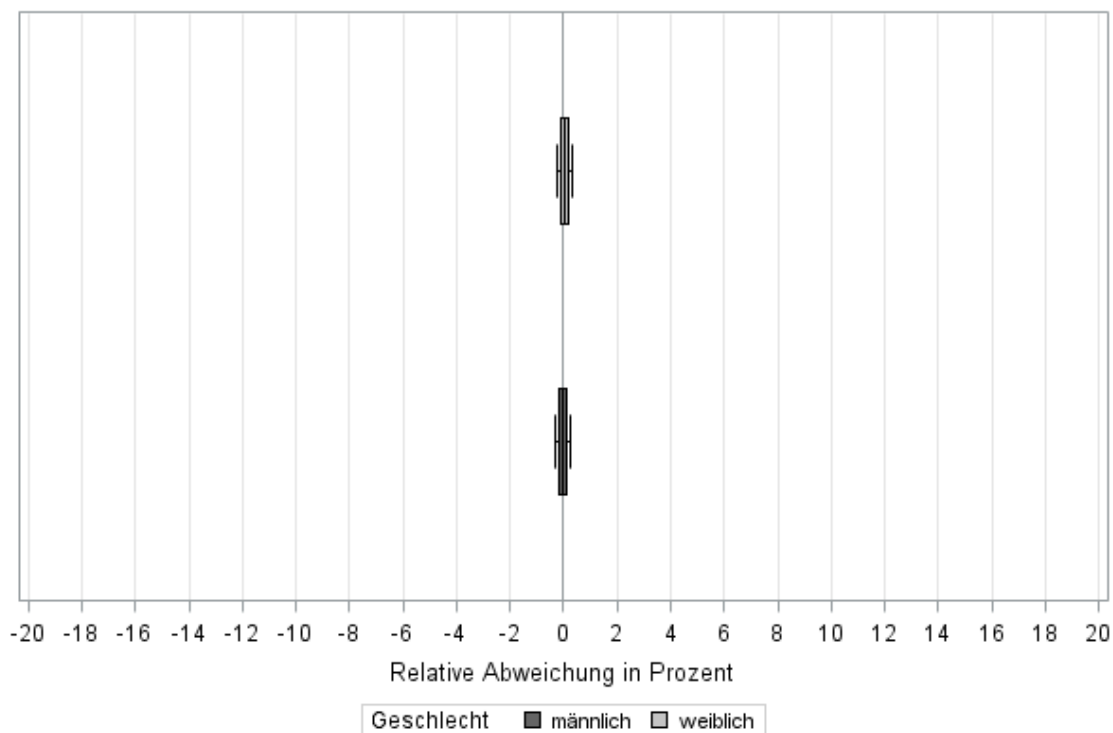
Diese Korrekturen im Rahmen der Haushaltegenerierung haben den Bestand an Gebäuden und Wohnungen zum Veröffentlichungstermin 2 im Vergleich zum Veröffentlichungstermin 1 verändert.

11. Ergebnisdifferenzen nach ausgewählten Merkmalen

Unterschiede zwischen den beiden Veröffentlichungszeitpunkten beim Zensus 2011 beruhen bei demografischen Merkmalen aus dem unterschiedlichen Berechnungsverfahren (Hochrechnung zum Veröffentlichungstermin 1, Auszählung zum Veröffentlichungstermin 2 sowie möglichen unterschiedlichen Ausprägungen zwischen Angaben in der Befragung und im Melderegister. Beim Geschlecht, dem Alter und der Staatsangehörigkeit nach „Deutsch“ und „Nicht deutsch“ sind die Unterschiede in der vorgenommenen groben Gliederung insgesamt zu vernachlässigen¹. Auffällig sind lediglich eine größere Streuung der Differenzen zwischen den Gemeinden in der Altersklasse der 18 bis 29-jährigen sowie eine leichte Tendenz zu mehr Nichtdeutschen und weniger Deutschen in der Auszählung zum zweiten Veröffentlichungstermin im Vergleich zur Hochrechnung beim ersten Veröffentlichungstermin. Es gibt daher bei diesen Merkmalen keine Anhaltspunkte zu möglichen systematischen Differenzen im Vergleich von Hochrechnungen und Auszählungen sowie von Angaben aus der Befragung und dem Melderegister.

¹ Die Aussage bezieht sich auf die hier vorgenommene grobe Gliederung in eindimensionaler Auswertung. Randverteilungen dieser Merkmale sind in den hier betrachteten Ausprägungen in das zuvor genannte Verfahren zur statistischen Korrektur der Melderegister um Über- und Untererfassungen eingeflossen. Von daher ist die gute Übereinstimmung nicht überraschend. Bei feineren Gliederungen und mehrdimensionalen Auswertungen kann es insbesondere auf der Gemeindeebene in Einzelfällen zu (relativ) größeren Abweichungen kommen.

Abbildung 1.1 Relative Abweichung der Ergebnisse des Veröffentlichungstermins 2 des Zensus 2011 zu denen des Veröffentlichungstermins 1 in Prozent² nach dem Geschlecht:



² Dargestellt werden hier wie auch in den folgenden Abbildungen die relativen Abweichungen des zweiten Veröffentlichungstermins zum ersten in den Gemeinden ab 10 000 Einwohnern. Die Boxes stellen den Bereich zwischen den Quartilen dar. Angezeigt sind das 10%- sowie das 90%-Quantil. In die Berechnungen sind die Ergebnisse nach der Datenveränderung durch das Geheimhaltungsverfahren SAFE eingeflossen, auch wenn diese aufgrund großer Abweichungen zum Originalwert in den Veröffentlichungen unterdrückt wurden.



Abbildung 1.2 Relative Abweichung der Ergebnisse des Veröffentlichungstermins 2 des Zensus 2011 zu denen des Veröffentlichungstermins 1 in Prozent nach Altersklassen:

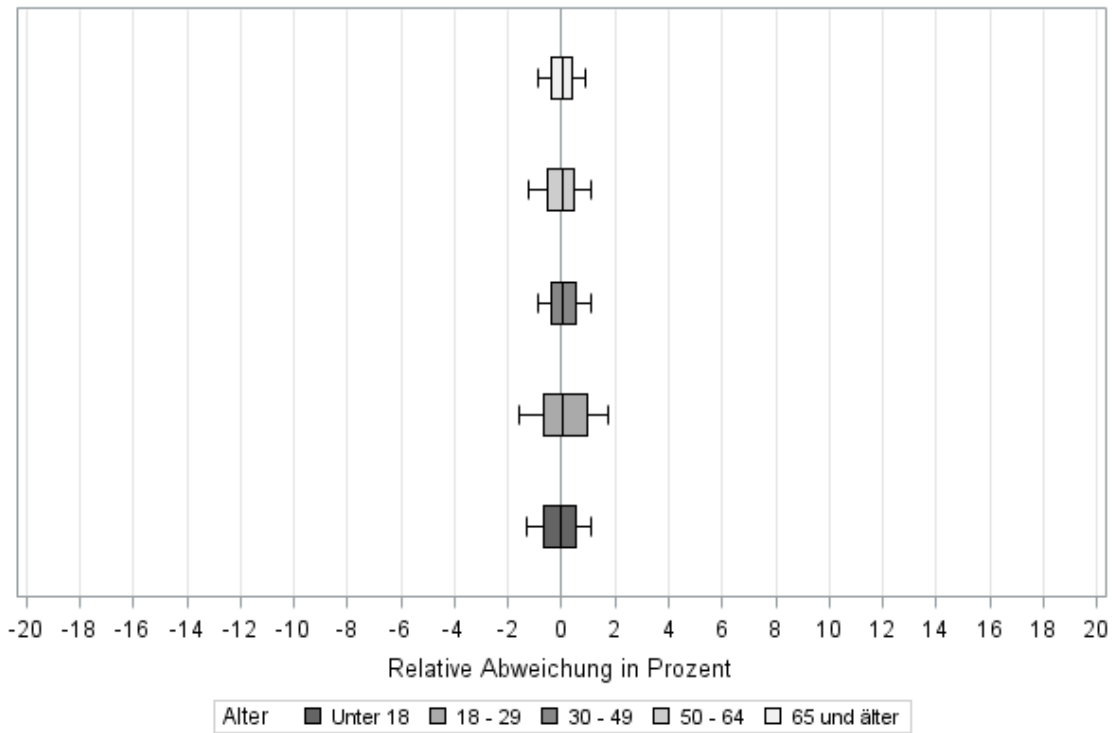
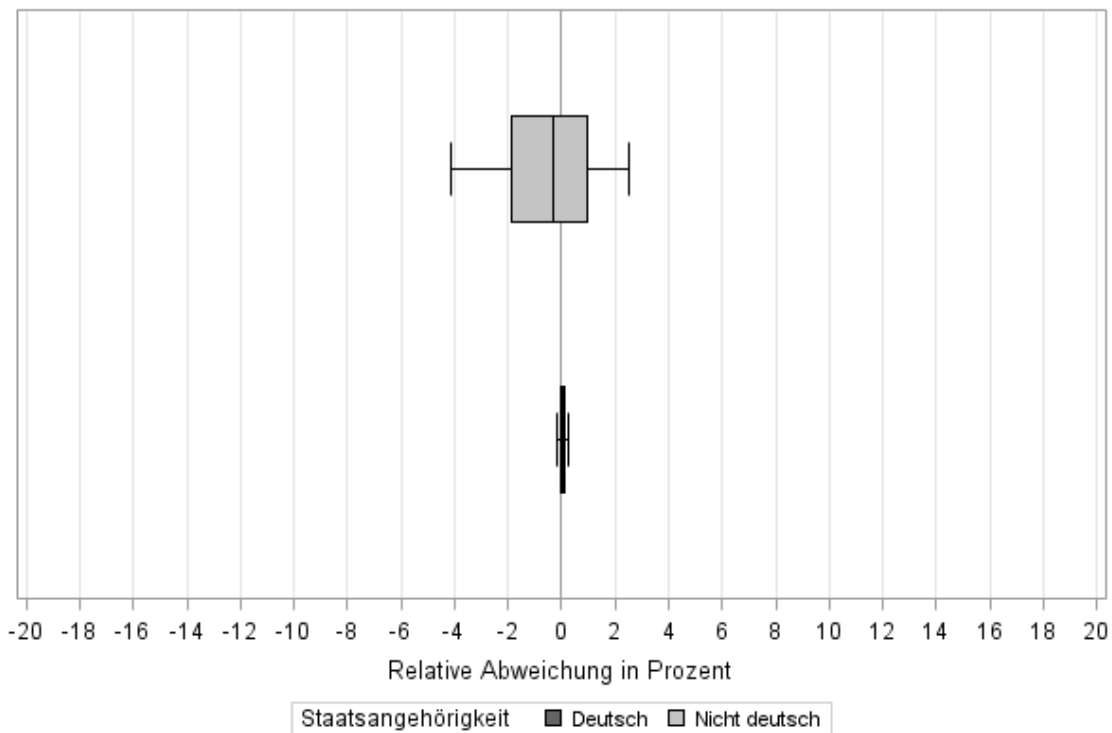


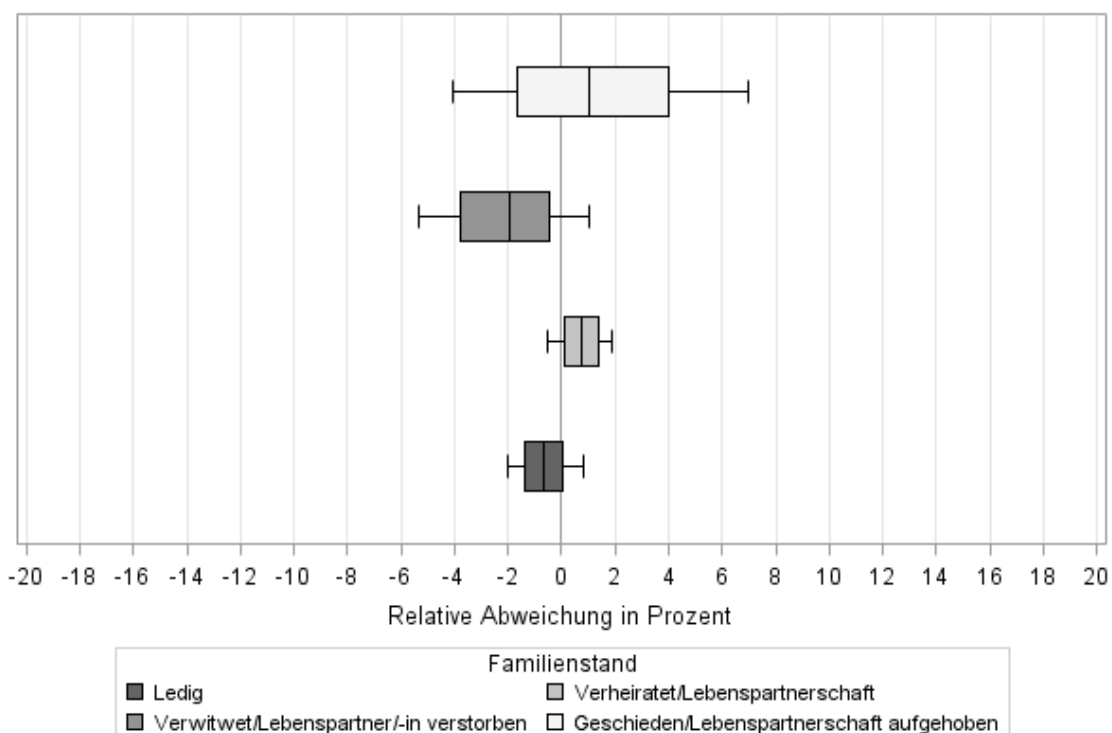
Abbildung 1.3 Relative Abweichung der Ergebnisse des Veröffentlichungstermins 2 des Zensus 2011 zu denen des Veröffentlichungstermins 1 in Prozent nach der Staatsangehörigkeit:



Beim Familienstand zeigen sich dagegen systematische Unterschiede. Die Ausprägungen „Ledig“ und „Verwitwet“ (einschließlich „eingeschriebene(r) gleichgeschlechtliche(r) Lebenspartner(in) verstorben“) weisen bei der Hochrechnung aus der Befragung (zum Veröffentlichungstermin 1) höhere Werte auf als bei der Auszählung aus dem Melderegister (zum Veröffentlichungstermin 2), bei den Ausprägungen „Verheiratet“ (einschließlich „eingetragene gleichgeschlechtliche Lebenspartnerschaft“) und „Geschieden“ (einschließlich „eingetragene gleichgeschlechtliche Lebenspartnerschaft aufgehoben“) ist es umgekehrt.

Da die unterschiedlichen Berechnungsverfahren (Hochrechnung versus Auszählung) zu keinen nennenswerten Unterschieden bei den Merkmalen Geschlecht, Alter und Staatsangehörigkeit geführt haben und auch der Familienstand im Rahmen des Verfahrens zur statistischen Korrektur der Melderegister berücksichtigt wurde, ist eine systematische Verzerrung aus diesem Grund auch beim Familienstand nicht zu erwarten. Als Erklärungsansatz bleiben daher unterschiedliche Angaben bei der Befragung und im Melderegister. Ursachen für unterschiedliche Angaben zum Familienstand können sowohl im subjektiven Verständnis der Befragten (Welchen Familienstand habe ich, wenn der geschiedene Partner verstorben ist? Sind getrennt Lebende subjektiv noch verheiratet?) als auch in verspäteten oder unvollständigen Meldungen an die Meldeämter (zum Beispiel zum Familienstandswechsel beim Sterben des Partners oder der Partnerin) liegen.

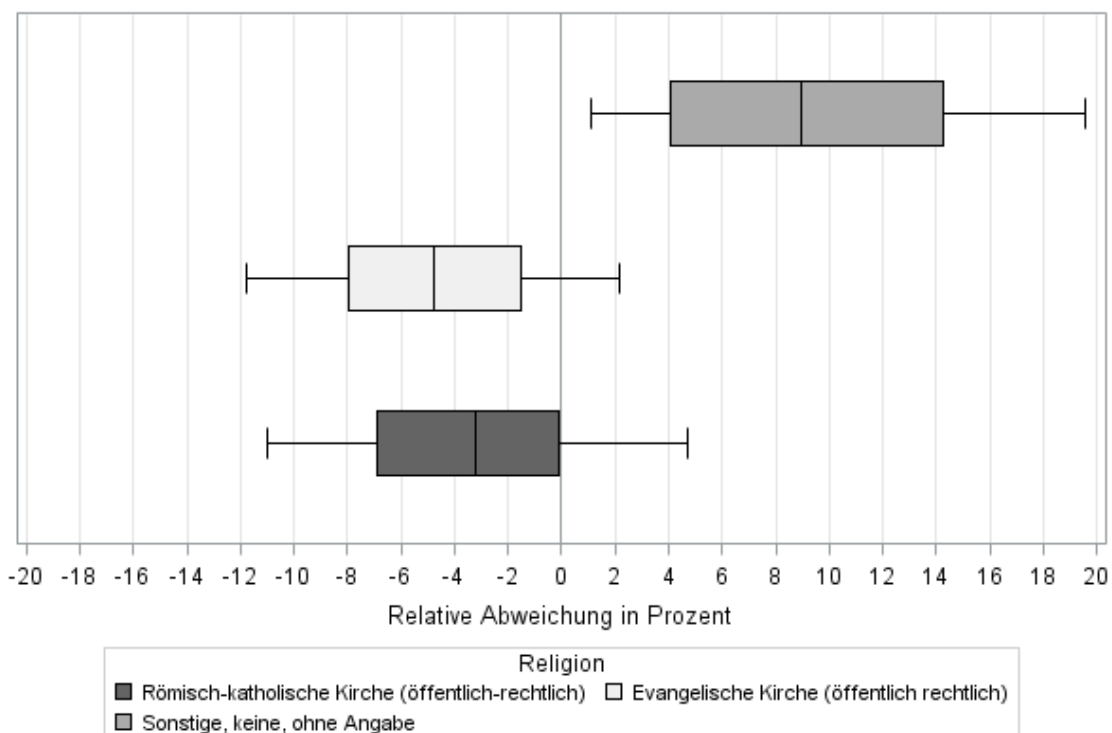
Abbildung 1.4 Relative Abweichung der Ergebnisse des Veröffentlichungstermins 2 des Zensus 2011 zu denen des Veröffentlichungstermins 1 in Prozent nach dem Familienstand:



Noch deutlicher als beim Familienstand sind die Unterschiede bei der Zugehörigkeit zu einer (öffentlich-rechtlichen) Religionsgesellschaft. Sowohl die römisch-katholische als auch die evangelische Kirche werden bei der Befragung sehr viel häufiger angegeben, als diese in den Melderegistern verzeichnet sind.

