DESTATIS
Statistisches Bundesamt

WISTA Wirtschaft und Statistik

Sonderheft Zensus 2021

Stefan Dittrich	**Der registergestützte Zensus 2021**
Corinna Bretschi | Dr. Kai Lorentz	**Präzisionsziele für die Ermittlung der Einwohnerzahl bei der Haushaltsstichprobe im Zensus 2021**
Dr. Jan Pablo Burgard | Prof. Dr. Ralf Münnich | Dr. Martin Rupp	**Die Entwicklung des Stichprobenkonzepts für den Zensus 2021**
Dr. Nilgun Massih-Tehrani | Adrian Reichert	**Die Haushaltegenerierung im Zensus 2021**
Benjamin Freier | Juliane Mosel	**Online First als Leitgedanke für effiziente Primärerhebungen beim Zensus 2021**
Thorsten Tümmler | Irina Meinke	**Aufbau des Qualitätsmanagements für den Zensus 2021**
Thomas Körner | Anja Krause | Kathrin Ramsauer	**Anforderungen und Perspektiven auf dem Weg zu einem künftigen Registerzensus**

S | 2019

wissen.nutzen.

ABKÜRZUNGEN

D	Durchschnitt (bei nicht addierfähigen Größen)
Vj	Vierteljahr
Hj	Halbjahr
a. n. g.	anderweitig nicht genannt
o. a. S.	ohne ausgeprägten Schwerpunkt
Mill.	Million
Mrd.	Milliarde

ZEICHENERKLÄRUNG

–	nichts vorhanden
0	weniger als die Hälfte von 1 in der letzten besetzten Stelle, jedoch mehr als nichts
.	Zahlenwert unbekannt oder geheim zu halten
. . .	Angabe fällt später an
X	Tabellenfach gesperrt, weil Aussage nicht sinnvoll
I oder —	grundsätzliche Änderung innerhalb einer Reihe, die den zeitlichen Vergleich beeinträchtigt
/	keine Angaben, da Zahlenwert nicht sicher genug
( )	Aussagewert eingeschränkt, da der Zahlenwert statistisch relativ unsicher ist
	Abweichungen in den Summen ergeben sich durch Runden der Zahlen.
	Tiefer gehende Internet-Verlinkungen sind in der Online-Ausgabe hinterlegt.

INHALT

3	Editorial
5	Stefan Dittrich **Der registergestützte Zensus 2021** *The register-based 2021 Census*
12	Corinna Bretschi, Dr. Kai Lorentz **Präzisionsziele für die Ermittlung der Einwohnerzahl bei der Haushaltsstichprobe im Zensus 2021** *Precision targets for determining the number of inhabitants in the household sample of the 2021 Census*
23	Dr. Jan Pablo Burgard, Prof. Dr. Ralf Münnich, Dr. Martin Rupp **Die Entwicklung des Stichprobenkonzepts für den Zensus 2021** *Developing the sample design for the 2021 Census*
36	Dr. Nilgun Massih-Tehrani, Adrian Reichert **Die Haushaltegenerierung im Zensus 2021** *The household generating procedure of the 2021 Census*
46	Benjamin Freier, Juliane Mosel **Online First als Leitgedanke für effiziente Primärerhebungen beim Zensus 2021** *Online first as the guiding principle for efficient primary surveys as part of the 2021 Census*
59	Thorsten Tümmler, Irina Meinke **Aufbau des Qualitätsmanagements für den Zensus 2021** *Building a quality management system for the 2021 Census*
74	Thomas Körner, Anja Krause, Kathrin Ramsauer **Anforderungen und Perspektiven auf dem Weg zu einem künftigen Registerzensus** *Towards a future register census – demands and prospects*

EDITORIAL

Dr. Georg Thiel

LIEBE LESERIN, LIEBER LESER,

am 19. September 2018 hat das Bundesverfassungsgericht den Zensus 2011 für rechtmäßig erklärt (BVerfG - 2 BvF 1/15 -). Das Verfahren und das Urteil haben die große Tragweite der verbindlich festgestellten Einwohnerzahlen verdeutlicht und die Arbeit der amtlichen Statistik bestätigt. Der Anspruch an die Einwohnerzahl ist für die Vergangenheit und für die Zukunft klar benannt: qualitativ hochwertig, was vom Gericht als hinreichend genau und zuverlässig definiert wird, und durch eine grundrechtsschonende Datenerhebung geprägt.

Auf dieser Grundlage und unter Berücksichtigung der Zielsetzung der Statistischen Ämter des Bundes und der Länder, die Ergebnisse des Zensus 2021 schon 18 Monate nach Stichtag zu veröffentlichen, wurden die bereits fortgeschrittenen Arbeiten für den Zensus 2021 mit Hochdruck weiterverfolgt.

Die Zustimmung des Bundestags zum Zensusdurchführungsgesetz 2021 (ZensG 2021) am 6. Juni 2019 nehmen wir zum Anlass, ein Sonderheft zum Zensus 2021 herauszugeben und Einblicke in die laufenden Arbeiten sowie das Qualitätsmanagement für den Zensus 2021 zu ermöglichen.

Drei Artikel beschäftigen sich konkret mit der Methodik des registergestützten Zensus. So werden die Präzisionsziele für die Ermittlung der Einwohnerzahl bei der Haushaltsstichprobe 2021 näher erläutert, die Stichprobenmethodik insgesamt dargestellt und die Veränderungen der Haushaltegenerierung beim Zensus 2011 gegenüber dem Zensus 2021 erklärt.

Neu beim Zensus 2021 ist der Grundsatz Online First bei der Datenerhebung. Er gilt sowohl bei der Gebäude- und Wohnungszählung als auch bei den Personenerhebungen für die sozio-demografischen Zusatzmerkmale. Für die Auskunftgebenden bedeutet dies insbesondere Zeitersparnis, Zeitsouveränität und einen größeren Befragungskomfort. Das Vorgehen ist grundrechtsschonender als das bisherige und wird zu schnelleren und plausibleren Ergebnissen führen.

Ein Blick Richtung Österreich weist den Weg für den Zensus 2031. Das Ziel ist, registerbasiert Bevölkerungszahlen zu ermitteln. Wie dieses Verfahren in Deutschland aussehen könnte und welche Voraussetzungen hierfür geschaffen werden müssen, zeigt der Artikel zum Integrierten Registerzensus.

Präsident des Statistischen Bundesamtes

DER REGISTERGESTÜTZTE ZENSUS 2021

Weiterentwicklung des Zensus 2011 und weitere Schritte hin zu einem registerbasierten Zensus

Stefan Dittrich

Stefan Dittrich

studierte Volkswirtschaftslehre in Göttingen und arbeitet seit 2000 in verschiedenen Funktionen und Fachbereichen im Statistischen Bundesamt. Seit 2014 leitet er die Gruppe „Zensus“ und ist in dieser Funktion fachlicher Projektleiter für den Zensus 2021.

Schlüsselwörter: Zensus 2021 – Online-First – Zensusgesetz – Melderegister – Stichprobe

ZUSAMMENFASSUNG

Deutschland beteiligt sich mit dem Zensus 2021 an der international vereinbarten Volkszählungsrunde 2020/2021. Der wiederum registergestützte Zensus 2021 baut inhaltlich auf dem Zensus 2011 auf. Aufgrund der Erkenntnisse des letzten Zensus gibt es Änderungen insbesondere hinsichtlich der Korrekturstichprobe zur Ermittlung der Einwohnerzahl in kleinen Gemeinden, der Zielsetzung Online First als primärem Erhebungsweg und der IT-technischen Umsetzung des Projekts.

Keywords: *2021 Census – Online First – Census Act – population register – sample survey*

ABSTRACT

Preparing the 2021 Census, Germany will participate in the internationally agreed round of censuses in 2020/2021. The 2021 Census will again be register-based and its content will be based on the 2011 Census. Taking account of the knowledge gained in the last census, there will be changes especially with regard to the corrective sample survey to be conducted to determine the number of inhabitants in small municipalities, the goal of Online First as the main survey method, and the project implementation in terms of IT.

1

Hintergrund des Zensus 2021

Der Zensus ist das größte Projekt der amtlichen Statistik zur Erhebung von Bevölkerungs-, Gebäude- und Wohnungsdaten. Der nächste Zensus wird aufgrund europarechtlicher Vorgaben im Jahr 2021 (Stichtag in Deutschland: 16. Mai 2021) stattfinden. Beim letzten Zensus im Jahr 2011 gab es zu den bisher in Deutschland durchgeführten Volkszählungen einen grundlegenden Methodenwechsel. Die traditionelle Form einer Vollbefragung der Bevölkerung wurde – vor allem aus Kosten- und Akzeptanzgründen – durch ein überwiegend registergestütztes Verfahren ersetzt. In diesem Verfahren wurden vorhandene Daten aus Verwaltungsregistern, unter anderem der Melderegister, genutzt und insbesondere mit einer ergänzenden Haushaltebefragung auf Stichprobenbasis, einer Erhebung an Anschriften mit Sonderbereichen, einer Gebäude- und Wohnungszählung sowie einer automatisierten Haushaltegenerierung kombiniert.

Die Vorbereitungsarbeiten der Statistischen Ämter des Bundes und der Länder für den Zensus 2021 haben unmittelbar nach Abschluss der Aufbereitung des Zensus 2011 begonnen. Dabei wurden sowohl die Erfahrungen des Zensus 2011 als auch Vorhaben in anderen Ländern und technische Neuerungen bei der methodischen Vorbereitung berücksichtigt. Von Nutzerseite wurden insbesondere Wünsche nach einer früheren Bereitstellung der Ergebnisse geäußert, auch die steigende Nachfrage nach kleinräumigen Zensusergebnissen floss in die Konzeption ein. Mit dem Urteil des Bundesverfassungsgerichts zur Normenkontrollklage der Städte Berlin und Hamburg liegt seit dem Herbst 2018 auch in rechtlicher Hinsicht Klarheit über die Anforderungen an die Ausgestaltung des Zensus vor (Bierschenk/Leischner, 2019).

2

Rechtliche Grundlagen

Die Rahmenverordnung für die gemeinschaftsweite Volks- und Wohnungszählung 2011 [Verordnung (EG) Nr. 763/2008] gilt für 2021 unverändert weiter. In den korrespondierenden Durchführungsverordnungen (Nr. 543/2017, Nr. 712/2017, Nr. 881/2017, Nr. 779/2018) wurden lediglich Anpassungen vorgenommen hinsichtlich der Merkmalsausprägungen und der an die Europäische Union zu liefernden regionalen Gliederungen. Der europarechtlich geforderte Merkmalskranz ist gegenüber dem Zensus 2011 unverändert geblieben.

Aufgrund der langen Vorbereitungszeit und der zu Beginn der Arbeiten noch ausstehenden Entscheidung des Bundesverfassungsgerichts wurde die nationale rechtliche Grundlage wie bereits beim Zensus 2011 in zwei Gesetze aufgeteilt, das Zensusvorbereitungsgesetz 2021 sowie das Zensusgesetz 2021.

Im März 2017 ist das Zensusvorbereitungsgesetz 2021 in Kraft getreten. Abweichend zum Zensus 2011 wird in diesem Gesetz die Verantwortung für die Entwicklung der für den Zensus benötigten technischen Anwendungen dem Statistischen Bundesamt übertragen sowie die Verantwortung für den IT-Betrieb in Zusammenarbeit mit dem Informationstechnikzentrum Bund geregelt. Mit dieser Regelung wurden zwei zentrale Empfehlungen der Evaluation des Zensus 2011 umgesetzt:

»*1. Die Aufteilung der IT-Verantwortlichkeiten auf das Statistische Bundesamt und einzelne Statistische Landesämter nach § 12 Abs. 7 Zensusgesetz 2011 hat Probleme bereitet und sollte in dieser Form nicht wiederholt werden.*

2. Stattdessen sollten alle systemrelevanten IT-Anwendungen in der Verantwortung des Statistischen Bundesamtes entwickelt und betrieben werden.« (BMI, 2016)

Aufgrund von Änderungen im XMeld-Standard[1] der Meldebehörden und neuer Konzepte wurde mit dem Zensusvorbereitungsänderungsgesetz eine Pilotdatenlieferung der Melderegisterdaten zum Januar 2019 angeordnet.

Das Zensusgesetz 2021 wurde vom Bundestag in seiner Sitzung vom 6. Juni 2019 beschlossen und dem Bundesrat zugeleitet. Aufgrund des Antrags auf Anrufung des Vermittlungsausschusses durch den Bundesrat wird sich das Inkrafttreten jedoch voraussichtlich bis Herbst 2019 verzögern. Mit dem Gesetz werden Umfang und Erfassung der zu erhebenden Daten sowie eine Reihe von Bestimmungen zu organisatorischen Fragen des

1 XMeld ist ein auf XML basierendes standardisiertes Datenaustauschformat für den Austausch von Meldedaten in der öffentlichen Verwaltung.

Zensus geregelt. So wird beispielsweise in § 23 Zensusgesetz 2021 angeordnet, dass die Auskunftserteilung grundsätzlich elektronisch erfolgen soll.

Aufgrund des durch § 13 Absatz 2 Bundesstatistikgesetz angeordneten dauerhaften Anschriftenregisters können für den Zensus 2021 auch Angaben zu den Anschriften aus dem Zensus 2011 verwendet werden; die Ergebnisse des Zensus 2021 zu den Anschriften können in dieses Register überführt werden.

3

Methodische Änderungen gegenüber dem Zensus 2011

3.1 Ermittlung der Einwohnerzahl

Beim Zensus 2011 wurden unterschiedliche Verfahren zur Überprüfung und Korrektur der Angaben aus den Melderegistern angewendet. In allen Gemeinden gab es – unabhängig von ihrer Einwohnerzahl – eine Bereinigung um unzulässige Dubletten (Mehrfachfallprüfung) sowie eine Vollerhebung an sogenannten Sonderbereichen (Gemeinschaftsunterkünfte und Wohnheime). In großen Gemeinden mit mindestens 10 000 Einwohnerinnen und Einwohnern wurden darüber hinaus die aus Registern stammenden Daten mithilfe einer Stichprobe bereinigt. Bei dieser sogenannten Korrekturstichprobe wurden etwa 10 % aller in Deutschland lebenden Personen persönlich befragt. Die dabei ermittelten sogenannten Fehlbestände und Karteileichen wurden anschließend auf die gesamte Gemeinde hochgerechnet. In kleinen Gemeinden unter 10 000 Einwohnerinnen und Einwohnern gab es eine solche Stichprobe nicht. Hier fand stattdessen eine Befragung zur Klärung von Unstimmigkeiten zwischen den Angaben aus den Melderegistern und den Angaben der Gebäude- und Wohnungszählung statt.

Mit den Ergebnissen des Zensus 2011 und deren Evaluierung liegen heute jedoch Erkenntnisse vor, dass in kleinen Gemeinden der tatsächliche, ex post festgestellte Korrekturbedarf 2011 höher war, als vorab aufgrund des Zensustests 2001 erwartet (Bretschi/Lorentz, 2019).

Deshalb wird es beim Zensus 2021 die Korrekturstichprobe zur Ermittlung der Einwohnerzahl generell in allen Gemeinden geben. Die Frage, wie die Stichprobe in kleinen Gemeinden durchgeführt werden soll, wurde intensiv und in unterschiedlichen Varianten untersucht. Würde auch für alle kleinen Gemeinden eine Präzision der Einwohnerzahl von 0,5 % angestrebt, müsste – rein rechnerisch – die Einwohnerzahl in Gemeinden mit weniger als 100 Einwohnerinnen und Einwohnern auf Bruchteile einer Person genau ermittelt werden. Eine solche Forderung ist aus fachlicher ebenso wie aus wissenschaftlicher Sicht weder sinnvoll noch hinsichtlich der Belastung der Auskunftspflichtigen vertretbar. Daher wird das angestrebte prozentuale Präzisionsziel in kleinen Gemeinden im Vergleich zu großen Gemeinden gelockert werden. Als Folge wird es in kleinen Gemeinden geringere Auswahlsätze als bei Forderung eines einheitlichen Präzisionsziels geben. Das wird sowohl die Belastung der Auskunftspflichtigen als auch die Kosten begrenzen und entspricht dem Ziel einer grundrechtsschonenden Erhebung. Ein gelockertes Präzisionsziel bei kleinen Gemeinden bedeutet bei einer verzerrungsfreien Hochrechnung allerdings keine „schlechtere“ Einwohnerzahl, sondern lediglich ein höheres Risiko, dass die ermittelte Einwohnerzahl von der tatsächlichen stärker abweicht. Dieses erhöhte Risiko ist aber immer in beide Richtungen symmetrisch, das heißt das Restrisiko einer um mehr als xy % zu niedrigen Einwohnerzahl geht immer einher mit einem gleich hohen Restrisiko einer um xy % zu hohen Einwohnerzahl. Um den Stichprobenumfang weiter zu begrenzen haben die Länder gleichzeitig die Möglichkeit, verbandsangehörige Gemeinden zu Gemeindeverbänden zusammenzufassen und das Präzisionsziel auf Gemeindeverbandsebene anzustreben. Dies gilt analog für Zusammenschlüsse verbandsangehöriger Gemeinden mit weniger als 2 000 Einwohnerinnen und Einwohnern. Bedingung ist eine landesrechtliche Fundierung der Gemeindeverbände und eine gemeinsame Melderegisterführung. Auch bei dieser Option werden durch die Anwendung geeigneter Verfahren Einwohnerzahlen für jede einzelne Gemeinde bereitgestellt werden.

Diese Änderung gegenüber 2011 bedeutet insbesondere, dass in allen Gemeinden unabhängig von ihrer Größe ein einheitliches methodisches Instrument – die Korrekturstichprobe – zur Ermittlung der Einwohnerzahl zum Einsatz kommt.

Die Untersuchungen zum Stichprobenumfang wurden wie bereits beim Zensus 2011 in Zusammenarbeit mit Professor Dr. Ralf Münnich und seinem Team der Universität Trier erarbeitet (Burgard und andere, 2019). Weitere wissenschaftliche Untersuchungen zur Vorbereitung des Zensus 2021 waren nicht in gleichem Umfang wie beim Zensus 2011 notwendig, da das Modell in weiten Teilen unverändert blieb.

3.2 Online First

Online First bedeutet beim Zensus 2021, dass Befragte im Rahmen der Personenerhebung sowie der Gebäude- und Wohnungszählung Fragebogen standardmäßig selbstständig online ausfüllen. Die Online-Zugangsdaten (Zugangsnummer und Aktivierungscode) werden bei der Gebäude- und Wohnungszählung den Befragten postalisch zugestellt. Bei der Personenerhebung in Privathaushalten werden sie von den Erhebungsbeauftragten persönlich übergeben, nachdem diese die für die reine Existenzfeststellung zur Einwohnerzahl notwendigen Merkmale erhoben haben. Sowohl bei der Gebäude- und Wohnungszählung als auch bei der Haushaltebefragung werden Papierfragebogen nur noch in Ausnahmefällen genutzt. Mit Ausnahme der Fragen zur Existenzfeststellung füllen die Befragten den Fragebogen selbstständig online aus (Freier/Mosel, 2019).

Der Erhebungsweg, also ob die Angaben auf Papier oder in einem Online-Formular erfolgen und wie dies konkret ausgestaltet wird, hat Auswirkungen sowohl für die statistischen Ämter als auch für die Befragten.

Für die statistischen Ämter bestehen aus methodischer Sicht die Vorteile darin, dass von vornherein plausibilisierte Daten erhoben werden, die Ergebnisse schneller zur Verfügung stehen und das Verfahren ressourcenschonender und kostengünstiger ist. Die Vorteile für die Befragten liegen insbesondere in einem geringeren Eingriff in das Recht auf informationelle Selbstbestimmung, in der Zeitersparnis und Zeitsouveränität sowie dem größeren Befragungskomfort.

Die Online-First-Strategie für den Zensus 2021 basiert unter anderem auf der Erkenntnis, dass bereits seit 2016 der flächendeckende Zugang zum Internet in der Bevölkerung vorausgesetzt werden kann. Erhebungen zeigen, dass seit diesem Zeitpunkt das Smartphone in Deutschland das am meisten verwendete Gerät für die Internetnutzung ist. Darauf hat sich die amtliche Statistik zwischenzeitlich eingestellt: Bis 2021 sollen die Online-Fragebogen des Zensus so bereitstehen, dass sie auch auf Smartphones bedient werden können, daneben aber natürlich weiterhin auf PC, Laptop oder Tablet.

3.3 Weitere Neuerungen

Bei dem aus mehreren primär- und sekundärstatistischen Quellen bestehenden Zensus ist das Zusammenspiel der einzelnen Quellen von entscheidender Bedeutung. Die Schnittstellen zwischen den Erhebungsteilen beziehungsweise Daten und Prozessen wurden für ein besseres Ineinandergreifen früher und klarer definiert.

Der Aufbau des Registers aller Anschriften mit Gebäuden mit Wohnraum hat zeitiger begonnen als beim letzten Zensus. Auf die Integration der Daten der Bundesagentur für Arbeit wurde dabei verzichtet, da diese Daten keinen Mehrwert für die Ermittlung von Anschriften mit Wohnraum erbracht hatten.

Zwar hat sich das Verfahren zum Erwerbsregister im Zensus 2011 hinsichtlich der Ergebnisqualität bewährt, allerdings war die Auswertung durch die Zuschätzung nicht im Erwerbsregister abgebildeter Teilpopulationen (zum Beispiel Selbstständige, Nichterwerbspersonen) aus der Haushaltsstichprobe sehr komplex. Sie brachte – da die Nutzerinnen und Nutzer in aller Regel an einem erwerbsstatistischen Gesamtbild interessiert waren – keinen Mehrwert gegenüber einer reinen Hochrechnung. Auch hier wird deshalb auf die Nutzung von Registerdaten zugunsten einer einheitlichen Auswertung der Haushaltebefragung verzichtet.

Bei der Hochrechnung des Korrekturbedarfs der Melderegister aus der Stichprobe wird ein erhöhtes Augenmerk darauf gerichtet, dass möglichst keine Verzerrungen in den demografischen Untergliederungen auftreten.

4

Aktueller Stand der Arbeiten

Im Jahr 2019 liegt der Fokus der Arbeiten der Statistischen Ämter des Bundes und der Länder zum einen auf den Analysen der Melderegisterdaten anhand der im Januar 2019 übermittelten Pilotdaten. Weitere Schwerpunkte sind zum anderen die Begleitung der Gesetzgebungsverfahren, die Pretests zur Haushaltebefragung und zur Gebäude- und Wohnungszählung sowie die technische Vorbereitung der Erhebungen.

4.1 Pilotdatenlieferung Melderegister

Die Pilotdatenlieferung der Melderegister im Januar 2019 umfasste zu Testzwecken den maximalen Merkmalskranz. Die technische Übermittlung der etwa 90 Millionen Datensätze aus den kommunalen Melderegistern verlief reibungslos. Der auf nur zehn Tage verkürzte Lieferzeitraum konnte hingegen nicht von allen Gemeinden eingehalten werden. Wichtige Erkenntnisse konnten insbesondere aufgrund von nicht auflösbaren Doppellieferungen gewonnen werden. Hier werden für die Echtdatenlieferungen entsprechende organisatorische Maßnahmen zu ergreifen sein. Mit ersten erfolgreichen Tests des für den Zensus 2021 vorgesehenen Konzepts des erweiterten Personenkreises konnte ein Ziel der Testdatenlieferung bereits bearbeitet werden. Das Konzept dient der präziseren Abbildung der aufgrund von Bearbeitungszeiten in den Melderegistern bestehenden Unschärfen insbesondere durch An- und Abmeldungen rund um den Stichtag. Weitere umfangreiche Untersuchungen zu den einzelnen Merkmalen und zur Verwendung der Daten für die Weiterentwicklung der Mehrfachfallprüfung und der Haushaltegenerierung wurden bereits begonnen.

4.2 Quantitative Pretests

Neben den bereits durch das Statistische Bundesamt durchgeführten qualitativen Pretests, das heißt Untersuchungen unter Laborbedingungen, sollen im Herbst 2019 umfangreiche quantitative Pretests durch die Statistischen Ämter des Bundes und der Länder durchgeführt werden. Die Pretests sollen möglichst alle Prozesse der späteren Befragung realitätsnah simulieren, wobei der Schwerpunkt auf der Verständlichkeit der Fragen und der Nutzungsfreundlichkeit der Befragungsinstrumente liegt. Gerade auf die elektronischen Fragebogen, die gemäß der Online-First-Strategie im Jahr 2021 als Standard angesehen werden, sind die Erfahrungen aus dem Zensus 2011 kaum mehr übertragbar. Zu sehr haben sich die Geräte für den Internetzugang verändert – vom Desktop-PC hin zum Smartphone – und damit auch die erwarteten Bedienkonzepte. Als Standardverfahren müssen die elektronischen Fragebogen im Jahr 2021 von einem sehr breiten Personenkreis bedient werden können, also auch von Personen, die weniger geübt sind bei der Nutzung von Online-Diensten.

4.3 Technische Vorbereitung der Erhebungen

Für die verschiedenen Erhebungs- und Aufbereitungsmodule des Zensus 2021 müssen die notwendigen IT-Anwendungen konzipiert werden. Aufgrund der Komplexität des Zensusmodells in Deutschland handelt es sich dabei um zahlreiche umfangreiche und leistungsfähige Neuentwicklungen. Hierbei stellen nicht zuletzt die sicherheitstechnischen Anforderungen zur Bearbeitung sensibler personenbezogener Daten eine besondere Herausforderung dar.

5

Zielsetzung der Statistischen Ämter des Bundes und der Länder

Ein Ziel der Statistischen Ämter des Bundes und der Länder für den Zensus 2021 ist die im Vergleich zum Zensus 2011 deutlich früher vorgesehene Bereitstellung der Ergebnisse. Vor dem Hintergrund des im Zeitverlauf kontinuierlich nachlassenden „Wertes“ der Daten ist das Ziel, 18 Monate nach Stichtag die Daten zu veröffentlichen, nicht nur anzustreben, sondern auch zu realisieren. Angesichts der hohen Kosten des Zensus muss eine maximale Nutzung möglichst aktueller Daten im Mittelpunkt stehen. Sämtliche Prozessoptimierungen müssen daher dieses prioritäre Ziel im Blick behalten. Weiterhin werden alle Prozesse so konzipiert und umgesetzt, dass

- eine hohe Genauigkeit der Zensusergebnisse sichergestellt wird,
- der Kostenrahmen eingehalten wird,
- Transparenz über die Methodik vorliegt,
- die Belastung der Auskunftspflichtigen minimiert wird und
- das gesamte Zensuspotenzial den Nutzerinnen und Nutzern zur Verfügung gestellt wird.

LITERATURVERZEICHNIS

Bierschenk, Michaela/Leischner, Sonja. *Zur Verfassungsmäßigkeit der Vorschriften über den Zensus 2011*. In: WISTA Wirtschaft und Statistik. Ausgabe 1/2019, Seite 11 ff.

Bretschi, Corinna/Lorentz, Kai. *Präzisionsziele für die Ermittlung der Einwohnerzahl bei der Haushaltsstichprobe im Zensus 2021*. In: WISTA Wirtschaft und Statistik. Sonderheft Zensus 2021. 2019, Seite 12 ff.

Bundesministerium des Innern (BMI). *Externe Evaluierung des Zensus 2011*. Bericht im Auftrag des BMI (auf Anfrage erhältlich). Berlin 2016.

Burgard, Jan Pablo/Münnich, Ralf/Rupp, Martin. *Die Entwicklung des Stichprobenkonzepts für den Zensus 2021*. In: WISTA Wirtschaft und Statistik. Sonderheft Zensus 2021. 2019, Seite 23 ff.

Freier, Benjamin/Mosel, Juliane. *Online First als Leitgedanke für effiziente Primärerhebungen beim Zensus 2021*. In: WISTA Wirtschaft und Statistik. Sonderheft Zensus 2021. 2019, Seite 46 ff.

RECHTSGRUNDLAGEN

Gesetz über die Statistik für Bundeszwecke (Bundesstatistikgesetz – BStatG) in der Fassung der Bekanntmachung vom 20. Oktober 2016 (BGBl. I Seite 2394), das zuletzt durch Artikel 10 Absatz 5 des Gesetzes vom 30. Oktober 2017 (BGBl. I Seite 3618) geändert worden ist.

Gesetz zur Vorbereitung eines registergestützten Zensus einschließlich einer Gebäude- und Wohnungszählung 2021 (Zensusvorbereitungsgesetz 2021 – ZensVorbG 2021) vom 3. März 2017 (BGBl. I Seite 388), das durch Artikel 1 des Gesetzes vom 27. November 2018 (BGBl. I Seite 2010) geändert worden ist.

Verordnung (EG) Nr. 763/2008 des Europäischen Parlaments und des Rates vom 9. Juli 2008 über Volks- und Wohnungszählungen (Amtsblatt der EU Nr. L 218, Seite 14).

Corinna Bretschi

studierte „Survey Methodology“ (MA) an der Universität Duisburg-Essen. Sie ist Referentin im Referat „Haushaltsstichprobe (Methodik)“ des Statistischen Bundesamtes. Dort ist sie schwerpunktmäßig für die methodische Konzipierung der Haushaltsstichprobe beim Zensus 2021 zuständig.

Dr. Kai Lorentz

studierte Mathematik an der Johannes Gutenberg-Universität Mainz. Er leitet das Referat „Mathematisch-statistische Verfahren für Bevölkerungs- und Sozialstatistiken“ des Statistischen Bundesamtes und befasst sich insbesondere mit Fragen der mathematisch-statistischen Methodik für den Zensus 2021.

PRÄZISIONSZIELE FÜR DIE ERMITTLUNG DER EINWOHNERZAHL BEI DER HAUSHALTSSTICHPROBE IM ZENSUS 2021

Corinna Bretschi, Dr. Kai Lorentz

↘ **Schlüsselwörter:** Zensus 2021 – Präzisionsziel – Stichprobe – Einwohnerzahl

ZUSAMMENFASSUNG

Nach dem erstmals registergestützt durchgeführten Zensus 2011 wurden die genutzten Methoden und Verfahren evaluiert. Ergebnis war unter anderem, dass das Verfahren zur Ermittlung der Einwohnerzahl für den kommenden Zensus 2021 angepasst werden muss. Beim kommenden Zensus wird es eine Haushaltsstichprobe in allen Gemeinden geben. Es muss dabei insbesondere festgelegt werden, welche Präzisionsziele angestrebt werden sollen. Der Beitrag beschreibt die dabei angewendete Methodik, die zu einem adäquaten Interessensausgleich zwischen den konkurrierenden Anforderungen Genauigkeit, Belastung und interkommunale Gleichbehandlung führt.

↘ ***Keywords:*** *2021 Census – precision target – sample – number of inhabitants*

ABSTRACT

After the first register-based census was conducted in 2011, the methods and procedures employed were evaluated. One result was that the procedure for determining the number of inhabitants would have to be adjusted for the forthcoming 2021 Census. In the coming census, a household sample survey will be carried out in all municipalities. To this end, it is necessary, in particular, to set the precision targets. This article describes the methodology applied to achieve an adequate balance between the conflicting requirements of accuracy, burden reduction and inter-municipal equality.

1

Einleitung

Beim Zensus 2011 wurde zur Ermittlung der Einwohnerzahl eine Haushaltsstichprobe nur in Gemeinden mit mindestens 10 000 Einwohnerinnen und Einwohnern (angestrebtes Präzisionsziel: 0,5 % einfacher Standardfehler) durchgeführt. In Gemeinden mit weniger als 10 000 Einwohnerinnen und Einwohnern gab es stattdessen eine „Befragung zur Klärung von Unstimmigkeiten“, von der ein vergleichbarer Bereinigungseffekt erwartet wurde. Der Zensus 2011 hat gezeigt, dass diese Befragung zwar den erwarteten Bereinigungsbedarf erfüllt hat, dieser jedoch unter dem tatsächlichen ex post feststellbaren Bedarf lag. Für den Zensus 2021 wird es daher eine Haushaltsstichprobe in allen Gemeinden geben. Damit stellt sich die Frage, mit welchen Präzisionszielen die Ermittlung der Einwohnerzahl erfolgen wird, um einer überproportionalen Belastung der Bevölkerung entgegenzuwirken.

Dieser Artikel beschreibt in Kapitel 2 zunächst die Methodik, die der Ermittlung der Einwohnerzahl beim Zensus 2011 zugrunde lag, und die Anpassungen, die für den Zensus 2021 vorgenommen werden. Kapitel 3 thematisiert die Herausforderungen bei der Festlegung von Präzisionszielen. Die Schritte des Entscheidungsprozesses zur Festlegung der Präzisionsziele für die Ermittlung der Einwohnerzahl im Rahmen der Haushaltsstichprobe beim Zensus 2021 sind in Kapitel 4 dargestellt, gefolgt von der Herleitung der Präzisionszielfunktion (Kapitel 5). Weitere Ideen zur Begrenzung des Stichprobenumfangs beim Zensus 2021 werden in Kapitel 6 entwickelt. Der Beitrag schließt mit einem Ausblick auf die Optionen der Länder zur Optimierung des Befragungsumfangs und den voraussichtlichen Stichprobenumfang.

2

Methodik für die Ermittlung der Einwohnerzahl beim Zensus 2011 und 2021

Für die Ermittlung der Einwohnerzahl hat sich Deutschland beim Zensus 2021 – wie bereits beim Zensus 2011 – für ein registergestütztes Modell entschieden. Die Grundlage für die Ermittlung der Einwohnerzahlen bilden sowohl beim Zensus 2011 als auch beim Zensus 2021 die Daten aus den Melderegistern der Kommunen. Allerdings sind nicht alle Angaben aus den Registern präzise und aktuell. Manche Personen sind zum Zensusstichtag an ihrem Wohnort gar nicht gemeldet (Fehlbestände beziehungsweise Untererfassungen), andere stehen zwar im Register, sind aber schon längst umgezogen oder bereits verstorben (Karteileichen beziehungsweise Übererfassungen). Deshalb ist eine einfache Auszählung aus den Melderegistern bei dem derzeitigen Qualitätsstand der Melderegister nicht ausreichend. Auch beim Zensus 2021 müssen daher die aus den Registern stammenden Angaben durch geeignete Verfahren statistisch überprüft und primärstatistisch korrigiert werden.[1]

Beim Zensus 2011 wurden unterschiedliche Verfahren zur Überprüfung und Korrektur der Angaben aus den Melderegistern angewandt. In allen Gemeinden gab es – unabhängig von ihrer Einwohnerzahl – eine weitgehend automatisierte Bereinigung um unzulässige Dubletten (Mehrfachfallprüfung) sowie eine Vollerhebung an sogenannten Sonderbereichen[2]. In großen Gemeinden mit mindestens 10 000 Einwohnerinnen und Einwohnern wurden darüber hinaus die aus den Melderegistern stammenden Daten mithilfe einer Stichprobe bereinigt. Dazu wurden etwa 10 % aller in Deutschland lebenden Personen persönlich befragt (Haushaltsstichprobe). Die dabei ermittelten Fehlbestände und Karteileichen wurden anschließend auf die gesamte Gemeinde hochgerechnet. Für die Ermittlung der Einwohnerzahl wurde ein Präzisionsziel von 0,5 % einfachem relativen Standardfehler angestrebt. Das bedeutet, dass bei einer festgestellten Einwohnerzahl von 20 000 Personen die tatsächliche, aber unbekannte Einwohnerzahl mit einer Wahrscheinlichkeit von 95 % zwischen 19 800 und 20 200 Personen liegt. Die Entwicklung des Stichprobendesigns der Haushaltsstichprobe und der Schätz-

1 Darüber hinaus dienen die primärstatistischen Erhebungen der Erfassung von Zusatzmerkmalen, die nicht aus den Verwaltungsdaten ermittelt werden können. Beispiele dafür sind Merkmale zur Bildung und zur Erwerbstätigkeit (sogenannte Ziel-2-Merkmale).

2 Sonderbereiche sind Anschriften mit Gemeinschafts- und Anstaltsunterkünften, in denen Personen in der Regel längerfristig untergebracht beziehungsweise wohnhaft sind. Dabei wird zwischen Gemeinschaftsunterkünften (zum Beispiel Behindertenwohnheime) und Wohnheimen (zum Beispiel Studentenwohnheime) unterschieden. Eine Hochrechnung der festgestellten Karteileichen und Fehlbestände ist nicht erforderlich, da für alle Personen an den betreffenden Anschriften Daten erhoben wurden (Vollerhebung).

methoden erfolgte auf Initiative des Statistischen Bundesamtes von einer Forschungsgruppe um Professor Dr. Ralf Münnich (Universität Trier) und Privatdozent Dr. Siegfried Gabler (GESIS Mannheim) im Rahmen eines „Stichprobenforschungsprojekts zum deutschen Zensus 2011“ (Münnich und andere, 2012). Die Ergebnisse und Empfehlungen aus dem Stichprobenforschungsprojekt dienten als Grundlage für die Stichprobenziehung und der Hochrechnung beim Zensus 2011 (Berg/Bihler, 2011).

In kleinen Gemeinden unter 10 000 Einwohnerinnen und Einwohnern gab es eine solche Stichprobe nicht. Stattdessen fand dort eine gezielte Befragung zur Klärung von Unstimmigkeiten statt. Bei dieser Befragung wurden nur Personen an Anschriften mit einer bewohnten Wohnung (in der Regel Einfamilienhäuser) persönlich befragt, bei denen es Unstimmigkeiten zwischen den Angaben aus den Melderegistern und den Angaben der Gebäude- und Wohnungszählung[3] gab (Statistische Ämter des Bundes und der Länder, 2015). Die Entscheidung für die Wahl zweier unterschiedlicher Methoden zur Qualitätssicherung der Einwohnerzahlen basierte auf den Ergebnissen der Simulationsrechnungen mit den Daten aus dem Zensustest 2001. Die Auswertung des Zensustests zeigte, dass ein großer Teil der Bevölkerung in kleinen Gemeinden in Ein- oder Zweifamilienhäusern wohnen. Auch war die Qualität der Melderegister in kleinen Gemeinden tendenziell besser als in den Gemeinden mit mindestens 10 000 Einwohnerinnen und Einwohnern (Statistische Ämter des Bundes und der Länder, 2004, hier: Seite 816 ff.). Auf Basis dieser Ergebnisse wurde eine gezielte Befragung zur Klärung von Unstimmigkeiten in Gemeinden mit weniger als 10 000 Einwohnerinnen und Einwohnern als geeignet bewertet.

Eine Haushaltsstichprobe in den kleinen Gemeinden mit den gleichen Qualitätsanforderungen wie in den Gemeinden mit mindestens 10 000 Einwohnerinnen und Einwohnern (0,5 % einfacher relativer Standardfehler) hätte Folgendes bewirkt: Der Anteil der zu befragenden Personen wäre in kleinen Gemeinden aufgrund des stichprobenmethodischen Zusammenhangs zwischen Standardfehler und Stichprobenumfang unverhältnismäßig hoch gewesen. Damit wäre die Belastung für die Bevölkerung in den kleinen Gemeinden im Vergleich zu den großen Gemeinden überproportional groß gewesen. Andererseits wäre eine Befragung zur Klärung von Unstimmigkeiten in Gemeinden mit mindestens 10 000 Einwohnerinnen und Einwohnern nicht effektiv gewesen, da sie auf Anschriften mit nur einer bewohnten Wohnung beschränkt gewesen wäre. Unter Berücksichtigung der Aspekte Belastung der Auskunftspflichtigen, Effizienz und Kosten wurden beim Zensus 2011 daher zwei unterschiedliche Methoden zur Qualitätssicherung der Einwohnerzahlen eingesetzt: Als geeignete und effiziente Methode zur Korrektur der Melderegisterangaben wurden die Befragung zur Klärung von Unstimmigkeiten in kleinen Gemeinden und die Haushaltsstichprobe in großen Gemeinden identifiziert (Statistische Ämter des Bundes und der Länder, 2015, hier: Seite 44).

Die Ergebnisse des Zensus 2011 und deren Evaluierung haben ergeben, dass die Befragung zur Klärung von Unstimmigkeiten den Melderegisterbestand in der Gesamtheit der kleinen Gemeinden um 0,2 % nach unten korrigiert hat. Dies war aufgrund des Zensustests 2001 zu erwarten. Es zeigte sich aber auch, dass der ex post festgestellte Korrekturbedarf mit 1,1 % deutlich größer war (um insgesamt 0,9 Prozentpunkte) als aufgrund des Zensustests zu erwarten.[4] Es ist daher davon auszugehen, dass beim Zensus 2021 die Korrektur der Registerangaben in Gemeinden mit weniger als 10 000 Einwohnerinnen und Einwohnern auf Basis der Befragung zur Klärung von Unstimmigkeiten erneut nicht ausreichend ausfallen würde. Der Einsatz der Befragung zur Klärung von Unstimmigkeiten als Korrekturinstrument lässt sich daher für den Zensus 2021 nicht mehr methodisch rechtfertigen. Zudem haben umfangreiche Untersuchungen durch die Statistischen Ämter des Bundes und der Länder[5] ergeben, dass auch eine Ausweitung des bisherigen Verfahrens in kleinen Gemeinden – beispielsweise durch Ausweitung auf Anschriften mit zwei oder drei bewohnten Wohnungen (Mehrfamilienhäu-

3 Die Gebäude- und Wohnungszählung ist ein weiterer Bestandteil der primärstatistischen Erhebungen des Zensus 2011 und des Zensus 2021; sie erfasst flächendeckend Daten zu Gebäuden und Wohnungen.

4 Es handelt sich hierbei um den bundesweiten Korrekturbedarf, der sich aus der Ziel-2-Stichprobe (zur Erfassung von Zusatzmerkmalen) des Zensus 2011 ermitteln lässt. Da die Zusatzmerkmale auch flächendeckend auf Kreisebene nachzuweisen waren, gab es eine Haushaltsstichprobe auch in den kleinen Gemeinden, allerdings mit einem nur auf die Kreisergebnisse abgestellten Stichprobenumfang. Dieser erlaubt keine Ermittlung eines gemeindescharfen Korrekturbedarfs beziehungsweise einer gemeindescharfen Einwohnerzahl.

5 Unter anderem durch das Bayerische Landesamt für Statistik, den Landesbetrieb Information und Technik Nordrhein-Westfalen, das Statistische Landesamt des Freistaates Sachsen sowie das Statistische Bundesamt.

ser) – nicht den benötigten Korrekturumfang abdecken würde. Aus diesen Gründen wird beim Zensus 2021 die Befragung zur Klärung von Unstimmigkeiten nicht mehr eingesetzt. Stattdessen wird es in allen Gemeinden eine Haushaltsstichprobe zur Qualitätssicherung der Einwohnerzahlen geben.

Die Entscheidung für eine einheitliche Methodik bei der Ermittlung der Einwohnerzahl beim Zensus 2021 entspricht auch den Forderungen des Urteils des Bundesverfassungsgerichts vom 19. September 2018 zum Zensusgesetz 2011. Das Bundesverfassungsgericht hat in seinem Urteil die beim Zensus 2011 eingesetzten Verfahren zur Korrektur der Registerangaben bestätigt.[6] Für den Zensus 2021 muss das Verfahren des Zensus 2011 jedoch angepasst werden, sofern die Statistischen Ämter des Bundes und der Länder – aufgrund neuer Erkenntnisse – Verbesserungsbedarf festgestellt haben. Dies ist wie dargestellt in Bezug auf die Befragung zur Klärung von Unstimmigkeiten der Fall.

Daher wurde entschieden, beim Zensus 2021 in den Gemeinden mit weniger als 10 000 Einwohnerinnen und Einwohnern – wie bereits beim Zensus 2011 in den Gemeinden mit mindestens 10 000 Einwohnerinnen und Einwohnern – eine Haushaltsstichprobe zur Korrektur der Über- und Untererfassungen in den Melderegistern zu verwenden. Die Entscheidung stellt sicher, dass ein einheitliches Verfahren zur Ermittlung der Einwohnerzahl beim Zensus 2021 angewandt wird. Damit wird einem der Hauptkritikpunkte am letzten Zensus Rechnung getragen.