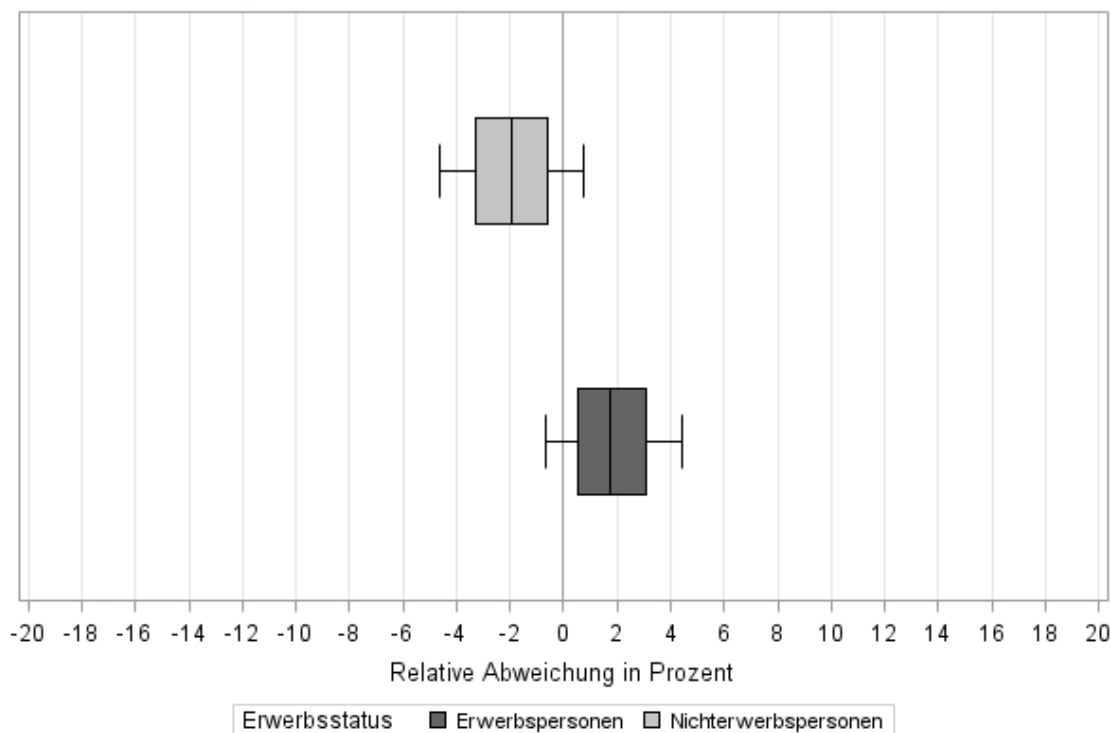
Der relativ große Unterschied zwischen ausgezählten und hochgerechneten Ergebnissen bei der Zugehörigkeit zu einer (öffentlich-rechtlichen) Religionsgesellschaft wirft Fragen auf. Gibt es einen nur formal aus steuerlichen Gründen vollzogenen Austritt aus einer Kirche, der man sich nach wie vor zugehörig fühlt? Wie häufig kommt dies vor? Wird die Angabe zur Frage nach der Religionszugehörigkeit bei der Anmeldung im Meldeamt vergessen? Gibt es Personen, die gegenüber der/dem Erhebungsbeauftragten nicht zugeben wollen, aus der Kirche ausgetreten zu sein? Eine Beantwortung dieser Fragen dürfte schwierig sein. Bezüglich Alter und Geschlecht zeigen sich beim Zensus 2011 keine Anhaltspunkte auf bestimmte Bevölkerungsgruppen, welche die unterschiedlichen Angaben zur Religion erklären könnten. Die Streuung der Ergebnisdifferenzen bei der Zugehörigkeit zu den beiden großen öffentlich-rechtlichen Kirchen ist groß. Möglicherweise ist dies ein Indiz für eine unterschiedliche Registerqualität bei diesem Merkmal.

Abbildung 1.5 Relative Abweichung der Ergebnisse des Veröffentlichungstermins 2 des Zensus 2011 zu denen des Veröffentlichungstermins 1 in Prozent nach der Zugehörigkeit zu einer (öffentlich-rechtlichen) Religionsgesellschaft:



Die Zahl der Erwerbspersonen nimmt unter der Einbeziehung der erwerbsstatistischen Register zum Veröffentlichungstermin 2 insgesamt leicht zu. Der „Baukasten“ zur Erwerbstätigkeit zählt zunächst alle in den erwerbsstatistischen Registern verzeichneten Personen als Erwerbspersonen. Erwerbstätige Personen außerhalb der erwerbsstatistischen Register (Selbstständige, mithelfende Familienangehörige, geringfügig Beschäftigte) werden aus der Haushaltsbefragung hinzugeschätzt. Sobald Angaben zur Erwerbstätigkeit entweder aus einem erwerbsstatistischen Register oder aus den Angaben zur Haushaltsbefragung vorliegen wird die Person im Zweifelsfall als Erwerbsperson gezählt. Das im Vergleich zur Hochrechnung aus der Haushaltsbefragung geringfügige Übergewicht an Erwerbspersonen aus dem „Baukasten“ ist daher erklärlich.

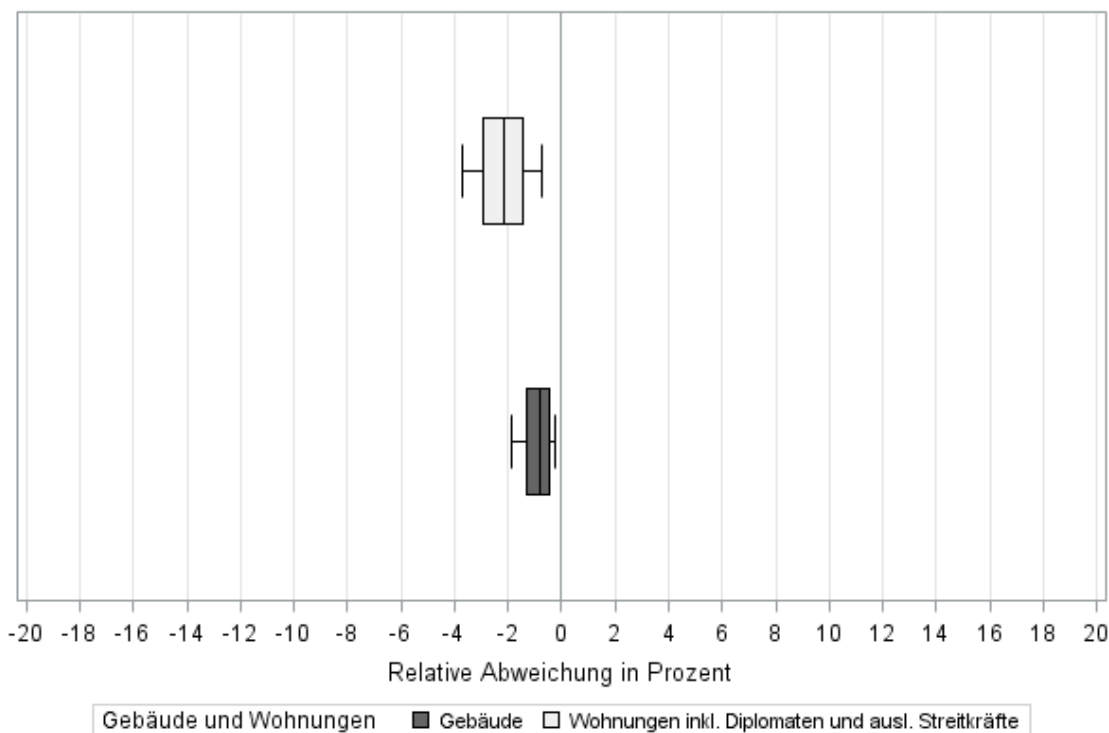
Abbildung 1.6 Relative Abweichung der Ergebnisse des Veröffentlichungstermins 2 des Zensus 2011 zu denen des Veröffentlichungstermins 1 in Prozent nach dem Erwerbsstatus:



Insgesamt führte die Berücksichtigung der Haushaltegenerierung bundesweit zu einer Reduzierung des Bestands um rund 140 000 Gebäude mit Wohnraum (0,8 Prozent) und mehr als 700 000 Wohnungen (1,8 Prozent; der Median der Abweichungen in der Zahl der Wohnung in Gemeinden ab 10 000 Einwohnern³ liegt bei 2,2 Prozent; vgl. Abbildung 1.7).

³ Zur besseren Vergleichbarkeit der relativen Abweichungen beruhen die Boxplots zu Gebäuden und Wohnungen wie, diejenigen zu den demografischen Merkmalen, auf den Gemeinden ab 10 000 Einwohnern, obwohl die Grenze von 10 000 Einwohnern für die Erhebung von Gebäuden und Wohnungen keine Bedeutung hat.

Abbildung 1.7 Relative Abweichung der Ergebnisse des Veröffentlichungstermins 2 des Zensus 2011 zu denen des Veröffentlichungstermins 1 in Prozent nach Gebäuden und Wohnungen:



12. Differenzen innerhalb des Veröffentlichungstermins 2

Demografische Merkmale (Alter, Geschlecht, Staatsangehörigkeit, Familienstand) sind bei den Datenquellen zum Zensus 2011 sowohl in den kommunalen Melderegistern als auch aus der Haushaltsbefragung vorhanden. Werden nur demografische Merkmale ausgewertet, erfolgt die Auswertung zum Veröffentlichungstermin 2 als Auszählung statistisch korrigierter Melderegister. Erfolgen kombinierte Auswertungen aus demografischen Merkmalen und Merkmalen, die nur über die Haushaltsbefragung erhoben wurden, z. B. nach Alter und Schulbildung, werden hierfür die Merkmalsausprägungen beider Merkmale aus der Haushaltsbefragung genommen und hochgerechnet. Zu beachten ist hierbei, dass Personen in Sonderbereichen, die nicht selbst befragt wurden (sogenannte sensible Sonderbereiche wie z. B. Justizvollzugsanstalten, psychiatrische Kliniken oder Meldeanschriften für Obdachlose, ferner Kasernen) bei diesen Auswertungen unberücksichtigt bleiben, was zu einer Verminderung der Gesamtzahl führt. Bundesweit umfasst dieser in den Hochrechnungen aus der Haushaltsbefragung unberücksichtigte Personenkreis etwa 0,7 Prozent der Gesamtbevölkerung.

Bei der Gegenüberstellung von ausgezählten und hochgerechneten demografischen Angaben zum Veröffentlichungstermin 2 erklärt der Ausschluss von Personen in den genannten Sonderbereichen im Wesentlichen die Ergebnisdifferenzen beim Geschlecht, beim Alter und bei der Staatsangehörigkeit nach „Deutsch“ und „Nicht deutsch“. Darüber hinaus sind leichte Verschiebungen zwischen den beiden unteren Altersklassen (unter 18 Jahre, 18-29 Jahre) sowie Deutschen und Nichtdeutschen erkennbar. Die Zahlen der 18-29-jährigen sowie der Nichtdeutschen sind als Ergebnis der Hochrechnung der Haushaltsbefragung niedriger als bei der Auszählung der Melderegister.

Abbildung 2.1 Relative Abweichung zwischen der Auszählung aus Registern und der Hochrechnung aus der Haushaltsbefragung in Prozent bei den Ergebnissen des Zensus 2011 (Veröffentlichungstermin 2) nach dem Geschlecht:

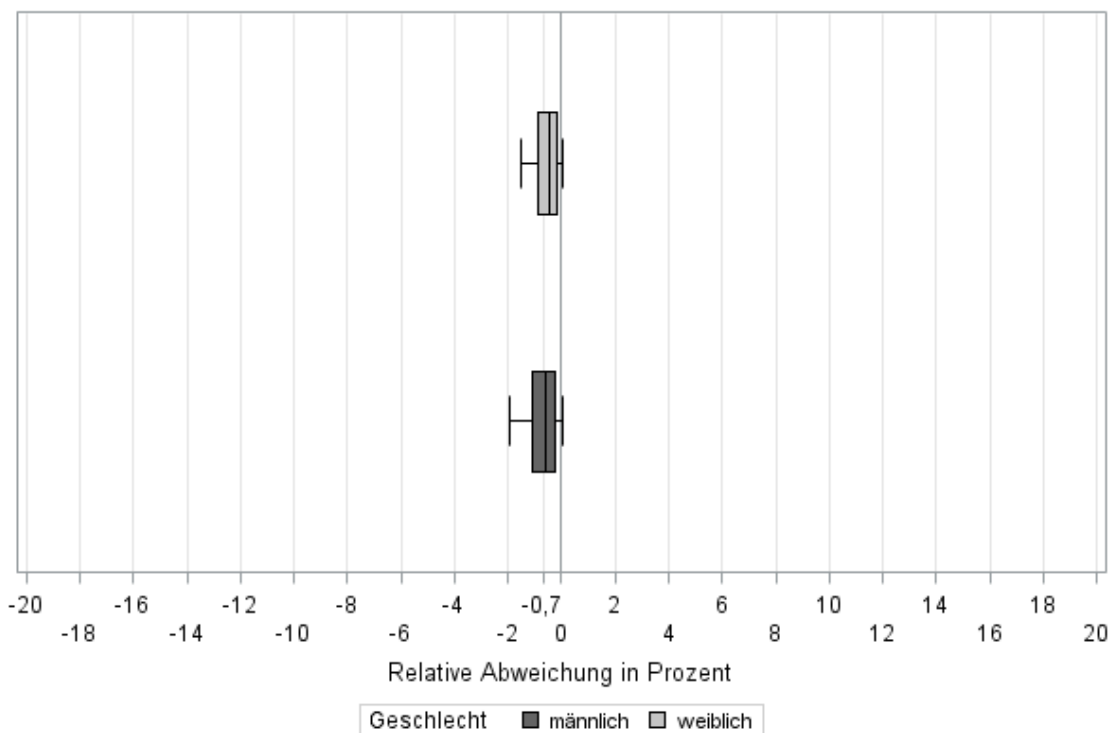


Abbildung 2.2 Relative Abweichung zwischen der Auszählung aus Registern und der Hochrechnung aus der Haushaltsbefragung in Prozent bei den Ergebnissen des Zensus 2011 (Veröffentlichungstermin 2) nach Altersklassen:

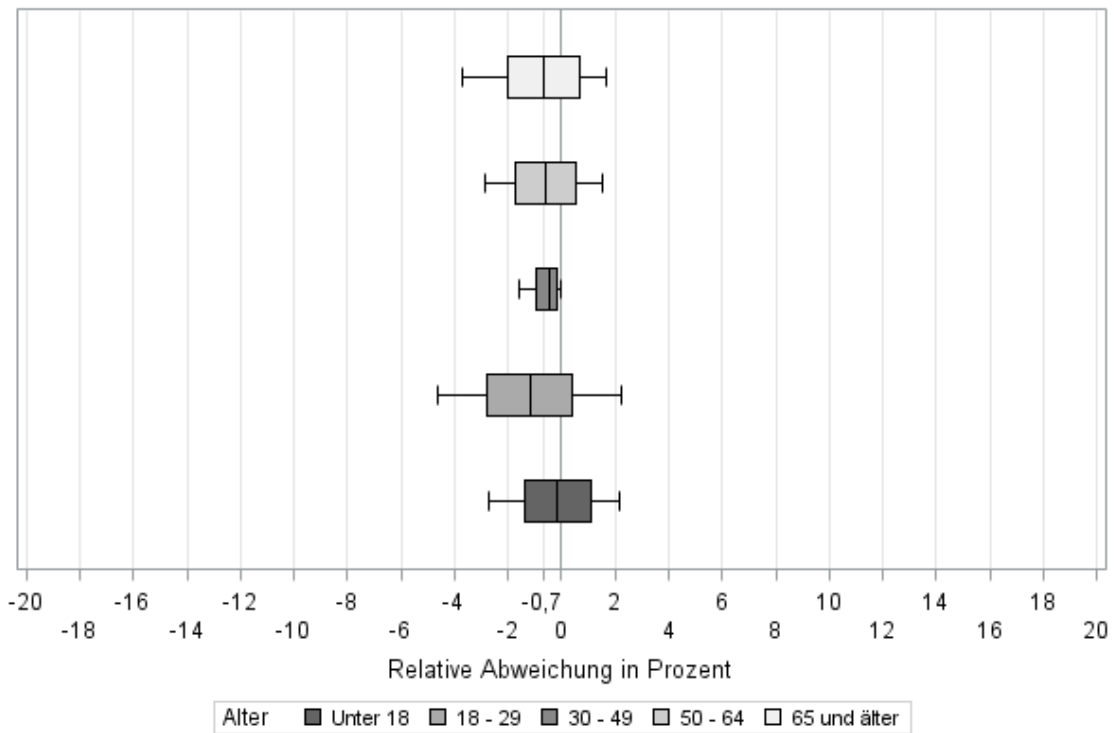
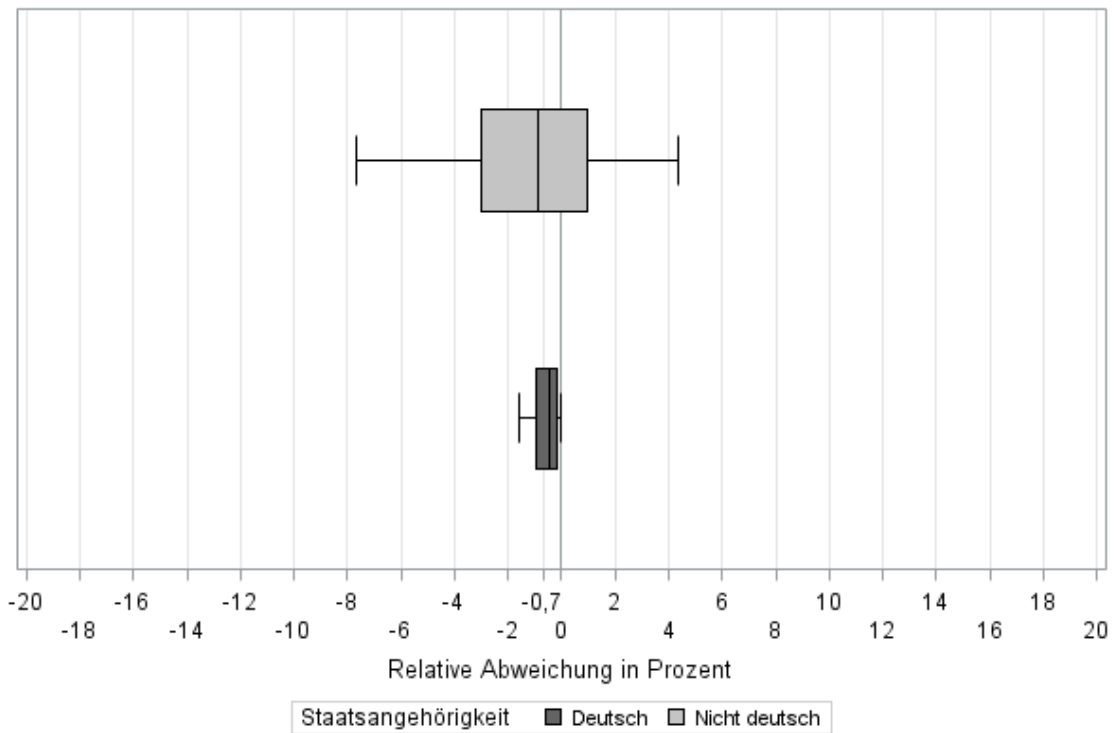
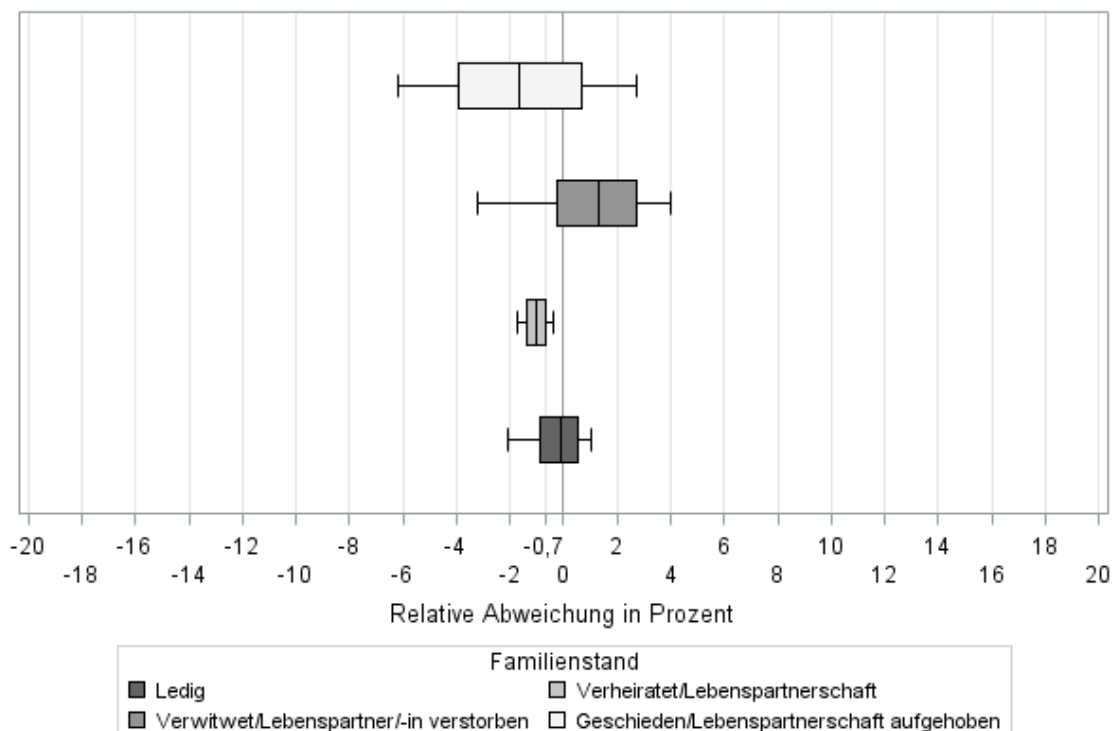