3

Herausforderungen bei der Festlegung von Präzisionszielen

Als Maß für die Präzision der aus der Haushaltsstichprobe hochgerechneten Ergebnisse und damit auch der Einwohnerzahlen wird beim Zensus der allgemein übliche einfache Standardfehler genutzt. Er lässt sich sowohl absolut (in Personen) als auch relativ (in Prozent) ausdrücken.

Mit der Festlegung eines anzustrebenden Präzisionsziels wird gleichzeitig eine Untergrenze für den erforderlichen Stichprobenumfang festgelegt. Würde ein identisches prozentuales Präzisionsziel von 0,5 % beim Zensus 2021 (wie 2011 in den großen Gemeinden) sowohl in großen als auch in kleinen Gemeinden angestrebt, so hätte dies einen sehr großen Auswahlsatz in kleinen Gemeinden zur Folge. Damit würde auch der Anteil der zu befragenden Personen sehr groß werden. In zahlreichen kleinen Gemeinden käme es sogar zu Vollerhebungen. Gleichzeitig steigen mit jeder auskunftspflichtigen Person, die zusätzlich befragt wird, die Kosten des Gesamtprojekts und die Belastung der Auskunftspflichtigen erheblich an.

Der Forderung nach einer möglichst geringen Belastung der Auskunftspflichtigen und begrenzten Projektkosten steht der Wunsch nach möglichst genauen Einwohnerzahlen entgegen. Zudem ist es sowohl unter juristischen Gesichtspunkten als auch unter dem Aspekt der Akzeptanz der Ergebnisse notwendig, die Gemeinden nach einheitlichen Verfahren zu behandeln. Bei der Festlegung eines Präzisionsziels gilt es somit, zwischen diesen teilweise konkurrierenden Zielen abzuwägen und einen Kompromiss zu finden.

4

Entscheidungsprozess zur Festlegung der Präzisionsziele

Es muss festgelegt werden, mit welchen Präzisionszielen die Ermittlung der Einwohnerzahl im Rahmen der Haushaltsstichprobe im Zensus 2021 erfolgen soll. Dazu sind die vier Entscheidungskriterien Belastung der Auskunftspflichtigen, Kosten, Gleichbehandlung der Gemeinden und Genauigkeit der Einwohnerzahlen zu berücksichtigen. Üblicherweise wird für große Ausweiseinheiten, zum Beispiel für große Gemeinden, der relative Standardfehler genutzt, wohingegen bei kleineren Ausweiseinheiten, wie bei sehr kleinen Gemeinden, der absolute Standardfehler zum Einsatz kommt.

Zur Entscheidungsfindung wurden empirische Testrechnungen zur Abschätzung des erforderlichen Stichprobenumfangs abhängig von der jeweils angestrebten Präzision der ermittelten Einwohnerzahl durch das Forschungsteam der Universität Trier in enger Zusammenarbeit mit dem Statistischen Bundesamt durchgeführt (Burgard und andere, 2019). Der Entscheidungsprozess

6 Bundesverfassungsgericht, Urteil vom 19. September 2018 – 2 BvF 1/15, 2 BvF 2/15 –.

zur Festlegung der Präzisionsziele für die Ermittlung der Einwohnerzahl im Rahmen der Haushaltsstichprobe beim Zensus 2021 umfasst im Wesentlichen drei Schritte:

Schritt 1: Festlegung eines Präzisionsziels in Gemeinden mit mindestens 10 000 Einwohnerinnen und Einwohnern

Das Urteil des Bundesverfassungsgerichts vom 19. September 2018 hat die beim Zensus 2011 getroffenen normativen Entscheidungen bestätigt. Beispiele dafür sind die „10 000er-Grenze“ bei der Ermittlung der Einwohnerzahl und das Präzisionsziel der Korrekturstichprobe von 0,5 % einfacher relativer Standardfehler in Gemeinden mit mindestens 10 000 Einwohnerinnen und Einwohnern. Auch die Evaluierung des Zensus 2011 hat keine Hinweise darauf geliefert, dass das Präzisionsziel von 0,5 % in Gemeinden mit mindestens 10 000 Einwohnerinnen und Einwohnern angepasst werden sollte. Aus fachlicher und wissenschaftlicher Sicht soll daher das Präzisionsziel beim Zensus 2021 in Gemeinden mit mindestens 10 000 Einwohnerinnen und Einwohnern nicht hinter dem Präzisionsziel von 0,5 % einfacher relativer Standardfehler zurückbleiben.

Schritt 2: Festlegung eines Präzisionsziels in Gemeinden mit weniger als 10 000 Einwohnerinnen und Einwohnern

Würde auch in diesen Gemeinden das gleiche Präzisionsziel von 0,5 % einfacher relativer Standardfehler angestrebt wie in Gemeinden mit mindestens 10 000 Einwohnerinnen und Einwohnern, so führte dies zu einer deutlichen Erhöhung des Befragungsumfangs. Statt mit einem Befragungsumfang von bundesweit 9,1 Millionen Personen[7] wie beim Zensus 2011 wäre 2021 mit einem bundesweiten Stichprobenumfang von insgesamt 11,9 Millionen Personen zu kalkulieren.[8]

Ferner müsste man – rein rechnerisch – die Einwohnerzahl in Gemeinden mit weniger als 100 Einwohnerinnen und Einwohnern auf Bruchteile einer Person genau ermitteln. Eine solche Forderung ist aus fachlicher und wissenschaftlicher Sicht weder sinnvoll noch hinsichtlich der Belastung der Auskunftspflichtigen vertretbar.

Aus wissenschaftlicher Sicht besteht eine Lösung darin, das angestrebte prozentuale Präzisionsziel in kleinen Gemeinden im Vergleich zu großen Gemeinden zu lockern. Dadurch wären in kleinen Gemeinden geringere Auswahlsätze möglich als bei Forderung eines einheitlichen prozentualen Präzisionsziels; sowohl die Belastung der Auskunftspflichtigen als auch die Kosten würden begrenzt. Ein gelockertes Präzisionsziel bedeutet bei einer verzerrungsfreien Hochrechnung auch nicht eine „schlechtere“ Einwohnerzahl, sondern lediglich ein höheres Risiko, dass die ermittelte Einwohnerzahl von der tatsächlichen Einwohnerzahl stärker abweicht. Dieses erhöhte Risiko ist aber immer in beide Richtungen symmetrisch, das heißt das Restrisiko einer um mehr als xy % zu niedrigen Einwohnerzahl geht immer einher mit einem gleich hohen Restrisiko einer um xy % zu hohen Einwohnerzahl. Die Lockerung des Präzisionsziels soll beim Zensus 2021 nicht sprunghaft, sondern gleitend erfolgen. Dies wird erreicht, indem das Präzisionsziel als mathematische Funktion mit der Gemeindegröße als abhängige Variable und Glattheitsforderung an der Stelle 10 000 Personen definiert wird (mittels einer sogenannten Präzisionszielfunktion). Je nach Ausgestaltung der Präzisionszielfunktion kann sich der Stichprobenumfang im Vergleich zur Vorgabe eines einheitlichen prozentualen Präzisionsziels in allen Gemeinden verringern.

Schritt 3: Festlegung eines Präzisionsziels in sehr kleinen Gemeinden mit weniger als 1 000 Einwohnerinnen und Einwohnern

Wie weiter oben bereits ausgeführt, ist es sinnvoll, einen einheitlichen relativen Standardfehler in der Regel nur bei großen Ausweiseinheiten anzustreben (siehe Schritt 1). Bei kleinen Ausweiseinheiten werden diese Vorgaben üblicherweise angepasst, um unverhältnismäßig viele Vollerhebungen zu reduzieren (siehe Schritt 2). Wie dargelegt, führt das Festlegen eines festen prozentualen Präzisionsziels in sehr kleinen

7 Einschließlich Hauptwohnsitz- und Nebenwohnsitz-Personen, einschließlich Stichprobe für Ziel 2 an Sonderbereichen und einschließlich Befragung zur Klärung von Unstimmigkeiten.

8 Einschließlich Hauptwohnsitz- und Nebenwohnsitz-Personen, einschließlich Stichprobe für Ziel 2 an Sonderbereichen; Kalkulation wurde unter der Leitung von Professor Dr. Münnich vorgenommen (Burgard und andere, 2019).

Gemeinden zu einer überproportionalen Belastung der Auskunftspflichtigen und ist auch aus wissenschaftlich-statistischer Sicht nicht sinnvoll. Der Ausweg für sehr kleine Gemeinden liegt also darin, ein absolutes Präzisionsziel für diese Gemeinden festzulegen. Hier stellt sich zunächst die Frage, was konkret „sehr klein“ in diesem Zusammenhang bedeuten soll. Das Team der Universität Trier hat anhand von Untersuchungen an ihrer dort entwickelten Simulationsgesamtheit hierzu eine Obergrenze von 1 000 Einwohnerinnen und Einwohnern empfohlen. Die wissenschaftliche Seite machte deutlich, dass sie zur Höhe des in sehr kleinen Gemeinden anzustrebenden absoluten Standardfehlers und zum Umgang bei Gemeinden mit weniger als 200 Anschriften keine quantitativen Empfehlungen aussprechen kann.

Das Statistische Bundesamt hat zusammen mit dem Team der Universität Trier drei einfache absolute Präzisionsziele für sehr kleine Gemeinden unter Anwendung der Präzisionszielfunktion in Gemeinden mit 1 000 bis 10 000 Einwohnerinnen und Einwohnern untersucht: 15, 25 und 40 Personen. Die Ergebnisse zeigen, dass je nach Wahl des absoluten Präzisionsziels der Stichprobenumfang erheblich reduziert werden kann. Der Stichprobenumfang bei Vorgabe eines absoluten Präzisionsziels von 15 Personen und bei Anwendung einer Präzisionszielfunktion in Gemeinden zwischen 1 000 und 10 000 Einwohnerinnen und Einwohnern beträgt 11,4 Millionen Personen. Er könnte bei Vorgabe eines absoluten Präzisionsziels von 25 Personen auf knapp 11 Millionen Personen und bei Vorgabe eines absoluten Präzisionsziels von 40 Personen sogar auf 10,4 Millionen Personen reduziert werden. Unter Berücksichtigung des Kriteriums der Genauigkeit haben sich die Statistischen Ämter des Bundes und der Länder für das strengste der untersuchten Präzisionsziele (15 Personen) entschieden.

Für Kleinstgemeinden mit weniger als 200 Anschriften ist zur Ermittlung der Einwohnerzahl aus fachlicher/wissenschaftlicher Sicht weder eine pauschale Vollerhebung noch die Übernahme des Melderegisterbestands ohne Stichprobenkorrektur zu empfehlen. Eine Vollerhebung in einzelnen Gemeinden entspricht nicht dem Anspruch einer grundrechtsschonenden Datenerhebungsmethode, wie es im Urteil des Bundesverfassungsgerichts vom 19. September 2018 zum Zensusgesetz 2011 gefordert ist. Der gänzliche Verzicht auf eine stichprobenmethodische Korrektur des Melderegisterbestands ist hingegen fachlich aufgrund des mit dem Zensus 2011 auch in diesen Gemeindegrößenklassen festgestellten Korrekturbedarfs nicht zu rechtfertigen.

Um dennoch Einwohnerzahlen in Gemeinden mit weniger als 200 Anschriften festzustellen, soll es auch in diesen Gemeinden eine Stichprobenerhebung geben. Für eine stichprobenbasierte Korrektur bedarf es jedoch mindestens 100 Anschriften in der Stichprobe. Das bedeutet, dass in Gemeinden zwischen 101 und 200 Einwohnerinnen und Einwohnern die Auswahlsätze teilweise deutlich über 50 % liegen werden. In Gemeinden mit 100 Anschriften und weniger wird es sogar zu Vollerhebungen kommen. Auf diese Weise ist gewährleistet, dass in allen Gemeinden eine Korrektur der Einwohnerzahlen erfolgen kann.

5

Präzisionszielfunktion

Das Forschungsteam der Universität Trier empfiehlt auf Basis seiner Untersuchungen, beim Zensus 2021 einen gleitenden Übergang der Präzisionsziele in Gemeinden zwischen 1 000 und 10 000 Einwohnerinnen und Einwohnern anzuwenden. Das bedeutet, in Gemeinden mit weniger als 10 000 Einwohnerinnen und Einwohnern sollte das Präzisionsziel bis zur Grenze von 1 000 gleitend gelockert werden. Ab dieser Grenze sollte jedoch ein konstantes absolutes Präzisionsziel fixiert und für kleinere Gemeinden nicht weiter verändert werden. Unterschreitet die Gemeindegröße jedoch eine bestimmte Anschriftengrenze, so ist eine stichprobenbasierte Korrektur grundsätzlich nicht sinnvoll, da eine belastbare Hochrechnung eine bestimmte Anzahl von Stichprobenanschriften voraussetzt. Die Grenze für solche Kleinstgemeinden liegt laut den Untersuchungsergebnissen bei 200 Anschriften.

Es soll nun im Folgenden die Entwicklung einer Funktionsvorschrift für die Präzisionsziele beschrieben werden. Im Bereich der Gemeinden bis 1 000 Einwohnerinnen und Einwohnern wird ein absolutes Präzisionsziel angestrebt (einfacher absoluter Standardfehler von 15 Personen, siehe oben), im Bereich der Gemeinden ab 10 000 Einwohnerinnen und Einwohnern ein relatives Präzisionsziel (0,5 % einfacher relativer Standardfehler). Das relative Präzisionsziel für die großen Gemeinden

wird zunächst als ein absolutes Präzisionsziel umformuliert, um eine einheitliche Zielgröße für die Gesamtfunktion, die sogenannte Präzisionszielfunktion, zu ermöglichen. Es fehlt noch die Festlegung einer Übergangsfunktion im Bereich der Gemeinden von 1 000 bis 10 000 Einwohnerinnen und Einwohnern. Dieser gleitende Übergang der Präzisionsziele soll folgende Bedingungen erfüllen:

- Es soll sich um einen glatten (stetig differenzierbaren) Übergang handeln, insbesondere an den „Nahtstellen" 1 000 und 10 000.
- Die Funktion soll streng monoton steigend sein.
- Die Funktion soll linksgekrümmt sein.

Mathematisch handelt es sich hierbei um ein sogenanntes numerisches Interpolationsproblem, das auf unterschiedliche Art und Weise behandelt und gelöst werden kann. Die prominenteste Herangehensweise bestünde in der Anwendung von sogenannten Spline-Funktionen. Hierzu wären allerdings die Festlegung weiterer Zwischenpunkte im Intervall zwischen 1 000 und 10 000 sowie die Vorgabe entsprechender Funktionswerte und Steigungen notwendig.

Von der Universität Trier wurde die Verwendung sogenannter Bézierkurven vorgeschlagen, die ohne diese zusätzlichen Festlegungen auskommen und die genannten Bedingungen erfüllen (Burgard und andere, 2019). Bézierkurven wurden Mitte des letzten Jahrhunderts definiert und finden heutzutage besonders im Bereich der Computergrafik Anwendung.

Für unsere Vorgaben (im Intervall [0,1000] soll der absolute Standardfehler $y = 15$ und im Bereich oberhalb 10 000 Einwohnerinnen und Einwohnern $0{,}005 \cdot x$ betragen [x = Einwohnerzahl]) wird die zugehörige Bézierkurve in Form einer Parameterdarstellung definiert durch

$$(1) \qquad \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = (1-t)^2 \begin{pmatrix} 1\,000 \\ 15 \end{pmatrix} + 2t(1-t) \begin{pmatrix} 3\,000 \\ 15 \end{pmatrix} + t^2 \begin{pmatrix} 10\,000 \\ 50 \end{pmatrix}, 0 \le t \le 1$$

(Darstellung in der sogenannten Bernsteinbasis; der mittlere Punkt ist dabei der Schnittpunkt der Tangenten in den Endpunkten).

Löst man $x(t)$ nach t auf und setzt dies in $y(t)$ ein, so kann man in unserem Fall auch direkt y als Funktion von x angeben.

Es ergibt sich

$$(2) \qquad y(x) = \frac{7}{1\,000}\,x - \frac{28}{10\sqrt{50}}\sqrt{x - 200} + \frac{96}{5}$$

Insgesamt lautet also die Präzisionszielfunktion (für den absoluten Standardfehler)

$$(3) \qquad y(x) = \begin{cases} 15, & \text{für } 0 \le x < 1\,000 \\ \frac{7}{1\,000}\,x - \frac{28}{10\sqrt{50}}\sqrt{x - 200} + \frac{96}{5}, & \text{für } 1\,000 \le x < 10\,000 \\ \frac{5}{1\,000}\,x, & \text{für } x \ge 10\,000 \end{cases}$$

6

Weitere Ideen zur Begrenzung des Stichprobenumfangs beim Zensus 2021

Der bundesweite Stichprobenumfang beträgt bei Vorgabe eines einheitlichen relativen Präzisionsziels von 0,5 % einfacher relativer Standardfehler in allen Gemeinden 11,9 Millionen Personen. Er kann durch die Verwendung einer Präzisionszielfunktion in Gemeinden zwischen 1 000 und 10 000 Einwohnerinnen und Einwohnern und einem absoluten Präzisionsziel von 15 Personen für Gemeinden unter 1 000 Einwohnerinnen und Einwohnern auf 11,4 Millionen Personen begrenzt werden. Aufgrund der sehr unterschiedlichen Gemeindestrukturen variieren die Auswahlsätze jedoch sehr stark zwischen den einzelnen Bundesländern, in einigen Ländern müsste sogar mehr als ein Viertel der Bevölkerung befragt werden.[9] In der Folge würde sich die relative Belastung der Auskunftspflichtigen sehr ungleich auf die Bundesländer verteilen.

9 Die Auswahlsätze liegen zwischen 5,4 % in Berlin und 25,2 % in Rheinland-Pfalz, im Durchschnitt beträgt der bundesweite Auswahlsatz 12,7 %.

Eine Möglichkeit, den Stichprobenumfang zu optimieren, ergibt sich daraus, dass in einigen Ländern die Melderegister für Gemeinden, die einem in der Landesgesetzgebung verankerten Gemeindeverband angehören, gemeinsam geführt werden. Dadurch kann für diese Gemeinden unterstellt werden, dass eine Fehlerquelle der Melderegister, nämlich Fehler bei der Registerführung, vergleichbar ist. Daher ist es naheliegend, verbandsangehörige Gemeinden mit einheitlicher Melderegisterführung zur Ermittlung der Einwohnerzahl gemeinsam zu betrachten. Sofern die genannten Bedingungen der landesrechtlichen Verankerung des Gemeindeverbands und der gemeinsamen Melderegisterführung für die Ermittlung der Einwohnerzahl zutreffen, können die Länder verbandsangehörige Gemeinden zu Gemeindeverbänden zusammenfassen und das Präzisionsziel auf Gemeindeverbandsebene anstreben („GV-Option“). Eine Abwandlung dieser Möglichkeit besteht darin, lediglich verbandsangehörige Gemeinden mit weniger als 2 000 Einwohnerinnen und Einwohnern zu sogenannten Gemeindeverbandsresten zu bündeln. Das Präzisionsziel ist dann nicht mehr für die Einzelgemeinde oder den Gemeindeverband insgesamt anzustreben, sondern für diesen Gemeindeverbandsrest („GV-Rest-Option“). Die größeren verbandsangehörigen Gemeinden mit mehr als 2 000 Einwohnerinnen und Einwohnern werden hierbei genauso wie Gemeinden behandelt werden, die keinem Gemeindeverband angehören. Durch die Anwendung der „GV-Rest-Option“ in allen Ländern mit Gemeindeverbänden kann der Stichprobenumfang bundesweit um 1,3 Millionen Personen auf knapp 10,1 Millionen Personen reduziert werden, durch die Anwendung der „GV-Option“ sogar auf bundesweit etwa 9 Millionen Personen.

Voraussetzung für die Zusammenfassung von verbandsangehörigen Gemeinden zu Gemeindeverbänden beziehungsweise Gemeindeverbandsresten ist allerdings zum einen, dass die Bildung von Gemeindeverbänden im Landesgesetz fundiert ist. Zum anderen müssen die Melderegister von verbandsangehörigen Gemeinden einer gemeinsamen Führung unterliegen.[10] Die Erfüllung dieser beiden Bedingungen ist nicht in allen Ländern gegeben. Durch eine Kombination der Alternativen werden räumliche Strukturen geschaffen, die von ihrer Größe her länderübergreifend eher vergleichbar sind.

10 Bei einer gemeinsamen Melderegisterführung kann vorausgesetzt werden, dass keine Unterschiede zwischen den verbandsangehörigen Gemeinden bezüglich der Melderegisterführung bestehen.

Unabhängig von der Nutzung einer der beiden Möglichkeiten zur Optimierung des Befragungsumfangs müssen für alle Gemeinden Einwohnerzahlen im Rahmen des Zensus 2021 ermittelt werden. In Ländern, die sich für die Gemeindeverbands-Option beziehungsweise die Gemeindeverbandsrest-Option entscheiden, erfolgt die Hochrechnung zur Ermittlung der Einwohnerzahl für den Gemeindeverband beziehungsweise für den Gemeindeverbandsrest. Die Einwohnerzahlen der verbandsangehörigen Gemeinden müssen anschließend bestimmt werden, indem die Einwohnerzahl des Gemeindeverbands beziehungsweise des Gemeindeverbandsrests heruntergebrochen wird. Die genaue Methode hierfür wird derzeit aufbauend auf folgendem Ansatz noch weiterentwickelt:

Schritt 1: Für den Gemeindeverband beziehungsweise den Gemeindeverbandsrest als Ganzes wird die Einwohnerzahl wie für die großen Gemeinden beim Zensus 2011 ermittelt (sogenannte GREG-Schätzung). Das heißt, hier sind Stichprobeninformationen und Registerinformationen aller Gemeinden des Gemeindeverbands beziehungsweise des Gemeindeverbandsrests beteiligt. Als Resultat ergeben sich Hochrechnungsgewichte für die einzelnen Stichprobenanschriften, die diese Informationen (Stichprobeninformationen und Registerinformationen aller Gemeinden des Gemeindeverbands) ebenfalls enthalten.

Schritt 2: Nun folgt die Schätzung der Einwohnerzahl einer verbandsangehörigen Gemeinde. Diese Schätzung besteht aus zwei Teilen:

(1) einer direkten Schätzung mittels der zuvor gebildeten Hochrechnungsgewichte

(2) einer indirekten Schätzung durch Herunterbrechen der Einwohnerzahlen des Gemeindeverbands beziehungsweise des Gemeindeverbandsrests (siehe GREG-Schätzung unter Schritt 1) auf die Gemeinde. Dies erfolgt nach dem proportionalen Anteil der gemeldeten Personen der Gemeinde an der Gesamtzahl der gemeldeten Personen des Gemeindeverbands beziehungsweise des Gemeindeverbandsrests.

Tabelle 1
Stichprobenumfang nach verschiedenen Varianten der Präzisionszielfunktion und der länderspezifischen Festlegung

	Präzisions-zielfunktion		Präzisionszielfunktion mit „Gemeindeverbandsrest-Option“ (GV-Rest-Option)			Präzisionszielfunktion mit „Gemeindeverbands-Option“ (GV-Option)			Länderspezifische Festlegung (GV-Rest-Option, GV-Option, keine)		
	1 000	%|1	Option	1 000	%|1	Option	1 000	%|1	Option	1 000	%|1
Deutschland	11 362	12,7	X	10 093	11,2	X	9 037	10,1	X	9 872	10,1
Schleswig-Holstein	657	20,5	GV-Rest-Option	428	13,4	GV-Option	335	10,5	GV-Rest-Option	431	13,5
Hamburg	103	5,6	/	103	5,6	/	105	5,7	/	104	5,6
Niedersachsen	1 059	12,0	GV-Rest-Option	957	10,9	GV-Option	821	9,3	GV-Option	793	9,0
Bremen	39	5,5	/	39	5,5	/	40	5,6	/	40	5,6
Nordrhein-Westfalen	1 360	7,0	/	1 366	7,0	/	1 406	7,2	/	1 371	7,0
Hessen	740	11,0	/	754	11,0	/	783	11,6	/	750	11,1
Rheinland-Pfalz	1 122	25,2	GV-Rest-Option	608	13,6	GV-Option	397	8,9	GV-Option	385	8,6
Baden-Württemberg	1 556	13,4	GV-Rest-Option	1 537	13,2	GV-Option	1 103	9,5	/	1 579	13,6
Bayern	2 206	16,1	GV-Rest-Option	2 095	15,2	GV-Option	2 004	14,6	/	2 238	16,3
Saarland	101	8,8	/	102	8,9	/	106	9,2	/	102	8,9
Berlin	197	5,4	/	198	5,4	/	200	5,4	/	198	5,4
Brandenburg	411	14,8	GV-Rest-Option	370	13,3	GV-Option	342	12,3	GV-Option	327	11,8
Mecklenburg-Vorpommern	456	24,6	GV-Rest-Option	300	16,2	GV-Option	224	12,0	GV-Rest-Option	302	16,3
Sachsen	600	13,1	GV-Rest-Option	597	13,1	GV-Option	562	12,3	/	608	13,3
Sachsen-Anhalt	298	11,4	GV-Rest-Option	276	10,6	GV-Option	253	9,7	/	301	11,6
Thüringen	457	18,3	GV-Rest-Option	373	15,0	GV-Option	357	14,3	GV-Option	343	13,8

Die dargestellten Stichprobenumfänge enthalten einen pauschalen Aufschlag für Hauptwohnsitz- und Nebenwohnsitz-Personen sowie für die Nachziehung, er umfasst die Stichprobe für Ziel 2 an Sonderbereichen. Die Kalkulation wurde unter der Leitung von Professor Dr. Ralf Münnich (Universität Trier) vorgenommen.
1 Auswahlsatz.

Schritt 3: Aus diesen beiden Teilen wird mittels eines generalisierten Kalibrierungsverfahrens eine gemeindespezifische Schätzung durchgeführt. Dabei wird in Gemeinden, für die keine Stichprobe existiert, der Registerwert ausgegeben. Mit steigender Zahl der Stichprobeneinheiten je Gemeinde gewährleistet das Verfahren einen dynamischen Übergang zu den direkten Schätzwerten. Dabei gilt: Je höher der Stichprobenanteil einer Gemeinde ist, umso höher wird der direkte Schätzanteil gewichtet.

↘ Tabelle 1 stellt die Stichprobenumfänge verschiedener Varianten dar. Bei der Variante „Präzisionszielfunktion mit GV-Option“ wurden alle verbandsangehörigen Gemeinden zu einem Gemeindeverband zusammengefasst, bei der Variante „Präzisionszielfunktion mit GV-Rest-Option“ alle verbandsangehörigen Gemeinden mit weniger als 2 000 Einwohnerinnen und Einwohnern zu einem Gemeindeverbandsrest. Bei der „Länderspezifischen Festlegung“ erfolgte eine Zusammenfassung von verbandsangehörigen Gemeinden nur in ausgewählten Ländern. Im Vergleich ist unter der Variante „Präzisionszielfunktion“ der Stichprobenumfang aufgeführt, wenn in allen Ländern von keiner dieser Optionen Gebrauch gemacht werden würde.

7

Ausblick

Mit den Ergebnissen des Zensus 2011 und deren Evaluation liegen Erkenntnisse vor, wonach das Verfahren zur Ermittlung der Einwohnerzahl für den kommenden Zensus 2021 angepasst werden muss. Beim Zensus 2021 wird es daher zur Ermittlung der Einwohnerzahl eine Haushaltsstichprobe in allen Gemeinden geben, nicht wie beim Zensus 2011 nur in Gemeinden mit mindestens 10 000 Einwohnerinnen und Einwohnern. Ein einheitlicher relativer Standardfehler für alle Gemeinden unabhängig von deren Größe als Qualitätsmaß für ein hochgerechnetes Ergebnis ist jedoch nicht sinnvoll. Zur Optimierung des Stichprobenumfangs wird daher

- in Gemeinden mit mindestens 10 000 Einwohnerinnen und Einwohnern ein Präzisionsziel von 0,5 % einfacher relativer Standardfehler angestrebt (wie beim Zensus 2011),
- in Gemeinden mit mindestens 1 000 und weniger als 10 000 Einwohnerinnen und Einwohnern das relative Präzisionsziel gleitend unter Hinzunahme einer Präzisionszielfunktion gelockert, und
- in Gemeinden mit weniger als 1 000 Einwohnerinnen und Einwohnern ein einfaches absolutes Präzisionsziel von 15 Personen angestrebt.

Die gleitende Überleitung von einem relativen zu einem absoluten Präzisionsziel über eine Präzisionszielfunktion führt zu einem adäquaten Interessensausgleich zwischen den konkurrierenden Anforderungen Genauigkeit, Belastung und interkommunale Gleichbehandlung.

Zielsetzung ist, den Stichprobenumfang noch weiter im Hinblick auf die Entlastung der Auskunftspflichtigen zu optimieren und die Spreizung der Auswahlsätze zu begrenzen. Im Zuge des Gesetzgebungsverfahrens zum Zensusgesetz 2021 haben die Länder daher die Möglichkeit, sich für eine der drei folgenden Optionen zu entscheiden:

1. Präzisionszielfunktion,
2. Präzisionszielfunktion mit „Gemeindeverbands-Option“ (das heißt verbandsangehörige Gemeinden werden zu einem Gemeindeverband zusammengefasst) oder
3. Präzisionszielfunktion mit „Gemeindeverbandsrest-Option“ (das heißt verbandsangehörige Gemeinden mit weniger als 2 000 Einwohnerinnen und Einwohnern werden zu einem sogenannten Gemeindeverbandsrest zusammengefasst)

Bei den Optionen 2 und 3 wird das Präzisionsziel jeweils für den Gemeindeverband beziehungsweise den Gemeindeverbandsrest angestrebt. Die Ermittlung der Einwohnerzahl im Rahmen des Zensus 2021 erfolgt für alle Gemeinden, unabhängig davon, ob einige Länder von einer der genannten Optionen zur Optimierung des Befragungsumfangs Gebrauch machen. Dazu wird die Einwohnerzahl des Gemeindeverbands(-rests) auf die zugehörigen Einzelgemeinden heruntergebrochen. Aktuell zeichnet sich eine länderspezifische Festlegung ab, bei der der bundesweite Stichprobenumfang voraussichtlich etwa 10 Millionen Personen beträgt.

LITERATURVERZEICHNIS

Berg, Andreas/Bihler, Wolf. *Das Stichprobendesign der Haushaltsstichprobe des Zensus 2011*. In: Wirtschaft und Statistik. Ausgabe 4/2011, Seite 317 ff.

Bundesverfassungsgericht (BVerfG). *Urteil vom 19. September 2018 – 2 BvF 1/15, 2 BvF 2/15*. In: Neue Zeitschrift für Verwaltungsrecht (NVwZ). 2018.

Burgard, Jan Pablo/Münnich, Ralf/Rupp, Martin. *Die Entwicklung des Stichprobenkonzepts für den Zensus 2021*. In: WISTA Wirtschaft und Statistik. Sonderheft Zensus 2021. 2019, Seite 23 ff.

Münnich, Ralf/Gabler, Siegfried/Ganninger, Matthias/Burgard, Jan Pablo/Kolb, Jan-Philipp. *Stichprobenoptimierung und Schätzung im Zensus 2011*. Band 21 der Schriftenreihe Statistik und Wissenschaft. Wiesbaden 2012.

Statistische Ämter des Bundes und der Länder. *Ergebnisse des Zensustests*. In: Wirtschaft und Statistik. Ausgabe 8/2004, Seite 813 ff.

Statistische Ämter des Bundes und der Länder. *Zensus 2011. Methoden und Verfahren*. 1. Auflage. Wiesbaden 2015.

DIE ENTWICKLUNG DES STICHPROBENKONZEPTS FÜR DEN ZENSUS 2021

Dr. habil. Jan Pablo Burgard, Prof. Dr. Ralf Münnich, Dr. Martin Rupp

↘ **Schlüsselwörter:** Zensus – Ermittlung der Einwohnerzahl – Gemeindestrukturen – Optimale Allokation – Regressionsschätzung

ZUSAMMENFASSUNG

Infolge des Urteils des Bundesverfassungsgerichts vom 19. September 2018 wird für den Zensus 2021 eine für alle Gemeinden einheitliche Erhebungs- und Schätzmethodik angestrebt. Dabei soll auf Vollerhebungsschichten möglichst verzichtet werden. Eine direkte Übertragung des bereits für den Zensus 2011 definierten Präzisionsziels von maximal 0,5 % relativem Standardfehler für Gemeinden ab 10 000 Einwohnerinnen und Einwohnern auf die kleineren Gemeinden ist jedoch nicht ohne Weiteres möglich. Aufgrund der Größenstruktur der Gemeinden würde dies zu einem deutlich erhöhten Befragungsaufwand und zu einer hohen Anzahl von Vollerhebungsschichten führen. In diesem Beitrag werden mehrere alternative Schichtungsansätze diskutiert, die eine einheitliche Erhebungs- und Schätzmethodik für alle Gemeinden zulassen.

↘ ***Keywords:*** *census – determining the number of inhabitants – municipality structures – optimum allocation – regression estimation*

ABSTRACT

As a result of the judgment given by the German Federal Constitutional Court on 19 September 2018, a uniform survey and estimation methodology for all municipalities is required for the 2021 Census. In this context, the aim is to avoid taking a full census in individual municipalities as far as possible. For the 2011 Census, the precision target of a relative standard error of at most 0.5 % in municipalities with 10,000 inhabitants or more was defined. A direct application of that precision target to smaller municipalities is however not easily possible. Due to the size structure of the municipalities, this would result in a significantly increased response burden and a large number of municipalities with a full census. This paper discusses several alternative stratification approaches which enable the application of a uniform survey and estimation methodology to all municipalities.

Dr. habil. Jan Pablo Burgard

ist akademischer Oberrat an der Professur für Wirtschafts- und Sozialstatistik der Universität Trier. Der Diplom-Volkswirt fertigte 2013 seine Dissertation über die Verwendung von Small-Area-Verfahren in der amtlichen Statistik an. Im Februar 2019 schloss er seine Habilitation an der Universität Trier erfolgreich ab.

Prof. Dr. Ralf Münnich

ist Professor für Wirtschafts- und Sozialstatistik an der Universität Trier. Nach dem Studium der Mathematik und Volkswirtschaftslehre promovierte und habilitierte er an der Universität Tübingen. Seit 2006 ist er Inhaber der Professur für Wirtschafts- und Sozialstatistik der Universität Trier. Zudem ist er seit 2018 Direktor des Research Institute for Official and Survey Statistics (RIFOSS) in Trier.

Dr. Martin Rupp

ist wissenschaftlicher Mitarbeiter an der Professur für Wirtschafts- und Sozialstatistik der Universität Trier. Nach dem Studium der Wirtschaftsmathematik (M. Sc.) promovierte er zwischen 2014 und 2018 an der Professur für Wirtschafts- und Sozialstatistik über die Optimierung für multivariate Methoden in der Survey Statistik.

1

Einleitung

Der Zensus 2011 in Deutschland wurde im Rahmen einer EU-weiten Volkszählungsrunde registergestützt durchgeführt. Dabei wurde zum einen das Einwohnermelderegister genutzt und zum anderen zusätzlich eine Stichprobe erhoben, die Informationen liefern sollte über mögliche Registerfehler (Ziel 1) sowie über im Register nicht vorhandene Merkmale (Ziel 2). Der Zensus ist üblicherweise die Basis der Bevölkerungsstatistik und skizziert idealerweise ein Gesamtbild der Gesellschaft (Grohmann, 2009). Die registergestützte Durchführung ersetzte die traditionelle Volkszählung mit einer Befragung aller Haushalte. Vorteile des registergestützten Zensus sind der im Vergleich zum vollen Zensus deutlich geringere Erhebungsaufwand und eine erheblich geringere Belastung der Bevölkerung. Die Stichproben- und Schätzmethodik des Zensus 2011 ist ausführlich in Münnich und andere (2011) sowie in Münnich und andere (2012a) dargestellt. Weitere Ausführungen sind in Statistische Ämter des Bundes und der Länder (2015) zu finden.

Mit einem Abstand von zehn Jahren ist für das Jahr 2021 der nächste Zensus geplant, der erneut auf den Melderegistern aufbaut und durch eine ergänzende stichprobenbasierte Befragung durchgeführt werden wird. Dabei sollen Grundbausteine der Methodik, wie Schichtung, Allokation und Schätzverfahren, grundsätzlich vom Zensus 2011 übernommen werden, wobei jedoch einige Anpassungen vorzunehmen sind. Diese Anpassungen resultieren auch aus dem Urteil des Bundesverfassungsgerichts zur Normenkontrollklage der Länder Berlin und Hamburg (BVerfG, 2018; Gößl, 2018). Gegenstand war eine Überprüfung der Methodik und insbesondere ein vermeintlicher «*Verstoß gegen das Gebot interföderaler und interkommunaler Gleichbehandlung durch den Methodenwechsel bei der Korrektur der Melderegister in Gemeinden über und unter 10 000 Einwohnern*» (Gößl, 2018, hier: Seite 685).

Dieser Methodenwechsel bei 10 000 Einwohnern bezieht sich auf eine unterschiedliche Behandlung von Gemeinden mit weniger als 10 000 Einwohnerinnen und Einwohnern (kleiner 10 000 Einwohner) und solchen mit mehr als 10 000 Einwohnerinnen und Einwohnern (größer 10 000 Einwohner). Während Gemeinden größer 10 000 Einwohner explizit als eigene Schicht im Design aufgefasst werden, wurden aus Zusammenschlüssen von Gemeinden kleiner 10 000 Einwohner sogenannte Sampling Points gebildet. Diese wurden anstelle der Gemeinden als Schichtungsmerkmal im Design berücksichtigt (Münnich und andere, 2012a, hier: Seite 22 ff.). Ein wesentlicher Aspekt der Klage war die Tatsache, dass Qualitätsanforderungen erst für Gemeinden größer 10 000 Einwohner definiert wurden. Unterhalb dieser Schwelle kam ein Alternativverfahren zum Einsatz, wodurch sich in diesen Gebieten eine erhebliche Kostenersparnis erzielen ließ. Nach der Veröffentlichung der Zensusergebnisse konnten allerdings auffällige Abweichungen der Registerkorrekturen an dieser Einwohnergrenze festgestellt werden (siehe die Diskussion in Christensen und andere, 2015, sowie Rendtel, 2015). Teilweise sind diese Abweichungen auf die unterschiedliche Methodik zurückzuführen. Untersuchungen ergaben, dass sie in Bundesländern mit zentral geführten Registern erheblich geringer ausfielen.

Die Klage vor dem Bundesverfassungsgericht wurde schließlich abgewiesen (BVerfG, 2018, sowie die Kommentare in Leischner und Weigelt, 2018, und Gößl, 2018). Vielmehr wurde die aktuelle statistische Methodik unter der gegebenen Informationslage als geeignet und wissenschaftlich fundiert bezeichnet. Gleichzeitig wurde in der Urteilsbegründung aber betont, dass die im Zensus 2011 gewonnenen Informationen bei der Weiterentwicklung der Methoden geeignet berücksichtigt werden sollen. Diese Weisung führte letztendlich zu Methodenkorrekturen, die in diesem Artikel beschrieben werden (siehe auch Bierschenk/Leischner, 2019). Diese werden zudem in ausführlicher Form in Burgard und andere (2019b) analysiert.

Im Zensus 2021 wird auf einen Wechsel der Methodik zwischen Gemeinden kleiner und größer als 10 000 Einwohner verzichtet. Sofern Gemeinden als rechtliche Einheiten weiterhin betrachtet werden, sollen diese im Stichprobendesign, in der Schätzmethodik und in den Präzisionsvorgaben für die Schätzungen gleichbehandelt werden. Dieser Grundsatz führt durch die föderale Organisation Deutschlands und die damit einhergehenden regional sehr unterschiedlichen Gemeindestrukturen zu weitgehenden Anpassungen der Methodik für den Zensus 2021. Die Anpassungen betreffen insbesondere eine veränderte Schichtung und eine über die Gemeinde-

größe definierte Präzisionszielfunktion. Neben dem Grundsatz der Vereinheitlichung der Methodik soll die Qualität der Schätzergebnisse für Gemeinden mit mehr als 10 000 Einwohnerinnen und Einwohnern mindestens das Niveau von 2011 erreichen.

In Kapitel 2 werden die neuen Präzisions- und Zielvorgaben für den Zensus 2021 beschrieben. Das Stichprobenkonzept wird in Kapitel 3 kurz vorgestellt, wobei der Fokus auf den alternativen Vorgehensweisen im Umgang mit kleinen Gemeinden liegt. In Kapitel 4 wird der benötigte Gesamtstichprobenumfang für diese Varianten dargestellt. Der Artikel schließt mit einem kurzen Fazit und Ausblick in Kapitel 5.

2

Präzisions- und Zielvorgaben für den Zensus 2021

Schätzmethodik

Die Schätzmethodik basiert wie 2011 auf der Verwendung eines Gemeinde- beziehungsweise Sampling-Point-separaten Regressionsschätzers und ist in Münnich und andere (2012a, hier: Seite 40 ff.) beziehungsweise Burgard und andere (2019b) detailliert beschrieben. Als Hilfsinformationen für die Schätzung dienen das Anschriften- und Gebäuderegister (Steuerungsregister), in dem die Anschriften als Stichprobeneinheiten ausgewiesen sind, sowie Melderegister der Meldebehörden (siehe § 7 Absatz 3 Zensusgesetz 2011 beziehungsweise Gesetzentwurf zum Zensusgesetz 2021, Bundestags-Drucksache 19/8693). Da die Melderegister von Grund auf sehr valide Informationen liefern, dienen die registrierten Personen in den Anschriften als hochkorreliertes Hilfsmerkmal für die Schätzung der tatsächlichen Einwohnerzahl in den interessierenden Gemeinden.

Zur Vereinfachung der Darstellung der Schätzungen wird nachfolgend zunächst auf den Index g der g-ten Gemeinde (beziehungsweise des Sampling Points) verzichtet. Mit $Y = (y_1, \ldots, y_N)'$ wird das Untersuchungsmerkmal Anzahl der Einwohner für alle $i = 1,\ldots,N$ Anschriften bezeichnet. X bezeichnet das Hilfsmerkmal. Im Zensus sind das für x_i die in der i-ten Anschrift registrierten Personen. Mit $\gamma_h = N_h/N$ wird der Anteil der Anschriften und mit $f_h = n_h/N_h$ der Auswahlsatz in der Schicht h bezeichnet. Wie zuvor erläutert, dienen die Anschriftengrößenklassen als Schichtungsmerkmal. Dann ist durch

$$(1) \qquad \hat{\tau}_{kR} = N \cdot \sum_{h=1}^{H} \gamma_h \cdot (\bar{y}_h + (\bar{X}_h - \bar{x}_h) \cdot \beta)$$

der Regressionsschätzwert gegeben, wobei $\bar{X}_h$ der Register- und $\bar{x}_h$ der Stichprobenmittelwert ist. β bezeichnet hierbei den Regressionskoeffizienten im linearen Regressionsmodell. Formal handelt es sich um einen kombinierten Regressionsschätzer, da die Regression gemeinsam über alle Schichten innerhalb der Gemeinde beziehungsweise des Sampling Points hinweg berechnet wird. Der Regressionsschätzer (1) kann auch unter Einsatz mehrerer Hilfsvariablen verwendet werden (siehe beispielsweise Cochran, 1977, Särndal und andere, 1992, Seite 225 ff., oder Lohr, 2010, Seite 372 ff.).