Abbildung 2.3 Relative Abweichung zwischen der Auszählung aus Registern und der Hochrechnung aus der Haushaltsbefragung in Prozent bei den Ergebnissen des Zensus 2011 (Veröffentlichungstermin 2) nach der Staatsangehörigkeit:



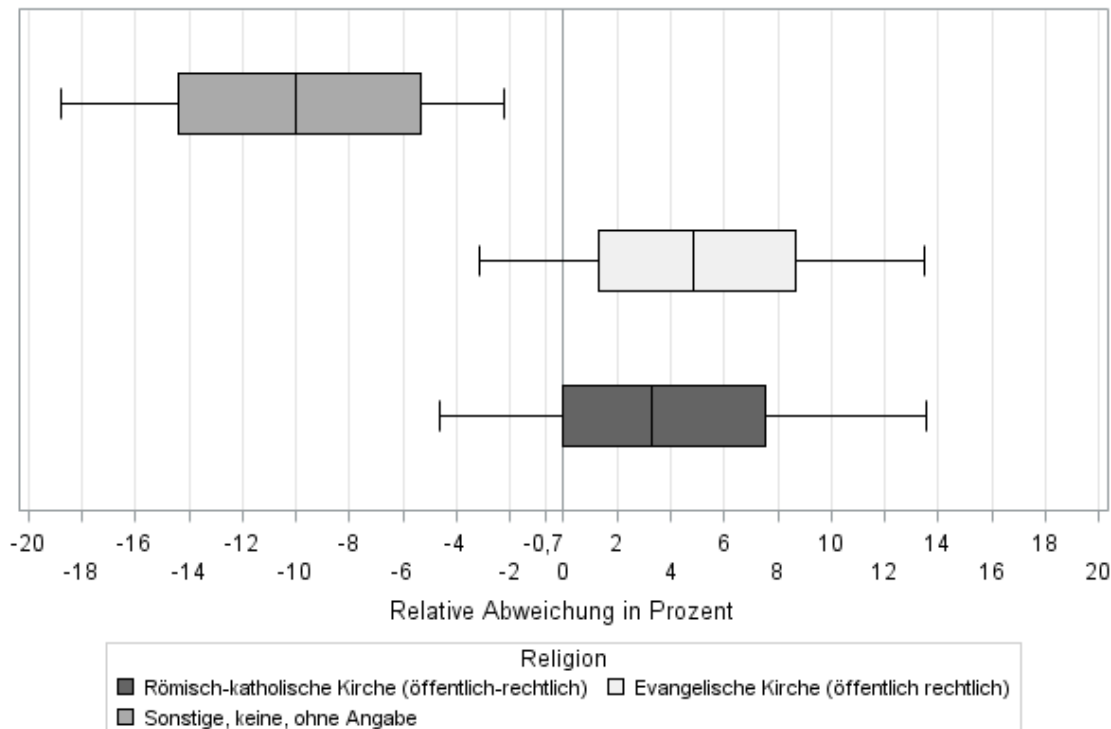
Beim Familienstand überlagern sich die Effekte des Wegfalls der nicht-befragten Bewohnerinnen und Bewohner bestimmter Sonderanschriften und des im Vergleich der Ergebnisse zwischen den Veröffentlichungsterminen 1 und 2 bereits auffällig gewordenen systematischen Unterschieds zwischen Befragungsergebnis und der Information aus dem Melderegister. Die im Mittel mehr bei „Null“ als bei „0,7 Prozent“ liegende Abweichung bei der Zahl lediger Personen ist hier ein Resultat sich gegenseitig aufhebender Effekte. Bei der Zahl der verwitweten Personen überwiegt der Effekt unterschiedlicher Angaben aus Befragung und Registern den Effekt des Wegfalls der nicht-befragten Personen. Hier weist die (reduzierte) Hochrechnung aus der Haushaltsbefragung mehr verwitwete Personen auf als die (vollständige) Auszählung des Melderegisters. Auf der anderen Seite verringert sich die Zahl der Verheirateten und Verwitweten bei der Hochrechnung in der Regel um mehr als die aufgrund der Nichtberücksichtigung der Bewohner(innen) bestimmter Sonderbereiche fehlenden 0,7 Prozent.

Abbildung 2.4 Relative Abweichung zwischen der Auszählung aus Registern und der Hochrechnung aus der Haushaltsbefragung in Prozent bei den Ergebnissen des Zensus 2011 (Veröffentlichungstermin 2) nach dem Familienstand:



Bei der Zugehörigkeit zu einer (öffentlich-rechtlichen) Religionsgesellschaft bleibt der Effekt einer deutlich höheren Zahl sowohl an Katholiken als auch an Protestanten aus der Haushaltsbefragung im Vergleich zum Melderegister bestehen. Der Effekt des Fehlens der nicht-befragten Bewohnerinnen und Bewohner bestimmter Sonderanschriften tritt gegenüber dem sehr viel größeren Unterschied der beiden im Zensus 2011 verwendeten Datenquellen Melderegister und Haushaltsbefragung zurück.