Die Varianz des kombinierten Regressionsschätzers (1) für eine geschichtete Hochrechnung unter Verwendung der Registerinformationen als Hilfsinformationen ergibt sich dann durch

$$(2) \qquad V(\hat{\tau}_{kR}) = N^2 \cdot \sum_{h=1}^{H} \frac{\gamma_h^2(1-f_h)}{n_h} \cdot S_{yh}^2 \cdot (1 - \vartheta^2).$$

Dabei stellt

(3)

$$(1 - \vartheta^2) = 1 - \frac{\left(\sum_{h=1}^{H} \frac{\gamma_h^2(1-f_h)}{n_h} \cdot S_{xyh} \right)^2}{\left(\sum_{h=1}^{H} \frac{\gamma_h^2(1-f_h)}{n_h} \cdot S_{xh}^2 \right) \cdot \left(\sum_{h=1}^{H} \frac{\gamma_h^2(1-f_h)}{n_h} \cdot S_{yh}^2 \right)}$$

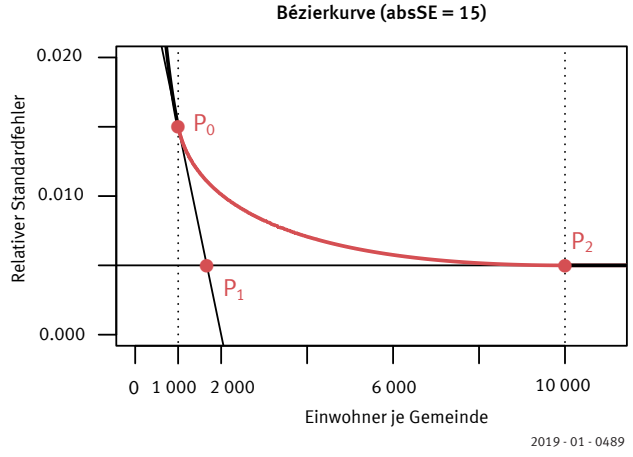
den Varianzreduktionsfaktor dar, der sich aus der Korrelation von abhängiger Variablen und Hilfsvariablen ergibt (Krug und andere, 2001, hier: Formel 7.1.40). Für eine detailliertere Herleitung von Formel (3) siehe Burgard und andere (2019b). Für den Zensus 2021 wird der Varianzreduktionsfaktor (3) je Gemeinde beziehungsweise Sampling Point aus der Zensusstichprobe von 2011 geschätzt. Eine durch die Schätzung entstehende Unsicherheit wird in der Berechnung des benötigten Gesamtstichprobenumfangs in Kapitel 4 berücksichtigt (Burgard und andere, 2019b).

Präzisionsvorgaben

Ausgangspunkt der Präzisionsvorgaben für die Schätzung mittels des Regressionsschätzers (GREG-Schätzung) aus Formel (1) ist die analog zum Zensus 2011 festgelegte Formulierung des § 7 Absatz 1 Zensusgesetz 2011, in welchem die «*Ermittlung der amtlichen Einwohnerzahl* (für Gemeinden mit mindestens 10 000 Einwohnerinnen und Einwohnern) *mit einer angestrebten Genauigkeit eines einfachen relativen Standardfehlers von höchstens 0,5 %*» festgelegt war. Dieser Wert sollte für den Zensus 2021 für die zu betrachtenden Ausweiseinheiten unangetastet bleiben, sodass sich daraus – unter Verwendung von Schichtgrößen, schichtspezifischen Varianzen und Varianzreduktionsfaktoren (Burgard und andere, 2019b) – Mindeststichprobenumfänge berechnen lassen. Für sehr kleine Gemeinden macht eine Formulierung einer relativen Präzisionsvorgabe jedoch keinen Sinn mehr, da eine klassische Stichprobenziehung nicht mehr möglich ist (Verma, 1991). Als Alternative zum relativen Standardfehler wird für die sehr kleinen Gemeinden (weniger als 1 000 Einwohnerinnen und Einwohner) eine Präzisionsvorgabe durch einen absoluten Standardfehler von 15 festgelegt. Für Gemeinden größer 10 000 Einwohner bleibt das relative Präzisionsziel von 0,5 % relativem Standardfehler bestehen. Für das Übergangsintervall von 1 000 bis 10 000 Einwohnerinnen und Einwohnern schlagen Burgard und andere (2019b) die Verwendung einer Bézierkurve vor, die einen glatten Übergang zwischen absoluter und relativer Präzisionsvorgabe ermöglicht (Weisstein, 2018).

Grafik 1

Bézierkurve zur Abhängigkeit des relativen Standardfehlers von der Gemeindegröße



Somit wird ein abrupter Sprung in den Präzisionsvorgaben vermieden. Ein weiterer Vorteil der Bézierkurve ist, dass sie ohne die weitere Festlegung von Parametern auskommt. Es werden nur Start- und Endpunkt (Punkte P_0 und P_2 in Grafik 1) des Übergangsintervalls, sowie die Steigungen an diesen beiden Punkten benötigt. ↘ Grafik 1

Der Verlauf der Bézierkurve ist in Grafik 1 dargestellt, die Funktionsvorschrift der Funktion des absoluten Standardfehlers *absSE* in Abhängigkeit der Gemeindegröße von Gemeinde N_h ist dann gegeben durch

$$(4)\ absSE(N_h) = \begin{cases} 15, & \text{für } 0 \leq N_h < 1\,000 \\ \frac{7}{1\,000} N_h - \frac{14}{25\sqrt{2}} \sqrt{N_h - 200} + \frac{96}{5}, & \text{für } 1\,000 \leq N_h < 10\,000 \\ \frac{5}{1\,000} N_h, & \text{für } N_h \geq 10\,000. \end{cases}$$

Somit wird durch die Verbindung von relativem und absolutem Standardfehler durch eine Bézierkurve eine glatte Präzisionszielfunktion erreicht. Durch die Anwendung für alle Gemeinden beziehungsweise Sampling Points und die Verwendung einer einheitlichen Schätzmethodik lässt sich ein Methodenbruch vermeiden.

Schichtung

Im Zensus 2011 wurden innerhalb jedes Sampling Points acht einwohnergleiche Schichten nach Anschriftengrößen gebildet, die sich nach der Anzahl der gemeldeten Personen je Anschrift definierten. Die Größe der Schichten wurde approximativ personengleich gewählt. Die Schichtung nach Anschriftengrößenklassen hat sich bewährt, da dadurch eine homogene Schichtstruktur entsteht, was die Effizienz der Schätzungen im Allgemeinen steigert. Um zum einen der sehr unterschiedlichen Größenstruktur der Gemeinden gerecht zu werden und zum anderen die Schichtung weiter zu optimieren, wird im Zensus 2021 eine variable Schichtzahl von 1 bis 16 Schichten je Gemeinde gebildet. Ausführliche Untersuchungen zeigen, dass diese Schichtung eine deutliche Effizienzsteigerung im Sinne der Verbesserung der Schätzgüte mit sich bringt. Letztendlich kann somit bei

gleicher Schätzqualität ein nicht zu vernachlässigender Teil des Stichprobenumfangs im Vergleich zu 2011 eingespart werden.

3

Gemeindestrukturabhängige Stichprobenkonzepte für den Zensus 2021

Vorweg ist anzumerken, dass die Untersuchungen und Ergebnisse auf einer synthetischen Grundgesamtheit aller Personen in Deutschland aus dem Jahr 2011 beruhen, die aus einem Registerabzug von 2008 erzeugt wurde. Die Gemeindestruktur wurde auf den Stand 1. Januar 2019 aktualisiert. Weitere Informationen zur synthetischen Grundgesamtheit sind in Münnich und andere (2012a) sowie Burgard und andere (2019b) nachzulesen.

Allokation des Gesamtstichprobenumfangs

Durch die Festlegung der Schichtung nach Gemeinden und einer variablen Anzahl von Anschriftengrößenklassen (1 bis 16 Schichten) kann nun unter Verwendung der Präzisionszielfunktion und der gemeindespezifischen Varianzreduktionsfaktoren ein benötigter Mindeststichprobenumfang berechnet werden. Dies wird numerisch durch ein iteratives Verfahren umgesetzt, das im Folgenden mit Amin bezeichnet wird. Weitere zuvor zu definierende Parameter sind untere und obere Schranken für die Auswahlsätze. Diese werden je Schicht auf 5 % beziehungsweise 50 % festgesetzt. Allerdings stellt sich nun die Frage, ob das Schätzverfahren mittels des GREG-Schätzers aus Formel (1) unabhängig von der gemeindespezifischen Stichprobengröße anwendbar ist. Dies ist nicht der Fall, da die Stabilität der GREG-Schätzungen nicht unabhängig von der Stichprobengröße ist. Sollen also in diesen kleinen Gemeinden statistische Auswertungen auf Basis einer Stichprobe erstellt werden, so ist dies kritisch zu diskutieren. Untersuchungen in Burgard und andere (2019b) zeigen, dass ein Mindeststichprobenumfang von 100 Anschriften je Gemeinde notwendig ist, um diese Stabilität zu gewährleisten. Als Folge dessen ist ein maximaler Auswahlsatz von 50 % für Gemeinden mit weniger als 200 Anschriften nicht mehr ausreichend. Dort wird die obere Schranke des Auswahlsatzes für die entsprechende Gemeinde erhöht, um einen Mindeststichprobenumfang von 100 Anschriften je Gemeinde zu erreichen.

Um eine kompensatorische optimale Allokation zu erzielen, wird im Anschluss an Amin der aus Amin resultierende Gesamtstichprobenumfang als Basis für die sogenannte Box-constraint optimale Allokation (BCopt) verwendet. Diese Allokation verteilt iterativ den Gesamtstichprobenumfang so, dass die gewichtete Summe der quadrierten RRMSEs[1] der Totalwertschätzer aller Gemeinden minimiert wird (Gabler und andere, 2012; Münnich und andere, 2012b; Friedrich und andere, 2015 und 2018). Dabei werden die geschätzten Varianzreduktionsfaktoren aus Formel (3) nicht je Gemeinde separat berücksichtigt, sondern über alle Gemeinden hinweg gemittelt. Dies soll eine ungleiche Behandlung einzelner Gemeinden aufgrund der konkreten aus der Stichprobe ermittelten Ergebnisse des Zensus 2011 verhindern. Die vorgegebenen Präzisionsvorgaben je Gemeinde, berechnet anhand der Präzisionszielfunktion [siehe Formel (4)], werden weiterhin im resultierenden Optimierungsproblem berücksichtigt. Diese Vorgehensweise hatte sich bereits beim Zensus 2011 bewährt (Münnich und andere, 2012a, Abschnitt 2.2.4).

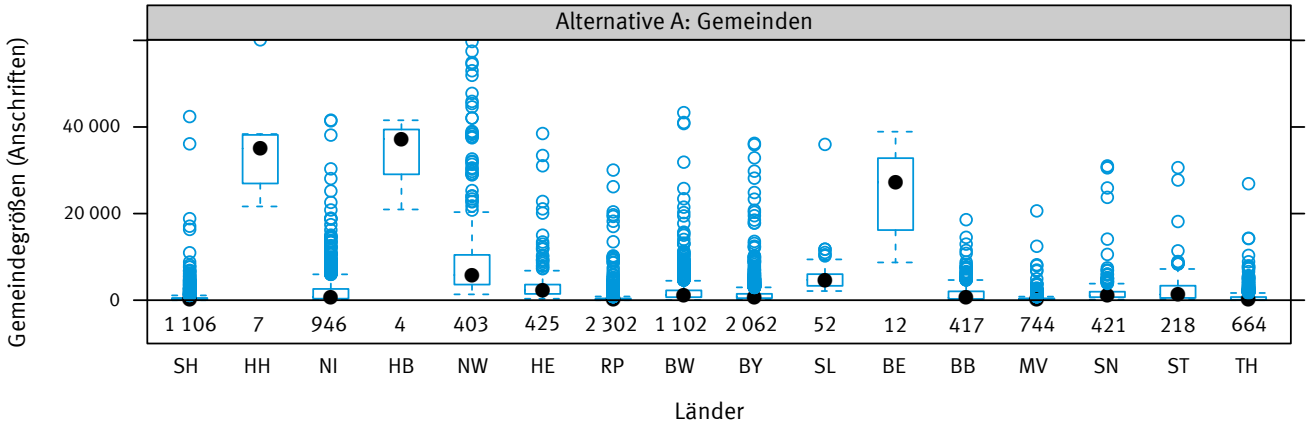
Für detaillierte Beschreibungen zu den verwendeten Allokationstechniken und Analysen zur Wahl der Parameter siehe Burgard und andere (2019b).

Problematik im Umgang mit kleinen Gemeinden

Im Vergleich zum Zensus 2011, bei dem 2 391 Sampling Points bei der Stichprobenziehung zu berücksichtigen waren, sind für den Zensus 2021 alle 10 885 Gemeinden Deutschlands einzubeziehen. Alle Gemeinden müssen unabhängig von ihrer Größe die Präzisionsanforderungen erfüllen; zudem wurde ein Mindeststichprobenumfang von 100 Anschriften je Gemeinde festgelegt. In ↘ Grafik 2 sind die Anzahl der Gemeinden (als numerische Zahl) sowie deren Größen (als Boxplot) je Land dargestellt. Es wird schnell ersichtlich, dass sich zum einen die Gemeindestrukturen innerhalb der Länder

1 Der Relative Root Mean-Squared Error wird als Wurzel aus der mittleren relativen quadratischen Abweichung zwischen dem Schätzwert und dem wahren Wert gebildet.

Grafik 2
Gemeinden (numerische Zahl) und Gemeindegrößen (Boxplot) aufgelistet nach Ländern



SH: Schleswig-Holstein; HH: Hamburg; NI: Niedersachsen; HB: Bremen; NW: Nordrhein-Westfalen; HE: Hessen; RP: Rheinland-Pfalz; BW: Baden-Württemberg; BY: Bayern; SL: Saarland; BE: Berlin; BB: Brandenburg; MV: Mecklenburg-Vorpommern; SN: Sachsen; ST: Sachsen-Anhalt; TH: Thüringen.

2019 - 01 - 0490

deutlich unterscheiden, und es zum anderen eine Vielzahl von sehr kleinen Gemeinden gibt. Dies trifft insbesondere in den Ländern Schleswig-Holstein, Rheinland-Pfalz, Mecklenburg-Vorpommern und Thüringen zu.

Als Folge der in Grafik 2 dokumentierten Vielzahl sehr kleiner Gemeinden würde die Allokation für dieses Stichprobenkonzept (in Folge bezeichnet als Alternative A) mit einem Mindeststichprobenumfang von 100 Anschriften eine Vollerhebung in insgesamt 1 076 der 10 885 Gemeinden ergeben (etwa 9,9 %). Zudem führen die über die Länder hinweg sehr heterogenen Gemeindestrukturen zu einer deutlichen Ungleichverteilung der erwähnten voll zu erhebenden Gemeinden. Die Aufteilung der Vollerhebungen nach den Ländern enthält ↘ Tabelle 1. Dies wirkt sich zudem stark auf den länderspezifischen Gesamtauswahlsatz aus. Dieser liegt für Thüringen bei rund 18 %, für Schleswig-Holstein bei rund 20 % und für Rheinland-Pfalz und Mecklenburg-Vorpommern sogar bei jeweils rund 25 %. Der über Deutschland gemessene Auswahlsatz liegt bei 12,7 % aller Einwohnerinnen und Einwohner. Darüber hinaus würden trotz dieser hohen Anzahl von Vollerhebungen die Präzisionszielvorgaben in insgesamt rund 12 % der Gemeinden nicht eingehalten werden können. Dabei ist anzumerken, dass die Überschreitung der Präzisionsvorgabe meist sehr gering ausfällt.

Tabelle 1
Vollerhebungen bei gemeindespezifischer Schichtung nach Ländern

	Anzahl
Deutschland	1 076
Schleswig-Holstein	226
Hamburg	0
Niedersachsen	0
Bremen	0
Nordrhein-Westfalen	0
Hessen	0
Rheinland-Pfalz	585
Baden-Württemberg	17
Bayern	1
Saarland	0
Berlin	0
Brandenburg	0
Mecklenburg-Vorpommern	82
Sachsen	0
Sachsen-Anhalt	0
Thüringen	165

Aus Tabelle 1 und den vorherigen Bemerkungen ist ersichtlich, dass dieses Stichprobenkonzept zu einer sehr ungleichen Belastung der Länder führt. In Bezug auf die teils hohe Anzahl von Vollerhebungen lässt sich aus dem Urteil des Bundesverfassungsgerichts (BVerfG, 2018) ableiten, dass eine Vollerhebung unter anderem aufgrund der informationellen Selbstbestimmung der Bürgerinnen und Bürger falls möglich vermieden werden soll. Dem gegenüber steht die Mindeststichprobengröße von 100 Anschriften je Gemeinde, die aus statistischer Sicht unumgänglich ist. Dieser Konflikt zwischen einem

vorgegebenen Mindeststichprobenumfang je Gemeinde und einer angemessenen Belastung der Bevölkerung kann nur durch eine (gegebenenfalls teilweise) Abweichung von der gemeindespezifischen Betrachtung gelöst werden. Im folgenden Abschnitt werden zwei Alternativen B und C aufgezeigt, die auf dieser Herangehensweise beruhen. Dabei basiert die Schichtung nicht mehr ausschließlich auf der Struktur der Gemeinden, sondern teilweise auch auf einer stärker aggregierten Ebene, den Gemeindeverbänden. In diesem Zusammenhang wird für die letztliche Umsetzung eine Betrachtung auf Gemeindeverbandsebene erlaubt, sofern diese gemeinsame Melderegister führen.[2] Dies ist nicht in allen Bundesländern der Fall. Gemeindeverbände entsprechen in vielen Bundesländern den großen Gemeinden, während kleinere Gemeinden häufig zu einem Gemeindeverband zusammengefasst werden. Zudem stimmen in einigen Bundesländern die Gemeindeverbände mit den Gemeinden vollkommen überein.

Um die Folgen der Betrachtung sehr kleiner Gemeinden zu verringern, wird für die Alternative B folgende Schichtung verwendet:

(i) Gemeinden mit mehr als 2 000 Einwohnerinnen und Einwohnern bilden jeweils eine eigene Schicht (unverändert zu Alternative A).

(ii) Gemeinden mit weniger als 2 000 Einwohnerinnen und Einwohnern werden zusammengefasst, sofern sie innerhalb eines Gemeindeverbandes liegen.

Die Wahl der Grenze von 2 000 Einwohnerinnen und Einwohnern basiert unter anderem auf der Analyse von Konfidenzintervallüberdeckungsraten in Burgard und andere (2019b). Die Schichtung innerhalb einer regionalen Ausweiseinheit bleibt unverändert variabel mit einer Anzahl von 1 bis 16 Schichten. Somit ergeben sich in Deutschland 6 393 Sampling Points, von denen rund 80 % einzelnen Gemeinden mit über 2 000 Einwohnerinnen und Einwohnern entsprechen. Die restlichen 20 % der Sampling Points ergeben sich aus Zusammenschlüssen von kleinen Gemeinden innerhalb eines Gemeindeverbands. Letztlich wird die Einhaltung des Präzisionsziels bei dieser Alternative nur für die neu entstandenen Sampling Points gefordert. Bei einer Verwendung dieses Stichprobenkonzepts ergeben sich nur noch 28 Sampling Points, in denen eine Vollerhebung durchgeführt werden muss.[3] Dies entspricht einem Anteil von 0,4 % und bedeutet im Vergleich zu dem Anteil von 9,9 % bei Alternative A eine sehr deutliche Verringerung. Zudem liegen alle Auswahlsätze unter 17 %; insbesondere in den vier oben genannten Bundesländern mit sehr kleinen Gemeinden wird deutlich Stichprobenumfang eingespart. Das senkt die Kosten in diesen Ländern und gleicht sie an die Kosten in den restlichen Ländern an. Der Auswahlsatz für Gesamtdeutschland liegt bei 11,2 %.

Die Alternative C greift die Vorgehensweise aus Alternative B auf, allerdings werden die Gemeinden ab 2 000 Einwohnerinnen und Einwohnern nicht mehr gesondert betrachtet; sie verbleiben im zugehörigen Gemeindeverband. Die übrige Vorgehensweise bleibt identisch. Für Deutschland ergeben sich dann 4 550 Sampling Points, die äquivalent zu den Gemeindeverbänden sind. Trotz weiterer Zusammenfassungen bleiben 28 Gemeindeverbände, in denen eine Vollerhebung durchgeführt werden muss.[4] Der Gesamtauswahlsatz für Deutschland beträgt bei Alternative C 10,1 %. Die Auswahlsätze für Schleswig-Holstein, Mecklenburg-Vorpommern und Thüringen liegen nun im Mittelfeld aller länderspezifischen Auswahlsätze, was zu einer ähnlichen finanziellen Belastung der Länder führen würde. Aufgrund der großen Gemeindeverbände in Rheinland-Pfalz befindet sich dessen Auswahlsatz allerdings nur noch im unteren Bereich aller Länder (8,9 %). Da die Qualitätsmessung auf den regionalen Ausweiseinheiten erfolgt, könnten sich bei der Betrachtung gemeindespezifischer Schätzungen vereinzelt Effizienzverluste ergeben. Darauf wird im folgenden Abschnitt ausführlicher eingegangen.

↘ Tabelle 2 listet die Anzahl der Gemeinden für Alternative A, die Anzahl der Sampling Points bei Alternative B und die Anzahl der Gemeindeverbände unter Alternative C je Land auf. Hier wird nochmals deutlich, dass hauptsächlich in den Bundesländern mit sehr kleinen Gemeinden sehr viele Gemeinden zusammengefasst wurden. Analog zu Grafik 2 sind die Anzahl der Sampling Points (Alternative B) und Gemeindeverbände (Alternative C) sowie deren Größen je Land in Grafik 3 auf Seite 31 dargestellt.

2 Zudem ist eine Besonderheit in Thüringen zu beachten: Dort dürfen einige teilweise sehr kleine Gemeinden trotz Gemeindeverbandsstruktur nicht zu Sampling Points zusammengefasst werden.

3 Diese liegen in Thüringen und repräsentieren kleine Gemeinden, die aufgrund der Sonderregel (siehe Fußnote 2) nicht zusammengefasst werden dürfen.

4 Siehe Fußnote 3.

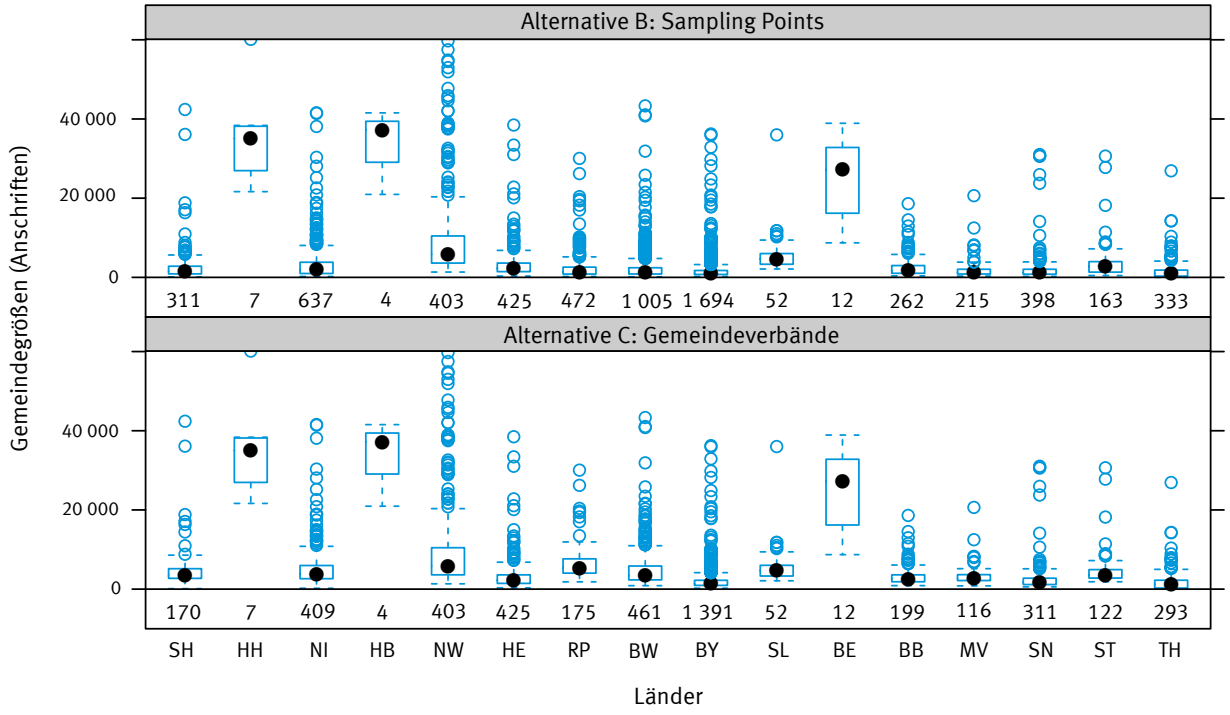
Tabelle 2
Gemeinden, Sampling Points und Gemeindeverbände je Land

	Gemeinden (Alternative A)	Sampling Points (Alternative B)	Gemeindeverbände (Alternative C)
Deutschland	10 885	6 393	4 550
Schleswig-Holstein	1 106	311	170
Hamburg	7	7	7
Niedersachsen	946	637	409
Bremen	4	4	4
Nordrhein-Westfalen	403	403	403
Hessen	425	425	425
Rheinland-Pfalz	2 302	472	175
Baden-Württemberg	1 102	1 005	461
Bayern	2 062	1 694	1 391
Saarland	52	52	52
Berlin	12	12	12
Brandenburg	417	262	199
Mecklenburg-Vorpommern	744	215	116
Sachsen	421	398	311
Sachsen-Anhalt	218	163	122
Thüringen	664	333	293

Trotz der Möglichkeit, Gemeinden zusammenzufassen beziehungsweise die Allokation und die Stichprobenziehung unter Berücksichtigung von Gemeindeverbänden statt Gemeinden durchzuführen, bleibt eine gemeindespezifische Schätzung der Einwohnerzahl und deren Qualitätsmessung unausweichlich. Die gemeindeseparate GREG-Schätzung mittels des in Kapitel 2 vorgestellten Schätzers [Formel (1)] ist dabei im Allgemeinen nicht mehr möglich. Der gemeindespezifische Stichprobenumfang ist nicht mehr fix, sondern eine Zufallszahl. Deshalb ist nicht mehr gewährleistet, dass der Stichprobenumfang in jeder Gemeinde ausreicht, um stabile GREG-Schätzungen zu erreichen. Im Extremfall kann es passieren, dass einer Gemeinde keine Stichprobeneinheit zugeordnet wird. Um trotzdem Ergebnisse für die kleinen Gemeinden auszuweisen, werden in Burgard und andere (2019b) synthetische Schätzmethoden (siehe auch Rao/Molina, 2015; Münnich und andere, 2013; Tzavidis und andere, 2018; Burgard und andere, 2019c) beziehungsweise modellassistierte Schätzungen mittels eines verallgemeinerten Kalibrierungsschätzers (Burgard und andere, 2019a) empfohlen, um kohärente Schätzungen zu erhalten. Zudem könnten sich die Schätzungen für kleiner werdende Gemeinden immer mehr an den Registerwerten orientieren. Wie in Rupp (2018, Kapitel 5) analysiert, ist dieser verallgemeinerte Kalibrierungsschätzer von hoher numerischer Stabilität und somit auch auf große Datenmengen – wie sie beim Zensus vorliegen – anwendbar. Durch dieses Verfahren würde die Schätzung auf Gemeindeebene durch eine nachträgliche Kalibrierung erfolgen, die sowohl Register- als auch Stichprobeninformationen nutzt. Dabei misst sie dem Register umso höheres Gewicht bei, je weniger Stichprobenanschriften in der betreffenden Gemeinde gezogen wurden. Auch eine Einbeziehung der Gemeindegröße ist denkbar. Somit ist ein glatter Übergang gewährleistet; die ausgewiesenen Schätzungen für sehr kleine Gemeinden werden sich umso mehr an den vorhandenen Registerwerten orientieren, je geringer der auf die Gemeinde entfallende Stichprobenumfang ist. ↘ Grafik 3

Aufgrund der insbesondere in Tabelle 2 abzulesenden stark heterogenen Struktur der Gemeinden über die Länder hinweg ist eine länderspezifische Wahl der Methodik aus wissenschaftlicher Sicht durchaus zu empfehlen. Eine somit entstehende Hybrid-Variante aus den Alternativen A, B und C kann auf die landesspezifischen Besonderheiten besonders gut abgestimmt sein und ermöglicht somit eine bestmögliche Schätzqualität für den Zensus 2021. Die Ergebnisse (Auswahlsätze, Gesamtstichprobenumfang und so weiter) für diese Hybrid-Variante werden im Allgemeinen zwischen den Ergebnissen für die Alternativen A, B und C liegen. Durch die über alle Länder hinweg durchgeführten optimalen Allokationen und die Verwendung der Varianzreduktionsfaktoren [Formel (3)] kann es jedoch zu geringfügigen Abweichungen kommen.

Grafik 3
Anzahl von Sampling Points und Gemeindeverbänden (numerische Zahl) und deren Größen (Boxplot) aufgelistet nach Ländern



SH: Schleswig-Holstein; HH: Hamburg; NI: Niedersachsen; HB: Bremen; NW: Nordrhein-Westfalen; HE: Hessen; RP: Rheinland-Pfalz; BW: Baden-Württemberg; BY: Bayern; SL: Saarland; BE: Berlin; BB: Brandenburg; MV: Mecklenburg-Vorpommern; SN: Sachsen; ST: Sachsen-Anhalt; TH: Thüringen.

2019 - 01 - 0491

4

Konservative Schätzung des Gesamtstichprobenumfangs

In diesem Kapitel werden die benötigten Gesamtstichprobenumfänge beziehungsweise die dazugehörenden Auswahlsätze verglichen und kritisch analysiert. Es ist hier anzumerken, dass diese Stichprobenumfänge sich durch die kompensatorische BCopt-Allokation ergeben, wie bereits zu Beginn von Kapitel 3 erläutert. Darüber hinaus sind die Ergebnisse einschließlich eines Sicherheitsaufschlagfaktors angegeben, der aufgrund der unsicheren Schätzung der gemeindespezifischen Varianzreduktionsfaktoren [Formel (3)] für eine konservative Schätzung notwendig ist. Für Informationen zu dessen Berechnung wird auf Burgard und andere (2019b) verwiesen. Aufgrund der zur Verfügung stehenden Daten können weitere Aufschlagfaktoren, wie etwa für Nebenwohnsitze, Nachziehungen und Gemeinschaftsunterkünfte, sowie für Sondereffekte durch Ausreißer, Schichtwechsler, Nullanschriften oder die Sondermigration nicht in den Ergebnissen berücksichtigt werden.

Als Determinanten für die berechneten Gesamtstichprobenumfänge sind zu nennen:

- die Präzisionszielfunktion (Kapitel 2)
- die BCopt-Allokation und die Schranken für die Auswahlsätze (Kapitel 3)
- der Umgang mit Kleinstgemeinden (Alternative A, B oder C) und die Schichtung nach Anschriftengrößenklassen (Kapitel 3)
- die Varianzreduktionsfaktoren (Kapitel 2)

↘ Tabelle 3 listet die länderspezifischen Stichprobenumfänge für die Alternativen A, B und C auf. In der ersten Zeile ist der Gesamtstichprobenumfang in Form der Anzahl der Personen für Deutschland abzulesen. Bei der Wahl der Alternative A – also einer gemeindespezi-

Tabelle 3
Länderspezifische Stichprobenumfänge

	Gemeinden (Alternative A)	Sampling Points (Alternative B)	Gemeindeverbände (Alternative C)
	Personen		
Deutschland	10 688 230	9 492 218	8 496 572
Schleswig-Holstein	619 161	403 100	315 151
Hamburg	96 784	96 927	98 539
Niedersachsen	996 619	900 303	772 069
Bremen	36 787	36 841	37 519
Nordrhein-Westfalen	1 277 574	1 283 114	1 321 127
Hessen	696 572	701 189	736 622
Rheinland-Pfalz	1 056 552	571 645	372 940
Baden-Württemberg	1 462 173	1 444 735	1 035 121
Bayern	2 075 491	1 970 643	1 884 380
Saarland	95 186	95 779	99 546
Berlin	185 010	185 285	187 811
Brandenburg	386 866	347 884	322 154
Mecklenburg-Vorpommern	429 490	282 442	210 352
Sachsen	564 059	561 512	528 867
Sachsen-Anhalt	279 983	260 022	238 404
Thüringen	429 923	350 797	335 970

fischen Allokation mit einem Mindeststichprobenumfang von 100 Anschriften – ergibt sich ein Gesamtstichprobenumfang von 10,7 Millionen Personen. Werden kleine Gemeinden aus Gemeindeverbänden zu Gemeindeverbandsresten zusammengefasst (Alternative B), so bedarf es nur noch einem Stichprobenumfang von 9,5 Millionen Personen. Dabei wird hauptsächlich in den Ländern der Stichprobenumfang reduziert, in denen viele kleine Gemeinden zusammengefasst werden – also vor allem in Schleswig-Holstein, Rheinland-Pfalz, Mecklenburg-Vorpommern und Thüringen. Dieser Effekt steigert sich nochmals bei der Betrachtung von vollständigen Gemeindeverbänden (Alternative C), wobei insgesamt 8,5 Millionen Personen gezogen werden sollten.

↘ Tabelle 4 enthält die zu Tabelle 3 gehörenden Auswahlsätze. Bei deren Analyse fällt neben der zuvor schon beschriebenen Reduktion des Auswahlsatzes für die Alternativen B und C eine extreme Ungleichverteilung der länderspezifischen Auswahlsätze bei Alternative A auf. Bei einem durchschnittlichen Auswahlsatz von 12,7 % aller Personen in Deutschland müssten in Rheinland-Pfalz beispielsweise 25,2 % aller Personen befragt werden, Berlin dagegen käme mit einem Auswahlsatz von 5,4 % aus. Der Hauptgrund dafür ist die große Anzahl kleiner Gemeinden in Rheinland-Pfalz. Neben dem sehr hohen finanziellen Aufwand für die betroffenen Länder widerspricht dieses Ergebnis auch dem Gleichbehandlungsgrundsatz, der für alle Länder gelten sollte. Zudem führt die in Tabelle 1 aufgezeigte hohe Anzahl von über 1 000 Gemeinden mit Vollerhebungen dazu, dass das Recht auf informationelle Selbstbestimmung der Bürgerinnen und Bürger teilweise eingeschränkt werden könnte. Das soll dem Urteil des Bundesverfassungsgerichts zufolge vermieden werden (BVerfG, 2018). Die Alternative B führt schon zu einer deutlichen Angleichung der länderspezifischen Auswahlsätze. Diese Angleichung wird bei der Alternative C, in der ausschließlich die Gemeindeverbandsstruktur genutzt wird, noch weiter vorangetrieben.

Tabelle 4
Länderspezifische Auswahlsätze

	Gemeinden (Alternative A)	Sampling Points (Alternative B)	Gemeinde-verbände (Alternative C)
	%		
Deutschland	12,7	11,2	10,1
Schleswig-Holstein	20,5	13,4	10,5
Hamburg	5,6	5,6	5,7
Niedersachsen	12,0	10,9	9,3
Bremen	5,5	5,5	5,6
Nordrhein-Westfalen	7,0	7,0	7,2
Hessen	11,0	11,0	11,6
Rheinland-Pfalz	25,2	13,6	8,9
Baden-Württemberg	13,4	13,2	9,5
Bayern	16,1	15,2	14,6
Saarland	8,8	8,9	9,2
Berlin	5,4	5,4	5,4
Brandenburg	14,8	13,3	12,3
Mecklenburg-Vorpommern	24,6	16,2	12,0
Sachsen	13,1	13,1	12,3
Sachsen-Anhalt	11,4	10,6	9,7
Thüringen	18,3	15,0	14,3

Trotz der Kostenersparnis bei den Alternativen B und C muss jedoch auch kritisch über die Zusammenfassungen geredet werden. Insbesondere bei Alternative C wird im Vergleich zur Alternative A der Stichprobenumfang in kleinen Gemeinden drastisch reduziert. Dieser Effekt ist am größten, wenn die Gemeindeverbände, die über viele Gemeinden hinweg gebildet werden, sehr groß sind. Zudem kann es bei Alternative C vorkommen, dass ein Gemeindeverband aus einer mittelgroßen Gemeinde und einigen sehr kleinen Gemeinden besteht. In diesen Fällen sinkt nicht nur der Stichprobenumfang, der auf die kleinen Gemeinden entfällt, sondern es gibt möglicherweise auch Gemeinden ohne Stichprobeneinheiten. Zudem ist es möglich, dass sogar in größeren Gemeinden die Qualitätsanforderung von 0,5 % relativem Standardfehler nicht mehr eingehalten werden kann. Dies ist bei den Alternativen B und C denkbar, da nicht mehr alle Gemeinden als eigene Schicht im Design berücksichtigt werden. Als Konsequenz werden die gemeindespezifischen Schätzungen – die weiterhin notwendig sind – trotz der Verwendung synthetischer Schätzmethoden und Kalibrierungstechniken deutlich erschwert. Dies kann zu vergleichsweise ineffizienten gemeindespezifischen Schätzungen führen. Letztendlich obliegt die Wahl zwischen den Alternativen A, B und C aber den Ländern. Zusammen ergibt sich dann ein Mix dieser drei Methoden, die bereits zuvor als Hybrid-Varianten bezeichnet wurden. Im Allgemeinen können je nach Gemeindegrößenstrukturen die Alternativen A, B oder C angewendet werden. Gleichwohl setzen die Alternativen B und C auf den Sampling Points beziehungsweise Gemeindeverbänden aus methodischen Gründen in etwa konstante Registereigenschaften voraus. Diese Annahme ist insbesondere dann plausibel, wenn mindestens auf den Ausweiseinheiten zentral geführte Register vorliegen.

5

Fazit und Ausblick

Infolge des Urteils des Bundesverfassungsgerichts (BVerfG, 2018) mussten die Informationen, die aus dem Zensus 2011 gewonnen wurden, berücksichtigt werden. Dies betraf insbesondere das Stichprobenverfahren in Bezug auf die 10 000-Einwohner-Grenze. Wie im Artikel ausführlich beschrieben, soll im Jahr 2021 eine für alle regionalen Schichten identische Stichprobenmethodik angewendet werden. Dabei soll das Präzisionsziel eine glatte Funktion darstellen, abhängig von der Größe der Gemeinde beziehungsweise der interessierenden Ausweiseinheit. Dieses Ziel wird durch die Verwendung der in Kapitel 2 beschriebenen Präzisionszielfunktion unter Verwendung einer Bézierkurve erreicht.

Ein Problem stellt dabei die Vielzahl sehr kleiner Gemeinden dar: Ein valider Ausweis von Schätzergebnissen inklusive Fehlermessung ist oft kaum noch möglich beziehungsweise statistisch ungeeignet. Eine Möglichkeit, dieses Problem zu lösen, ist die Verwendung einer Mindeststichprobengröße von 100 Anschriften je Gemeinde (Alternative A). Dadurch steigt zum einen der Gesamtstichprobenumfang, zum anderen muss in etwa 10 % aller Gemeinden dazu eine Vollerhebung durchgeführt werden. Als Alternative werden zwei Ansätze vorgeschlagen (Alternativen B und C), bei denen nicht mehr auf die Gemeinden als Regionen, sondern auf weiter aggregierte Gemeindeverbände zurückgegriffen wird. Dies reduziert den Gesamtstichprobenumfang und die Anzahl an Vollerhebungsschichten, macht jedoch in kleinen Gemeinden eine angepasste modellassistierte Schätzung auf Gemeindeebene erforderlich. Diese basiert auf sehr flexiblen Kalibrierungstechniken und synthetischen Schätzverfahren, die eine stabile gemeindespezifische Punktschätzung ermöglichen können.

Generell ist eine länderspezifische Wahl der drei Alternativen A, B und C im Umgang mit den Kleinstgemeinden möglich. Diese kann allerdings an Bedingungen gebunden sein, wie ein gemeinsam geführtes Melderegister je Sampling Point beziehungsweise Gemeindeverband.

Danksagung

Die Forschungsarbeiten zu diesem Artikel wurden im Rahmen des Forschungsprojekts Research Innovation for Official and Survey Statistics (RIFOSS) durchgeführt, das vom Statistischen Bundesamt gefördert wird. Wir danken den Kolleginnen und Kollegen, die in ausführlichen Diskussionen zur Erstellung dieses Artikels beigetragen haben.

LITERATURVERZEICHNIS

Bierschenk, Michaela/Leischner, Sonja. *Zur Verfassungsmäßigkeit der Vorschriften über den Zensus 2011*. In: WISTA Wirtschaft und Statistik. Ausgabe 1/2019, Seite 11 ff.

Burgard, Jan Pablo/Münnich, Ralf/Rupp, Martin. *A generalized calibration approach ensuring coherent estimates with small area constraints*. Arbeitspapier. Research Papers in Economics, No. 10/19, Universität Trier. 2019a.

Burgard, Jan Pablo/Münnich, Ralf/Rupp, Martin. *Qualitätszielfunktionen für stark variierende Gemeindegrößen im Zensus 2021*. Artikel mit Diskussion. In: AStA Wirtschafts- und Sozialstatistisches Archiv. Im Erscheinen. 2019b.

Burgard, Jan Pablo/Esteban, Maria Dolores/Morales, Domingo/Pérez, Agustin. *A Fay-Herriot model when auxiliary variables are measured with error*. In: TEST. An Official Journal of the Spanish Society of Statistics and Operations Research. Jahrgang 28. https://doi.org/10.1007/s11749-019-00649-3. 2019c.

Bundesverfassungsgericht (BVerfG). *Urteil des Zweiten Senats vom 19. September 2018 - 2 BvF 1/15 - Rn. (1-357)*. 2018. Verfügbar unter: www.bverfg.de

Christensen, Björn/Christensen, Sören/Hoppe, Tim/Spandel, Michael. *Everything counts!* In: AStA Wirtschafts- und Sozialstatistisches Archiv. Jahrgang 9. Ausgabe 3-4/2015, Seite 215 ff.

Cochran, William G. *Sampling Techniques, Third Edition.* New York 1977.

Friedrich, Ulf/Münnich, Ralf/Rupp, Martin. *Multivariate Optimal Allocation with Box-Constraints*. In: Austrian Journal of Statistics. Jahrgang 47. Ausgabe 2/2018, Seite 33 ff.

Friedrich, Ulf/Münnich, Ralf/de Vries, Sven/Wagner, Matthias. *Fast integer-valued algorithms for optimal allocations under constraints in stratified sampling*. In: Computational Statistics and Data Analysis. Band 92. Dezember 2015, Seite 1 ff.

Gabler, Siegfried/Ganninger, Matthias/Münnich, Ralf. *Optimal allocation of the sample size to strata under box constraints*. In: Metrika. Jahrgang 75. Ausgabe 2/2012, Seite 151 ff.

Gößl, Thomas. *Der Zensus vor dem Bundesverfassungsgericht*. In: Bayern in Zahlen. Ausgabe 10/2018, Seite 684 ff.

Grohmann, Heinz. *Von der Volkszählung zum Registerzensus: Paradigmenwechsel in der deutschen amtlichen Statistik*. In: AStA Wirtschafts- und Sozialstatistisches Archiv. Jahrgang 3. Ausgabe 1/2009, Seite 3 ff.

Krug, Walter/Nourney, Martin/Schmidt, Jürgen. *Wirtschafts- und Sozialstatistik: Gewinnung von Daten*. 6. Auflage. Berlin 2001.

LITERATURVERZEICHNIS

Leischner, Sonja/Weigelt, Sabine. *Vorschriften über Volkszählung 2011 verfassungsgemäß, Anmerkung zu BVerfG, Urteil vom 19. September 2018, 2 bvf1/15, 2 bvf2/15*. In: NVwZ Neue Zeitschrift für Verwaltungsrecht. Jahrgang 37. Heft 22/2018, Seite 1703, 1731 ff.

Lohr, Sharon L. *Sampling: Design and Analysis*. Boston 2010.

Münnich, Ralf/Burgard, Jan Pablo/Vogt, Martin. *Small Area-Statistik: Methoden und Anwendungen*. In: AStA Wirtschafts- und Sozialstatistisches Archiv. Jahrgang 6. Ausgabe 3-4/2013, Seite 149 ff.

Münnich, Ralf/Gabler, Siegfried/Ganninger, Matthias/Burgard, Jan Pablo/Kolb, Jan-Philipp. *Das Stichprobendesign des registergestützten Zensus 2011*. In: Methoden – Daten – Analysen. Jahrgang 5. Heft 1/2011, Seite 37 ff.

Münnich, Ralf/Gabler, Siegfried/Ganninger, Matthias/Burgard, Jan Pablo/Kolb, Jan-Philipp. *Stichprobenoptimierung und Schätzung im Zensus 2011*. Band 21 der Schriftenreihe Statistik und Wissenschaft. Statistisches Bundesamt (Herausgeber). Wiesbaden, 2012a.

Münnich, Ralf/Sachs, Ekkehard/Wagner, Matthias. *Numerical solution of optimal allocation problems in stratified sampling under box constraints*. In: Advances in Statistical Analysis. Jahrgang 96. Ausgabe 3/2012, Seite 435 ff. 2012b.

Rao, J. N. K./Molina, Isabel. *Small Area Estimation, 2nd Edition*. Wiley Series in Survey Methodology. Hoboken, New Jersey, 2015.

Rendtel, Ulrich. *Warum im Zensus die Ergebnisse der Stichprobenmethode keine Benachteiligung der großen Gemeinden darstellen. Eine Detektivarbeit*. In: AStA Wirtschafts- und Sozialstatistisches Archiv. Jahrgang 9. Ausgabe 3-4/2015, Seite 233 ff.

Rupp, Martin. *Optimization for Multivariate and Multi-domain Methods in Survey Statistics*. Dissertation, Universität Trier. 2018. Verfügbar unter: https://ubt.opus.hbz-nrw.de/frontdoor/index/index/searchtype/collection/id/382/docId/985/start/0/rows/10

Särndal, Carl-Erik/Swensson, Bengt/Wretman, Jan. *Model Assisted Survey Sampling*. New York 1992.

Statistische Ämter des Bundes und der Länder (Herausgeber). *Zensus 2011 – Methoden und Verfahren*. Wiesbaden 2015. Verfügbar unter: www.zensus2011.de

Verma, Vijay. *Sampling Methods: Training Handbook*. Tokyo: Statistical Institute for Asia and the Pacific (SIAP). 1991.

Weisstein, Eric W. *Bézier Curve. Von MathWorld–A Wolfram Web Resource*. 2018. [Zugriff am 18. Juni 2019].
Verfügbar unter: http://mathworld.wolfram.com/BezierCurve.html

Dr. Nilgun Massih-Tehrani

ist Diplom-Soziologin und seit 2018 als Referentin im Sachgebiet „Zensus“ des Bayerischen Landesamts für Statistik tätig. Sie ist Mitarbeiterin im Teilprojekt „Haushaltegenerierung“ des Zensus 2021.

Adrian Reichert

hat in München Volkswirtschaftslehre und Statistik studiert. Im Rahmen des Zensus 2021 leitet er den Projektbereich „Gebäude- und Wohnungszählung, Haushaltegenerierung“ im Bayerischen Landesamt für Statistik.

DIE HAUSHALTEGENERIERUNG IM ZENSUS 2021

Dr. Nilgun Massih-Tehrani, Adrian Reichert

↘ **Schlüsselwörter:** Zensus 2021 – Haushaltegenerierung – registergestützter Zensus – Einwohnerzahl – Wohnhaushalt

ZUSAMMENFASSUNG

Ziel des Zensus 2021 ist neben der Ermittlung der Einwohnerzahl auch die Bereitstellung von Daten zu Anzahl und Struktur von Haushalten und deren Wohnsituation. Hierfür erstellt das Teilprojekt Haushaltegenerierung einen bundesweiten Einzeldatenbestand auf Personenebene, den sogenannten zensustypischen Datensatz. Dieser ermöglicht erhebungsteilübergreifende Auswertungen zu demografischen, haushaltsstatistischen und erwerbsstatistischen Daten sowie zu Wohnungs- und Gebäudedaten auf kleinräumiger Ebene. Auch im Zensus 2021 wird ein registergestütztes Verfahren umgesetzt. Für die Erstellung des zensustypischen Datensatzes ist daher die Integration unterschiedlicher Register- und Erhebungsdaten sowie die Korrektur der Über- und Untererfassungen der Melderegister auf Einzelpersonenebene erforderlich. Die für den Zensus 2011 entwickelten Verfahren der Haushaltegenerierung wurden für den Zensus 2021 maßgeblich weiterentwickelt.

↘ ***Keywords:*** *2021 Census – household generating procedure – register-based census – number of inhabitants – residential household*

ABSTRACT

The objective of the 2021 Census is not only to determine the number of inhabitants, but also to provide data on the number and structure of households and their housing situation. To this end, a nationwide database at the level of individuals – the so-called census-typical data set – is created in the “household generating procedure” subproject. This census-typical data set permits the intertwined analysis of demographic, household and employment statistics, and of dwelling and building data, all at a small-area level. The 2021 Census, too, uses a register-based method. This means that the creation of the census-typical data set requires the integration of different register and survey data as well as the correction of overcoverage and undercoverage in the population register. The household generating procedure developed for the 2011 Census has been significantly refined for the 2021 Census.

1

Einleitung

Der folgende Beitrag gibt einen Einblick in die geplanten methodischen Weiterentwicklungen der Haushaltegenerierung im Rahmen des Zensus 2021. Dazu erfolgt zunächst eine kurze Einführung in die Problemstellung und die grundsätzliche Arbeitsweise der Haushaltegenerierung. Im Anschluss daran werden die wichtigsten methodischen Neuerungen in der Haushaltegenerierung skizziert.

Der letzte Zensus in Deutschland wurde zum Stichtag 9. Mai 2011 durchgeführt. Zur Vermeidung hoher Kosten, zur Entlastung der Befragten und im Hinblick auf die Aspekte Datenschutz und Datensparsamkeit entschied sich der Gesetzgeber dazu, die Volkszählung 2011 erstmals nicht mehr als Befragung aller Einwohnerinnen und Einwohner durchzuführen. Vielmehr wurden die Daten registergestützt – das heißt durch Auswertungen der Melderegister und anderer Verwaltungsregister – ermittelt. Befragungen sollten nur noch ergänzend erfolgen.[1] Auch der Zensus 2021 wird registergestützt durchgeführt werden.[2]

Der Zensus dient neben der Ermittlung der Einwohnerzahl auch der Bereitstellung fachlich und regional tief gegliederter Ergebnisse. Damit differenzierte Auswertungen auf kleinräumiger Ebene möglich werden, bedarf es der Zusammenführung der verschiedenen Datenquellen des Zensus. Die Datenintegration ist unter anderem Aufgabe der Haushaltegenerierung. Erst durch dieses Verfahren kann ein bundesweiter Einzeldatenbestand auf Personenebene zur Verfügung gestellt werden, der erhebungsteilübergreifende Auswertungen zu demografischen, haushaltsstatistischen und erwerbsstatistischen Daten sowie zu Wohnungs- und Gebäudedaten ermöglicht. So können beispielsweise nicht nur reine Haushaltsergebnisse (wie die Anzahl der Einpersonen- und Mehrpersonenhaushalte), sondern auch erhebungsteilübergreifende Ergebnisse (wie die durchschnittliche Wohnungsgröße von Haushalten mit Kindern) ausgewiesen werden (Kreuzmair/Reisch, 2013).

Zunächst wird auf die allgemeine Zielsetzung und Arbeitsweise der Haushaltegenerierung eingegangen, bevor die methodischen Neuerungen der Haushaltegenerierung im Zensus 2021 erläutert werden.

2

Zielsetzung der Haushaltegenerierung

Zur Erstellung fachlich und regional tief gegliederter Ergebnisse führt die Haushaltegenerierung zum einen eine Integration verschiedener Datenquellen und zum anderen eine Synchronisation der integrierten Daten mit der berechneten Einwohnerzahl durch. Die Integration der Melderegister und der Gebäude- und Wohnungszählung ermöglicht es, Informationen über die Zahl und Struktur von Haushalten und deren Wohnsituation zu gewinnen. Zudem führt die Haushaltegenerierung die Korrektur der Über- und Untererfassungen (Karteileichen und Fehlbestände) der Melderegister auf Einzeldatenbasis durch. Datenintegration und -synchronisation unterliegen unterschiedlichen Herausforderungen:

Mit einem registergestützten Verfahren werden Daten zur Struktur und Anzahl von Wohnhaushalten nicht mehr primärstatistisch erhoben, sondern durch die Integration verschiedener Datenquellen (insbesondere Melderegister und Gebäude- und Wohnungszählung) ermittelt. Die Information zu Haushaltszusammenhängen liegt in den Melderegistern jedoch nicht vollständig vor. Gebäude- und wohnungsstatistische Informationen werden zwar primärstatistisch erhoben, liefern jedoch keine Information zu den konkreten Wohnverhältnissen von Haushalten in Deutschland. Kleinräumig gegliederte Haushalts- und Wohnungsbelegungsdaten können demnach nur durch die anschriftenweise Zusammenführung der Personendaten aus den Melderegistern zu Haushalten und durch die Verknüpfung mit den Gebäude- und Wohnungsdaten der Gebäude- und Wohnungszählung erreicht werden (Vorndran, 2004, hier: Seite 1245). Im Mittelpunkt der Haushaltegenerierung steht die statistische Einheit des Wohnhaushalts, der jedoch weder im Melderegister noch in der Gebäude- und Wohnungszählung direkt erfasst wird. Deshalb werden Verfahren

1 Entwurf eines Gesetzes zur Vorbereitung eines registergestützten Zensus einschließlich einer Gebäude- und Wohnungszählung 2011 (Zensusvorbereitungsgesetz 2011 – ZensVorbG 2011) vom 30. Mai 2007 (BT-Drucksache 16/5525); Gesetz über den registergestützten Zensus im Jahre 2011 (Zensusgesetz 2011 – ZensG 2011) vom 8. Juli 2009 (BGBl. I Seite 1781).

2 Entwurf eines Gesetzes zur Durchführung des Zensus im Jahr 2021 (Zensusgesetz 2021 – ZensG 2021) vom 25. März 2019 (BT-Drucksache 19/8693).

eingesetzt, mit denen Haushaltszusammenhänge auf Basis von Registerangaben nachgebildet und mit den primärstatistisch erhobenen Daten der Gebäude- und Wohnungszählung verknüpft werden.

Eine besondere Herausforderung hierbei liegt darin, mit der Unvollständigkeit sowie Mehrdeutigkeit der verschiedenen Datenquellen umzugehen. Daher ist es nicht ausreichend, eine reine Datenzusammenführung im Sinne eines Record Linkage durchzuführen, sondern es werden zusätzlich statistische Schätzverfahren eingesetzt, um Haushalte zu bilden. Außerdem lassen sich die Datenquellen nicht immer eindeutig zusammenführen – beispielsweise entspricht die aus den Melderegistern generierte Anzahl der Haushalte nicht immer der durch Befragung erhobenen Anzahl der Wohnungen einer Anschrift. Hierfür wurde eine Vorgehensweise entwickelt, die eine eindeutige Lösung findet, zugleich aber den primärstatistisch erhobenen Angaben Vorrang gewährt.

Auch für die Datensynchronisation zwischen der durch Hochrechnung festgelegten Einwohnerzahl und dem Melderegisterbestand auf Einzeldatenebene werden in der Haushaltegenerierung neue Verfahren entwickelt. An Stichprobenanschriften kann die Korrektur von Karteileichen und Fehlbeständen auf Basis der primärstatistisch festgestellten Informationen unmittelbar durchgeführt werden. An allen anderen Anschriften wird hingegen ein statistisches Korrekturverfahren angewandt, um einen auf Einzelpersonenebene auswertbaren Datensatz zu erstellen und die Personendaten mit den durch Hochrechnung festgelegten Einwohnerzahlen abzustimmen. Der Umfang des zu korrigierenden Personenbestands wird aus der Stichprobe geschätzt. Zudem werden weitere Stichprobeninformationen zu Über- und Untererfassungen genutzt, um im Personenbestand des Melderegisters geeignete Personen zu bestimmen, die als Karteileiche gelöscht oder als Fehlbestand imputiert werden. Auch hierbei handelt es sich nicht um eine reine Datenzusammenführung, sondern um den Einsatz statistischer Schätzverfahren, um die Über- und Untererfassungen im Einzeldatenbestand möglichst plausibel zu korrigieren.

Auf Basis der Erfahrungen aus dem Zensus 2011 konnte das Verfahren der Haushaltegenerierung weiterentwickelt werden. Die Haushaltegenerierung im Zensus 2021 soll eine höhere Ausschöpfung der Registerinformationen sowie der Gebäude- und Wohnungsdaten erreichen. Außerdem ermöglichen die Weiterentwicklungen, stärker auf die Plausibilität der Ergebnisse auf Personen-, Haushalts- und Anschriftenebene zu achten. Hierfür wird in manchen Bearbeitungsschritten der Haushaltegenerierung auf zu detaillierte Hochrechnungen verzichtet und stärker auf modernere statistische Schätzverfahren zurückgegriffen. Dabei werden diverse Merkmale aus dem Melderegister sowie primärstatistisch erhobene Informationen einbezogen, die dabei helfen, plausiblere Haushaltszusammenhänge auf kleinräumiger Ebene zu erreichen. Einerseits wird darauf geachtet, dass die Gesamtzahl der Haushalte, Karteileichen und Fehlbestände auf der Gemeindeebene plausibel ist. Andererseits wird auch stärker berücksichtigt, an welchen Anschriften bestimmte Haushaltstypen zu bilden sind oder an welchen Anschriften und bei welchen Personen die Korrektur der Karteileichen und Fehlbestände vorzunehmen ist. Weiterhin werden im Vergleich zur Haushaltegenerierung 2011 vermehrt globale Lösungsansätze (zum Beispiel durch Einsatz von Optimierungsverfahren) umgesetzt. Das zeitgleiche Bearbeiten aller Haushalte und Wohnungen einer Anschrift (oder im Falle des Korrekturverfahrens sogar ein zeitgleiches Bearbeiten aller Personen einer Gemeinde) erzielt verbesserte Ergebnisse. Außerdem werden Unplausibilitäten vermieden, die im Zuge einer schrittweisen Bearbeitung einzelner Haushalte und Wohnungen nicht zu verhindern wären.

Die Änderungen werden in den Kapiteln 4 und 5 ausführlicher behandelt, zunächst erfolgt jedoch eine kurze Übersicht über die Arbeitsweise der Haushaltegenerierung im Zensus 2021.

3

Die Haushaltegenerierung im Überblick

Die Haushaltegenerierung im Zensus 2021 ist ein Verfahren, mit dessen Hilfe Daten, die aus verschiedenen Quellen stammen und sich auf verschiedene, für den Zensus relevante statistische Einheiten beziehen, zusammengeführt, miteinander verknüpft und bereinigt werden. Im Zentrum des Verfahrens steht die Konstruktion einer im Zensus nicht direkt beobachteten statistischen Einheit, nämlich des Wohnhaushalts. Darüber hinaus dient die Haushaltegenerierung dazu, die ein-

fließenden amtlichen Registerdaten durch Abgleich mit Daten aus statistischen Primärerhebungen zu korrigieren. Ziel ist die Erstellung des zensustypischen Datensatzes, der auf Personenebene und in fachlich sowie regional tiefer Gliederung ausgewertet werden kann (Kreuzmair/Reisch, 2013, hier: Seite 14).

Ihren Ursprung hat die Haushaltegenerierung in einem bereits in den 1980er-Jahren etablierten Verfahren aus der Städtestatistik, mit dessen Hilfe Haushaltszusammenhänge aus Melderegisterdaten konstruiert werden. Im Rahmen des Zensustests 2001 wurde die Tauglichkeit eines solchen Verfahrens für die Durchführung eines Zensus überprüft. Zusätzlich zu den Angaben des Melderegisters wurden hierfür auch gebäude- und wohnungsstatistische Informationen herangezogen (Vorndran, 2004, hier: Seite 1246). Schließlich wurde das basale Verfahren der Haushaltegenerierung im Zensus 2011 gegenüber dem Zensustest erheblich ausgebaut und weiterentwickelt, insbesondere im Hinblick auf die Verarbeitung und den gegenseitigen Abgleich verschiedener Datenbestände.[3] Die Haushaltegenerierung im Zensus 2021 baut auf den Erfahrungen aus dem Zensus 2011 auf, steuert an geeigneten Stellen methodisch nach und setzt modernere statistische Verfahren ein.

Die grundlegende Operation der Haushaltegenerierung besteht darin, Beziehungen zwischen statistischen Einheiten zu erzeugen. Dabei sind viele verschiedene Arten von Beziehungen zu berücksichtigen: Hierarchische Beziehungen verknüpfen über- und untergeordnete statistische Einheiten, etwa Haushalte mit Wohnungen und Personen mit Haushalten. Alle Beziehungen dieser Art sind eindeutig, das heißt jede Person ist eindeutig einer Anschrift und einem Haushalt zugeordnet. Da die Anzahl und Struktur von Wohnhaushalten erfasst werden soll, ist jeder Haushalt mit nur einer Wohnung verknüpft. Weitere, nicht hierarchische Beziehungen verknüpfen Personen mit Personen (beispielsweise Eltern-Kind-Beziehungen, Ehen, Lebenspartnerschaften), um Beziehungsstrukturen innerhalb von Wohnhaushalten abzubilden.

Die in der Haushaltegenerierung verarbeiteten Daten beziehen sich auf Gemeinden, Anschriften, Gebäude, Wohnungen, Haushalte, Beziehungen und Personen. Die Daten stammen zum Teil aus amtlichen Quellen, etwa aus den Registern der Einwohnermeldeämter, zum Teil auch aus primärstatistischen Erhebungen, etwa der Gebäude- und Wohnungszählung und den Personenerhebungen an Haushaltsstichproben- und Sonderbereichsanschriften. ↘ Grafik 1 auf Seite 40

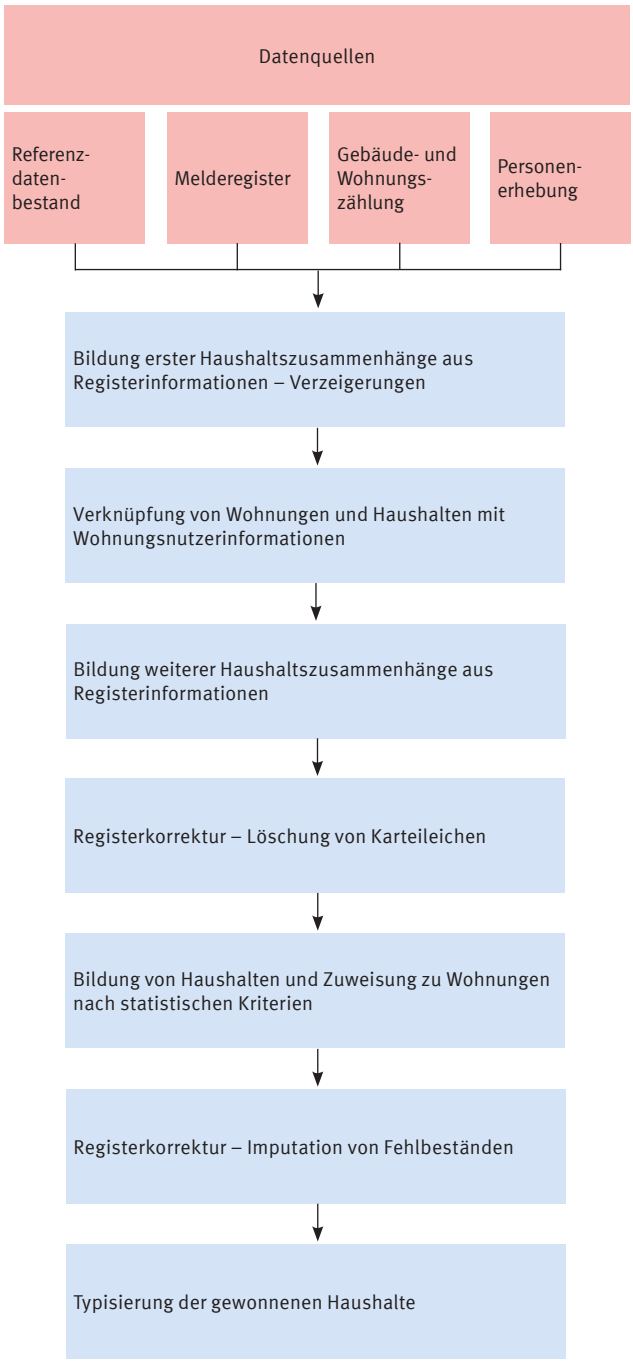
Die wichtigsten amtlichen Datenquellen sind die Register der Einwohnermeldeämter in den Kommunen. Aus diesen werden Daten gewonnen, mit deren Hilfe sich Personen bestimmten Gemeinden und Anschriften zuordnen lassen und die zudem Informationen über gesetzliche Beziehungen zwischen Personen enthalten, etwa Ehen, Lebenspartnerschaften oder Eltern-Kind-Beziehungen. Die Gebäude- und Wohnungszählung liefert Daten, die eine Beschreibung der Anschriften im Hinblick auf die dort vorliegenden Gebäude und Wohnungen ermöglichen. Sie erlauben zudem eine Zuordnung von Personen und Haushalten zu Wohnungen. Die Personenerhebungen an Anschriften der Haushaltsstichprobe und an Sonderbereichsanschriften liefern für einen Teil der Anschriften Daten über die dort tatsächlich wohnenden Personen und deren Haushaltszusammenhänge. Darüber hinaus werden auf der Grundlage der Haushaltsstichprobe Hochrechnungen erstellt, die primär dazu dienen, die Einwohnerzahl der Gemeinden zu bestimmen. In der Haushaltegenerierung werden weitere Hochrechnungen aus der Haushaltsstichprobe für die Bereinigung der amtlichen Registerdaten verwendet, die Fehler in Form von Karteileichen und Fehlbeständen aufweisen.

Die Haushaltegenerierung führt in einem mehrstufigen Verfahren Daten aus all diesen Quellen zusammen, konstruiert und beschreibt Haushaltszusammenhänge, bereinigt die Registerdaten mit Rückgriff auf primärstatistische Erkenntnisse und erstellt einen zensustypischen Datensatz.

Wie in Grafik 1 abgebildet, dient der erste Bearbeitungsschritt der Bildung von Haushaltszusammenhängen über Verzeigerungsinformationen aus den Melderegistern. Verzeigerungen sind in den Melderegistern enthaltene Informationen zu gesetzlichen Beziehungen zwischen zwei gemeldeten Personen, zum Beispiel Ehen oder Eltern-Kind-Beziehungen. In der Haushaltegenerierung wird eine Beziehung auf Basis einer Verzeigerung daher

3 Für detaillierte Informationen zur Entwicklung der Haushaltegenerierung im Zensustest und zum Verfahren im Zensus 2011 siehe Braun, 2004; Statistische Ämter des Bundes und der Länder, 2004; Vorndran, 2004; Kreuzmair/Reisch, 2013; Hofmeister/Fürnrohr, 2014.

Grafik 1
Datenquellen und Ablauf der Haushaltegenerierung



immer dann zwischen zwei Personen einer Anschrift gebildet, wenn deren im Melderegister geführte Ordnungsnummern oder andere eindeutige personenbezogene Merkmale gegenseitig aufeinander verweisen.

Im Bearbeitungsschritt „Verknüpfung von Wohnungen und Haushalten mit Wohnungsnutzerinformationen“ werden erste Verknüpfungen von Wohnungen und Haushalten durchgeführt. Mithilfe eines maschinellen Namensabgleichs werden an einer Anschrift gemeldete Personen mit Wohnungsnutzerinformationen der Gebäude- und Wohnungszählung derselben Anschrift abgeglichen. Die im Melderegister gefundenen Wohnungsnutzer und ihre Haushalte werden mit der entsprechenden Wohnung verknüpft und ihre Haushalte zusammengelegt. Darüber lässt sich beispielsweise der Haushaltszusammenhang von nichtehelichen Lebensgemeinschaften, über die keine Angaben in den Melderegistern vorliegen, rekonstruieren. Bei Nicht-Eindeutigkeit der Zuordnungen werden Haushaltstrennungen oder Wohnungszusammenlegungen durchgeführt, da jeder Wohnhaushalt grundsätzlich nur mit einer Wohnung verknüpft sein darf.

Im Schritt „Bildung weiterer Haushaltszusammenhänge aus Registerinformationen“ werden Registerinformationen genutzt, um weitere (weniger eindeutige) Haushaltszusammenhänge zu bilden. Anders als bei der Verarbeitung von Verzeigerungsinformationen werden in diesem Schritt auch nicht gesetzliche Beziehungen gebildet, wie etwa nicht eingetragene Lebensgemeinschaften oder Großeltern-Kind-Beziehungen.

Nachdem die Informationen aus den Melderegistern durch die vorherigen Schritte weitestgehend ausgeschöpft sind, erfolgt anschließend mit der Bereinigung des Einzeldatenbestands um hochgerechnete Übererfassungen der erste Teil der Registerkorrektur.[4] Grundprinzip der Korrektur von Karteileichen ist das Herunterbrechen der im Zuge der Einwohnerzahlhochrechnung festgestellten Gesamtkarteileichenzahl einer Gemeinde auf einzelne Anschriften und Personen (Datensynchronisation). Ziel ist dabei, durch die Auswahl der Karteileichen nach statistischen Kriterien keine räumlichen und demografischen Verzerrungen im Einzeldatensatz zu erzeugen. Eine detailliertere Darstellung dieses Verfahrens enthält Kapitel 5 dieses Aufsatzes.

Danach erfolgt der Bearbeitungsschritt „Bildung von Haushalten und Zuweisung zu Wohnungen nach statistischen Kriterien“. Ein Großteil der Haushalts- und

4 Der zweite Teil der Registerkorrektur erfolgt erst zu einem späteren Zeitpunkt im Bearbeitungsschritt „Registerkorrektur – Imputation von Untererfassungen des Melderegisters“.

Wohnungszusammenhänge kann in den vorhergehenden Schritten der Haushaltegenerierung bereits gebildet werden. Wurden alle Informationen aus dem Melderegister ausgeschöpft, werden noch unverknüpfte Personen und Wohnungen in diesem Bearbeitungsschritt nach statistischen Kriterien verknüpft. Bestehen an einer Anschrift mindestens genauso viele Wohnungen wie Haushalte, werden die Haushalte unter Berücksichtigung ihres Wohnflächenbedarfs mit Wohnungen der Anschrift verknüpft. An Anschriften, an denen mehr unverknüpfte Haushalte als Wohnungen bestehen, werden unverknüpfte Haushalte mithilfe eines statistischen Schätzverfahrens und eines Optimierungsverfahrens auf bereits verknüpfte Haushalte sowie unverknüpfte Wohnungen verteilt.

Als letzter Schritt vor der Haushaltstypisierung erfolgt der zweite Teil der Registerkorrektur, die Imputation der hochgerechneten Untererfassungen in den Melderegistern. Dazu werden bereits bestehende Personen aus dem Melderegister, bei denen davon ausgegangen werden kann, dass sie potenziellen Fehlbeständen möglichst ähnlich sind, kopiert und als neue und eigenständige Personen wieder in den Datensatz eingefügt. Welche Personen zur Kopie ausgewählt werden, wird auf Basis statistischer Schätzverfahren entschieden. Anschließend müssen die kopierten Personen konkreten Anschriften zugeordnet und mit den dort bestehenden Haushalten und Wohnungen verknüpft werden. Dieser zweite Teil der Registerkorrektur wird als Abschlussschritt der Bildung und Verknüpfung von Haushalten durchgeführt. Damit können bewohnte, aber bisher unverknüpfte Wohnungen vorrangig mit Fehlbeständen verknüpft werden. Diesem Verfahren liegt die Annahme zugrunde, dass eine in der Gebäude- und Wohnungszählung gemeldete bewohnte Wohnung tatsächlich besteht. Sie ist insbesondere deshalb zu diesem Zeitpunkt der Haushaltegenerierung noch unverknüpft, weil der Wohnungsnutzer der Wohnung ein Fehlbestand ist. Auch zur Korrektur um die hochgerechneten Fehlbestände erfolgt in Kapitel 5 eine detailliertere Darstellung.

Im letzten Schritt der Haushaltegenerierung wird die Haushaltstypisierung vorgenommen, die Informationen zur Anzahl und Struktur von Wohnhaushalten liefert. Die Ermittlung privater Haushalte nach dem „Konzept des gemeinsamen Wohnens“ sowie die Typisierung von Haushalten und Familien sind Vorgabe der Europäischen Union (EU) und in der Durchführungsverordnung 2017/543[5] genauer geregelt. Die Typisierung der in der Haushaltegenerierung gewonnenen Haushalte erfolgt nach drei verschiedenen Typisierungen: die EU-Typisierung von Hauptwohnsitzpersonen, die sogenannte nationale EU-Typisierung von Haupt- und Nebenwohnsitzpersonen sowie die nationale Typisierung von Haupt- und Nebenwohnsitzpersonen, die nah an den Kriterien des Mikrozensus angelegt ist.

4

Integration verschiedener Datenquellen im zensustypischen Datensatz

Welche Haushaltszusammenhänge bestehen und in welchen Wohnungen und Gebäuden diese Haushalte leben, wird durch die Zusammenführung von Registerangaben mit wohnungsstatistischen Informationen und den Angaben zu Wohnungsnutzern der Gebäude- und Wohnungszählung ermittelt. Ein Großteil der Haushaltszusammenhänge kann über Verzeigerungen aus den Melderegistern eindeutig zugewiesen werden. Ein weiterer großer Anteil der Haushalte kann über die Angabe von bis zu zwei Wohnungsnutzern aus der Gebäude- und Wohnungszählung zusammengeführt werden. Die Registerinformationen liefern jedoch keine vollständigen Haushaltsinformationen und es werden auch nur bis zu zwei Wohnungsnutzer je Wohnung erfragt. Daher können nicht alle Personen in den ersten beiden Bearbeitungsschritten Haushalten und Wohnungen zugeordnet werden. Eine eindeutige Zuordnung ist insbesondere meist dann nicht möglich, wenn in einem Haushalt mehrere Personen leben, die nicht verzeigert sind (zum Beispiel Wohngemeinschaften oder nichteheliche Lebensgemeinschaften).

Im Verfahren der Haushaltegenerierung 2011 erfolgte in diesen Fällen auf Basis von weiteren (weniger sicheren) Registerinformationen die Bildung weiterer Haushaltszusammenhänge aus Registerinformationen. Hierbei wurden Registerangaben jedoch nur für die Suche wei-

5 Durchführungsverordnung (EU) 2017/543 der Kommission vom 22. März 2017 zur Festlegung der Regeln für die Anwendung der Verordnung (EG) Nr. 763/2008 des Europäischen Parlaments und des Rates über Volks- und Wohnungszählungen in Bezug auf die technischen Spezifikationen für die Themen sowie für deren Untergliederungen (Amtsblatt der EU Nr. L 78, Seite 13).

terer Verwandtschafts- und Partnerschaftsbeziehungen genutzt. Konnten keine weiteren Haushaltszusammenhänge an einer Anschrift gefunden werden, dann erfolgte die Haushaltegenerierung nach statistischen Kriterien mithilfe einer gemeinde- oder kreisweiten Haushaltshochrechnung von 14 verschiedenen Haushaltstypen. Ziel der statistischen Generierung war es, durch Haushaltszusammenlegungen an einer Anschrift die Anzahl der hochgerechneten Haushaltstypen auf Gemeinde- oder Kreisebene zu erreichen (Kreuzmair/Reisch, 2013).

In der Haushaltegenerierung 2021 werden hingegen zunächst die Informationen der Melderegister stärker ausgeschöpft, um weitere Haushaltszusammenhänge zu bilden. Es wird nicht nur nach Verwandtschafts- und Partnerschaftsbeziehungen, sondern auch nach sonstigen Haushaltszusammenhängen gesucht. Beispielsweise werden Merkmale wie die gleiche vorherige Anschrift, das gleiche Einzugsdatum, der Familienname oder auch die Staatsangehörigkeit in unterschiedlichen Kombinationen einbezogen, um Personen an einer Anschrift zu Haushalten zusammenzuführen.

Sind alle Möglichkeiten der Haushaltsbildung auf Basis von Registerangaben erschöpft, erfolgt die statistische Generierung von Haushalten. Im Gegensatz zum Vorgehen der Haushaltegenerierung 2011 wird auf die detaillierte gemeinde- und kreisweite Haushaltshochrechnung verzichtet. Die Erfahrung zeigte, dass zu detaillierte Zielvorgaben auf Gemeindeebene nur schwer zu erreichen waren und der plausiblen Verteilung von Personen auf Haushalte und Wohnungen einer Anschrift zu wenig Augenmerk geschenkt wurde. Deshalb soll im neuen Verfahren dazu übergegangen werden, die Zielvorgaben der statistischen Generierung zu vergröbern (3 Haushaltsgrößenklassen statt 14 Haushaltstypen) und diese auf Anschriften- und nicht auf Gemeindeebene zu beziehen. Jede Anschrift erhält damit eine Vorgabe, welcher Haushaltsgrößenverteilung sich nach der statistischen Generierung möglichst anzunähern ist. Die Bestimmung dieser Haushaltsgrößenklassen erfolgt für jede Anschrift auf Basis eines statistischen Schätzverfahrens, das sich auf unterschiedliche Personen-, Wohnungs-, Gebäude- und Anschriftenmerkmale bezieht.

Auf Basis der geschätzten Zielvorgaben soll an Anschriften, an denen zu diesem Zeitpunkt noch mehr unverknüpfte Haushalte als unverknüpfte Wohnungen bestehen, eine plausible Verteilung der unverknüpften Personen auf verknüpfte Haushalte und noch unverknüpfte Wohnungen gefunden werden. Hierbei wird anstatt eines schrittweisen Vorgehens der Haushaltegenerierung wie beim Zensus 2011 ein Optimierungsverfahren eingesetzt, um ein zeitgleiches Bearbeiten aller Haushalte und Wohnungen an einer Anschrift zu gewährleisten. Unplausibilitäten, die durch eine schrittweise Bearbeitung entstehen, können dadurch verhindert werden. Im Zensus 2011 konnte die schrittweise Bearbeitung in manchen Fällen nämlich dazu führen, dass große Haushalte mit Wohnungen mit kleiner Wohnfläche verknüpft wurden und kleine Haushalte mit Wohnungen mit großer Wohnfläche. Die zeitgleiche Bearbeitung aller Haushalte und Wohnungen einer Anschrift im Zensus 2021 ermöglicht hingegen die bestmögliche Verteilung der unverknüpften Haushalte auf Wohnungen.

Aufgrund des Verzichts auf sehr detaillierte Hochrechnungen auf Gemeinde- oder Kreisebene, der stärkeren Ausschöpfung der Personen-, Wohnungs- und Anschriftenmerkmale sowie der Umsetzung eines globalen Lösungsansatzes, kann besser gewährleistet werden, dass nicht nur die Gesamtzahl der gebildeten Haushalte auf Gemeindeebene plausibel ist, sondern auch eine plausible Verteilung der Personen auf Haushalte und Wohnungen an einer Anschrift erreicht wird. Dadurch verbessert sich die Qualität erhebungsteilübergreifender Auswertungen zu haushalts- und wohnungsstatistischen Informationen auf kleinräumiger Ebene.

5

Synchronisation der integrierten Daten mit der hochgerechneten Einwohnerzahl: das Korrekturverfahren

Aufgrund der Konzeption des Zensus 2021 als registergestützter Zensus erfolgt die Haushaltebefragung – und damit die Existenzfeststellung der im Melderegister verzeichneten Personen – nur auf Stichprobenbasis. Damit ist nur an Anschriften, die für die Stichprobe ausgewählt wurden, bekannt, wie viele und welche Personen dort tatsächlich leben. Für die Ermittlung der Gesamteinwohnerzahl einer jeden Gemeinde ist daher eine Hochrechnung aus den Ergebnissen der Haushaltsstichprobe erforderlich. Diese Hochrechnung bestimmt nicht nur die Gesamteinwohnerzahl einer Gemeinde, sondern auch die nicht durch die Befragung aufgedeckten Übererfassungen (Karteileichen) und Untererfassungen (Fehlbestände) sowie deren Verteilungen hinsichtlich demografischer Merkmale.

Wie in den vorherigen Kapiteln dargelegt, besteht eine der Hauptaufgaben der Haushaltegenerierung darin, einen auf Einzeldatensatzebene auswertbaren Datenbestand zu erzeugen. Um dieses Ziel zu erreichen, müssen daher die auf Gemeindeebene hochgerechneten Karteileichen und Fehlbestände auf Einzelpersonenebene heruntergebrochen werden. Das bedeutet, dass konkrete Melderegistereinträge als potenzielle Karteileichen und zu imputierende Fehlbestände zu markieren sind. Diese Aufgabe der Datensynchronisation übernimmt das Korrekturverfahren in der Haushaltegenerierung. Maßgabe ist dabei, die demografische Struktur des Personenbestands durch die Korrekturen möglichst wenig zu verfälschen. Gegenüber dem Korrekturverfahren, das in der Haushaltegenerierung des Zensus 2011 zum Einsatz kam (Hofmeister/Fürnrohr, 2014), werden für den Zensus 2021 drei Änderungen grundsätzlicher Art vorgenommen:

1. Anders als im Zensus 2011 kommt das Korrekturverfahren in allen Gemeinden zum Einsatz. Durch die Ausweitung der Haushaltsbefragungen auf Stichprobenbasis auf Gemeinden mit weniger als 10 000 Einwohnerinnen und Einwohnern kann im Zensus 2021 auch in diesen Gemeinden eine Korrektur um hochgerechnete Karteileichen und Fehlbestände im Rahmen des Korrekturverfahrens erfolgen. Da jedoch mit sinkender Gemeindegröße auch der Detailgrad möglicher Hochrechnungen sinkt, erfolgt bei kleineren Gemeinden eine Verringerung des Umfangs der berücksichtigten demografischen Merkmale.

2. Bei der Auswahl konkreter Melderegistereinträge zur Löschung von Karteileichen und zur Kopie für die Imputation wird stärker auf eine plausible Verteilung hinsichtlich der wichtigsten demografischen Merkmale geachtet, um Verfälschungen durch das Korrekturverfahren zu minimieren. Eine Hochrechnung von Karteileichen und Fehlbeständen hinsichtlich der wichtigsten demografischen Merkmale (Geschlecht, Nationalität und Alter) ist nur univariat, also für jedes Merkmal einzeln möglich. Daher muss durch das Korrekturverfahren vor der Auswahl konkreter Personen zunächst die gemeinsame multivariate Verteilung dieser Merkmale, jeweils getrennt nach Karteileichen und Fehlbeständen, bestimmt werden. Im Zensus 2011 erfolgte die Bestimmung der Verteilung mithilfe einer Monte-Carlo-Simulation unter Annahme von Unabhängigkeit zwischen den demografischen Merkmalen. Durch eine Formulierung als Optimierungsproblem nach der Vorgehensweise des IPF-Verfahrens (Agresti, 2013, hier: Seite 364 ff.) kann auf diese nur hilfsweise verwendete Unabhängigkeitsannahme verzichtet werden. Die hochgerechneten Karteileichen- und Fehlbestandszahlen je Gesamtgemeinde und je demografischer Merkmalsausprägung gehen dabei als zwingend einzuhaltende Randbedingungen ein. Gleichzeitig kann eine Referenzstruktur definiert werden, der sich das Ergebnis, also die multivariaten Verteilungen von Karteileichen und Fehlbeständen, so gut wie möglich annähern soll. Als Referenz wird hier die demografische Bevölkerungsstruktur aus den Melderegistern verwendet. Durch das Optimierungsverfahren wird dann unter allen zulässigen Lösungen, die die Nebenbedingungen einhalten, diejenige Lösung mit der geringsten Distanz zur Referenzstruktur gewählt. Zusätzlich zur Bestimmung der multivariaten Verteilungen von Karteileichen und Fehlbeständen auf Gemeindeebene erfolgt eine Schätzung der individuellen Karteileichen- und Fehlbestandswahrscheinlichkeiten auf Einzelpersonenebene. Dazu werden statistische Modelle aus dem Bereich des Machine Learning (maschinelles Lernen) eingesetzt. Sie sollen anhand der Muster in den Ergebnissen der Haushaltsstichprobe prädizieren, wie hoch

eben jene Wahrscheinlichkeiten für jede nicht befragte Person ausfallen. Die konkrete Auswahl von Melderegistereinträgen zur Löschung beziehungsweise Imputation erfolgt dann durch Kombination der Informationen auf Gemeindeebene und auf Personenebene. Dadurch ist im Vergleich zum Zensus 2011 eine gezieltere Auswahl von Karteileichen und Fehlbeständen möglich, durch die eine geringere demografische und regionale Verzerrung im Ergebnisdatensatz zu erwarten ist.

3. Bei der Löschung von Karteileichen und der Imputation von Fehlbeständen wird im Zensus 2021 auf die Berücksichtigung von haushaltsstatistischen Hochrechnungen verzichtet. Für das Korrekturverfahren im Zensus 2011 wurde aus der Stichprobe die jeweilige Anzahl von Karteileichen- und Fehlbestandshaushalten untergliedert nach Größe und Typ je Gemeinde hochgerechnet. Diese gingen neben den demografischen Randbedingungen als zusätzliche Nebenbedingungen bei der Auswahl von konkreten Personen zur Löschung und Imputation ein. Auf derartige haushaltsstatistische Hochrechnungen wird im Zensus 2021 verzichtet: Einerseits hat sich gezeigt, dass die zu starke Untergliederung zu Qualitätsproblemen in der Hochrechnung geführt hat; andererseits war in der Regel keine gleichzeitige Erfüllung der demografischen und haushaltsstatistischen Nebenbedingungen möglich. Haushaltszusammenhänge von Karteileichen und Fehlbeständen sollen daher im Korrekturverfahren im Zensus 2021 unmittelbar in die statistischen Modelle einbezogen werden; sie finden so einen noch direkteren Eingang in die Auswahl konkreter Personen. Damit werden bevorzugt Personen als Karteileichen markiert, die in Haushaltskonstellationen leben, in denen auch in der Stichprobe gehäuft Karteileichen identifiziert werden. Gleichermaßen werden bevorzugt Personen zur Kopie als Fehlbestände ausgewählt, die hinsichtlich ihres Haushaltszusammenhangs Ähnlichkeiten zu typischen Fehlbeständen aus der Stichprobe aufweisen.

6

Fazit

Auch im Zensus 2021 kommt der Haushaltegenerierung wieder eine zentrale Bedeutung zu. Durch Datenintegration verschiedener Register- und Erhebungsdaten wird die Anzahl und Struktur von Haushalten und deren Wohnsituation ermittelt. Zudem wird der Personenbestand auf Einzeldatenbasis an die hochgerechnete Einwohnerzahl (und damit an die ermittelte Anzahl der Karteileichen und Fehlbestände) einer Gemeinde angepasst. Dadurch kann ein bundesweiter Einzeldatenbestand auf Personenebene zur Verfügung gestellt werden, der erhebungsteilübergreifende Auswertungen zu demografischen, haushaltsstatistischen, erwerbsstatistischen sowie zu Wohnungs- und Gebäudedaten ermöglicht.