Abbildung 2.5 Relative Abweichung zwischen der Auszählung aus Registern und der Hochrechnung aus der Haushaltsbefragung in Prozent bei den Ergebnissen des Zensus 2011 (Veröffentlichungstermin 2) nach der Zugehörigkeit zu einer (öffentlich-rechtlichen) Religionsgesellschaft:




13. Fazit

Der Zensus 2011 hat in vielfacher Hinsicht Neuland betreten. Er hat dabei unterschiedliche Datenquellen genutzt und unterschiedliche Erhebungsverfahren eingesetzt. Damit hat sich ein breites Spektrum an Auswertungsmöglichkeiten ergeben. Daten aus unterschiedlichen Quellen und unterschiedlichen Erhebungsverfahren, wie zum Beispiel Melderegister und Haushaltsbefragung, wurden nach jeweils anderen Verfahren (Auszählung, Hochrechnung) ausgewertet.

Die Haushaltgenerierung führte unterschiedliche Datenbestände zusammen, was zu Korrekturen am Datenbestand, zum Beispiel bei Gebäuden und Wohnungen, geführt hat.

Mit einem auf den Zensus 2011 abgestimmten Verfahren der statistischen Geheimhaltung waren Auswertungen aus komplexen Datenquadrern möglich. Sperrungen von Tabellenfeldern mit Angaben zu ein oder zwei Fällen mit entsprechenden Gegensperrungen wurden vermieden, allerdings führen hohe Abweichungen der Werte nach Geheimhaltung vom Originalwert hier ebenfalls zu einer Unterdrückung von Angaben, wenn auch aus anderen Gründen.



Erste Ergebnisse wurden im Mai 2013 bereits vor Abschluss aller Arbeiten am Zensus 2011 ausgewertet und veröffentlicht. Die endgültigen Ergebnisse nutzten weitere Informationen, zum Beispiel aus der Haushaltegenerierung, und verwendeten zum Teil andere Datenquellen, zum Beispiel Auszählungen aus statistisch korrigierten Melderegistern anstatt Hochrechnungen aus der Haushaltsbefragung.

Kombinierte Auswertungen unterschiedlicher Merkmale, zum Beispiel nach der Schulbildung und dem Alter, nutzten teilweise andere Datenquellen (im Beispiel Alter gemäß Haushaltsbefragung anstatt Alter gemäß Melderegister) als Auswertungen eines Merkmals allein (im Beispiel dem Alter).

All dies führte zu unterschiedlichen Ergebnissen zum gleichen Sachverhalt.

Bei vielen Merkmalen (wie zum Beispiel dem Alter, dem Geschlecht und der Staatsangehörigkeit) sind die Differenzen aufgrund unterschiedlicher Datenquellen und Berechnungsverfahren in der Regel vernachlässigbar. Bei anderen Merkmalen gibt es dagegen auffällige systematische Differenzen zwischen den beiden Veröffentlichungsterminen oder zwischen Auszählungen aus Registern und Hochrechnungen aus der Haushaltsbefragung. Auffällig ist dies vor allem bei

- den Gebäuden und Wohnungen,
- dem Familienstand,
- der Zahl der Erwerbspersonen und
- der Religionszugehörigkeit.

Gründe der Abweichungen liegen bei den Gebäuden und Wohnungen aus Erkenntnissen der Haushaltegenerierung, beim Familienstand und der Religionszugehörigkeit in unterschiedlichen Angaben der Befragten und der Register, bei der Zahl der Erwerbspersonen aus dem Zusammenzählen von Erwerbsregistern und hochgerechneten Angaben, die eine Person im Zweifelsfall als Erwerbsperson zählen.

Bezüglich dieser Differenzen wirft der Zensus 2011 interessante Fragen auf. Zu deren Klärung wären weitere Analysen erforderlich.

Literatur

Eppmann, H.; Schäfer, J. (2006): *Der Weg zu einem ersten registergestützten Zensus in Deutschland im Jahre 2011* in Statistische Analysen und Studien Nordrhein-Westfalen, Band 27

Gießing, S.; Heinzl, F.; Kleber, B.; Wilke, A. (2014): *Geheimhaltung beim Zensus 2011* in Wirtschaft und Statistik 11/2014

Grunwald, S.; Krause, A. (2014): *Umgang mit fehlenden Angaben in der Gebäude- und Wohnungszählung 2011* in Wirtschaft und Statistik 8/2014

Hofmeister, K.; Fürnrohr, M. (2014): *Das Korrekturverfahren beim Zensus 2011 in Bayern* in Zahlen 6/2014

Höhne, J. (2011): *SAFE – A Method for Anonymizing the German Census*, präsentiert auf der Joint UNECE/Eurostat Work Session on Statistical Data Confidentiality, Tarragona, Spanien 26. - 28. Oktober 2011

Kreuzmair, I. Reisch, M. (2012): *Zensus 2011: Ablauf der Haushaltegenerierung in Bayern* in Zahlen 9/2012

Münnich, R.; Gabler, S.; Ganninger, M.; Burgard, P.; Kolb, J.-P. (2012): *Stichprobenoptimierung und Schätzung im Zensus 2011* in Statistik und Wissenschaft, Band 21, Herausgeber: Statistisches Bundesamt

Schäfer, J. (2004): *Ergänzende Verfahren für einen künftigen registergestützten Zensus* in Statistische Analysen und Studien Nordrhein-Westfalen, Band 17

Scholze, S. (1996): *Gebäude- Wohnungszählung 1995 in den neuen Bundesländern und im Ostteil Berlins* in Statistisches Monatsheft April 1996, Thüringer Landesamt für Statistik

Sedmihradsky, D.; Eisenmenger, M.; Burmeister, F. (2012): *Erhebung, Aufbereitung und Zusammenführung der Erwerbsregisterdaten im Zensus 2011* in Wirtschaft und Statistik, 12/2012

Statistische Ämter des Bundes und der Länder (2004): *Ergebnisse des Zensustests*, u. a. in Statistische Analysen und Studien Nordrhein-Westfalen, Band 17 oder Wirtschaft und Statistik 8/2004

Diplom-Statistiker Josef Schäfer ist Projektleiter des Projektes „Zensus 2021“ im Geschäftsbereich Statistik bei Information und Technik Nordrhein-Westfalen (IT.NRW).