Das weiterentwickelte Verfahren der Haushaltegenerierung im Zensus 2021 stellt eine höhere Datenqualität durch den Einsatz modernerer Schätzverfahren und Optimierungsverfahren in Aussicht. Es erfolgt eine stärkere Ausschöpfung der Personen-, Wohnungs- und Anschriftenmerkmale zur Bildung von Haushaltszusammenhängen; im Fokus steht auch eine gezieltere Auswahl bei der Löschung von Karteileichen und der Imputation von Fehlbeständen. Durch diese Weiterentwicklungen ist im Vergleich zur Haushaltegenerierung im Zensus 2011 eine plausiblere Abbildung der tatsächlichen Haushalts- und Wohnungszusammenhänge auf Anschriftenebene sowie eine bessere Auswahl von Karteileichen und Fehlbeständen hinsichtlich demografischer und anschriftenbezogener Merkmale zu erwarten.

LITERATURVERZEICHNIS

Agresti, Alan. *Categorical Data Analysis. 3. Auflage*. Hoboken, New Jersey 2013.

Braun, Ralph. *Zensustest 2001: Ergebnisse der Testerhebungen und Ausblick auf ein zukünftiges Zensusmodell*. In: Statistisches Monatsheft Baden-Württemberg. Ausgabe 5/2004, Seite 3 ff.

Hofmeister, Katrin/Fürnrohr, Michael. *Das Korrekturverfahren beim Zensus 2011*. In: Bayern in Zahlen. Ausgabe 6/2014, Seite 310 ff.

Kreuzmair, Ingrid/Reisch, Marco. *ZENSUS 2011: Ablauf der Haushaltegenerierung*. In: Statistisches Monatsheft Baden-Württemberg. Ausgabe 4/2013, Seite 13 ff.

Statistische Ämter des Bundes und der Länder 2004. *Ergebnisse des Zensustests*. In: Wirtschaft und Statistik. Ausgabe 8/2004, Seite 813 ff.

Vorndran, Ingeborg. *Zensustest – Verfahren und Ergebnisse der Haushaltegenerierung*. In: Wirtschaft und Statistik. Ausgabe 11/2004, Seite 1245 ff.

RECHTSGRUNDLAGEN

Durchführungsverordnung (EU) 2017/543 der Kommission vom 22. März 2017 zur Festlegung der Regeln für die Anwendung der Verordnung (EG) Nr. 763/2008 des Europäischen Parlaments und des Rates über Volks- und Wohnungszählungen in Bezug auf die technischen Spezifikationen für die Themen sowie für deren Untergliederungen (Amtsblatt der EU Nr. L 78, Seite 13).

Entwurf eines Gesetzes zur Vorbereitung eines registergestützten Zensus einschließlich einer Gebäude- und Wohnungszählung 2011 (Zensusvorbereitungsgesetz 2011 – ZensVorbG 2011) vom 30. Mai 2007 (Drucksache 16/5525).

Entwurf eines Gesetzes zur Durchführung des Zensus im Jahr 2021 (Zensusgesetz 2021 – ZensG 2021) vom 25. März 2019 (BT-Drucksache 19/8693).

Gesetz über den registergestützten Zensus im Jahre 2011 (Zensusgesetz 2011 – ZensG 2011) vom 8. Juli 2009 (BGBl. I Seite 1781).

Benjamin Freier

hat Soziologie und urbane Kultur, Gesellschaft und Raum studiert und ist seit Mai 2017 als wissenschaftlicher Mitarbeiter im Referat „Gebäude- und Wohnungszählung, Veröffentlichungskonzept für den Zensus“ beim Statistischen Bundesamt tätig. Schwerpunkt seiner Tätigkeit ist die Arbeit an der Online-First-Strategie für den Zensus 2021, darüber hinaus beschäftigt er sich mit der Öffentlichkeitsarbeit.

Juliane Mosel

ist Diplom-Soziologin und konzipiert als wissenschaftliche Mitarbeiterin im Referat „Haushaltsstichprobe (Konzeption und Aufbereitung)“ des Statistischen Bundesamtes die Haushaltebefragung des kommenden Zensus 2021. Dabei beschäftigt sie sich vor allem mit der Online-First-Strategie und der Öffentlichkeitsarbeit.

ONLINE FIRST ALS LEITGEDANKE FÜR EFFIZIENTE PRIMÄRERHEBUNGEN BEIM ZENSUS 2021

Benjamin Freier, Juliane Mosel

↘ **Schlüsselwörter:** Zensus – Volkszählung – Online-Fragebogen – Primärerhebung – Gebäude- und Wohnungszählung

ZUSAMMENFASSUNG

Seit dem Zensus 2011 hat sich ein rasanter technologischer Wandel vollzogen. Immer mehr Menschen nutzen das Internet und touchbasierte mobile Endgeräte haben sich in nahezu allen Bevölkerungsschichten etabliert. Der Zensus 2021 greift diese Entwicklung auf: Die Online-First-Strategie soll erreichen, dass ein möglichst hoher Anteil der Daten online übermittelt wird. Dadurch verbessert sich die Datenqualität und die Ergebnisse sind schneller auszuwerten. Dieser Aufsatz beschreibt die gesellschaftlichen Trends der Internet- und Technologienutzung dieses Jahrzehnts, die Maßnahmen der Online-First-Strategie in Deutschland als Antwort auf diese Trends sowie Zensusbeispiele aus anderen Ländern.

↘ ***Keywords:*** *census – population census – online questionnaire – primary survey – census of buildings and housing*

ABSTRACT

Since the last census in 2011, rapid technological change has taken place. More and more people are using the internet, and touch-based mobile devices are now common in almost all sections of the population. The 2021 Census picks up on this development: the Online First strategy aims to achieve the highest possible proportion of online-transmitted data. This will improve data quality, and results can be evaluated more quickly. This article describes the social trends of internet and technology use in this decade, the measures for the Online First strategy in Germany in response to these trends, and census examples from other countries.

1

Einleitung

Wie lässt sich die Qualität der Zensusergebnisse verbessern? Wie kann die Belastung von Auskunftgebenden gering gehalten werden? Wie lassen sich zeitintensive Arbeitsschritte verringern und die Zensusergebnisse schneller veröffentlichen? Diese Fragen stellten sich die Beteiligten im Zuge der Evaluierung des Zensus 2011. Wesentliche Ideen für Verbesserungen sind:

- Die Erhebungsinstrumente für die Bevölkerung sollen einfacher und verständlicher gestaltet werden.
- Die Verarbeitung der Daten für die Statistischen Ämter des Bundes und der Länder soll schneller und unkomplizierter werden.
- Der Zensus 2021 soll digitaler werden.

Im gemeinsamen Management-Handbuch für den Zensus 2021 verständigten sich die Statistischen Ämter des Bundes und der Länder auf gemeinsame Projektziele, unter anderem auch auf „Online First“ als Leitgedanke für effiziente Primärerhebungen beim Zensus 2021. Dieser Artikel dokumentiert den Weg von Online First – von der Idee bis zur Umsetzung.

Das anschließende Kapitel 2 stellt zunächst die Befragungen vor, die es im kommenden Zensus geben wird. Kapitel 3 beleuchtet die Rahmenbedingungen des kommenden Zensus und beschreibt den digitalen Wandel, der seit dem letzten Zensus 2011 eingetreten ist. Daneben gibt das Kapitel Einblick in die Online-Strategien anderer Länder bei der Durchführung ihrer Zensus und zeigt auf, welche Maßnahmen zu einer hohen Online-Beteiligung geführt haben. Kapitel 4 erläutert die Ziele der Online-Strategie für den kommenden Zensus 2021 in Deutschland sowie die angestrebten Vorteile. Das darauffolgende Kapitel beschreibt die Maßnahmen der Online-Strategie, die von den Statistischen Ämtern des Bundes und der Länder umgesetzt werden. Abgeschlossen wird der Artikel mit einem Ausblick, der den Nutzen der Online-First-Strategie des Zensus für andere Fachstatistiken aufzeigt.

2

Primärerhebungen im Kontext des Zensusmodells 2021

Ein Hauptziel des Zensus 2021 ist die Ermittlung der Einwohnerzahlen für unterschiedliche Gebietseinheiten. Darüber hinaus sollen qualitativ hochwertige und regional tiefgegliederte Daten über die Bevölkerung bereitgestellt werden, beispielsweise zur Erwerbstätigkeit oder zur Wohnsituation (sozio-demografische Daten). Das Zensusmodell 2021 baut auf dem Verfahren des registergestützten Zensus von 2011 auf. Grundlage sind Daten aus den Melderegistern und weiteren Verwaltungsverzeichnissen (zum Beispiel Grundsteuerdaten, Daten des amtlichen Liegenschaftskatasters). Da die späteren Zensusergebnisse nicht oder nicht in der erforderlichen Qualität in Verwaltungsregistern vorliegen, ist es notwendig, sie um Angaben aus Primärstatistiken zu ergänzen. Diese Primärstatistiken sind:

- die Haushaltebefragung auf Stichprobenbasis,
- die Gebäude- und Wohnungszählung,
- die Erhebung an Gemeinschaftsunterkünften sowie
- die Erhebung an Wohnheimen.

Außerdem findet eine Wiederholungsbefragung zum Zensus 2021 als Kontrollerhebung auf Stichprobenbasis statt. Mit ihr wird die Qualität der ermittelten Einwohnerzahlen geprüft.

Die Gebäude- und Wohnungszählung und die Erhebung an Gemeinschaftsunterkünften sind Vollerhebungen. Für die Gebäude- und Wohnungszählung werden alle Eigentümerinnen sowie Eigentümer von Wohnraum und für die Erhebung an Gemeinschaftsunterkünften die Leitungen aller Einrichtungen postalisch gebeten, an der Befragung teilzunehmen. Bei den beiden anderen Primärerhebungen muss die Grundgesamtheit ermittelt werden. Dies erfolgt bei der Haushaltebefragung an zufällig ausgewählten Stichprobenanschriften, während an Wohnheimen im Rahmen einer Vollerhebung alle dort wohnhaften Personen befragt werden. Dabei werden zum einen Personendaten für die Einwohnerzahl, zum anderen sozio-demografische Daten erhoben.

Zum Ermitteln der Einwohnerzahlen werden in der Befragung an Wohnheimen, der Haushaltebefragung

auf Stichprobenbasis sowie der Wiederholungsbefragung (nachfolgend zusammenfassend „Personenerhebungen“ genannt) Erhebungsbeauftragte eingesetzt. Um eine hohe Qualität zu gewährleisten, ist für diesen Teil der Befragung ein persönlicher Kontakt von Erhebungsbeauftragten mit den Personen aus der gezogenen Stichprobe notwendig. Die Erhebungsbeauftragten überprüfen, welche Personen an den jeweiligen Anschriften wohnen. Somit erfolgt die Erfassung dieser Personenmerkmale in der Regel im persönlichen Kontakt, die Befragung zu den sozio-demografischen Merkmalen jedoch schriftlich (online oder auf Papier) oder im Rahmen eines persönlichen Interviews.

3

Wegbereiter für einen digitalen Zensus

3.1 Rahmenbedingungen des Zensus 2021

Im Zensus 2011 konnten die Befragten in den Personenerhebungen ihre sozio-demografischen Angaben entweder im persönlichen Interview, auf dem Papierfragebogen oder im Online-Fragebogen machen. Rund 79 % der Befragten wählten das persönliche Interview, 21 % füllten den Fragebogen eigenständig aus – ein Drittel davon online.

Für die Gebäude- und Wohnungszählung wurden die zu Befragenden postalisch angeschrieben. Sie hatten die Möglichkeit, ihre Daten online abzugeben oder in einen beigelegten Papierfragebogen einzutragen und zurückzusenden. Den Papierfragebogen füllten zwei Drittel der Personen aus. Auch hier nutzte ein Drittel den Online-Fragebogen, obwohl die Online-Meldung kaum beworben wurde, viele Zugangshürden bestanden und der Online-Fragebogen für die Bürgerinnen und Bürger im Privatbereich nicht optimal war.

Für den Zensus 2021 haben sich die Rahmenbedingungen verändert: Der Stichprobenumfang für die Ermittlung der Einwohnerzahl wird sich voraussichtlich auf ungefähr 10 Millionen Personen in den Personenerhebungen vergrößern, davon werden voraussichtlich 6,7 Millionen auch zu den sozio-demografischen Merkmalen befragt werden (Bretschi/Lorentz, 2019). Zudem haben sich die Statistischen Ämter des Bundes und der Länder das Ziel gesetzt, die Ergebnisse 18 Monate nach dem Stichtag (16. Mai 2021) zu veröffentlichen. Dies erfordert eine stringente Arbeits- und Zeitplanung für die Durchführungs- und Aufbereitungsarbeiten. Darüber hinaus ist bei der Konzeption des Zensus 2021 die Forderung des Bundesverfassungsgerichts nach einer grundrechtsschonenden Verfahrensweise zu berücksichtigen.[1]

Der Zensus 2011 wurde durch einen externen Gutachter evaluiert. Mit Blick auf die Datenverarbeitung und Datenqualität empfahl er, den Anteil der online übermittelten Antworten sowohl bei der Gebäude- und Wohnungszählung als auch bei der Haushaltebefragung deutlich zu erhöhen. Die Online-First-Strategie bei den Primärerhebungen im Zensus 2021 ist hierbei ein Verfahrensvorschlag, der die genannten Rahmenbedingungen aufgreift, zur Zielerreichung beiträgt und dabei gleichzeitig Einsparpotenzial bietet. Die Statistischen Ämter des Bundes und der Länder haben sich daher unter anderem im Rahmen des Management-Handbuchs Zensus 2021 für eine Online-First-Strategie ausgesprochen.

3.2 Digitaler Wandel seit der Zensuserhebung 2011

Seit der letzten Zensuserhebung hat sich die weltweite Internet- und Technologienutzung enorm gewandelt. Einige Jahre zuvor wurde das erste Smartphone vorgestellt, was den Start in ein neues technologisches Zeitalter markierte. Die ersten Jahre nach der Einführung avancierte das Smartphone noch nicht zum alltäglichen Gebrauchsgegenstand. Deshalb war der Online-Fragebogen für den Zensus 2011 nur für die Nutzung auf Desktop-Geräten optimiert. Mittlerweile besitzen fast 90 % der Personen in Deutschland ein Smartphone – innerhalb der letzten zehn Jahre hat sich dieser Anteil um fast 73 Prozentpunkte gesteigert. Ähnlich sieht es bei der Nutzung von Tablets aus: Besaßen 2011 nur knapp 4 % der Bevölkerung ein Tablet, stieg dieser Anteil im Jahr 2017 auf über 50 % an. Touchbasierte Geräte haben sich gegenwärtig in nahezu allen Bevölkerungsschichten etabliert. Mit der Online-First-Strategie des Zensus

1 Bundesverfassungsgericht (BVerfG). Urteil des Zweiten Senats vom 19. September 2018. 2018. [Zugriff am 18. Juli 2019]. Verfügbar unter: www.bverfg.de

2021 reagiert die amtliche Statistik auf diesen rasanten technologischen Wandel.

Die steigende Anzahl von mobilen Endgeräten und deren Nutzung trägt dazu bei, dass insgesamt immer mehr Menschen an das Internet angebunden sind. Im Jahr 2011 besaßen immerhin 75 % der Bevölkerung in Deutschland einen Internetanschluss, im Jahr 2018 waren es bereits 90 %. Sollte die Internetverfügbarkeit wie in den vergangenen Jahren um durchschnittlich 2 % jährlich zunehmen, werden 2021 rund 95 % der Bevölkerung in Deutschland online sein. Von den 90 % der Bevölkerung, die 2018 das Internet nutzten, waren im vergangenen Jahr laut IKT-Erhebung[2] wiederum 89 % jeden Tag oder fast jeden Tag online. Nur in der Altersgruppe der über 65-Jährigen betrug der Anteil lediglich 68 %. Allerdings nutzten auch in dieser Bevölkerungsgruppe nur 9 % das Internet seltener als einmal wöchentlich. Die durchschnittliche tägliche Online-Zeit in der Gesamtbevölkerung betrug im Jahr 2011 ungefähr 137 Minuten je Person und Tag und ist in den letzten Jahren um knapp 60 Minuten auf 196 Minuten im Jahr 2018 angestiegen. ↘ Tabelle 1

Tabelle 1
Durchschnittliche Nutzung des Internets durch Personen, die das Internet im ersten Quartal 2018 genutzt haben, nach Altersgruppen

	Jeden Tag oder fast jeden Tag	Mindestens einmal in der Woche	Weniger als einmal in der Woche
	%		
Insgesamt	89	8	3
10 bis 15 Jahre	90	8	/
16 bis 24 Jahre	98	/	/
25 bis 44 Jahre	97	3	/
45 bis 64 Jahre	87	10	3
65 Jahre und älter	68	23	9

Ergebnisse aus der Erhebung zur Nutzung von Informations- und Kommunikationstechnologien in privaten Haushalten (IKT).

In den 2000er-Jahren prägte die Wissenschaft den Begriff „digital divide“. Der Begriff liefert einen Erklärungsansatz dafür, warum bestimmte Bevölkerungsteile das Internet nutzen oder davon absehen. Ausschlaggebend für die Internetnutzung sind demnach Merkmale wie Geschlecht, Alter, Einkommen oder Bildungsgrad. Junge Männer mit hohem Bildungsstand und Einkommen beispielsweise nutzten das Internet häufiger als arbeitslose Männer gleichen Alters. Mittlerweile lassen sich anhand spezifischer Merkmale kaum mehr Unterschiede in der Internetnutzung feststellen. Lediglich das Alter spielt noch eine entscheidende Rolle. Allerdings nimmt der Anteil der Internetnutzerinnen und -nutzer über 65 Jahre jährlich stark zu.

Auch der Blick auf das Konsumverhalten im Internet bestätigt den Digitalisierungstrend in breiten Gesellschaftsschichten: 73 % der Internetnutzerinnen und -nutzer zwischen 16 und 74 Jahren nutzten das Internet im ersten Quartal 2018 für Online-Käufe. Knapp 20 % von ihnen kauften sogar mehrmals wöchentlich im Internet ein.

Behördengänge werden immer noch am häufigsten persönlich erledigt. Dies geschieht unter anderem, weil viele behördliche Angelegenheiten nicht – wie im Privatbereich üblich – über Benutzername und Passwort oder PIN/TAN-Verfahren zu erledigen sind oder auch, weil die Dienstleistungen in elektronischer Form kommunal noch nicht angeboten werden. Der Trend zur E-Government-Nutzung in Deutschland ist gar rückläufig: Nutzten 2012 noch 45 % der Bürgerinnen und Bürger elektronische Dienste, ist der Anteil im Jahr 2018 auf 40 % gesunken. Viele Bürgerinnen und Bürger haben auch Bedenken hinsichtlich des Schutzes und der Sicherheit ihrer persönlichen Daten – dies ist der am häufigsten angegebene Grund auf die Frage, warum behördliche Formulare nicht über das Internet ausgefüllt werden.

Seit dem Zensus 2011 hat sich das Internet immer mehr als alltägliche Infrastruktur etabliert. Aber auch die Bedenken der Bürgerinnen und Bürger müssen im Rahmen der Online-First-Strategie berücksichtigt werden, um erfolgreich zu sein. Statistische Ämter anderer Länder haben bereits einen konsequenten Online-First-Ansatz verfolgt und damit auch Erfolge im Sinne einer hohen Online-Quote erzielen können.

2 Erhebung über die private Nutzung von Informations- und Kommunikationstechnologien.

3.3 Online-Strategien in ausländischen Zensusverfahren

Ein Blick auf andere Länder, die ebenfalls Primärerhebungen durchführen, zeigt, dass bei Befragungen im Rahmen des Zensus zunehmend digitale Instrumente – wie Online-Fragebogen – eingesetzt werden. Ein Vergleich der UNECE-Länder[3] der letzten Zensusrunde 2010 mit dem aktuellen Stand der Planung für den kommenden Zensus 2020 zeigt, dass immer mehr Länder ein Online-Erhebungsinstrument im Rahmen von Primärerhebungen einsetzen. Waren es 2010 noch 17 von 56 Ländern, die einen Online-Fragebogen anboten, planen für die kommende Zensusrunde 20 von 57 Ländern[4] diesen fest ein. Weitere neun Länder ziehen die Nutzung eines Online-Fragebogens in Betracht. Diese Online-Strategien haben in der Regel folgende Ziele:

- Erhöhung/Aufrechterhaltung der Teilnahmebereitschaft,
- Verbesserung der Datenqualität,
- Kostenreduktion beziehungsweise -effizienz,
- Anpassung an die Erwartungen der Öffentlichkeit,
- mehr Bürgerfreundlichkeit, da diese entscheiden, wann sie den Online-Fragebogen ausfüllen.

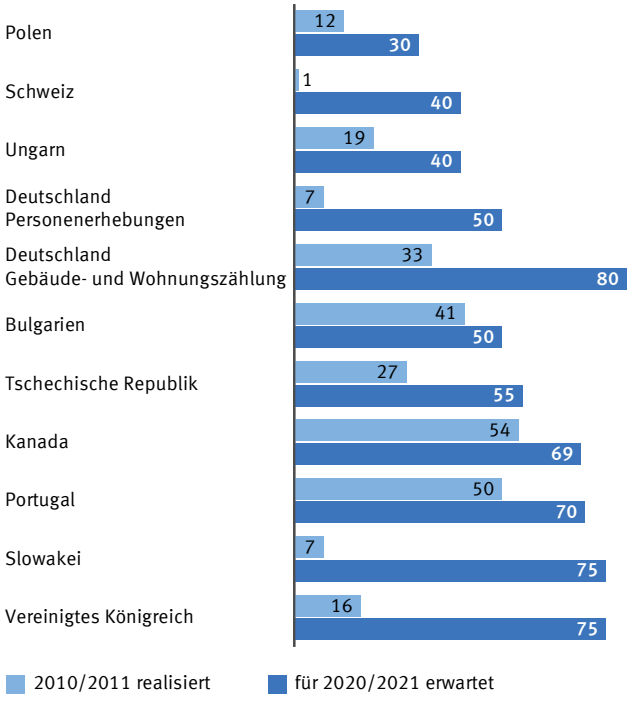
Ein Blick auf die Online-Rücklaufquoten belegt, dass es erfolgreich ist, zunehmend auf Online-Fragebogen zu setzen. In der Zensusrunde 2010 haben bereits 17 UNECE-Länder Online-Fragebogen eingesetzt. Diese Länder gehen für die kommende Zensusrunde 2020 ausnahmslos von einer Erhöhung der Online-Quote aus. Die statistischen Ämter der Slowakei, des Vereinigten Königreichs, Portugals und Kanadas rechnen mit Online-Rückläufen von bis zu 75 %. ↘ Grafik 1

Die meisten Länder erwarten, dass mindestens 30 % der Daten online eingehen werden. Im zuletzt durchgeführten Zensus 2016 konnte Kanada bereits einen Online-Dateneingang von 67 % verzeichnen. Ein weiteres Beispiel ist Neuseeland, das kein Mitglied in der UNECE ist: Dort wurde bereits für den Zensus im Jahr 2018 eine Online-Rücklaufquote von 70 % angestrebt, tatsächlich gingen über 82 % der Daten online ein.

3 Der Wirtschaftskommission für Europa der Vereinten Nationen (UNECE) gehören 56 Länder aus Europa, Nordamerika und Asien an. Die Kommission bietet unter anderem ein Forum zum Austausch über den Zensus in den Mitgliedsländern an und fördert gemeinsame Standards bei der Durchführung des Zensus.

4 56 Länder und Mexiko, das an den Zensusaktivitäten der UNECE teilnimmt, ohne Mitglied zu sein.

Grafik 1
Bei der Zensusrunde 2010/2011 realisierte Online-Quote und Erwartungen für die Zensusrunde 2020/2021[1]
in %



1 Für Deutschland sind die Prozentangaben differenziert nach Personenerhebungen sowie Gebäude- und Wohnungszählung ausgewiesen, da Erhebungsorganisation und -methodik der Werte eine Zusammenfassung nicht zulassen.

2019 - 01 - 0493

3.4 Steigerung der Online-Rücklaufquote in ausgewählten UNECE-Ländern

Die Strategien und Maßnahmen zur Erhöhung der Online-Quoten ausgewählter UNECE-Länder können auch für die Online-First-Strategie in Deutschland hilfreich sein. Eine Analyse der öffentlich verfügbaren Strategie- und Maßnahmenpapiere von Ländern mit hoher Online-Rücklaufquote zeigt folgende Erfolgsfaktoren:

Das Office for National Statistics (Vereinigtes Königreich) legt seinen Schwerpunkt auf die Entwicklung eines responsiven Online-Fragebogens. Responsiv bedeutet, dass der Fragebogen auf dem stationären PC sowie auf allen gängigen mobilen Endgeräten dargestellt und ausgefüllt

werden kann. Dabei erfolgt eine automatisierte Plausibilisierung der Angaben bereits während der Dateneingabe. Durch eine automatische Filterführung zeigt das Programm den Befragten nur die für sie relevanten Fragen. Auch die Erhebungsorganisation wird optimiert: In einem umfangreichen Pretest im Jahr 2017 wurden die Online-Rücklaufquoten ermittelt, je nachdem, ob dem Anschreiben zusätzlich ein Papierfragebogen beilag oder nicht. Der höchste Online-Dateneingang (84 %) kam zustande, wenn man den Papierfragebogen gar nicht erst beilegte und eine Bestellung notwendig war. War der Fragebogen hingegen dem Anschreiben beigelegt, meldeten nur 30 % online.

Das Amt für Statistik in Neuseeland, Stats NZ Tatauranga Aotearoa, bot für seinen Zensus 2018 neben dem Online-Erhebungsinstrument auch eine Online-Hilfe an. In sozialen Medien wurde eine umfangreiche Öffentlichkeitsarbeit betrieben. Die Befragten erhielten im Erstkontakt keinen Papierfragebogen. Sie konnten Papierfragebogen anfordern, die dann im sogenannten Print-on-Demand-Verfahren gedruckt und versandt wurden. Eine breit angelegte Medienkampagne informierte die Öffentlichkeit über die Veränderungen im Zensus.

Auch Statistics Canada verzichtete im Zensus 2016 beim Erstkontakt mit den Befragten auf die Beigabe von Papierfragebogen und setzte einen Online-Fragebogen ein. Einige Felder des Fragebogens waren schon vorausgefüllt, um den Befragten Arbeit zu ersparen und gleichzeitig sicherzustellen, dass die Formulare nur von den vorgesehenen Befragten ausgefüllt wurden. Die Befragten konnten zudem einen Zwischenstand abspeichern und den Fragebogen später weiter ausfüllen. Statistics Canada berichtet, dass bei Nutzung des Online-Fragebogens weniger unbeantwortete Fragen verzeichnet wurden und die Datenqualität der online eingegangenen Daten deutlich höher war als bei Papierfragebogen.

Natürlich weichen die Rahmenbedingungen in anderen Ländern ab und eine 1:1-Übertragung der Maßnahmen ist nicht ohne Weiteres möglich. Dennoch kann die amtliche Statistik in Deutschland von den Erfahrungen anderer Länder mit ihren Strategien und Maßnahmen profitieren, erfolgreiche Maßnahmen für den kommenden Zensus übernehmen und auf den Bedarf in Deutschland anpassen.

4

Die Online-First-Strategie des Zensus 2021

Die Online-First-Strategie des Zensus 2021 basiert auf der bei der Bevölkerung vorhandenen Infrastruktur zur Internetnutzung. Die Erhebungsprozesse werden so konzipiert, dass der Fragebogen standardmäßig online ausgefüllt wird – das gilt für die Personenerhebung wie für die Gebäude- und Wohnungszählung. Denn Daten, die online eingehen, sind verglichen mit Daten von Papierfragebogen genauer und schneller aufzubereiten. Die Online-First-Strategie trägt somit maßgeblich dazu bei, wesentliche Teile der internationalen Qualitätsziele der Statistik zu erreichen.

4.1 Beitrag der Online-Erhebung zur Erreichung der Projektziele im Zensus 2021

Die Statistischen Ämter des Bundes und der Länder haben die Erkenntnisse der Evaluierung des Zensus von 2011 aufgegriffen. Im gemeinsamen Management-Handbuch für den Zensus 2021 verständigten sie sich auf folgende sechs gemeinsame Projektziele:

- Genauigkeit,
- Aktualität,
- Wirtschaftlichkeit,
- Methodentransparenz,
- weniger Belastung der Auskunftspflichtigen und
- Nutzung des Zensuspotenzials.

Der für den Zensus 2021 wohl wichtigste Qualitätsgewinn ist die verbesserte Ergebnisgenauigkeit, die unter anderem durch einen höheren Anteil an Online-Meldungen erreicht werden soll. Das Antworten mittels elektronischer Fragebogen führt zu plausibleren Ergebnissen. Schon beim Eingeben durch die Befragten prüft das Programm die Angaben auf Plausibilität und Vollständigkeit und zeigt Tippfehler an. Bei der Gebäude- und Wohnungszählung 2011 waren im Nachgang der Erhebung fehlerhafte Angaben zu korrigieren, und

zwar fast doppelt so viele bei der Beantwortung mittels Papierbogen wie bei der Online-Meldung.

Wird ein Papierfragebogen eingesetzt, muss dieser immer komplett für alle Befragten vorliegen, obwohl viele Personen nur einen Teil ausfüllen müssen. Anweisungen (zum Beispiel „wenn 1 zutrifft springen Sie bitte zu Frage Nummer 8“) sind ein wichtiger Bestandteil, um die befragten Personen durch den Fragebogen zu führen. Dies ist notwendig, erzeugt aber auch Unklarheiten und stellt eine bedeutende Hürde der papierbasierten Befragung dar. Im Online-Fragebogen des Zensus erfolgt die Filterführung im Hintergrund, sodass die Befragten im Vergleich zum Papierfragebogen nur die Fragen sehen und beantworten müssen, die sie auch wirklich betreffen. Ebenso passen sich die Fragetexte und Antwortkategorien teilweise an, je nachdem, welche Antwort zuvor ausgewählt wurde. Somit entfällt die häufige Fehlerquelle, dass Befragte zur falschen Frage springen. Der elektronische Fragebogen des Zensus liegt in mehreren Sprachen vor. Dies steigert die korrekte Beantwortung und deckt die sprachliche Diversität in Deutschland gut ab.

Die Maßnahmen der Online-First-Strategie tragen dazu bei, dass die zu bearbeitende Anzahl an Papierfragebogen deutlich abnimmt. Dadurch ist ein großer Teil der Angaben medienbruchfrei und vollautomatisiert zu verarbeiten, was zu einer höheren Genauigkeit der Daten führt.

Eine wichtige Erkenntnis der Evaluierung des Zensus 2011 war die relativ späte Veröffentlichung der Ergebnisse. Die Online-First-Strategie ist geeignet, die Ergebnisse zeitnäher zum Stichtag zu veröffentlichen. Im Wesentlichen lässt sich das Ziel der früheren Ergebnisveröffentlichung durch eine kürzere Datenerhebungsphase erreichen. Bei der papierbasierten Befragung fallen viele manuelle Prozessschritte an, die die Ver- und Bearbeitungsdauer erhöhen, zum Beispiel:

- Postannahme,
- Vorsortieren der Sendungen,
- maschinelle Brieföffnung,
- manuelle Entnahme der Briefinhalte,
- Belegschneiden,
- maschinelle Belegerfassung und
- manuelle Nachbearbeitung der Belegerfassung.

Allein beim Bayerischen Landesamt für Statistik wurden beim Zensus 2011 «*an insgesamt 216 Anlieferungstagen* […] *735 Transportwagen von der Deutschen Post angeliefert. Diese Wagen waren mit zusammengenommen 9496 Postkisten gefüllt* […]. *In diesen Postkisten wiederum waren rund 2,4 Millionen Sendungen. Dies entspricht einem Gewicht von rund 80 Tonnen*» (Forster/Schmidl, 2012).

Die aufgeführten Prozessschritte und die dadurch entstehenden Zeitverluste durch Herstellung, Transport und manuelle Verarbeitung von Fragebogen entfallen bei der Online-Meldung. Die Angaben sind nach erfolgter Meldung unmittelbar innerhalb der IT-Umgebung für die anschließenden Aufbereitungsprozesse verfügbar.

Der verminderte Papierbogenbedarf trägt schließlich wesentlich zur Ressourcenschonung bei – wirtschaftlich wie auch aus Perspektive des Umweltschutzes. Allein die ausgefüllten Fragebogen der Gebäude- und Wohnungszählung beim Zensus 2011 wären übereinandergestapelt höher als der Mount Everest mit seinen 8 848 Metern Höhe. Die Verminderung der Transportwege für die Papierfragebogen verbessert die CO_2-Bilanz. Kosteneinsparungen ergeben sich vor allem durch verringerte Druckkosten und Einsparungen beim Porto. Des Weiteren wird die kostenintensive manuelle Nachbearbeitung der Papierfragebogen stark reduziert.

Ein wichtiges Ziel des Zensus ist es, die Auskunftspflichtigen so wenig wie möglich mit der Befragung zu belasten. Die Filterführung des Online-Fragebogens trägt wesentlich dazu bei, dass zum Beispiel die größte Gruppe der Befragten bei der Gebäude- und Wohnungszählung – nämlich die rund 14 Millionen Eigenheimbesitzerinnen und -besitzer – nur wenige Fragen des gesamten Fragenkatalogs lesen und beantworten müssen. Sie können alle Daten nun innerhalb von fünf bis zehn Minuten eintragen. Die Beantwortung eines Papierbogens hingegen benötigt mehr Zeit, da die jeweilige Filterführung erst verstanden und die betreffenden Fragen eindeutig identifiziert werden müssen.

5

Maßnahmen der Online-First-Strategie

5.1 Übergeordnete Maßnahmen

Zur Umsetzung der Online-First-Strategie sind bisherige Verwaltungsabläufe und Instrumente auf den Prüfstand zu stellen und bei Bedarf zu modernisieren. Beim Zensus 2011 zielten alle Prozesse darauf, einen Papierfragebogen zu entwickeln, zu testen und davon ein elektronisches Abbild zu erstellen. Die damals unerwartet hohe Online-Quote bei der Gebäude- und Wohnungszählung mit rund 30 % kam zustande, obwohl die Möglichkeit der Online-Meldung kaum beworben wurde, Zugangshürden bestanden und der Online-Fragebogen nicht ausreichend optimiert war. Für den Zensus 2021 ist von einem deutlich höheren Online-Rücklauf auszugehen, wenn diese Prozesse von Grund auf digital konzipiert und die Entwicklungen im digitalen Bereich einbezogen werden.

Die Entwicklung eines leicht verständlichen Online-Fragebogens, der den technischen Fortschritt seit 2011 berücksichtigt und der es auch weniger internetaffinen Personen erlaubt, online am Zensus teilzunehmen, ist die vordringlichste Maßnahme der Online-First-Strategie. Daher wurde der Online-Fragebogen für den Zensus 2021 neu konzipiert. Berücksichtigt wurden dabei die Anregungen der Befragten aus dem Zensus 2011 sowie aktuelle wissenschaftliche Erkenntnisse aus der Online-Forschung. Nicht zuletzt hatte auch der Blick auf die Bedienung großer Internetplattformen, die viele Millionen Menschen im Alltag nutzen, Einfluss. Die Befragten des Zensus 2021 sollen alle gängigen Bedienkomforts im Online-Fragebogen des Zensus wiederfinden. Beispielsweise gehören dazu:

- ein responsives Design, das den Fragebogen unabhängig von der Geräteklasse immer optimal darstellt,
- eine moderne grafische Darstellung,
- intelligente Filterführung, sodass jeder Person nur die für sie relevanten Fragen angezeigt werden,
- übersichtlich gestaltete Fragebogenseiten mit maximal drei Fragen je Seite, sodass Scrolling möglichst vermieden wird,
- gute Lesbarkeit, Barrierefreiheit und ausreichend große Schriftgröße, damit auch Menschen mit Sehschwäche oder anderen Behinderungen die gleiche Möglichkeit haben, den Fragebogen korrekt auszufüllen,
- individuell angepasste Fehlertexte durch sofortige Plausibilitätsprüfungen, damit die Befragten falsche Eintragungen unkompliziert verstehen und ändern können, sowie
- das Angebot des Fragebogens in unterschiedlichen Sprachen und allgemein verständlicher Sprache.

↘ Grafik 2 auf Seite 54

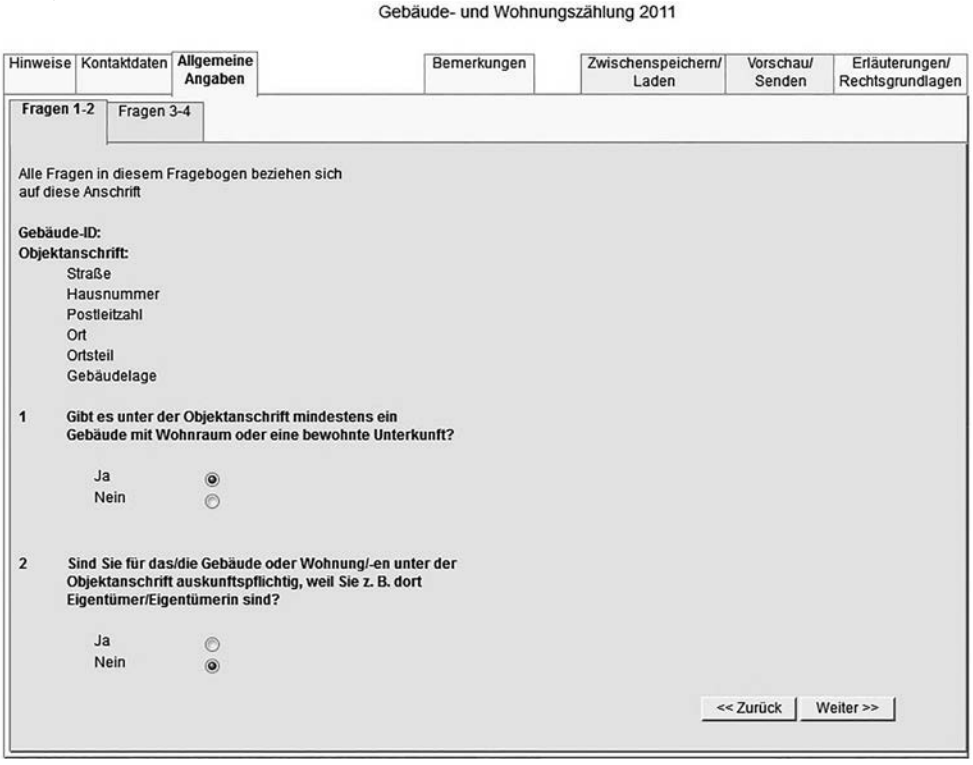
Nicht nur der Online-Fragebogen als Erhebungsinstrument wurde kritisch auf den Prüfstand gestellt, sondern auch, ob der Zugang für die Befragten zum Online-Fragebogen möglichst leicht ist. Dass diese Prüfung berechtigt ist, zeigen die Erfahrungen aus dem letzten Zensus. Beim Zensus 2011 war der Hinweis zur Online-Meldung auf dem Papierfragebogen leicht zu übersehen – das lag an der Platzierung der Zugangsdaten und der kleinen Schriftgröße. Rief man im Internet die Startseite des Zensus auf, mussten erst mehrere Seiten geöffnet und jeweils die richtige Auswahl (Bundesland und Erhebung) getroffen werden, ehe man zum Login des Online-Fragebogens gelangte. Die vielen Rückmeldungen der Bürgerinnen und Bürger zeigten deutlich, dass die Nutzerfreundlichkeit verbessert werden muss. Daher ist der Zugang zum Online-Fragebogen beim Zensus 2021 stark vereinfacht. Beim Aufruf der Zensus-Website wird der Login-Bereich zum Online-Fragebogen auf allen Geräteklassen gut sichtbar auf der Startseite platziert. Nach dem Eingeben der Zugangsdaten nimmt das System die Zuteilung in die richtige Erhebung und somit in den relevanten Online-Fragebogen automatisch vor. Somit erhalten auch Personen, die sowohl an der Personenerhebung als auch an der Gebäude- und Wohnungszählung teilnehmen müssen, einen leichten Zugang zum Fragebogen. Auch die neue Zensus-2021-Website weist moderne Bedienelemente auf und stellt Themen interaktiv und leicht verständlich dar. Für alle Interessierten gibt es auch weiterhin die vertiefenden Inhalte (Qualitätsberichte, tiefgegliederte Ergebnisse, Gemeindeblätter und so weiter).

Eine weitere Zugangshürde für die Online-Befragung beim Zensus 2011 war die damalige Verwendung kryptischer Zeichenfolgen für die Online-Zugangsdaten

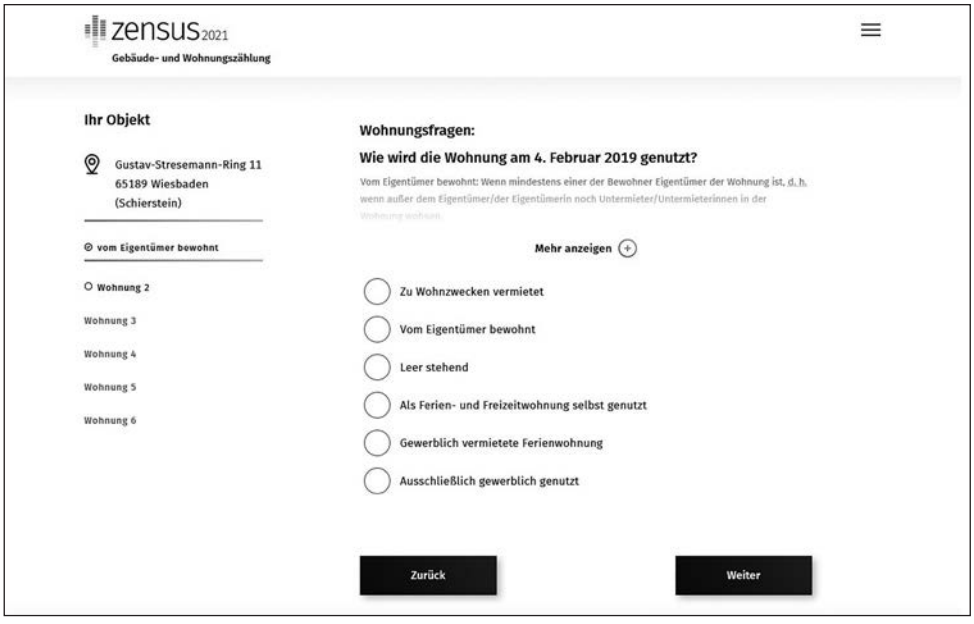
Grafik 2

Gegenüberstellung der Online-Fragebogen 2011 und 2021 zur Gebäude- und Wohnungszählung (Ausschnitt)

Zensus 2011



Zensus 2021



2019 - 01 - 0497

(Login-Vorgang). Vielen Menschen fällt das Entziffern kryptischer Zeichenfolgen aufgrund verminderten Sehvermögens schwer. Hinzu kommt, dass Sonderzeichen auf den Tastaturen von Touchgeräten selbst für erfahrene Internetnutzende teilweise nur schwer aufzufinden sind. Durch das Auslassen von Sonderzeichen und sich stark ähnelnden Nummern oder Buchstaben, die ausschließliche Kleinschreibung und die Wahl einer gut leserlichen Schriftart sind die Online-Zugangsdaten nun deutlich leichter einzugeben, ohne dass die Sicherheit leidet.

Neu im Vergleich zum Zensus 2011 ist auch, dass im Pretestlabor des Statistischen Bundesamtes alle beschriebenen Prozessschritte und Produkte des Online-Verfahrens ausgiebig getestet wurden: Bisher haben über 100 Testpersonen an qualitativen Interviews teilgenommen. Wertvolle Erkenntnisse zum Bedienverhalten der Testpersonen liefern hier die ausgewerteten Kameraaufzeichnungen. Im Zuge der Zensusvorbereitung führen die statistischen Ämter darüber hinaus zwei quantitative Pretests durch, in denen rund 140 000 Personen um freiwillige Teilnahme am Online-Verfahren gebeten werden.

Die Medienkampagne und die Bilder, die die Öffentlichkeit mit dem Zensus in Verbindung bringt, tragen entscheidend dazu bei, welchen Meldeweg die Befragten nutzen. Die Aufmerksamkeit in den Medien um den Zensusstichtag herum wird hoch sein. Passend zur Online-First-Strategie werden Bilder erzeugt, die die Online-Teilnahme in den Fokus rücken und motivierend für die Online-Auskunft sind. Die Eingabe von „Zensus 2011“ in eine Bildersuchmaschine zeigt beispielsweise derzeit hauptsächlich Kugelschreiber und Papierformulare. Beim Zensus 2021 sollen diese Ergebnisse im besten Fall Tablets und Smartphones sowie den Online-Fragebogen zeigen. Auch soll der Zensus auf Social-Media-Kanälen präsent sein und es wird eine höhere Anzahl digitaler Veröffentlichungen von Ergebnissen angestrebt.

5.2 Kontaktaufnahme bei der Gebäude- und Wohnungszählung

Der Medienbruch von Papier zur Online-Meldung lässt sich auch beim kommenden Zensus 2021 nicht vermeiden, denn die Befragten können nicht per E-Mail kontaktiert werden. Die auskunftspflichtigen Eigentümerinnen und Eigentümer erhalten bei der Gebäude- und Wohnungszählung über den postalischen Weg ein Anschreiben. Im Sinne einer Push2Web-Strategie, also dem Setzen von Anreizen für eine Online-Meldung, wird allerdings auf die Beigabe eines Papierfragebogens – wie es beim Zensus 2011 noch der Fall war – verzichtet. Wissenschaftliche Studien, aber auch der Zensustest 2017 im Vereinigten Königreich belegen für diese Strategie eine deutlich höhere Online-Teilnahmebereitschaft (siehe Abschnitt 3.4). Somit nimmt das Anschreiben als Teil des Beantwortungsprozesses eine sehr wichtige Funktion ein, denn daraus muss klar ersichtlich werden, wie die Meldung online erfolgt.

5.3 Kontaktaufnahme bei den Personenerhebungen

Die Besonderheit bei den Personenerhebungen ist eine Erhebungsorganisation, die sich durch die Mischung aus notwendigem persönlichem Interview für die Erfassung der Personenmerkmale und der präferierten selbstständigen, digitalen Angabe der sozio-demografischen Merkmale ergibt. Der Fokus der Kontaktaufnahme liegt daher sowohl auf der Information der Befragten, als auch in der Ausgestaltung der Erhebungsorganisation. Die Erhebungsbeauftragten sind hierbei die Ansprechpersonen für die Befragten: Sie können bei Unklarheiten informieren und den Personen, die den Online-Meldeweg nicht nutzen können oder wollen, Unterstützung oder Alternativen anbieten.

Um eine hohe Online-Selbstausfüllerquote zu erreichen, werden die Meldewege (nach Nutzen für die Online-Quote und den Qualitätszielen) priorisiert. Alle möglichen Meldewege werden den Auskunftgebenden angeboten, jedoch nicht gleichzeitig beworben. Beispielsweise soll wie bei der Gebäude- und Wohnungszählung ein Papierfragebogen nur auf ausdrücklichen Wunsch ausgehändigt werden.

Die Online-First-Strategie bliebe unvollständig, würde sie nicht auch die Belange derjenigen Befragten ernst nehmen, die den Online-Zugang nicht nutzen können oder wollen. Daher können bei Bedarf weiterhin die sozio-demografischen Merkmale als papierbasierte Meldung sowie im persönlichen Interview erhoben werden.

6

Ausblick

Die geplanten und zum Teil schon umgesetzten Maßnahmen der Online-First-Strategie bieten großes Innovationspotenzial, um die Primärerhebungen im kommenden Zensus 2021 effizienter zu gestalten und die selbst gesteckten Qualitätsziele zu erreichen. Sie knüpfen an nationale und internationale Standards an und werden dem digitalen Wandel in der Gesellschaft gerecht. Des Weiteren fügen sie sich ein in die Ziele der Digitalen Agenda des Statistischen Bundesamtes sowie des Regierungsprogramms „Digitale Verwaltung 2020“. Die jüngsten Beispiele des Zensus in Kanada und Neuseeland zeigen, dass bereits andere Länder erfolgreich den digitalen Wandel mitgestalten.

Der Zensus 2021 in Deutschland soll der letzte registergestützte Zensus mit primärstatistischen Erhebungen werden. In Zukunft soll ein rein registerbasierter Zensus stattfinden, ohne Notwendigkeit, Bürgerinnen und Bürger zu befragen (Körner und andere, 2019). Aber die Erkenntnisse, wie eine Online-First-Strategie eine Primärerhebung effizienter gestaltet, sind nicht nur für den Zensus relevant. Auch andere Fachstatistiken werden von den Erfahrungen des Zensus 2021 profitieren, Online First als Leitgedanken für effizientere Primärerhebungen implementieren und zur besseren Datenqualität und Entlastung der Bürgerinnen und Bürger beitragen.

LITERATURVERZEICHNIS

Bitkom e.V. *Jeder Fünfte shoppt mehrmals pro Woche online.* 2018. [Zugriff am 18. Juli 2019]. Verfügbar unter: www.bitkom.org

Bretschi, Corinna/Lorentz, Kai. *Präzisionsziele für die Ermittlung der Einwohnerzahl bei der Haushaltsstichprobe im Zensus 2021*. In: WISTA Wirtschaft und Statistik. Sonderheft Zensus 2021. 2019, Seite 12 ff.

Bundesverfassungsgericht (BVerfG). *Urteil des Zweiten Senats vom 19. September 2018.* 2018. [Zugriff am 18. Juli 2019]. Verfügbar unter: www.bverfg.de

Dillman, Don A. *The promise and challenge of pushing respondents to the Web in mixed-mode surveys.* In: Survey Methodology. Ausgabe 1/2017, Seite 3 ff.

Eurostat. *Verhaltenskodex für europäische Statistiken.* 2018. [Zugriff am 18. Juli 2019]. Verfügbar unter: www.destatis.de

Fittkau, Susanne/Maaß, Holger. *W3B-Report Trends im Nutzerverhalten Herbst/Winter 2017/18. Aktuelle Internet-Trends aus der Nutzerperspektive 45.* WWW-Benutzer-Analyse W3B. 1. Auflage. Hamburg 2018.

Forster, Daniel/Schmidl, Blasius. *ZENSUS 2011: Produktionsablauf der Belegverarbeitung am Beispiel der Gebäude- und Wohnungszählung (GWZ).* In: Bayern in Zahlen. Ausgabe 4/2012, Seite 219 ff.

Frees, Beate/Eimeren, Birgit. *ARD/ZDF-Onlinestudie 2011: Drei von vier Deutschen im Netz – ein Ende des digitalen Grabens in Sicht?* 2011. [Zugriff am 25. Juli 2018]. Verfügbar unter: www.ard-zdf-onlinestudie.de

Frees, Beate/Koch, Wolfgang. *ARD/ZDF-Onlinestudie 2018: Zuwachs bei medialer Internetnutzung und Kommunikation.* 2018. [Zugriff am 18. Juli 2019]. Verfügbar unter: www.ard-zdf-onlinestudie.de

Grunwald, Sven/Krause, Anja. *Umgang mit fehlenden Angaben in der Gebäude- und Wohnungszählung 2011.* In: Wirtschaft und Statistik. Ausgabe 8/2014, Seite 437 ff.

Initiative D21 e.V., fortiss gGmbH. *eGovernment MONITOR 2018, Nutzung und Akzeptanz digitaler Verwaltungsangebote – Deutschland, Österreich und Schweiz im Vergleich.* 2018. [Zugriff am 18. Juli 2019]. Verfügbar unter: www.egovernment-monitor.de

Körner, Thomas/Krause, Anja/Ramsauer, Kathrin. *Anforderungen und Perspektiven auf dem Weg zu einem künftigen Registerzensus*. In: WISTA Wirtschaft und Statistik. Sonderheft Zensus 2021. 2019, Seite 74 ff.

Moore, Tracy/Bailie, Lorna/Gilmour, Graeme. *Building a Business Case for Census Internet Data Collection.* 2009. [Zugriff am 18. Juli 2019]. Verfügbar unter: www.statcan.gc.ca

Neutze, Michael/Pfahl, Miriam/Schweinert-Albinus, Stefan. *Das Innovationspotenzial des Zensus 2021.* In: WISTA Wirtschaft und Statistik. Ausgabe 5/2017, Seite 37 ff.

LITERATURVERZEICHNIS

Office for National Statistics. *2017 Census Test Report.* 2017. [Zugriff am: 18. Juli 2019]. Verfügbar unter: www.ons.gov.uk

Statistisches Bundesamt. *Private Haushalte in der Informationsgesellschaft – Nutzung von Informations- und Kommunikationstechnologien.* 2018. [Zugriff am 18. Juli 2019]. Verfügbar unter: www.destatis.de

Statistisches Bundesamt. *Wie digital ist Deutschland? – Machen Sie mit bei der IKT-Befragung 2019!* Pressemitteilung Nr. 119 vom 28. März 2019. [Zugriff am 18. Juli 2019]. Verfügbar unter: www.destatis.de

UNECE. *2010 Population Census Round.* 2017. [Zugriff am 18. Juli 2019]. Verfügbar unter: www.statswiki.unece.org

UNECE. *2020 Population Census Round*. 2019. [Zugriff am 18. Juli 2019]. Verfügbar unter: www.statswiki.unece.org

AUFBAU DES QUALITÄTSMANAGEMENTS FÜR DEN ZENSUS 2021

Thorsten Tümmler, Irina Meinke

↘ **Schlüsselwörter:** Zensus 2021 – Qualitätsrichtlinien – Qualitätsdatenblätter im Verbund – Verhaltenskodex für europäische Statistiken – ISO

ZUSAMMENFASSUNG

Für den Zensus 2021 wird seit Oktober 2018 im Statistischen Bundesamt ein Qualitätsmanagementsystem aufgebaut. Ein systematisches Qualitätsmanagement soll die Qualität der Prozesse und Ergebnisse steuern, bewerten und dokumentieren. Politik, Verwaltung und Wirtschaft benötigen die Daten des Zensus als verlässliche Entscheidungs- und Planungsgrundlage. Daher muss sichergestellt werden, dass die Ergebnisse des Zensus höchsten Qualitätsansprüchen genügen. Dieser Aufsatz stellt die derzeitige Ausgestaltung des Qualitätsmanagements für den Zensus 2021 dar. Im ersten Teil werden die Akteure, die Ziele und der Umfang des Qualitätsmanagements thematisiert. Der zweite Teil beschreibt die vier Dimensionen des Qualitätsmanagements.

↘ ***Keywords:*** *2021 Census – quality guidelines – quality data sheets in the official statistics network – European Statistics Code of Practice – ISO*

ABSTRACT

Since October 2018 a quality management system has been set up for the 2021 Census at the Federal Statistical Office. Systematic quality management is designed to control, assess and document the quality of the processes and results. Politicians, public authorities and the business community need the Census data as a reliable basis for decision-making and planning. For this reason, it must be ensured that the Census results meet the highest quality requirements. This article describes the current development of quality management processes for the 2021 Census. The first part of the article introduces the parties involved, goals and extent of quality management, while the second part describes the four dimensions of quality management.

Thorsten Tümmler

hat in München Volkswirtschaft studiert und ist im Statistischen Bundesamt mit der Projektleitung für das Qualitätsmanagement für den Zensus 2021 beauftragt. Er war von 2004 bis 2010 im Bereich Unternehmensstatistik und Unternehmensregister tätig. Von 2010 bis 2013 leitete er das Referat „Produktion der Industrie“, seit 2013 das Referat „Methoden der Qualitätssicherung, Digitale Agenda“.

Irina Meinke

studierte Politikwissenschaft sowie Technologie- und Innovationsmanagement. Seit November 2018 unterstützt sie im Statistischen Bundesamt die Projektleitung Qualitätsmanagement Zensus 2021. Zuvor war sie seit 2013 in den Referaten „Methoden der Qualitätssicherung“ und „Nationale Koordinierung, Programmplanung“ tätig und befasste sich mit der nationalen Umsetzung des Verhaltenskodex für europäische Statistiken sowie der Einführung und Entwicklung von Instrumenten zur Qualitätssicherung.

1 Einleitung

Mit dem Zensus 2021 führen die Statistischen Ämter des Bundes und der Länder in Deutschland wieder eine Volks-, Gebäude- und Wohnungszählung durch. Gewonnen werden sollen die amtliche Einwohnerzahl sowie weitere tief gegliederte Daten zur Bevölkerung, Erwerbstätigkeit und zum Gebäude- und Wohnungsbestand. Politik, Verwaltung und Wirtschaft benötigen die Daten des Zensus als verlässliche Entscheidungs- und Planungsgrundlage. Für Wissenschaft, Markt- und Meinungsforschung sowie für die amtliche Statistik selbst liefert der Zensus neue, wichtige Basisdaten.

Zum Einsatz kommen wird auch 2021 wieder das Grundmodell des registergestützten Zensus 2011. Grundlage werden somit Daten aus den Melderegistern und weiteren Verwaltungsregistern sein. Allerdings liegen die benötigten Daten nicht vollständig oder nicht in der erforderlichen Qualität in Verwaltungsregistern vor. Daher werden die Registerauswertungen um Angaben aus Personenerhebungen (Haushaltebefragung auf Stichprobenbasis und Erhebung an Anschriften mit Sonderbereichen) und einer Gebäude- und Wohnungszählung ergänzt.

Um sicherzustellen, dass die Ergebnisse des Zensus 2021 höchsten Qualitätsansprüchen genügen, wird seit Oktober 2018 ein Qualitätsmanagementsystem im Statistischen Bundesamt aufgebaut, das in diesem Artikel erläutert wird. Das Qualitätsmanagement (QM) für den Zensus 2021 dient der zielgerichteten Steuerung und Dokumentation der Qualität der Prozesse und der Ergebnisse des Zensus, also der für die Nutzerinnen und Nutzer bereitzustellenden Daten. Kapitel 2 stellt den Rahmen des Qualitätsmanagements für den Zensus 2021 vor, den Akteure, Ziele, Umfang, das Zusammenspiel von Qualitätsmanagement und Qualitätssicherung sowie das Qualitätscontrolling bilden. Die vier Dimensionen des Qualitätsmanagements werden in Kapitel 3 beschrieben, Kapitel 4 zieht ein kurzes Fazit der bisherigen Aktivitäten zum Qualitätsmanagement für den Zensus 2021.

2 Rahmen des Qualitätsmanagements

Die folgenden Abschnitte beschreiben die Akteure, die Ziele, den Umfang des Qualitätsmanagements, das Zusammenspiel zwischen Qualitätsmanagement und Qualitätssicherung (QS) sowie das Qualitätscontrolling.

2.1 Akteure

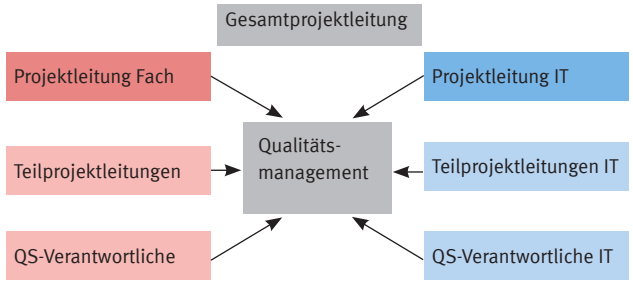
Im Oktober 2018 wurde im Statistischen Bundesamt eine Projektleitung für das Qualitätsmanagement (PL QM) für den Zensus 2021 etabliert. Diese verantwortet die Erstellung und Pflege der Qualitätsmanagementvorschriften. Sie ist zuständig für alle übergreifenden Qualitätsbelange sowie den Inhalt, die Wirtschaftlichkeit und Wirksamkeit des Qualitätsmanagementsystems sowie seine permanente Fortschreibung. Daneben gibt es die fachliche Projektleitung (PL Fach), die Projektleitung Informationstechnik (PL IT) und die Projektleitung für die externe Zensuskommunikation (PL KOM). Die Gesamtverantwortung für den Zensus 2021 trägt die Gesamtprojektleitung (GPL).

Das Großprojekt Zensus 2021 wird von den Statistischen Ämtern des Bundes und der Länder gemeinsam durchgeführt. Es gliedert sich in etwa 20 Teilprojekte, die sich mit der fachlich-methodischen Konzeption, der Durchführung der Erhebungen und der Aufbereitung der Ergebnisse befassen.[1] Zu diesen Erhebungen zählen die Haushaltebefragung auf Stichprobenbasis sowie die Erhebung an Sonderbereichen (unter anderem an allen Anschriften mit Wohnheimen und Gemeinschaftsunterkünften) zur statistischen Korrektur der Melderegister und Ermittlung personenbezogener und soziodemografischer Strukturmerkmale. Dazu zählt auch die Gebäude- und Wohnungszählung, bei der Daten zur Art des Gebäudes, zur Anzahl der Wohnungen im Gebäude, zur Nutzungsart und zur Fläche der Wohnung ermittelt werden. Die Datenlieferungen der Datenbestände aus den Melderegistern, die neben personenbezogenen Daten auch Angaben zu den Anschriften enthalten, sind ebenfalls als Erhebungen zu behandeln. Fast alle

1 Außerdem gibt es etwa 20 weitere technische Teilprojekte mit dem Fokus auf der Entwicklung notwendiger IT-Komponenten.

Teilprojekte umfassen auch die Entwicklung entsprechender IT-Anwendungen für die Statistikproduktion. Teilweise wird die Software-Entwicklung an Externe vergeben. Letztlich gibt es Teilprojekte mit dem Schwerpunkt Facharbeit und Teilprojekte mit dem Schwerpunkt IT-Arbeit. ↘ Grafik 1

Grafik 1
Wichtige Akteure im Qualitätsmanagement für den Zensus 2021



Fachbereich Informationstechnik (IT)
QS = Qualitätssicherung

2019 - 01 - 0494

In diesen Teilprojekten wird die eigentliche Arbeit zur Qualitätssicherung sowohl fach- als auch IT-seitig geleistet. Das systematische und umfassende Qualitätsmanagement hat zum Ziel, diese Arbeiten zu koordinieren und nach einheitlichen Vorgaben ablaufen zu lassen. Zudem soll es die für die Projektsteuerung notwendige Transparenz schaffen. Wichtige Partner für die Projektleitung Qualitätsmanagement beim Aufbau und Betrieb eines Qualitätsmanagementsystems sind daher die Qualitätssicherungs-Verantwortlichen (QS-Verantwortliche) aus dem Fachbereich und aus der IT, die Teilprojektleitungen aus dem Fachbereich und der IT sowie die verantwortlichen PL Fach, PL IT und schließlich die Gesamtprojektleitung (siehe Grafik 1).

2.2 Ziele

Die Ergebnisse des Zensus 2021 sollen pünktlich zum geplanten Termin 18 Monate nach dem Stichtag veröffentlicht werden, nach wissenschaftlichen und internationalen Standards erstellt werden und eine definierte hohe Genauigkeit aufweisen. Das Zustandekommen der Daten soll nachvollziehbar dokumentiert und durch aussagekräftige Begleitdokumente erläutert werden. Die Zensusergebnisse sollen für alle Nutzergruppen zugänglich sein und insgesamt deren Bedarf entsprechen. Das systematische Qualitätsmanagement soll das Erreichen dieser Ziele sicherstellen.

Die Qualitätsziele, die das Qualitätsmanagement verfolgt, sind im Allgemeinen am Verhaltenskodex für europäische Statistiken ausgerichtet. Der Verhaltenskodex definiert in insgesamt 16 Grundsätzen die Qualitätskriterien in drei Bereichen:

a) zum institutionellen Rahmen der Statistikerstellung,

b) zu den statistischen Prozessen und

c) zu den statistischen Produkten.

Alle Aktivitäten bei der Entwicklung, Erstellung und Verbreitung von Statistiken – und damit auch im Zensus 2021 – sorgen für die Sicherstellung der Qualität der statistischen Endprodukte und somit der Ergebnisse, die den Nutzerinnen und Nutzern bereitgestellt werden. Im Verhaltenskodex für europäische Statistiken sind insgesamt fünf Grundsätze für die Qualität der Ergebnisse festgelegt: (1) Relevanz, (2) Genauigkeit und Zuverlässigkeit, (3) Aktualität und Pünktlichkeit, (4) Kohärenz und Vergleichbarkeit sowie (5) Zugänglichkeit und Klarheit.[2] Insbesondere für die Genauigkeit (beispielsweise anhand des Stichprobenfehlers) und die Aktualität (anhand des Veröffentlichungstermins) werden konkrete Zielwerte definiert.

2 Parallel zum Verhaltenskodex werden in Artikel 12 der Verordnung (EG) Nr. 223/2009 des Europäischen Parlaments und des Rates vom 11. März 2009 über europäische Statistiken (Amtsblatt der EU Nr. L 87, Seite 164) – die die rechtliche Grundlage für die Entwicklung, Erstellung und Verbreitung von europäischen Statistiken und somit auch für den Zensus 2021 bildet – genau diese fünf Qualitätskriterien rechtlich festgeschrieben. Dies verdeutlicht die große Bedeutung dieser fünf Qualitätskriterien. Auch im Artikel 6 der Verordnung (EG) Nr. 763/2008 des Europäischen Parlaments und des Rates vom 9. Juli 2008 über Volks- und Wohnungszählungen (Amtsblatt der EU Nr. L 218, Seite 14) – die die rechtliche Grundlage für den Zensus 2021 bildet – werden genau diese Qualitätskriterien festgeschrieben.

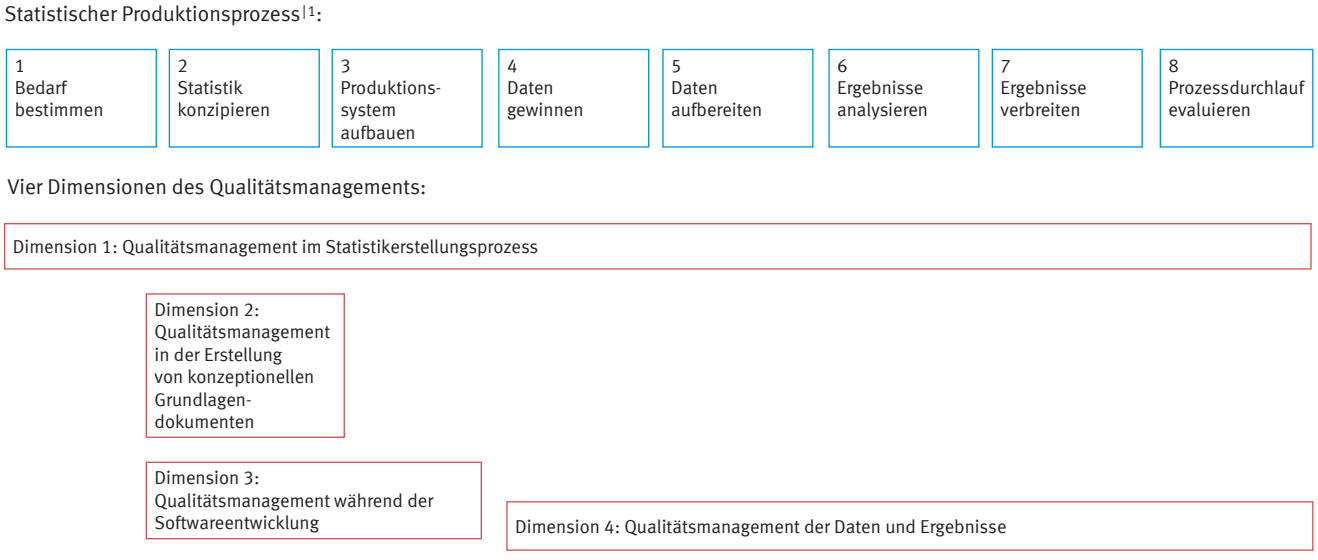
2.3 Umfang

Die Vorgaben zum Qualitätsmanagement decken den gesamten statistischen Produktionsprozess ab und umfassen vier Dimensionen: ↘ Grafik 2

1) Qualitätsmanagement im Statistikerstellungsprozess,
2) Qualitätsmanagement in der Erstellung von konzeptionellen Grundlagendokumenten,
3) Qualitätsmanagement während der Softwareentwicklung und
4) Qualitätsmanagement der Daten und Ergebnisse.

Ziel ist es, in den vier Dimensionen des Qualitätsmanagements die Voraussetzungen zu schaffen und spezifische Vorgaben festzulegen, damit durch eine hohe Qualität der Prozesse die Qualität der Ergebnisse systematisch, zuverlässig und nachweisbar sichergestellt wird. Die Qualität der letztlich veröffentlichten Zensusergebnisse wird durch die vorgelagerten Konzeptions- und Produktionsprozesse maßgeblich beeinflusst. Daher ist es für das Qualitätsmanagement für den Zensus 2021 wichtig, den gesamten Prozess der Statistikerstellung im Blick zu haben und in allen Arbeitsphasen entsprechende Qualitätssicherungsmaßnahmen zu implementieren (Dimension 1). Das Fundament für qualitativ hochwertige Ergebnisse wird bereits bei der Konzeption einer Statistik gelegt: Konzeptionelle Grundlagendokumente, die Vorgaben für einzelne Schritte der Statistikerstellung festlegen (beispielsweise zur Datengewinnung, zur Durchführung von Datenprüfungen und Fehlerkorrekturen und so weiter) stellen eine wichtige Dimension des Qualitätsmanagements dar (Dimension 2). Aber auch während der Softwareentwicklung setzt das Qualitätsmanagement für den Zensus 2021 an. So ist es wichtig, eine korrekt laufende und ausgereifte Software für die Datengewinnung, -aufbereitung und -analyse zu verwenden, um die Qualität der „am Ende“ zu veröffentlichenden Ergebnisse sicherzustellen (Dimension 3). Weiterhin beeinflussen der Erhebungs- und der Aufbereitungsprozess die Qualität der zu veröffentlichenden Ergebnisse: Werden (Roh-)Daten (aus verschiedenen Gründen) bereits fehlerhaft erfasst und können diese Fehler nur bedingt oder nicht identifiziert und korrigiert werden, führt das zu einer qualitativ minderwertigen Basis für die Endergebnisse (Dimension 4).

Grafik 2
Die vier Dimensionen des Qualitätsmanagements



1 Nach dem Geschäftsprozessmodell Amtliche Statistik (GMAS).

2019-01-0495

In jeder Dimension werden Qualitätssicherungsmaßnahmen (QS-Maßnahmen) durchgeführt, um die festgelegten Qualitätsziele zu erreichen. Hierfür stellt die Projektleitung Qualitätsmanagement in Zusammenarbeit mit der fachlichen Projektleitung und der Projektleitung Informationstechnik beziehungsweise mit den entsprechenden QS-Verantwortlichen aus dem Fachbereich und aus dem IT-Bereich für die jeweilige Dimension entsprechende Qualitätssicherungsinstrumente (QS-Instrumente) zur Verfügung. Für die Qualitätssicherung spielen viele Akteure mit unterschiedlichen Aufgabenschwerpunkten auf der Fach-, IT- und QM-Seite eine wichtige Rolle, die es im Rahmen des systematischen Qualitätsmanagements miteinander zu verzahnen gilt. Diese Verzahnung und Zusammenführung von Instrumenten und Akteuren erfolgt im Rahmen der Beschreibung und Modellierung von qualitätsrelevanten Prozessen beziehungsweise Abläufen und stellt eine wichtige Aufgabe im Qualitätsmanagement dar.

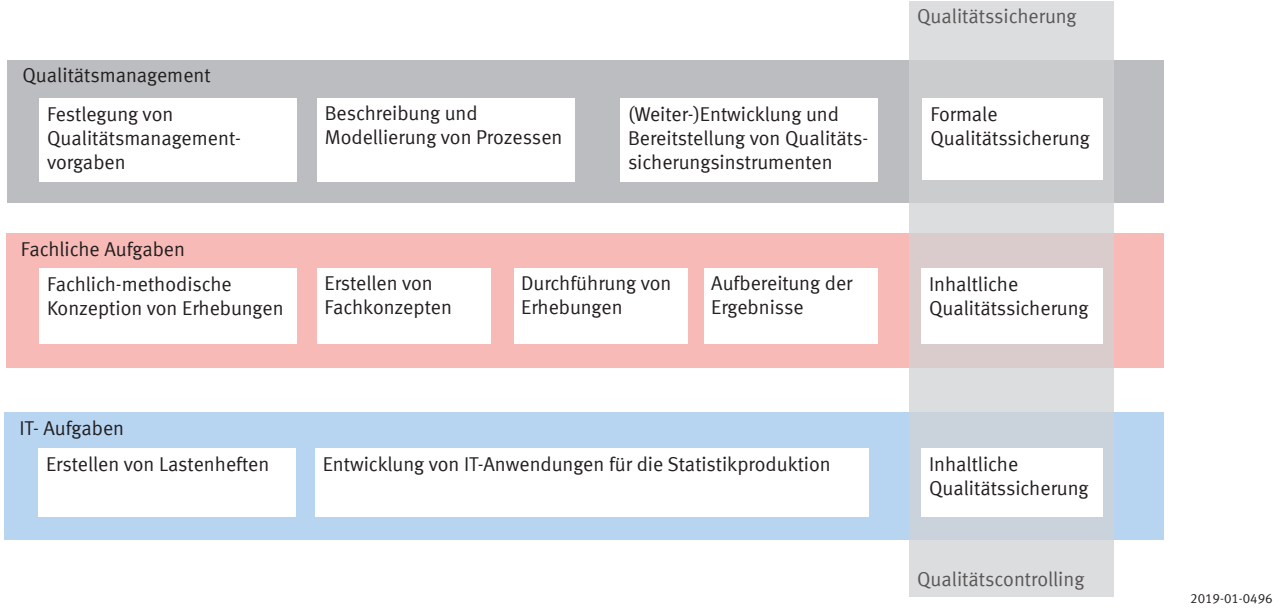
2.4 Zusammenspiel Qualitätsmanagement und Qualitätssicherung

Grundsätzlich wird zwischen Qualitätsmanagement und Qualitätssicherung unterschieden. ↘ Grafik 3

Das Qualitätsmanagement erfüllt verschiedene Aufgaben. Die Projektleitung Qualitätsmanagement legt für die vier Dimensionen des Qualitätsmanagements in Abstimmung mit der Gesamtprojektleitung, der fachlichen Projektleitung und der Projektleitung Informationstechnik Qualitätsmanagementvorschriften fest. Diese leiten die inhaltliche Qualitätssicherung der verschiedenen Akteure auf Fach- und IT-Seite, wo die eigentliche Qualitätsarbeit stattfindet. Im Rahmen der formalen Qualitätssicherung (siehe unten) wird die Einhaltung dieser Vorgaben durch die Projektleitung Qualitätsmanagement überprüft und sichergestellt. Weitere wichtige Aufgaben im Qualitätsmanagement sind die Beschreibung und Modellierung qualitätsrelevanter Prozesse sowie die (Weiter-)Entwicklung und Bereitstellung von QS-Instrumenten, die die entsprechenden Akteure in ihrer Qualitätsarbeit unterstützen sollen. In Kapitel 3 wird für jede Dimension erläutert, welche Prozesse durchlaufen werden müssen, welche QS-Instrumente in der entsprechenden Dimension relevant sind, welche QM-Vorgaben im Vordergrund stehen und wie die inhaltliche sowie formale Qualitätssicherung zusammenspielen.

Bei der Qualitätssicherung geht es um die eigentlichen Aktivitäten, um eine hohe Prozess- und Ergebnisquali-

Grafik 3
Zusammenspiel Qualitätsmanagement und Qualitätssicherung



tät zu erreichen. Dabei wird im Zensus 2021 zwischen zwei Arten der Qualitätssicherung unterschieden: Das sind zum einen die sogenannte inhaltliche Qualitätssicherung (die durch das jeweilige Teilprojekt fach- beziehungsweise IT-seitig erfolgt) und zum anderen die formale Qualitätssicherung (die durch die Projektleitung Qualitätsmanagement erfolgt).

Bei der inhaltlichen Qualitätssicherung geht es um die Durchführung von QS-Maßnahmen in den Statistischen Ämtern des Bundes und der Länder, die fach- beziehungsweise IT-seitig zu leisten sind und die die Qualität der Prozesse und Ergebnisse fachlich oder IT-seitig beeinflussen. Für die qualitätsgerechte Durchführung der Qualitätsmanagementvorschriften (Qualitätssicherung) sind (neben anderen Aufgaben) die Projektleitung Fach und die Projektleitung IT einschließlich der jeweils untergeordneten Teilprojektleitungen verantwortlich. In diesem Kontext wird von „inhaltlicher QS Fach“ beziehungsweise „inhaltlicher QS IT“ gesprochen. Dabei berät und unterstützt die Projektleitung Qualitätsmanagement die (Teil-)Projektleitungen bei allen Fragen, die das Qualitätsmanagement betreffen.

Bei der formalen Qualitätssicherung geht es insbesondere darum, die Einhaltung der Vorgaben des Qualitätsmanagements sicherzustellen. Die formal korrekte Dokumentation der in den Teilprojekten durchgeführten QS-Maßnahmen erfolgt in den Teilprojekten selbst und wird entsprechend der festgelegten Qualitätsmanagementvorschriften durch die Projektleitung Qualitätsmanagement geprüft. In diesem Kontext wird von „formaler Qualitätssicherung“ gesprochen.

2.5 Qualitätscontrolling

Die Umsetzung der Vorgaben des Qualitätsmanagements in den vier Dimensionen in den Statistischen Ämtern des Bundes und der Länder wird mit einem Qualitätscontrolling systematisch überprüft. Das Qualitätscontrolling ist neben Termin-, Ressourcen- und Risikocontrolling ein Teilaspekt des Projektcontrollings.

Das Qualitätscontrolling soll den Stand der Erreichung der Qualitätsziele beziehungsweise der Vorgaben des Qualitätsmanagements messbar und so die Zielerreichung sowie eventuell auftretende Abweichungen sichtbar machen. Hieraus ergeben sich Ansatzpunkte für die verschiedenen Leitungsebenen in den Statistischen Ämtern des Bundes und der Länder, um steuernd einzugreifen.

Zu allen Teilaspekten des Projektcontrollings erfolgen eine regelmäßige Berichterstattung sowie eine Statuseinschätzung der einzelnen Teilprojekte. Das Projektcontrolling wird durchgeführt und koordiniert durch das Projektmanagementoffice[3] des Statistischen Bundesamtes.

Für zusätzliche Planungs- und Umsetzungssicherheit der strategischen Zensusziele wurde außerdem ein vierteljährliches externes Controlling eingerichtet. Um regelmäßig eine neutrale und unabhängige Sicht von außen auf den Projektplan, den Projektfortschritt sowie die Qualität der Projektergebnisse zu erhalten, wurde dafür das Wirtschaftsprüfungs- und Beratungsunternehmen KPMG beauftragt.

3

Vier Dimensionen des Qualitätsmanagements

Das Qualitätsmanagementsystem deckt den gesamten statistischen Produktionsprozess ab (siehe Grafik 2), wobei in den vier Dimensionen unterschiedliche Instrumente der Qualitätssicherung zum Einsatz kommen. Im Folgenden werden die Dimensionen und die entsprechenden Instrumente näher beschrieben.

3.1 Dimension 1: Statistikerstellungsprozess

Der Statistikerstellungsprozess kann typischerweise in acht Phasen eingeteilt werden. Sie decken von den vorbereitenden Phasen über die eigentliche Statistikproduktion bis hin zur Veröffentlichung der Ergebnisse und der Evaluierung des Prozessdurchlaufs alle für die Statistikerstellung relevanten Prozessschritte ab.

3 Das Projektmanagementoffice ist das Bindeglied zwischen der Gesamtprojektleitung, den Projektleitungen sowie den fachlichen und IT-Teilprojektleitungen.

↘ Geschäftsprozessmodell Amtliche Statistik (GMAS)

Referenzrahmen für die Beschreibung und Darstellung eines solchen „typischen“ Ablaufs einer Statistik – unabhängig von der Methodik, den Erhebungsinhalten und Besonderheiten der einzelnen Statistik – bildet das sogenannte Geschäftsprozessmodell Amtliche Statistik (GMAS), das auch auf den Zensus 2021 anwendbar ist. Vorbereitende Phasen: Phase 1 „Bedarf bestimmen“, Phase 2 „Statistik konzipieren“ und Phase 3 „Produktionssystem aufbauen“. Eigentliche Statistikproduktion: Phase 4 „Daten gewinnen“, Phase 5 „Daten aufbereiten“ und Phase 6 „Ergebnisse analysieren“. Veröffentlichung der Ergebnisse: Phase 7 „Ergebnisse verbreiten“. Evaluierung des Prozessdurchlaufs: Phase 8 „Prozessdurchlauf evaluieren“.

Um die Qualität des Erstellungsprozesses der Ergebnisse des Zensus 2021 sicherzustellen, sollen die im Qualitätshandbuch (Statistische Ämter des Bundes und der Länder, 2018) beschriebenen Qualitätsrichtlinien umgesetzt werden. Die Qualitätsrichtlinien dienen der Erfüllung des oben genannten Verhaltenskodex für europäische Statistiken und sind eine umfassende Sammlung von Maßnahmen zur Qualitätssicherung, die den gesamten Statistikerstellungsprozess nach dem GMAS abdecken. Sie bilden den Standard für die systematische Qualitätssicherung im Statistikerstellungsprozess in der amtlichen Statistik. ↘ Übersicht 1 auf Seite 66

Die Qualitätsrichtlinien des Qualitätshandbuchs der Statistischen Ämter des Bundes und der Länder unterstützen die inhaltliche QS-Arbeit des Fachbereichs. Indem für jeden Schritt des Statistikerstellungsprozesses notwendige Qualitätsrichtlinien beziehungsweise QS-Maßnahmen aufgelistet werden, fördern sie zudem, dass die Statistikerstellung das angestrebte Qualitätsniveau erreicht. Die Qualitätsrichtlinien sind in den Statistischen Ämtern des Bundes und der Länder ein etabliertes Instrument der Qualitätssicherung, das seit 2015 für alle Statistiken angewandt wird. Für die Anwendung beim Zensus 2021 wurden die Qualitätsrichtlinien an dessen Rahmenbedingungen angepasst. Insgesamt ist beim Zensus 2021 die Durchführung mehrerer Hundert solcher Qualitätssicherungsmaßnahmen in den fachlichen Teilprojekten vorgesehen.

Die Qualitätsrichtlinien erfüllen im Zensus 2021 damit die Funktion eines umfassenden QS-Plans[4], der die inhaltliche Qualitätssicherung im Statistikerstellungsprozess leitet. Im Rahmen eines solchen QS-Plans wird bestimmt, welche QS-Maßnahmen wann umgesetzt werden. Dies hilft dabei, in einem über mehrere Jahre angelegten Großprojekt wie dem Zensus 2021 den Überblick über die zu erfüllenden QS-Maßnahmen zu behalten, die Qualitätsarbeiten im Statistikerstellungsprozess gezielt voranzubringen und Defizite systematisiert zu erkennen. Da das Großprojekt Zensus 2021 in mehrere Teilprojekte untergliedert ist, hat jedes Teilprojekt einen solchen spezifischen QS-Plan ausgearbeitet.

Im Rahmen des Qualitätscontrollings wird die Durchführung des hier beschriebenen QS-Plans für jedes Teilprojekt überprüft. Die erfolgte Durchführung der QS-Maßnahmen wird von den Teilprojekten durch eine aussagekräftige und nachvollziehbare Dokumentation belegt. Diese Dokumentation wird im Rahmen der formalen Qualitätssicherung durch die Projektleitung Qualitätsmanagement kritisch überprüft. Im Fall einer Beanstandung wird dies an das verantwortliche Teilprojekt zurückgespiegelt, um entweder die Verbesserung einer mangelhaften Dokumentation über die erfüllte QS-Maßnahme zu erwirken oder um die Umsetzung einer überfälligen QS-Maßnahme in die Wege zu leiten.

4 In Anlehnung an die Qualitätsmanagement-Normenreihe DIN EN ISO 9001:2015. Die Qualitätsmanagement-Normenreihe DIN EN ISO 9000 ff. wurde im Jahr 1987 eingeführt und bildet weltweit für eine Vielzahl von Organisationen der verschiedensten Branchen die Basis für Qualitätssicherung und Qualitätsmanagement. Sie ist das am weitesten verbreitete Normenwerk der Internationalen Organisation für Normung (ISO – International Organization for Standardization).

Übersicht 1

Beispiele für Qualitätsrichtlinien (strukturiert nach GMAS)[1]

Vorbereitung
Phase 1: Bedarf bestimmen
Neben dem eigentlichen Bedarf nach Informationen werden auch Nutzeranforderungen hinsichtlich der fachlichen und regionalen Gliederungstiefe, der geforderten Genauigkeit und Aktualität der Ergebnisse sowie der Periodizität der Ergebnisveröffentlichung ermittelt und dokumentiert.
Phase 2: Statistik konzipieren
Ausgestaltung der Datengewinnung: Beim Einsatz mehrerer Wege der Datengewinnung, beispielsweise indem für die Befragung ein Papierfragebogen oder ein elektronischer Fragebogen angeboten wird (sogenannter mixed mode), wird getestet, welche Auswirkungen daraus auf die Qualität der Daten resultieren (sogenannte mode effects). Die in diesen Tests gewonnenen Erkenntnisse werden bei der endgültigen Ausgestaltung der Datengewinnung berücksichtigt.
Phase 3: Produktionssystem aufbauen
Datengewinnung aufbauen: Grundsätzlich wird ein sogenannter Pretest durchgeführt, bei dem das Ausfüllen des Fragebogens vorab simuliert wird. […] Fragebogen werden dabei beispielsweise fachlich-inhaltlich sowie auf Verständlichkeit und neutrale Formulierung geprüft. […]
Verbreitungskomponenten aufbauen: Es wird festgelegt, welche Metadaten (zum Beispiel zu Definitionen, Methodik, Qualität) und Erläuterungen bei der Veröffentlichung der Ergebnisse bereitgestellt werden, um den Nutzerinnen und Nutzern eine korrekte Interpretation und aussagekräftige Vergleiche zu erleichtern.
Eigentliche Statistikproduktion
Phase 4: Daten gewinnen
Datengewinnung vorbereiten: Werden Interviewerinnen und Interviewer eingesetzt, so werden diese vorab geschult. Dabei wird darauf hingewirkt, dass jede Interviewerin beziehungsweise jeder Interviewer die Befragung methodisch korrekt durchführt. Inhalte, Umfang und Termine absolvierter Interviewer-Schulungen werden dokumentiert.
Datengewinnung durchführen: Um die angestrebte Genauigkeit und damit aussagekräftige Ergebnisse erreichen zu können, ist eine hinreichende Datenmenge (im Hinblick auf Vollzähligkeit und Vollständigkeit) erforderlich. Aus diesem Grund wird während der Datengewinnung darauf geachtet, dass der dafür notwendige Rücklauf erreicht wird.
Phase 5: Daten aufbereiten
Daten plausibilisieren und imputieren: Bei jeder Datenveränderung wird automatisch protokolliert, aufgrund welche(n/r) Fehlerschlüssel(s) diese durchgeführt wurde. Dabei werden manuelle Datenveränderungen aufgrund von Rückfragen beim Auskunftgebenden von Datenveränderungen durch Algorithmen (automatische Imputation) und von manuellen Datenveränderungen durch Expertenschätzung (manuelle Imputation) unterschieden.
Phase 6: Ergebnisse analysieren
Ergebnisse validieren: Antwortausfälle (auf Ebene von Einheiten oder Merkmalen; also Unit- beziehungsweise Item-Non-Response) werden im Hinblick auf systematische Antwortausfälle und die resultierende Belastbarkeit der Ergebnisse untersucht.
Ergebnisse fertigstellen: Die Qualität der Ergebnisse wird im Qualitätsbericht anhand der durch die europäische Statistikverordnung […] definierten Qualitätskriterien bewertet. […] Alle für ein statistisches Produkt relevanten Definitionen, Konzepte und Klassifikationen werden im Qualitätsbericht erläutert. Dies betrifft insbesondere den Gegenstand der Statistik, die statistischen Einheiten, die erhobenen und abgeleiteten Merkmale, die berechneten Maßzahlen und Indikatoren, die verwendeten Klassifikationen sowie regionale und sonstige Gliederungen.
Veröffentlichung und Evaluierung
Phase 7: Ergebnisse verbreiten
Die statistischen Ergebnisse werden durch […] Metadaten (zum Beispiel zu Definitionen, Methodik, Qualität) und Erläuterungen ergänzt, um den Nutzerinnen und Nutzern das Verständnis und die Interpretation der Ergebnisse zu erleichtern.
Phase 8: Prozessdurchlauf evaluieren
Die Qualität der Statistik wird anhand der im Verhaltenskodex für europäische Statistiken festgelegten Qualitätsdimensionen zur Prozess- und Produktqualität regelmäßig unter Einbeziehung aller relevanten Akteure evaluiert […].

GMAS: Geschäftsprozessmodell Amtliche Statistik.

1 Auszug aus dem Qualitätshandbuch (Statistische Ämter des Bundes und der Länder, 2018). Für eine bessere Verständlichkeit wurden manche Qualitätsrichtlinien in ihrem Wortlaut leicht angepasst.

3.2 Dimension 2: Erstellung von konzeptionellen Grundlagendokumenten

Zu Beginn spielen bei der Konzeption einer Statistik[5] konzeptionelle Grundlagendokumente (beispielsweise sogenannte Fachkonzepte, Lastenhefte und Pflichtenhefte beziehungsweise Spezifikationen) eine wichtige Rolle.

In Fachkonzepten werden alle für die Erstellung der Statistik relevanten fachlichen Abläufe, Tätigkeiten und Zuständigkeiten beschrieben. Die Fachkonzepte werden von den fachlichen Teilprojekten ausgearbeitet und richten sich insbesondere an die IT-Organisatorinnen und -Organisatoren sowie Anforderungsanalytikerinnen und -analytiker aus der Informationstechnik. Die Fachkonzepte stellen die Basis für die Erstellung von Lastenheften dar, die wiederum als Grundlage für die Softwareentwicklung[6] dienen (siehe Dimension 3).

Damit die zu erstellende Software den Bedürfnissen ihrer Nutzerinnen und Nutzer entspricht, müssen die Fachbereiche im Fachkonzept die fachlichen Anforderungen an ein neues IT-Fachverfahren ausreichend detailliert beschreiben, sodass Anforderungsanalytiker/-innen auf dieser Basis das Lastenheft[7] erstellen können. Für eine wirksame Anforderungsanalyse ist eine enge Zusammenarbeit mit den Fachkolleginnen und -kollegen sowie Nutzerinnen und Nutzern zwingend notwendig. Die fachlichen Mitarbeiterinnen und Mitarbeiter sowie die Nutzerinnen und Nutzer der entsprechenden Teilprojekte der Statistischen Ämter des Bundes und der Länder erarbeiten wesentliche Anforderungen mit und unterstützen die Anforderungsanalysten. Das Lastenheft beschreibt die funktionalen Anforderungen – also das, was das neue IT-Produkt leisten soll – und die nicht funktionalen Anforderungen, zum Beispiel Verfügbarkeit, Administration und Sicherheitskonzept. Für Rückfragen und Spezifizierungen stehen die fachlichen Mitarbeiterinnen und Mitarbeiter der entsprechenden Teilprojekte der Statistischen Ämter des Bundes und der Länder zur Verfügung.

Im Anschluss an die Lastenhefte werden das Pflichtenheft und die Spezifikationen erstellt. Dies erfolgt durch den Auftragnehmer der Softwareentwicklung (beim Zensus 2021 intern das Statistische Bundesamt selbst oder ein externer IT-Dienstleister) in Zusammenarbeit mit dem Auftraggeber (vertreten durch Anforderungsanalyse und Fachbereich). Das Pflichtenheft enthält folgende Ergebnisse: Datenflussplan, Datensatzbeschreibungen, Schnittstellen, Beschreibung der Lösung, Realisierung der Datenhaltung, Programmierumgebung und Leistungsanforderungen an Systeme. Gemeinsam mit dem Lastenheft bildet das Pflichtenheft die Grundlage für die Programmierung.

Folglich hängen Fachkonzepte, Lastenhefte, Pflichtenhefte und Spezifikationen, aber auch die Softwareentwicklung (siehe hierzu Abschnitt 3.3) eng miteinander zusammen und bauen aufeinander auf. Hierbei sind eine Vielzahl von Akteuren aus Fachbereich, Informationstechnik (bei interner Programmierung im Statistischen Bundesamt) und teilweise auch externen Gewerken (bei externer Programmierung durch IT-Dienstleistungsunternehmen) involviert. Dies stellt eine Herausforderung für das Qualitätsmanagement dar – insbesondere hinsichtlich folgender Frage: Wie kann der Bearbeitungszustand und der Stand der Qualitätssicherung nachvollziehbar dokumentiert sowie die Verständlichkeit von beispielsweise Fachkonzepten und Lastenheften sichergestellt werden, damit die Softwareentwicklung möglichst effektiv und effizient erfolgen kann?

Eine wichtige Aufgabe des Qualitätsmanagements in der Dimension 2 ist es, qualitätsrelevante Prozesse beziehungsweise Abläufe[8] hinsichtlich der Erstellung, der Qualitätssicherung und der Abnahme von konzeptionellen Grundlagendokumenten (beispielsweise Fachkonzepten und Lastenheften) zu beschreiben und zu

5 GMAS-Phase 2 „Statistik konzipieren".

6 Teilweise werden Fachkonzepte erstellt, die nicht zwangsläufig für die Softwareentwicklung bestimmt sind, wie beispielsweise im Falle von Pretests. Die in diesem Kapitel beschriebene Vorgehensweise bleibt auch in diesen Fällen in ihren Grundzügen ähnlich. Fachkonzepte richten sich nicht ausschließlich an die Informationstechnik, sie dienen auch der fachlichen Darstellung gegenüber anderen am Projekt Beteiligten (zum Beispiel Fachbereiche der Statistischen Ämter der Länder) und anderen Stakeholdern (zum Beispiel Ministerien).

7 Bei der Softwareentwicklung werden auch agile Methoden, beispielsweise Scrum, eingesetzt. In diesem Fall wird statt eines Lastenhefts bei Scrum ein sogenanntes Backlog erstellt und gepflegt. Das Backlog dient (wie das Lastenheft) der Beschreibung der Anforderungen an das zu entwickelnde IT-Werkzeug. Für die Erstellung von im Backlog enthaltenen User Stories ist der sogenannte Product Owner beziehungsweise ein Product Owner Committee zuständig. Auch das Backlog beziehungsweise die dort enthaltenen User Stories sollen Gegenstand einer Qualitätssicherung sein.

8 Teilweise werden bereits bestehende Prozesse nachvollzogen, geprüft und gegebenenfalls an die Vorgaben der Projektleitung Qualitätsmanagement angepasst. Alternativ werden noch nicht bestehende Prozesse den Vorgaben der Projektleitung Qualitätsmanagement entsprechend etabliert.

modellieren[9] und mit den entsprechenden Akteuren sowie QS-Instrumenten zu verzahnen. Bei der Prozessmodellierung werden sogenannte Schleusen der Qualitätssicherung (Quality Gates) dargestellt. Quality Gates sind Punkte im Ablauf, bei denen anhand von im Voraus festgelegten Qualitätskriterien beziehungsweise Qualitätsmanagementvorschriften über die Freigabe für den nächsten Prozessschritt entschieden wird.

Für die Quality Gates bei der Erstellung von konzeptionellen Grundlagendokumenten hat die Projektleitung Qualitätsmanagement für den Zensus 2021 verschiedene Dokumentenvorlagen (in Anlehnung an die Qualitätsmanagement-Normenreihe DIN EN ISO 9000 ff.) etabliert: zum einen für die Erstellung von Fachkonzepten und Lastenheften und zum anderen für Prüfberichte und -protokolle, die bei der Umsetzung der festgelegten inhaltlichen und formalen Qualitätskriterien beziehungsweise Qualitätsmanagementvorschriften unterstützen. Der verpflichtende Einsatz dieser Dokumentenvorlagen (und das Führen einer entsprechenden Nachweisakte mithilfe dieser Dokumentenvorlagen) dokumentiert den Bearbeitungszustand und den Stand der Qualitätssicherung nachvollziehbar und stellt die Verständlichkeit des Dokuments sicher. Außerdem werden die getätigten inhaltlichen Qualitätsarbeiten seitens des Fachbereichs und der Informationstechnik dokumentiert und damit sichtbar gemacht sowie über die verschiedenen Teilprojekte hinweg standardisiert.

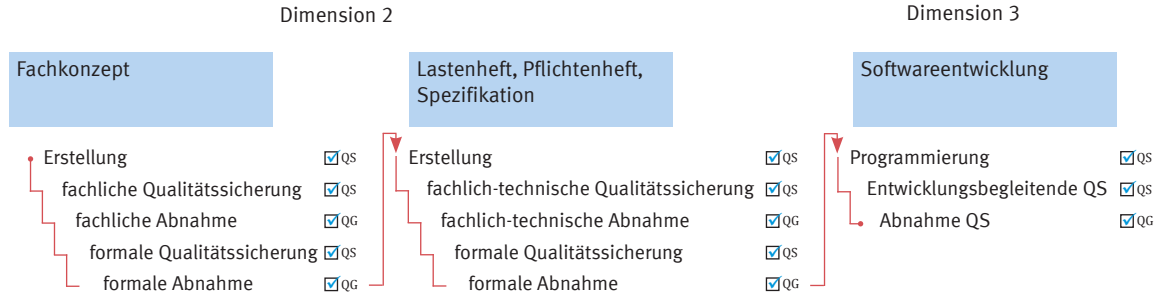
Wichtige Schritte im Qualitätsmanagement bei der Erstellung sowohl von Fachkonzepten als auch von Lastenheften sind die Ersterstellung, die fachliche beziehungsweise IT-seitige Qualitätssicherung sowie die fachliche beziehungsweise IT-seitige Abnahme des jeweiligen Dokuments. Insgesamt werden jeweils fünf Schritte „durchlaufen“, wobei es sich bei den Schritten drei und fünf um Quality Gates handelt. ↘ Grafik 4 Bei diesen wird geprüft, ob die jeweiligen inhaltlichen und formalen Qualitätskriterien und Qualitätsmanagementvorschriften eingehalten wurden. Erst wenn das der Fall ist, wird der nächste Prozessschritt initiiert. Das heißt erst wenn das Fachkonzept inhaltlich und formal abgenommen ist, kann das Fachkonzept zur Lastenhefterstellung übergeben werden. Gleiches gilt für das Lastenheft, das an die Softwareentwicklung übergeben werden soll (siehe Grafik 4 und den folgenden Exkurs).[10]

9 Prozesse zu modellieren bedeutet, dass Prozesse (abstrahiert) grafisch dargestellt werden, um eine einheitliche Sicht auf die Abläufe zu erhalten. Bei der Prozessmodellierung wird sich zum einen an der Abbildung der Realität beziehungsweise etablierten Vorgehensweise orientiert. Zum anderen spielen auch regulatorische Vorgaben zur Dokumentation von Prozessen und das Durchlaufen sogenannter Schleusen der Qualität (Quality Gates) eine wichtige Rolle. Außerdem werden klare Rollen beziehungsweise Verantwortlichkeiten festgeschrieben.

10 Im Rahmen dieses Abschnitts werden lediglich die in Grafik 4 dargestellten Quality Gates bei der Erstellung der Fachkonzepte und Lastenhefte behandelt. Das Qualitätsmanagement während der Softwareentwicklung wird in Abschnitt 3.3 behandelt.

Grafik 4

Quality Gates bei der Erstellung von konzeptionellen Grundlagendokumenten (Dimension 2) sowie bei der Softwareentwicklung (Dimension 3)



QS: Qualitätssicherung – QG: Quality Gate.

Zuständigkeiten: fachlich – Fachbereich, technisch – Informationstechnik, formal – Projektleitung Qualitätsmanagement

2019-01-0498

Beschreibung des Prozesses „Erstellung von Fachkonzepten und Lastenheften“

In einem ersten Schritt erstellen die Mitarbeiterinnen und Mitarbeiter aus dem Fachbereich oder der Informationstechnik des jeweiligen Teilprojekts das Fachkonzept und das Lastenheft. Hierfür hat die Projektleitung Qualitätsmanagement ein Template ausgearbeitet. Mit dessen Hilfe werden den Teilprojekten wichtige formale Aspekte der Dokumentenlenkung (wie ein Änderungs- und Prüfverzeichnis), Dokumenteneigenschaften (wie Autor, Version und Datum), Glossar und Abkürzungsverzeichnis vorstrukturiert und über alle Teilprojekte hinweg standardisiert an die Hand gegeben.

In einem zweiten Schritt erfolgt die inhaltliche Qualitätssicherung auf der Fachseite, wenn es sich um das Fachkonzept handelt beziehungsweise federführend auf der IT-Seite, wenn es sich um ein Lastenheft handelt. Hier wird die eigentliche inhaltliche Qualitätsarbeit durchgeführt. Die inhaltliche Qualitätssicherung wird in einem Prüfbericht dokumentiert, dafür hat die Projektleitung Qualitätsmanagement eine Vorlage ausgearbeitet. Mithilfe der Vorlage für den Prüfbericht soll die inhaltliche Prüfung von Fachkonzepten und Lastenheften nach einheitlichen Prüfkriterien erfolgen. Resultieren aus der inhaltlichen Qualitätssicherung keine weiteren Verbesserungsbedarfe, erfolgt anschließend in einem dritten Schritt die inhaltliche Abnahme des Fachkonzepts beziehungsweise Lastenhefts (Quality Gate).

Im vierten Schritt erfolgt die formale Qualitätssicherung durch die Projektleitung Qualitätsmanagement. Hierbei wird geprüft, ob die Vorgaben aus dem Template für die Erstellung eines Fachkonzepts beziehungsweise Lastenhefts durch den Fachbereich beziehungsweise die Informationstechnik eingehalten wurden. Ebenfalls wird geprüft, ob eine inhaltliche Qualitätssicherung und Abnahme durch den Fachbereich beziehungsweise die Informationstechnik erfolgt ist. Die formale Qualitätssicherung wird in einem sogenannten QS-Prüfprotokoll dokumentiert. Hierfür hat die Projektleitung Qualitätsmanagement eine Vorlage ausgearbeitet, die sicherstellt, dass die formale Prüfung von Fachkonzepten und Lastenheften nach einheitlichen Prüfkriterien erfolgt. Ist die formale Qualitätssicherung erfolgreich, erfolgt anschließend in einem fünften und letzten Schritt die formale Abnahme durch die Projektleitung Qualitätsmanagement (Quality Gate).

3.3 Dimension 3: Softwareentwicklung

Die Statistikproduktion läuft in den Statistischen Ämtern des Bundes und der Länder IT-gestützt ab. Für den Zensus 2021 sind IT-Fachanwendungen zu entwickeln, die die Datenerhebung, -prüfung, -korrektur und -analyse sowie die Veröffentlichung unterstützen.

Die Qualitätssicherung in der Softwareentwicklung ist ein mehrstufiger, ineinandergreifender Prozess, in dem mehrere Rollen involviert sind und miteinander kooperieren. Dazu gehören Anforderungsanalyse, Systemarchitektur, Softwarearchitektur, Konfigurationsmanagement, Softwareentwicklung, IT-Betrieb, Qualitätssicherung in der Informationstechnik und der Fachbereich Zensus. Basierend auf dem abgenommenen Fachkonzept, dem Lasten- und Pflichtenheft sowie Spezifikationen werden Softwarekomponenten für die Statistikproduktion entwickelt, wobei auch agile Methoden (zum Beispiel Scrum) eingesetzt werden können. In diesem Fall wird statt Lastenheft und Pflichtenheft ein sogenanntes Backlog[11] (bei Scrum) erstellt und gepflegt. Im Allgemeinen kann dabei der Qualitätssicherungsprozess in zwei Stufen untergliedert werden:

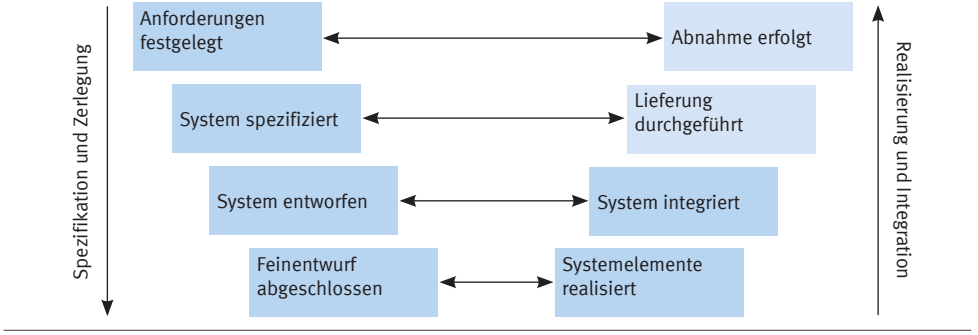
1) Qualitätssicherung während der Softwareentwicklung (Entwickler-Qualitätssicherung) und
2) Abnahme-, Systemtests und User-Acceptance-Tests (Abnahme-Qualitätssicherung)[12].

Referenzrahmen für den Qualitätssicherungsprozess im Software-Engineering-Prozess bildet dabei das für den Zensus angepasste V-Modell XT Bund, das den Softwareentwicklungsprozess in Phasen organisiert und das Vorgehen zur Qualitätssicherung (Testen) definiert. Die einzelnen Entwicklungsphasen werden den Testphasen gegenübergestellt, sodass das Namensgebende „V“ entsteht. Auf der linken Seite des „V“ wird mit einer funktionalen/fachlichen Spezifikation begonnen, die immer tiefer detailliert zu einer technischen Spezifikation und Implementierungsgrundlage ausgebaut wird. In der

11 Das Backlog dient (wie das Lastenheft) der Beschreibung der Anforderungen an das zu entwickelnde IT-Werkzeug; diese Beschreibungen erfolgen im Rahmen von sogenannten User Stories.

12 Im Fall einer Softwareentwicklung durch einen externen IT-Dienstleister wird eine zusätzliche Stufe der Entwickler-Qualitätssicherung im Statistischen Bundesamt durchgeführt. Dies erfolgt, um die Abnahme mit Qualitätskennzahlen der statischen Codeanalyse zu ergänzen und die Entwickler-Qualitätssicherung der Auftragnehmerin oder des Auftragnehmers zu verifizieren.

Grafik 5
V-Modell XT Bund[1]



1 Eigene Darstellung nach V-Modell XT Bund, Version 2.1, Seite 32.

2019-01-0499

Spitze erfolgt die Implementierung, die anschließend auf der rechten Seite gegen die entsprechenden Spezifikationen der linken Seite getestet wird. ↘ Grafik 5

Qualitätssicherung während der Softwareentwicklung

Bei der Softwareentwicklung wird ein Entwicklerteam durch den Fachbereich, die Software- und Systemarchitektur sowie das Konfigurationsmanagement und die Qualitätssicherung unterstützt. Bei der Umsetzung von Software-Anforderungen/User-Stories werden Style-Guides und Codier-Richtlinien angewandt, die Standards für Oberflächengestaltung und Programmierregeln bereitstellen.

Ein generisches QS-Handbuch für die Softwareentwicklung[13] enthält insbesondere die Festlegung der zu prüfenden Produkte und Vorgaben für Teststufen und Durchführung der Qualitätssicherung im Gesamtprojekt sowie weitere Vorgaben. Dazu zählen Vorgaben für die Qualitätssicherung von Lieferungen und für die Erstellung von Qualitätsberichten.

Bei der entwicklungsbegleitenden Qualitätssicherung wird über die Einhaltung der Codier-Richtlinien gewacht, es werden statische Code-Analysen durchgeführt sowie Programmfehler identifiziert und nach Schweregrad sortiert angezeigt. Für die programmierten Komponenten werden entsprechende Tests erstellt.

Bei Anwendung agiler Methoden erfolgt die Sicherstellung einer hohen Softwarequalität in der Abnahme der User Story durch einen Peer Review der Entwicklerinnen und Entwickler. Dabei wird der Prüfumfang gemäß einer Checkliste (Definition of Done) abgearbeitet. Außerdem werden die Statistischen Ämter der Länder (als Stakeholder beziehungsweise als Anwender) in den sogenannten Sprint-Reviews durch zwei Stellvertretungen (sogenannte Proxies) vertreten. Im Sprint-Review präsentiert das Scrum-Team die Ergebnisse des letzten Sprints. Die Proxies haben im Rahmen von Sprint-Reviews die Möglichkeit, den jeweils aktuellen Stand des Programms kritisch zu prüfen und Verbesserungsvorschläge zu machen. Die Statistischen Ämter der Länder geben Rückmeldung zu den Funktionalitäten der entsprechenden Fachanwendung beziehungsweise des entsprechenden Programmteils. Es wird evaluiert, welche Stories erfolgreich umgesetzt wurden. Ferner haben die Statistischen Ämter der Länder die Möglichkeit, an der Priorisierung der Umsetzung von User Stories und der Formulierung von Programmanforderungen mitzuwirken.

Abnahme- und Systemtests

Für Abnahme-, Integrations- und Systemtests werden Testfälle erstellt, die zur Verfügung stehende Funktionen abtesten. Mithilfe eines solchen Testfallkatalogs werden anwendernahe Testfälle beschrieben, die eine hochwertige und möglichst realitätsnahe Qualitätssicherung ermöglichen sollen. Bei der Erstellung der erforderlichen Testfälle ist die Mitwirkung der Proxyländer von wesentlicher Bedeutung.[14] Die strukturierten Testfälle werden durch explorative Tests ergänzt, die sowohl von der Fachseite als auch von der IT-Qualitätssicherung durchgeführt werden.

13 Das generische QS-Handbuch befindet sich in Entwicklung.

14 Die Mitwirkung von Proxyländern (die die Statistischen Ämter der Länder repräsentieren und die eigentlichen Nutzer der Software sind) bei der Durchführung von Tests als ein Standardfall wird angestrebt. Derzeit wird dies nur in einigen Teilprojekten praktiziert. Der Fachbereich im Statistischen Bundesamt wirkt an der Durchführung von Tests generell mit; das ist wesentlich für die Validierung der Testfälle.

Sowohl bei der internen als auch bei der externen Programmierung werden in vorher von der IT-Teilprojektleitung in Abstimmung mit der fachlichen Teilprojektleitung festgelegten Entwicklungsschritten Programmteile geliefert. Die gelieferten Programmteile werden einer Qualitätsprüfung unterzogen, indem die erstellten Testfälle ausgeführt werden. Die Einbindung der Statistischen Ämter der Länder erfolgt auf freiwilliger Basis. Im Rahmen von sogenannten Integrationstests wird das Zusammenwirken mit anderen Komponenten und Schnittstellen sowie die Integration in die Software-Architektur geprüft. Auftretende Fehlwirkungen werden als Fehlerticket erfasst. Die Ergebnisse dieser Testvorgänge werden in einem Testbericht für die Teilprojektleitungen (fachlich/technisch) zusammengefasst und eine Abnahmeempfehlung ausgesprochen.

3.4 Dimension 4: Daten und Ergebnisse

In den Dimensionen 1 bis 3 steht die Qualität der vorbereitenden Arbeiten im Fokus. Beim Qualitätsmanagement der Daten und Ergebnisse geht es nun um den Kern des Qualitätsmanagements, nämlich um die Qualität der erhobenen Daten. Dabei setzt das Qualitätsmanagement während des Erhebungs- und Aufbereitungsprozesses[15] an, indem die entstehende Qualität der Prozesse und Daten gemessen wird.

Das zur Messung der Datenqualität in den Statistischen Ämtern des Bundes und der Länder eingesetzte Instrument wird auch für den Zensus 2021 verwendet: das Konzept der sogenannten Qualitätsdatenblätter im Verbund (QuiV). Sie sollen im laufenden Erhebungs- und Aufbereitungsprozess umfassende prozessbezogene Informationen, Informationen zum Arbeitsfortschritt und insbesondere zur Datenqualität bereits bei der Entstehung automatisiert festhalten, daraus Kennzahlen ermitteln und diese dokumentieren.

Im Fokus der Qualitätsdatenblätter im Verbund stehen bestimmte Schritte der Erhebungs- und Aufbereitungsphase[16]:

1) Erhebungsgrundlage erstellen und Stichprobe ziehen,
2) Datengewinnung durchführen,
3) Daten prüfen und validieren,
4) Daten plausibilisieren und imputieren sowie
5) Ergebnisse fertigstellen und Ergebnisqualität bewerten.

Je nach Arbeitsschritt sind unterschiedliche QuiV-Kennzahlen relevant, die verschiedene Hinweise zur Einschätzung der Qualität der Daten und Ergebnisse sowie zum Arbeitsfortschritt geben. Die verschiedenen Kennzahlen decken somit verschiedene Arbeitsschritte des Erhebungs- und Aufbereitungsprozesses ab.

Die QuiV-Kennzahlen sollen die Verantwortlichen in den Statistischen Ämtern des Bundes und der Länder bei der Arbeitssteuerung, bei der Bewertung der Datenqualität und bei der Dokumentation der erreichten Qualität unterstützen. Sie sind somit ein Instrument zum Messen, Steuern und Dokumentieren der Prozess- und Ergebnisqualität:

› Qualitätsdatenblätter im Verbund als ein Instrument zum Messen und Steuern:
Die QuiV-Kennzahlen sollen während der Erhebungs- und Aufbereitungsphase Transparenz hinsichtlich der bisher erreichten Datenqualität und Arbeitsfortschritte gewährleisten. Diese Transparenz ist für eine effektive Steuerung im Statistischen Bundesamt sowie in den Statistischen Ämtern der Länder notwendig. Das bedeutet, dass das Statistische Bundesamt solche QuiV-Kennzahlen gegliedert nach Bundesländern und die Statistischen Ämter der Länder gegliedert nach Gemeinden beziehungsweise Erhebungsstellen[17] brauchen. Es soll für Gemeinden beziehungsweise Erhebungsstellen, Bundesländer und für die Bundesebene insgesamt erkennbar werden, wann Qualitätsvorgaben erreicht sind und wo noch Defizite bestehen: Wie erfolgreich ist die Datengewinnung? Gibt es hierbei Auffälligkeiten bei einzelnen Erhebungsstellen? Wie sind die Arbeitsfort-

15 Die Erhebungs- und Aufbereitungsphasen umfassen GMAS 4 „Daten gewinnen“, GMAS 5 „Daten aufbereiten“ und GMAS 6 „Daten analysieren“.

16 Eine ausführliche Übersicht über die Inhalte der Qualitätsdatenblätter im Verbund ist in der Onlineversion dieses Artikels enthalten.

17 Erhebungsstellen werden in den Kommunen extra für den Zensus eingerichtet. Aufgabe der Erhebungsstellen ist es zum Beispiel, die Einsätze der Interviewerinnen und Interviewer vor Ort zu organisieren, die Befragungsunterlagen zusammenzustellen oder die Vollständigkeit der Antworten zu prüfen.

Inhalte der Qualitätsdatenblätter im Verbund (QuiV)

	Erhebungsgrundlage erstellen und Stichprobe ziehen	Datengewinnung durchführen	Daten prüfen und validieren	Daten plausibilisieren und imputieren	Ergebnisse fertigstellen und Ergebnisqualität bewerten
Was passiert bei den für die QuiV-Kennzahlen relevanten Schritten der Statistikproduktion?	Zur Durchführung von Erhebungen wird eine geeignete Datenbasis als Erhebungsgrundlage benötigt. Die Erhebungsgrundlage umfasst alle Einheiten (beispielsweise meldepflichtige Personen ab 15 Jahren), die zur Durchführung von Erhebungen herangezogen werden können. Die Erhebungsgrundlage sollte möglichst vollständig sein. Wenn es sich bei der Erhebung um eine Stichprobe handelt, werden aus der Erhebungsgrundlage Stichprobeneinheiten ausgewählt und damit eine Stichprobe gezogen. Wenn es sich bei der Erhebung um eine Vollerhebung handelt, werden alle Einheiten aus der Erhebungsgrundlage zur Erhebung herangezogen.	Um die Datengewinnung durchzuführen, kommen verschiedene Instrumente zur Sammlung oder Erfassung statistischer Informationen zum Einsatz. Beispielsweise kann die Befragung mithilfe eines Papierfragebogens begleitet durch Interviewer/-innen oder mithilfe eines elektronisch beziehungsweise über das Internet übermittelten Fragebogens zum Selbstausfüllen ohne Interviewer/-innen erfolgen. Oder die direkte Befragung von einzelnen Personen wird ersetzt, indem beispielsweise sogenannte Verwaltungsdaten[2] für die Statistikerstellung herangezogen werden, die über die einzelnen Personen Auskunft geben. In diesem Fall kommt für die Datenübermittlung spezielle Software zum Einsatz.	Um potenzielle Probleme und Diskrepanzen, beispielsweise Ausreißer und Antwortausfälle in den Daten beziehungsweise Datensätzen zu erkennen, sind Datenprüfungen notwendig. Dabei werden Daten mithilfe vorgegebener Plausibilisierungsregeln validiert.	Bei falschen, fehlenden oder unzuverlässigen Daten können im Rahmen der sogenannten Plausibilisierung beziehungsweise Imputation neue Werte mithilfe eines regelbasierten Vorgehens eingefügt oder unzuverlässige Daten geändert werden.	Anhand von QuiV-Kennzahlen (und vorab festzulegenden Zielwerten für die Kennzahlen) wird das am Ende erreichte Qualitätsniveau der Zensusergebnisse (die Ergebnisqualität) bewertet und unter anderem in einem Qualitätsbericht dokumentiert und den Nutzerinnen und Nutzern der veröffentlichten Zensusergebnisse erläutert.
Über welche Aspekte geben die QuiV-Kennzahlen Auskunft?	Die QuiV können Hinweise über die Qualität der Erhebungsgrundlage liefern: • Wie viele Einheiten wurden zur Befragung ausgewählt? • Wie viele davon sind ausgewählt worden, obwohl sie nicht zu der Grundgesamtheit[1] gehören und damit nicht hätten für die Befragung ausgewählt werden dürfen (sogenannte „unechte Antwortausfälle“; beispielsweise Kinder unter 15 Jahren oder nicht meldepflichtige Personen ab 15 Jahren wie diplomatische Vertretungen)?	Die QuiV können Hinweise über Qualität und Erfolg der durchgeführten Datengewinnung liefern: • Wie erfolgreich war die Datengewinnung? Sprich: Wie viele Einheiten mussten gemahnt beziehungsweise erinnert werden und welche Rücklaufquote wurde erzielt? • Wie viele Einheiten haben welchen Meldeweg (beispielsweise postalisch, über das Internet, …) verwendet?	Die QuiV können Hinweise über die Qualität der eingegangenen Originaldaten sowie über die Qualität der Daten während und nach Abschluss der Datenprüfung und -validierung liefern: • Wie ist die Qualität der eingegangenen Originaldaten? Wie viele Antwortausfälle auf Ebene von Einheiten und Merkmalen gibt es[3]? • Wie ist die Einschätzung zur Qualität der Daten, nachdem der Arbeitsschritt „Datenprüfung“ durchgeführt wurde? • Wie viele Fehler wurden identifiziert?	Die QuiV können Hinweise über die Qualität der Daten während und nach Abschluss der Plausibilisierung und Imputation liefern: • Wie ist die Qualität der Daten, nachdem der Arbeitsschritt „Daten plausibilisieren und imputieren“ durchgeführt wurde? • Wie viele Fehler wurden behandelt beziehungsweise abgearbeitet? • Wie wurden die Fehler behandelt beziehungsweise abgearbeitet? Welches Imputationsverfahren wurde wie oft angewandt? • Wie viele behandelte Fehler waren sogenannte Muss- oder Kannfehler[4]? • Wie viele Fehler wurden nicht bearbeitet?	Die QuiV können Hinweise über die am Ende erreichte Qualität der veröffentlichten Ergebnisse liefern: • Wie ist die Qualität der am Ende veröffentlichten Ergebnisse? • Wie hoch ist der Ergebnisanteil, der nicht vom Auskunftgebenden stammt? (Imputationsquote) • Wie hoch ist die Anzahl der Einheiten, für die keine oder keine verwertbaren Daten gesammelt wurden, im Vergleich zur Gesamtzahl der Einheiten, für die Daten gewonnen werden sollten? (Quote der Antwortausfälle bei Einheiten) • …

1 Die Grundgesamtheit bezeichnet die Menge aller statistischen Einheiten (beispielsweise meldepflichtige Personen ab 15 Jahren), zu denen eine Aussage getroffen werden soll.
2 Auf Grundlage von Einzelstatistikgesetzen können Verwaltungsdaten für die Statistikerstellung verwendet werden. Verwaltungsdaten entstehen in Behörden im Rahmen der eigenen Verwaltungsprozesse und werden nicht primär für die amtliche Statistik generiert. Daher werden solche Erhebungen Sekundärerhebungen genannt. Bei der Befragung von Personen beispielsweise mithilfe von Interviewerinnen und Interviewern werden eigens beziehungsweise primär für die Zwecke der Statistik statistische Informationen direkt von den Befragten mithilfe eines Fragebogens ermittelt. Daher werden solche Erhebungen Primärerhebungen genannt.
3 Antwortausfälle bei Einheiten: zum Beispiel wie viele Personen haben keine Rückmeldung gegeben? Antwortausfälle bei Merkmalen: zum Beispiel wie viele Personen haben zu bestimmten Fragen keine Auskunft gegeben?
4 Mussfehler bezeichnen solche Fehler, die sicher Fehler sind, während Kannfehler Fehler sein können, aber nicht unbedingt Fehler sein müssen.

schritte bei der Fehlerprüfung und -korrektur? Werden für diese Arbeitsschritte weitere Ressourcen benötigt? So können gezielt und systematisch Handlungsbedarfe durch die verschiedenen Leitungsebenen in den Statistischen Ämtern der Länder und im Statistischen Bundesamt identifiziert und gemeinsam Maßnahmen eingeleitet werden, um Qualitätsdefizite zu beheben. Das ermöglicht eine bessere Prozesssteuerung.

› Qualitätsdatenblätter im Verbund als ein Instrument zur Dokumentation:
Ein weiteres wichtiges Ziel ist die Dokumentation der erreichten Qualität zu bestimmten Zeitpunkten des Produktionsprozesses, beispielsweise jeweils nach Durchführung der Datengewinnung, nach Abschluss der Fehlerprüfung, nach Abschluss der Fehlerkorrektur sowie nach Abschluss der Aufbereitung der Ergebnisse. Dies dient einer nachvollziehbaren Dokumentation der Entstehung der einzelnen Ergebnisse. Auch sind die Qualitätsdatenblätter im Verbund notwendig, um zum Abschluss des Zensus 2021 einen aussagekräftigen Qualitätsbericht zu erstellen, der die Nutzerinnen und Nutzer über die letztlich erreichte Qualität der veröffentlichten Zensusergebnisse informiert.

Im Fokus der Qualitätsdatenblätter im Verbund stehen die sogenannten Erhebungs-Teilprojekte, wie die Erhebung der Datenbestände aus Melderegistern, Personenerhebung, Erhebung an Sonderbereichen sowie die Gebäude- und Wohnungszählung. Bei diesen Teilprojekten spielen die Datenerhebung und die Datenaufbereitung zentrale Rollen. In der Erhebungs- und Aufbereitungsphase müssen die Arbeiten in Bund, Ländern und Gemeinden beziehungsweise Erhebungsstellen koordiniert und gesteuert werden. Datenerhebung und -aufbereitung beeinflussen die Qualität der Zensusergebnisse maßgeblich. Ein Katalog an aussagekräftigen Kennzahlen wird derzeit erarbeitet und muss anschließend in den IT-Fachverfahren so implementiert werden, dass die Kennzahlen jederzeit angezeigt werden können.

4

Fazit

Beim Zensus 2021 wird bereits für wesentliche Komponenten des Qualitätsmanagements (wie Statistikerstellungsprozess, Fachkonzepte, Lastenhefte und Softwareentwicklung) eine systematische Qualitätssicherung durchgeführt, die in ein strukturiertes und übergreifendes Qualitätsmanagementsystem integriert ist. Weitere Komponenten befinden sich derzeit in der Entwicklungsphase (beispielsweise das Qualitätsmanagement der Daten und Ergebnisse) und müssen anschließend in das Qualitätsmanagementsystem überführt werden. Das Qualitätsmanagement für den Zensus 2021 wird fortlaufend weiterentwickelt, ergänzt und optimiert.

LITERATURVERZEICHNIS

Bundesministerium des Innern, für Bau und Heimat. *V-Modell XT Bund*. 2009. [Zugriff am 14. Juni 2019]. Verfügbar unter: www.cio.bund.de

Eurostat/Ausschuss für das Europäische Statistische System. *Verhaltenskodex für europäische Statistiken*. 2017. [Zugriff am 27. Juni 2019]. Verfügbar unter: www.destatis.de

Statistische Ämter des Bundes und der Länder. *Vorgehensmodell amtliche Statistik bei der Softwareentwicklung und -pflege im statistischen Verbund*. VMAS-SE Version 3.0. 2016. (internes Dokument).

Statistische Ämter des Bundes und der Länder (Herausgeber). *Qualitätshandbuch der Statistischen Ämter des Bundes und der Länder. Version 1.1*. März 2018. [Zugriff am 25. Juni 2019]. Verfügbar unter: www.destatis.de

Statistikverbund (Statistische Ämter des Bundes und der Länder). *Geschäftsprozessmodell Amtliche Statistik (GMAS)*. (internes Dokument).

Wirtschaftskommission für Europa der Vereinten Nationen (UNECE). *Conference of European Statisticans. Recommendations for the 2020 Censuses of Population and Housing*. New York, Genf 2015. [Zugriff am 27. Juni 2019]. Verfügbar unter: www.unece.org

RECHTSGRUNDLAGEN

Verordnung (EG) Nr. 223/2009 des Europäischen Parlaments und des Rates vom 11. März 2009 über europäische Statistiken und zur Aufhebung der Verordnung (EG, Euratom) Nr. 1101/2008 des Europäischen Parlaments und des Rates über die Übermittlung von unter die Geheimhaltungspflicht fallenden Informationen an das Statistische Amt der Europäischen Gemeinschaften, der Verordnung (EG) Nr. 322/97 des Rates über die Gemeinschaftsstatistiken und des Beschlusses 89/382/EWG, Euratom des Rates zur Einsetzung eines Ausschusses für das Statistische Programm der Europäischen Gemeinschaften (Amtsblatt der EU Nr. L 87, Seite 164).

Verordnung (EG) Nr. 763/2008 des Europäischen Parlaments und des Rates vom 9. Juli 2008 über Volks- und Wohnungszählungen (Amtsblatt der EU Nr. L 218, Seite 14).

Thomas Körner

hat in Mainz und Caen (Frankreich) Soziologie studiert und in verschiedenen Bereichen des Statistischen Bundesamtes gearbeitet. Seit 2017 leitet er das Referat „Registerbasierte Ermittlung der Bevölkerungszahlen“, das für die fachliche und methodische Ausgestaltung eines künftigen Registerzensus verantwortlich ist.

Anja Krause

ist Geografin und arbeitet im Referat „Registerbasierte Ermittlung der Bevölkerungszahlen“ im Statistischen Bundesamt. Ihr Arbeitsschwerpunkt liegt dort auf der Entwicklung von Methoden zur registerbasierten Ermittlung von Angaben zu Gebäuden und Wohnungen.

Kathrin Ramsauer

hat Volkswirtschaftslehre in Freiburg studiert. Im Referat „Registerbasierte Ermittlung der Bevölkerungszahlen“ des Statistischen Bundesamtes beschäftigt sie sich insbesondere mit der Entwicklung von Methoden zur registerbasierten Ermittlung von Bevölkerungszahlen.

ANFORDERUNGEN UND PERSPEKTIVEN AUF DEM WEG ZU EINEM KÜNFTIGEN REGISTERZENSUS

Thomas Körner, Anja Krause, Kathrin Ramsauer

↘ **Schlüsselwörter:** Zensus – Bevölkerungsstatistik – Gebäude und Wohnungen – Registernutzung – Bevölkerungsfortschreibung

ZUSAMMENFASSUNG

Die Anforderungen an Zensus und Bevölkerungsstatistik unterliegen aktuell einem starken Wandel. Inzwischen werden Ergebnisse von den Nutzerinnen und Nutzern häufiger, aktueller und geokodiert gefordert. Diese sich abzeichnenden Anforderungen können das Modell des Zensus 2021 und die darauf aufbauende Bevölkerungsfortschreibung perspektivisch nicht mehr erfüllen. Daher haben schon heute die Überlegungen für die Zeit nach der Zensusrunde 2021 begonnen. Der vorliegende Beitrag stellt die Anforderungen an ein künftiges Zensusmodell vor und skizziert die Voraussetzungen und Umsetzungsmöglichkeiten auf dem Weg zu einem künftigen Registerzensus.

↘ ***Keywords:*** *census – population statistics – buildings and dwellings – register use – intercensal population updates*

ABSTRACT

The demands on the census and on population statistics are changing considerably. Users are now requiring results to be more timely, geocoded and available at shorter intervals. In the medium term, the model of the 2021 Census and the intercensal population updates based on it will not be able to meet these emerging requirements. This is why preparations for the post-2021 censuses have already started. This article presents the requirements to be met by a future census model and outlines the prerequisites and implementation options of a future register census.

1

Vom Zensus 2021 zum Registerzensus

Die Daten des Zensus und der Bevölkerungsstatistiken zählen zu den wichtigsten Produkten der amtlichen Statistik. Sie dienen als Entscheidungsgrundlage in vielen Bereichen, unter anderem für den Finanzausgleich zwischen Bund, Ländern und Kommunen und für die Festlegung der Wahlkreiszuschnitte. Daneben werden sie als Planungsgrundlage für Politik, Verwaltung und Wirtschaft sowie zur Adjustierung diverser Statistiken genutzt. Aufgrund der herausgehobenen Bedeutung unterliegen die Ergebnisse höchsten Qualitätsanforderungen, was nicht zuletzt das Urteil des Bundesverfassungsgerichts vom 19. September 2018 zum Zensus 2011 dokumentiert.

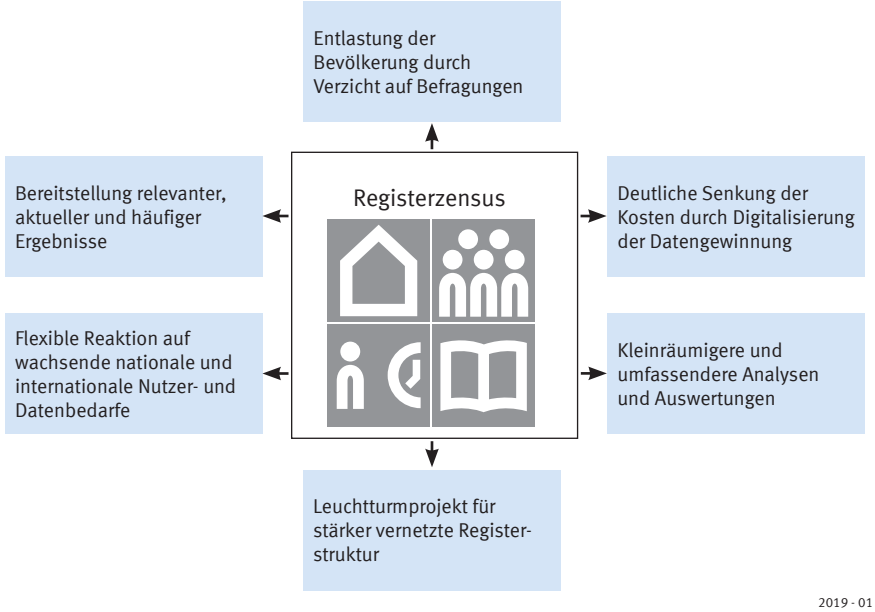
Die letzte traditionelle Volkszählung fand im früheren Bundesgebiet im Jahr 1987 statt, in der ehemaligen DDR bereits 1981. Der kommende Zensus 2021 wird hingegen – wie bereits der Zensus 2011 – in einem registergestützten Verfahren mit zusätzlichen Befragungen durchgeführt. Die Grundlage des registergestützten Verfahrens bilden mit Angaben aus primärstatistischen Erhebungen ergänzte Registerdaten. Die Angaben aus primärstatistischen Erhebungen dienen zum einen der Korrektur von Fehlern in den Registerdaten in der Statistik. Zum anderen werden sie zur Datengewinnung in Themengebieten genutzt, in denen die erforderlichen Register und Registerverknüpfungen bislang nicht zur Verfügung stehen (Dittrich, 2019).

Wie in der Verordnung (EG) Nr. 763/2008 des Europäischen Parlaments und des Rates vom 9. Juli 2008 über Volks- und Wohnungszählungen festgelegt, hat ein Zensus Daten zu folgenden Themenbereichen bereitzustellen:

- Bevölkerung (demografische Merkmale)
- Gebäude und Wohnungen
- Haushalte, Familien und deren Wohnsituation
- Arbeitsmarktbeteiligung der Bevölkerung
- Bildungsstand der Bevölkerung

Zwischen den Zensusrunden werden die Ergebnisse zum Bevölkerungsstand durch die Bevölkerungsfortschreibung gewonnen. Diese schreibt die Ergebnisse des Zensus auf Basis von Registerangaben zu Geburten, Sterbefällen, Fort- und Zuzügen monatlich fort. Die Ergebnisse

Grafik 1
Gründe für den Umstieg auf einen Registerzensus



2019 - 01 - 0486

der Gebäude- und Wohnungszählung werden jährlich im Rahmen der Fortschreibung des Wohngebäude- und Wohnungsbestands aktualisiert.

Das in den Jahren 2011 und 2021 genutzte registergestützte Verfahren kann die künftigen Anforderungen nach aktuelleren, häufigeren und zunehmend auch geokodiert[1] erforderlichen Daten in Zukunft nicht mehr erfüllen. Daher gibt es Bestrebungen in Deutschland, nach 2021 auf einen Registerzensus umzusteigen, in dem auf Befragungen verzichtet wird. ↘ Grafik 1

Dieser Beitrag stellt zunächst die vielfältigen Anforderungen dar, die ein Zensusmodell erfüllen muss, um künftig den Bedarfen der Nutzerinnen und Nutzer, aber auch den Erwartungen von Öffentlichkeit und Politik zu genügen (Kapitel 2). Im Anschluss daran werden die aktuellen internationalen Entwicklungen beschrieben (Kapitel 3) und die Grundvoraussetzungen für die Vision eines reinen Registerzensus skizziert (Kapitel 4). Aufbauend hierauf folgen erste Überlegungen dazu, wie die Datengewinnung in den fachlichen Themenbereichen des Zensus gestaltet werden könnte (Kapitel 5).

2

Anforderungen an ein künftiges Zensusmodell

Der Zensus sticht hinsichtlich seiner Bedeutung und seines Aufwands unter den Aufgaben der amtlichen Statistik heraus. Dementsprechend muss ein Zensus eine Vielzahl an Anforderungen erfüllen, die sich neben der fachlichen Eignung der Daten auch auf Fragen der Kosten, die Belastung der Bürgerinnen und Bürger sowie Abhängigkeiten von anderen Akteuren erstrecken. Im Folgenden werden die wichtigsten Anforderungen an ein zukunftsfähiges Zensusmodell im Überblick dargestellt.

2.1 Realitätsgerechte Ermittlung der Einwohnerzahlen

Mit seinem Urteil vom 19. September 2018 hat das Bundesverfassungsgericht den Rahmen dafür abgesteckt, wie ein Zensus aus verfassungsrechtlicher Sicht langfristig weiterzuentwickeln ist (2 BvF 1/15, 2 BvF 2/15; siehe Bierschenk/Leischner, 2019). Grundlegend hat das Bundesverfassungsgericht festgestellt, dass der Gesetzgeber die realitätsgerechte Ermittlung der Einwohnerzahlen sicherstellen muss. Er muss deshalb insbesondere die für deren Ermittlung erforderlichen Grundlagen schaffen. Als wichtige Grundsätze hat das Bundesverfassungsgericht die Nutzung aller ausschöpfbaren Erkenntnisquellen sowie eine grundrechtsschonende Datenerhebung auf dem Stand der statistischen Wissenschaft formuliert. Dabei muss das Verfahren geeignet sein, das zur Erfüllung der verfassungsrechtlichen Zwecke erforderliche Maß an Genauigkeit zu erreichen.

Das Bundesverfassungsgericht hat den Vorzug der Datenübermittlungen aus Verwaltungsregistern gegenüber der Primärerhebung zur Ermittlung der Zensusergebnisse unter dem Gesichtspunkt der geringeren Eingriffsintensität in die Grundrechte ausdrücklich hervorgehoben (Bierschenk/Leischner, 2019, hier: Seite 13). Darüber hinaus hat das Gericht das Rückspielverbot als grundlegend betont, sodass ein Rückfluss aus dem statistischen Fachverfahren in den Verwaltungsvollzug (weiterhin) ausgeschlossen werden muss.

2.2 Perspektivische Anforderungen der Europäischen Union

Neben dem Bedarf des nationalen Gesetzgebers müssen Zensus und Bevölkerungsstatistiken auch den Anforderungen der Europäischen Union (EU) gerecht werden. Der Datenbedarf für die politischen Programme der Europäischen Kommission unterliegt derzeit einem Wandel und verlangt insbesondere nach aktuelleren, häufiger vorliegenden und geokodierten Daten. Vor diesem Hintergrund hat die Europäische Kommission unter dem Titel European Statistics on Population (ESOP) eine EU-Rahmenverordnung für Zensus und Bevölkerungsstatistiken initiiert. Sie soll die derzeit geltenden Verordnungen zu Volks- und Wohnungszählungen [Verordnung (EG) Nr. 763/2008], zu demografischen Statistiken (Bevölkerungsstatistik) [Verordnung (EU) Nr. 1260/2013] sowie zur Wanderungs- und Migrationsstatistik [Artikel 3 der Verordnung (EG) Nr. 862/2007: Statistiken über internationale Wanderung, Wohnbevölkerung und den Erwerb der Staatsangehörigkeit] ersetzen und fortentwickeln.

1 Das heißt die Ergebnisse werden auf Ebene von geografischen Gitterzellen zur Verfügung gestellt.

Die Integration der bestehenden Verordnungen in die neue EU-Rahmenverordnung soll die Kohärenz der Ergebnisse verbessern und angesichts der steigenden Anforderungen Synergien nutzen. Kernforderung der neuen Rahmenverordnung ist die jährliche Bereitstellung geokodierter Bevölkerungszahlen ab dem Berichtsjahr 2024, die perspektivisch auch auf Untergliederungen nach sozio-ökonomischen Merkmalen ausgeweitet werden soll. In allen Themenbereichen des Zensus besteht darüber hinaus ab der Zensusrunde 2031 die Anforderung, dass Ergebnisse zumindest teilweise häufiger als alle zehn Jahre zur Verfügung stehen sollen.

2.3 Entlastung

Das aktuelle registergestützte Zensusmodell ist mit erheblichen Belastungen für die Bevölkerung verbunden: Die auskunftspflichtige Befragung von mehr als 8 Millionen Einwohnerinnen und Einwohnern und rund 25 Millionen Gebäude- und Wohnungseigentümerinnen und -eigentümern (im Zensus 2011) ist für die Betroffenen mit einem zeitlichen und bürokratischen Aufwand verbunden. Dieser könnte beim Umstieg auf ein rein registerbasiertes Verfahren ganz entfallen. Angesichts der vielfach bereits vorliegenden Verwaltungsdaten wird immer wieder kritisch hinterfragt (zum Beispiel vom Nationalen Normenkontrollrat, 2016; 2019), ob die Durchführung von Befragungen eigens für Zwecke des Zensus noch gerechtfertigt werden kann.

2.4 Senkung der finanziellen Belastungen

Wie Beispiele anderer europäischer Länder zeigen, bietet der Umstieg auf ein registerbasiertes Zensusmodell erhebliche finanzielle Einsparungspotenziale (UNECE, 2014, hier: Seite 64 ff.). Auf Basis der Kostenangaben der UNECE (2014) kann angenommen werden, dass kombinierte Modelle (wie die registergestützten Zensus 2011 und 2021) durchschnittlich um rund die Hälfte günstiger sind als Vollerhebungen. Ein ausschließlich registerbasierter Zensus erzielt demnach gegenüber kombinierten Zensusmodellen nochmals Einsparungen von über 90 %. Entsprechend weist auch der Nationale Normenkontrollrat in seinem Gutachten zum Zensusgesetz 2021 darauf hin, dass «*Zensusbefragungen durch registerbasierte Auswertungen zu einem Bruchteil der jetzt entstehenden Kosten gestaltet werden können*» (Nationaler Normenkontrollrat, 2019).

2.5 Flexibilität

Nutzerbedarfe unterliegen sowohl hinsichtlich der Merkmale als auch der interessierenden Untergliederungen einem immer schnelleren Wandel. Die derzeit für die Gewinnung der Bevölkerungszahlen zwischen den Zensusjahren genutzte Methode der Bevölkerungsfortschreibung benötigt als Basis immer erst einen Zensus mit den erforderlichen Merkmalen und Untergliederungen als Ausgangsbestand. Somit sind Änderungen der Merkmale und Untergliederungen jeweils erst zum nächsten Zensus möglich. Zudem stößt die Nutzung der Kohorten mit steigender Anzahl an Merkmalen und regionalen Untergliederungen hinsichtlich der Komplexität des Systems an methodische Grenzen. Darüber hinaus sind bei jeder Änderung aufwendige Anpassungen der IT-Verfahren nötig. Die Bevölkerungsfortschreibung ist daher künftig nicht mehr ausreichend flexibel.

2.6 Erhöhte Aktualität und Kohärenz

Für politische Entscheidungen werden die auf Basis des Zensus ermittelten Ergebnisse zunehmend häufiger als nur alle zehn Jahre benötigt. Vollerhebungen oder das in den Zensus 2011 und 2021 zum Einsatz kommende registergestützte Verfahren sind aber schon wegen des hohen Aufwands praktisch nur in größeren zeitlichen Abständen durchführbar.

Die bisherigen Befragungen zum Zensus erfordern umfangreiche Vorbereitungen und eine komplexe Organisation. Dies setzt ebenso wie die Dauer der Feldarbeit der angestrebten Aktualität bei der Bereitstellung der Ergebnisse Grenzen. Angesichts der Bedeutung der auf den Zensusdaten basierenden Entscheidungen erwarten die Nutzerinnen und Nutzer jedoch aktuelle Daten, die nur ein Registerzensus liefern kann.

Da die in der Haushaltsstichprobe des Zensus ermittelten Korrekturen der Melderegisterdaten nicht an die Melderegister zurückgespielt werden dürfen, weichen die Auszählungen der Melderegister und der amtlichen Bevölkerungszahl voneinander ab. Diese Abweichungen werden umso größer, je länger der letzte Zensus zurückliegt; sie führen zu regelmäßigen Zeitreihenbrüchen in den Bevölkerungszahlen. Dies verunsichert die Entscheidungsträgerinnen und Entscheidungsträger auf den verschiedenen Ebenen und birgt erhebliches Konflikt-

potenzial. Das hat sich nicht zuletzt an den zahlreichen gerichtlichen Auseinandersetzungen im Zusammenhang mit der Feststellung der amtlichen Einwohnerzahlen nach dem Zensus 2011 gezeigt.

2.7 Bereitstellung kleinräumiger und geokodierter Daten

Auf kommunaler, regionaler, nationaler und europäischer Ebene setzen immer mehr Fragestellungen voraus, dass die Ergebnisse auf kleinräumigen Gliederungsebenen (zum Beispiel Blockseiten) oder auf Ebene geografischer Gitterzellen bereitgestellt werden. Der Ergebnisgliederung unterhalb der Gemeindeebene sind jedoch von vornherein enge Grenzen gesetzt, wenn bei der Datengewinnung wie bisher Stichproben genutzt werden.

Gitterzellen haben den Vorteil, dass es sich bei ihnen um räumliche Einheiten handelt, die sowohl international als auch im Zeitablauf miteinander vergleichbar sind: Sie haben im Gegensatz zu administrativen Einheiten wie Gemeinden überall die konstante, gleiche Größe und Lage. Sie ermöglichen eine Auswertung von statistischen Ergebnissen für beliebige (auch nicht administrative) Gebietseinheiten (wie Wahlkreise, Schulsprengel oder Überschwemmungsgebiete), weil die Daten mithilfe geografischer Informationssysteme solchen Gebietseinheiten zugeordnet werden können (Verschneidung von Gitterzellen mit beliebigen Polygonen). Des Weiteren können externe Informationen an vorhandene Daten über Gitterzellen angespielt werden, die zwar kleinräumig auf Gitterzellenebene, aber nicht auf Anschriftenebene vorliegen (zum Beispiel Breitbandverfügbarkeit).

2.8 Berücksichtigung weiterer Rahmenbedingungen

Empfehlungen der Wissenschaft

Der Rat für Sozial- und Wirtschaftsdaten (RatSWD) hat im Jahr 2016 umfangreiche Empfehlungen für künftige Zensusrunden erarbeitet (RatSWD, 2016). Darin spricht er sich kurzfristig (mit Blick auf die Zensusrunde 2021) dafür aus, die Melderegister zu optimieren. Langfristiges Ziel soll die Einbindung der Melderegister in eine verknüpfbare Registerinfrastruktur sein, in der die Daten der Melderegister anhand der übrigen Register plausibilisiert werden können. Ebenso wird der Aufbau eines Gebäude- und Wohnungsregisters angeregt.

Digitalisierung der Verwaltung und Modernisierung der Registerlandschaft

Ein künftiger registerbasierter Zensus kann stark von einer Modernisierung der Registerlandschaft im Kontext des Ausbaus digitaler Verwaltungsleistungen profitieren. Der aktuelle Diskussionsstand zur allgemeinen Register- und Verwaltungsmodernisierung lässt dabei allerdings noch keine Rückschlüsse darauf zu, wie sich der Registerzensus in die künftige Registerlandschaft einfügen wird.

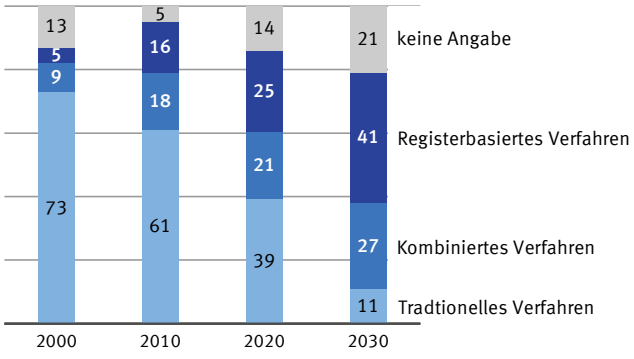
3

Internationale Entwicklungen

Auch international zeigt sich ein Trend hin zur Nutzung registerbasierter Zensusverfahren. Setzten bei der Zensusrunde 2000 noch 73 % der damals 55 Mitgliedstaaten im Gebiet der Wirtschaftskommission der Vereinten Nationen für Europa (UNECE) ein traditionelles Zensusmodell mit primärstatistischer Vollerhebung ein, so sind dies in der Zensusrunde um das Jahr 2020 nur noch 39 %. Im gleichen Zeitraum stieg der Anteil der Mitgliedstaaten, die ein ausschließlich registerbasiertes Verfahren nutzen, von 5 % im Jahr 2000 auf 25 % im Jahr 2020 (Valente, 2015). Eine Befragung zu den aktuellen Entwicklungsplänen in Hinblick auf die Zensusrunde um das Jahr 2030 lässt erwarten, dass sich dieser Trend noch deutlich beschleunigen wird (Valente, 2018). So wird der Anteil der Staaten mit traditioneller Vollerhebung auf nur noch 11 % der Länder sinken, während ein rein registerbasierter Zensus dann in voraussichtlich 41 % der UNECE-Staaten durchgeführt wird. Gleichzeitig steigt der Anteil der Mitgliedstaaten mit registergestützten (kombinierten) Verfahren weiter auf 27 % (Zensusrunde 2020: 21 %). ↘ Grafik 2

Diese Entwicklung findet ihre Entsprechung in den geänderten Schwerpunkten der internationalen Empfehlungen über Bevölkerungs- sowie Gebäude- und Wohnungszählungen. So wurden im Zuge der Vorbereitung auf die

Grafik 2
Wandel der Zensusmodelle in den Mitgliedstaaten der UNECE
Anteile an den UNECE-Mitgliedstaaten in %



UNECE: Wirtschaftskommission für Europa der Vereinten Nationen
Quellen: Valente (2015); Valente (2018); eigene Berechnungen
2019 - 01 - 0487

Zensusrunde 2020 dezidierte Leitlinien für den Umstieg auf registerbasierte und registergestützte Zensusmodelle erarbeitet und von der Konferenz Europäischer Statistiker im Juni 2018 verabschiedet (UNECE, 2018).

Auf die Frage nach den darüber hinaus bestehenden wichtigsten Änderungsbereichen im Hinblick auf die Zeit nach der Zensusrunde 2020 antwortet gut die Hälfte der Mitgliedstaaten der UNECE, die Periodizität verkürzen zu wollen. Gut ein Drittel nennt die Erhöhung der geografischen Gliederungstiefe, zum Beispiel durch Veröffentlichung von Ergebnissen auf der Ebene geografischer Gitterzellen. ↘ Grafik 3

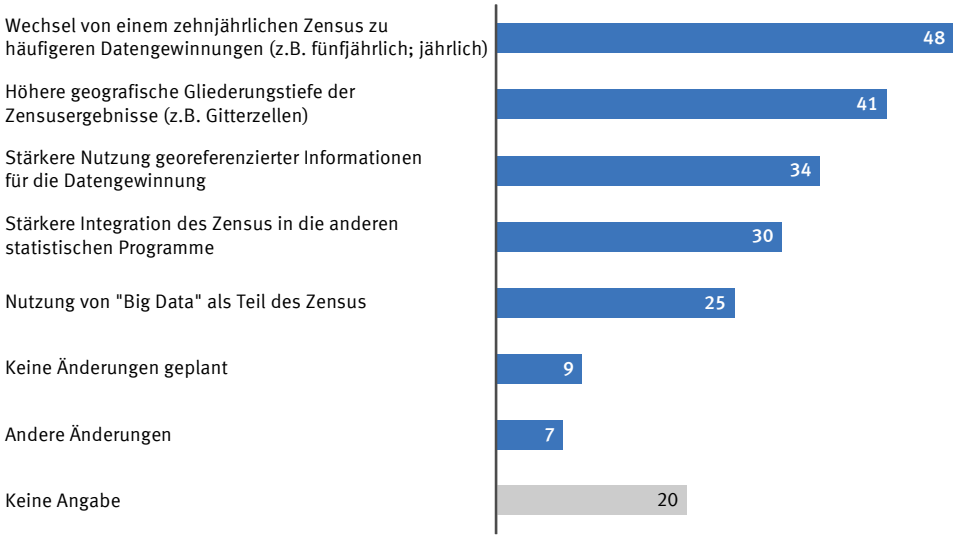
4

Grundvoraussetzungen

Mit der Vision eines Registerzensus geht das Ziel einher, die in Kapitel 2 aufgezeigten neuen Herausforderungen und Anforderungen an Zensus und Bevölkerungsstatistiken zu erfüllen. Die Grundidee eines Registerzensus ist es, die Ergebnisse des Zensus und der Bevölkerungsstatistik basierend auf verknüpften Verwaltungs- und Statistikregistern zu ermitteln und auf den Einsatz primärstatistischer Befragungen zu verzichten.

Für den Umstieg vom derzeitigen registergestützten System auf einen Registerzensus müssen einige Grundvoraussetzungen erfüllt sein. Hier sind insbesondere die Datengrundlagen, die Datenqualität sowie Möglichkeiten zur Registerverknüpfung zu nennen.

Grafik 3
Weitere Veränderungsthemen für den Zensus nach 2021
Anteile an den UNECE-Mitgliedstaaten in %



Mehrfachnennungen möglich. UNECE: Wirtschaftskommission für Europa der Vereinten Nationen.
Quelle: Valente (2018)
2019 - 01 - 0488

4.1 Datengrundlagen

Ein registerbasierter Zensus greift als Datengrundlage zur Gewinnung der statistischen Ergebnisse ausschließlich auf Registerdaten zurück. Da die erforderlichen Informationen in unterschiedlichen Verwaltungs- oder Statistikregistern vorliegen, müssen unterschiedliche Quellen herangezogen werden. Einige Themenbereiche können mit bereits bestehenden Registern abgedeckt werden, für die der Zugang zu den erforderlichen Daten für die amtliche Statistik rechtlich, technisch und organisatorisch geregelt werden müsste. In Bereichen, in denen derzeit noch keine Register vorhanden sind, müssten darüber hinaus neue Register aufgebaut werden, beispielsweise ein Gebäude- und Wohnungsregister oder ein Bildungsregister. Des Weiteren kann es erforderlich sein, bereits vorhandene Register zu ertüchtigen, also Anpassungen vorzunehmen, die für eine statistische Nutzung erforderlich sind. Erste Überlegungen zu den Möglichkeiten der registerbasierten Datengewinnung werden in Kapitel 5 gegliedert nach den einzelnen fachlichen Themenbereichen dargestellt.

4.2 Datenqualität

Qualitativ hochwertige Ergebnisse und damit die Einhaltung eines bundesweit vergleichbaren Qualitätsstandards sind zentrale Anforderungen an einen Registerzensus. Diesem Anspruch kann nur Rechnung getragen werden, wenn bereits die Datengrundlage eine hohe Datenqualität aufweist. Dies erleichtert den statistischen Aufbereitungsprozess, trägt dazu bei, notwendige Korrekturen zu verringern, und unterstützt eine schnelle und aktuelle Ergebnisbereitstellung. Dies gilt umso mehr, da – anders als beim Zensus 2011 oder 2021 – keine Daten aus Befragungen zur Qualitätssicherung der verwendeten Registerinformationen genutzt werden sollen. Gleichwohl gilt es zu berücksichtigen, dass Verwaltungsregister originär für Verwaltungszwecke geführt werden. Entsprechend der zugrunde liegenden Rechtsvorschriften kann die Verwaltung teilweise andere Anforderungen an die Qualität ihrer Daten stellen oder differierende Konzepte abbilden. Beispielsweise weichen die Definitionen von Arbeitslosigkeit nach dem Sozialgesetzbuch und die dem Zensus zugrunde liegende Definition von Erwerbslosigkeit der Internationalen Arbeitsorganisation (ILO) voneinander ab. Daher gehen eine qualitativ hochwertige Datengrundlage und die Möglichkeit zu statistischen Anpassungen beziehungsweise Qualitätssicherungsmaßnahmen im Rahmen der statistischen Aufbereitung Hand in Hand bei der Nutzung von Verwaltungsdaten zur Statistikerstellung. Zu berücksichtigen ist hierbei allerdings, dass Bereinigungen, die durch die Statistik erfolgen, nicht an die Verwaltung zurückfließen dürfen. Sie können daher keinen Beitrag zu Verbesserungen in den zugrunde liegenden Verwaltungsregistern leisten.

4.3 Möglichkeiten zur Registerverknüpfung

Da die zur Datengewinnung und zur statistischen Aufbereitung erforderlichen Registerinformationen nicht einem einzelnen Register entnommen werden können, sind Möglichkeiten der Registerverknüpfung in einem Registerzensus von zentraler Bedeutung. Die Informationen liegen in unterschiedlichen Registern vor. Sie müssen zur registerbasierten Gewinnung der erforderlichen Merkmale und zur Qualitätssicherung im statistischen Aufbereitungsprozess auf Personenebene verknüpft werden. Prinzipiell sind hierfür zwei Verknüpfungsverfahren oder deren Kombination denkbar:

1) Verknüpfung über (gegebenenfalls verschlüsselte) konstante Identifikatoren[2]
2) Verknüpfung über (gegebenenfalls pseudonymisierte) Personenmerkmale, wie Name, Geburtsdatum, Geburtsort und -staat oder Staatsangehörigkeit, wobei die Belange des Datenschutzes zu gewährleisten sind.
3) Kombination aus konstanten Identifikatoren und identifizierenden Personenmerkmalen, sofern durch mindestens ein Register eine Überschneidung der Verknüpfungskriterien gewährleistet ist.

Aufgrund der Eindeutigkeit und Unveränderlichkeit sind (verschlüsselte) konstante Identifikatoren am besten geeignet, um Registerdaten miteinander zu verknüpfen. Neben der Verknüpfung von Personenangaben aus

2 Ein konstanter Identifikator ist definiert als eine über die Zeit hinweg stabile Zeichenfolge, die einer Person einmalig zugewiesen wird. Beispiele für konstante Identifikatoren stellen in Deutschland die Steuer-Identifikationsnummer oder die Sozialversicherungsnummer dar. Beide dürfen allerdings nur für den definierten Verwaltungszweck genutzt werden.

verschiedenen Registern ist die Zusammenführung von Informationen zu Personen mit Gebäude- und Wohnungsangaben zur Gewinnung von Merkmalen zu Haushalten und Familien unerlässlich. Hierfür ist die Einführung einer Gebäude- und Wohnungsnummer in die Melderegister erforderlich. Im Gegensatz zu den unterschiedlichen Möglichkeiten, die für die Verknüpfung von Personenangaben bestehen, ist bei Gebäuden und Wohnungen nur die Nutzung konstanter Identifikatoren möglich. Konstante Identifikatoren sind erforderlich, da in Gebäude- und Wohnungsdatensätzen und Personendatensätzen in den Melderegistern keine identischen Merkmale enthalten sind, die für eine Verknüpfung genutzt werden könnten (siehe Abschnitte 5.2 und 5.3).

Grundsätzlich sollte bei der Verknüpfung von Registern ein Vorgehen gewählt werden, das den verfassungsrechtlichen (Grundrecht auf informationelle Selbstbestimmung) und datenschutzrechtlichen Vorgaben genügt. Es muss geeignet, angemessen, sicher und zugleich aufwandsschonend sein.

5

Erste Überlegungen zur Datengewinnung in den einzelnen Themenbereichen

5.1 Bevölkerung

Neben der Bereitstellung realitätsgerechter demografischer Basisdaten bildet der Zensus die Grundlage für die Bevölkerungsfortschreibung, die derzeit als Methode zur Gewinnung der Bevölkerungszahlen zwischen den Zensusjahren genutzt wird. In einem Registerzensus sollen die Ermittlung der Bevölkerungszahlen aus dem Zensus und die darauf aufbauende Bevölkerungsfortschreibung integriert werden. Der Registerzensus kann dann sowohl monatliche als auch jährliche Bevölkerungszahlen liefern.

Wesentliche Datengrundlage für eine registerbasierte Ermittlung der Bevölkerungszahlen sollten, wie auch im derzeitigen System aus Zensus und Bevölkerungsfortschreibung, die lokalen Melderegister sein. Ein statistikinterner bevölkerungsstatistischer Datenbestand könnte durch anlassbezogene Datenlieferungen der lokalen Meldebehörden (beispielsweise bei einem Umzug) an die amtliche Statistik gepflegt und so zur Ermittlung der Bevölkerungszahlen genutzt werden.

Um realitätsgerechte Bevölkerungszahlen ermitteln und bundesweit vergleichbare Qualitätsstandards garantieren zu können, müssen die Daten der Melderegister in der Statistik auf Unstimmigkeiten hin überprüft und Unplausibilitäten bereinigt werden. Der Zensustest 2001 und der Zensus 2011 haben gezeigt, dass die Daten der Melderegister aufgrund von Über- und Untererfassungen unter anderem infolge von Verletzungen der Meldepflicht durch Bürgerinnen und Bürger zum Teil fehlerhaft waren. Daher wird auch im Zensus 2021 eine Haushaltsstichprobe zur statistischen Korrektur dieser Fehler in den Melderegistern durchgeführt. Im Registerzensus soll nicht mehr auf primärstatistische Befragungen zurückgegriffen werden, um Über- und Untererfassungen in den Melderegistern festzustellen. Dies hat zur Folge, dass neue Instrumente zur Prüfung der Daten entwickelt werden müssen. Eine Option zur registerbasierten Ermittlung von Fehlerfassungen liefert der sogenannte Lebenszeichenansatz, der durch Registerabgleiche potenzielle Über- und Untererfassungen in den Melderegistern feststellt. Neben den Daten aus dem Meldewesen werden dabei weitere Register, sogenannte Vergleichsregister, für einen Registerabgleich herangezogen. Beispiele für Vergleichsregister könnten unter anderem das Zentrale Fahrzeugregister, die Datenbestände der Sozialversicherungen oder der Finanzverwaltung sein. Bei Personen, die neben dem Melderegister in keinem weiteren Vergleichsregister geführt werden, kann davon ausgegangen werden, dass es sich um Übererfassungen in den Melderegistern handelt.[3] Aus verfassungs- und datenschutzrechtlichen Gründen müssen die Register sorgfältig ausgewählt werden. So können beispielsweise keine Register als Vergleichsregister herangezogen werden, bei denen aufgrund der Eintragung der Person auf sensible Informationen geschlossen werden kann. Die Umsetzung des Lebenszeichenansatzes setzt eine verlässliche und datenschutzkonforme Registerverknüpfung voraus.

3 Ob der Lebenszeichenansatz auch zur Ermittlung von Untererfassungen geeignet ist, bleibt zu prüfen.

5.2 Gebäude und Wohnungen

Neben der Ermittlung von demografischen und sozioökonomischen Daten zur Bevölkerung ist die zweite Hauptaufgabe eines Zensus, Informationen zu Gebäuden und Wohnungen sowie zur Wohnsituation der Bevölkerung bereitzustellen. Bisher werden Zensusergebnisse zu Gebäuden und Wohnungen alle zehn Jahre im Rahmen einer Befragung von allen Eigentümerinnen und Eigentümern gewonnen (Gebäude- und Wohnungszählung). Eine registerbasierte Ermittlung dieser Daten eröffnet jedoch die Möglichkeit, Ergebnisse weniger belastungsintensiv zu ermitteln und auch in kürzeren Abständen zu veröffentlichen.

Bisher gibt es in Deutschland keine flächendeckenden Register mit den erforderlichen Merkmalen. Daher werden vonseiten der Statistik unterschiedliche Möglichkeiten geprüft, wie Gebäude- und Wohnungsangaben dennoch registerbasiert gewonnen werden können. Im Fokus der Überlegungen stand dabei zunächst die Grundstücksdatenbank, die von der Finanzverwaltung im Zuge der Grundsteuerreform neu aufgebaut wird. Hintergrund dafür ist das Urteil des Bundesverfassungsgerichts vom 10. April 2018[4] , das die Ermittlung der Grundsteuer auf der Basis von veralteten Einheitswerten für verfassungswidrig erklärt hat und eine Neuregelung des Verfahrens zur Bewertung von Grundbesitz verlangt.

In der Grundstücksdatenbank sollen künftig die Merkmale geführt werden, die für die Berechnung der Grundsteuer erforderlich sind. Entscheidend für eine Bewertung aus Sicht eines Registerzensus ist, ob sowohl Angaben zu Gebäuden und Wohnungen als auch die für Zensuszwecke erforderlichen Merkmale flächendeckend enthalten sein werden. Nach dem aktuellen Stand der Diskussion[5] sehen die geplanten Gesetzesvorhaben zur Grundsteuerreform jedoch eine Öffnungsklausel vor, die es den Ländern ermöglicht, von dem im Grundsteuergesetz vorgesehenen Modell zur Berechnung der Grundsteuer abzuweichen und die Ermittlung der Steuer jeweils selbst zu regeln. Da bereits einige Länder angekündigt haben, von dieser Öffnungsklausel Gebrauch zu machen, ist eine bundeseinheitliche Führung der Grundstücksdatenbank nicht gewährleistet. Aus diesem Grund wird es voraussichtlich nicht möglich sein, alle im Registerzensus erforderlichen Gebäude- und Wohnungsmerkmale flächendeckend aus der Grundstücksdatenbank zu gewinnen. Daher wäre es für eine vollständig registerbasierte Ermittlung der Gebäude- und Wohnungsmerkmale erforderlich, ein separates Gebäude- und Wohnungsregister als Verwaltungsregister aufzubauen.

Ein solches Gebäude- und Wohnungsregister müsste alle Gebäude mit Wohnraum, sogenannte bewohnte Unterkünfte[6] und die darin befindlichen Wohnungen mit den entsprechenden Merkmalen führen. Darüber hinaus ist es naheliegend, ein solches Register auch für Verwaltungszwecke zu nutzen. So könnte es zum Beispiel die Stadt- und Regionalplanung sowie das Monitoring von Entwicklungen auf Wohnungs- und Immobilienmärkten unterstützen. Zusätzlich wäre es auch denkbar, über ein solches Gebäude- und Wohnungsregister eine in der Statistik bestehende Datenlücke zu schließen, indem Informationen zum Bestand der gewerblich und administrativ genutzten Gebäude zur Verfügung gestellt werden könnten.

Für die eindeutige Identifizierung von Gebäuden und Wohnungen und für die regelmäßige Aktualisierung der Angaben wäre es darüber hinaus erforderlich, Gebäude- und Wohnungsnummern in einem solchen Gebäude- und Wohnungsregister zu führen. Diese sollten nach einer einheitlichen Systematik vergeben werden und nach Möglichkeit auch die Lage einer Wohnung im Gebäude kennzeichnen. Gebäude- und Wohnungsnummern werden auch benötigt, um registerbasiert Haushalte zu bilden und diese Wohnungen zuordnen zu können (siehe Abschnitt 5.3).

Im Zuge der Konzeption eines möglichen Gebäude- und Wohnungsregisters sind drei Aspekte besonders zu berücksichtigen:

- Ermittlung der Angaben für den Gebäude- und Wohnungsbestand: Es ist festzulegen, wie die Angaben für bestehende Gebäude und Wohnungen in Deutschland ermittelt werden können.
- Aktualisierung und Pflege des Registers: Um die Aktualität und Datenqualität des Registers zu gewährleisten, müssen Prozesse für eine regelmäßige Pflege der Angaben etabliert werden.

4 BVerfG, Urteil des Ersten Senats vom 10. April 2018 -1 BvL 11/14, 1 BvR 889/12, 1 BvR 639/11, 1 BvL 1/15, 1 BvL 12/14 -Rn. (1 -181), http://www.bverfg.de/e/ls20180410_1bvl001114.html

5 Stand Juli 2019.

6 Behelfsmäßige Bauten, wie bewohnte Bauwagen, Wohncontainer oder Jagdhütten, die im Zensus ebenfalls erhoben werden müssen, wenn sie bewohnt sind.

› Bereitstellung von Schnittstellen: Schließlich müssen Schnittstellen eine regelmäßige Übermittlung der Gebäude- und Wohnungsangaben an die Statistik beziehungsweise eine Nutzung des Registers für Verwaltungszwecke ermöglichen.

Derzeit werden vonseiten der Statistik zwar erste Überlegungen zu diesen Themenkomplexen angestellt, diese müssen aber von der Verwaltungsstelle, die das neue Gebäude- und Wohnungsregister führen müsste, vertieft und anschließend umgesetzt werden.

5.3 Haushalte und Familien

Informationen zur Größe und Struktur von Haushalten und Familien sind essenzielle Grundlagen für die Einschätzung der Lebensverhältnisse der Bevölkerung in Deutschland. In Verbindung mit Wohnungen informieren sie darüber hinaus über die Wohnverhältnisse von Haushalten und Familien und unterstützen damit die aktuelle und künftige Ermittlung des Wohnungsbedarfs.

In einem Registerzensus können aus methodischen Gründen grundsätzlich nur Angaben zu Wohnhaushalten ermittelt werden. Einen Wohnhaushalt bilden alle Personen, die gemeinsam in einer Wohnung leben, unabhängig davon, ob sie auch ihre Einnahmen und Ausgaben miteinander bestreiten. Ohne Befragungen kann aus Registern nur diese Haushaltsform abgeleitet werden.

Angaben zu Haushalten, Familien und den entsprechenden Merkmalen können nicht direkt aus einem Register gewonnen werden, sondern müssen mithilfe von Daten aus verschiedenen Registern abgeleitet werden. Die meisten Angaben sind in den Melderegistern enthalten. Allerdings fehlt dort eine entscheidende Komponente. Um alle Personen zu Haushalten zusammenfassen und konkreten Wohnungen zuordnen zu können, ist es erforderlich, jeder Person die Nummern des Gebäudes und der Wohnung zuzuordnen, in der sie lebt. Die Einführung der Gebäude- und Wohnungsnummern ist Grundvoraussetzung für die registerbasierte Bildung von Haushalten. Allerdings können weitere Angaben in den Melderegistern die Bildung von Haushalten unterstützen und zur Qualitätssicherung der Ergebnisse beitragen. So lässt sich beispielsweise – analog zur Vorgehensweise im Zensus 2011 und 2021 – durch sogenannte beigeschriebene Personen an einer Anschrift (Verknüpfungen zwischen Ehepartnerinnen und -partnern sowie zwischen Eltern und minderjährigen Kindern) ableiten, welche Personen in einem Haushalt leben. Auch das gleiche Einzugs- oder Anmeldedatum oder weitere Merkmale aus dem Melderegister geben Hinweise auf Haushaltszusammenhänge (siehe auch Massih-Tehrani/Reichert, 2019).

Entsprechend der im Zensus genutzten Definitionen ist zu berücksichtigen, dass es in Deutschland auch Personen gibt, die keinen eigenen Haushalt führen. Dies sind Personen, die in Anstalten (zum Beispiel Justizvollzugsanstalten) oder in Gemeinschaftsunterkünften (zum Beispiel Kinder- oder Pflegeheimen) leben. Um solche Personen in der Statistik identifizieren zu können, ist es erforderlich, einen dauerhaften Datenbestand aufzubauen, der Informationen zu Anstalten und Gemeinschaftsunterkünften enthält (zum Beispiel ein Einrichtungsregister). In Verbindung mit den Melderegisterangaben können so Personen identifiziert werden, die in solchen Einrichtungen gemeldet sind, keinen eigenen Haushalt führen und daher von der statistischen Haushaltsbildung ausgeschlossen werden müssen.

5.4 Arbeitsmarktbeteiligung

Neben demografischen Merkmalen sind auch eine Reihe sozio-ökonomischer Merkmale wichtige Bestandteile eines Zensus, unter anderem Merkmale zur Arbeitsmarktbeteiligung (UNECE, 2015). Zuletzt haben die Leiterinnen und Leiter der europäischen statistischen Ämter im Rahmen des sogenannten Budapester Memorandums (Eurostat, 2017) die wachsende Bedeutung sozio-ökonomischer Merkmale in Bevölkerungsdaten betont. Es ist daher davon auszugehen, dass auch solche Angaben perspektivisch häufiger als zehnjährlich sowie in geokodierter Form bereitgestellt werden müssen.[7]

Grundlegend für die Datenanforderungen im Bereich Arbeitsmarkt ist das Merkmal „Derzeitiger Erwerbsstatus“. Es orientiert sich an den Konzepten der Internationalen Arbeitsorganisation (ILO) und unterteilt die Gesamtbevölkerung in die drei überschneidungsfreien Gruppen der Erwerbstätigen, Erwerbslosen und Nicht-

7 Die Durchführungsverordnung (EU) 2018/1799 schreibt bereits für den Zensus 2021 die Lieferung von Informationen zur Zahl der Erwerbstätigen auf Ebene geokodierter Gitterzellen fest, sieht allerdings eine Ausnahmeregelung vor.

erwerbspersonen. Abgesehen von einer Untergliederung der Nichterwerbspersonen in Personen unter 15 Jahren, Personen, die von Ruhegehalt oder Kapitalerträgen leben, Schülerinnen und Schüler sowie Studierende, stellen alle weiteren Merkmale im Bereich Arbeitsmarkt Untergliederungen der Gruppe der Erwerbstätigen dar. Für die Erwerbstätigen sind neben dem Merkmal „Derzeitiger Erwerbsstatus“ die Stellung im Beruf, der ausgeübte Beruf, der Wirtschaftszweig und der Arbeitsort zu ermitteln, wobei jeweils die international vereinbarten Klassifikationen maßgeblich sind.

Zur registerbasierten Gewinnung der Daten zur Arbeitsmarktbeteiligung kann voraussichtlich auf überwiegend vorhandene – wenn auch bislang nicht harmonisierte – Register zurückgegriffen werden. Eine besondere Herausforderung ist die Gewinnung von Daten unter anderem für die Gruppen der Selbstständigen und der mithelfenden Familienangehörigen. Darüber hinaus müsste eine Vorgehensweise dafür entwickelt werden, wie die Angaben aus unterschiedlichen Quellen, die sich teilweise in den Definitionen und den zugrunde liegenden Konzepten unterscheiden, in einen konsistenten Datensatz integriert werden können. Mögliche Register für die Ermittlung der erforderlichen Informationen sind insbesondere die Daten der Bundesagentur für Arbeit und gegebenenfalls weiterer Sozialversicherungsträger, die Daten der öffentlichen Arbeitgeber, die Daten der Finanzverwaltung, die Register von Stellen der berufsständischen Selbstverwaltung sowie ein noch aufzubauendes Bildungsregister (siehe Abschnitt 5.5). Die Umsetzungsmöglichkeiten wären in einem nächsten Schritt gemeinsam mit den jeweils registerführenden Stellen näher zu untersuchen.

Grundvoraussetzung für die registerbasierte Ableitung der Angaben zur Arbeitsmarktbeteiligung ist insbesondere ein zuverlässiges, automatisiertes Verfahren zur Verknüpfung der Registerbestände.

5.5 Bildungsstand

Neben den Daten zur Arbeitsmarktbeteiligung werden auch kleinräumige Angaben zur Bildungsbeteiligung immer wichtiger. Insbesondere in Verbindung mit demografischen Merkmalen besteht ein wachsender Bedarf von Nutzerinnen und Nutzern nach diesen Ergebnissen. Im Zensus werden allerdings nur einige grundlegende Angaben zum Bildungsstand der Bevölkerung (unter anderem der höchste Bildungsabschluss) ermittelt.

Derzeit existieren in Deutschland allerdings keine Register, die die erforderlichen Angaben zum Bildungsstand für alle in Deutschland lebenden Personen bereitstellen. Daher prüft die amtliche Statistik, welche Möglichkeiten es gibt, diese Angaben registerbasiert zu gewinnen. Bestehende Datenlücken könnten durch ein Bildungsregister geschlossen werden. Ziel eines solchen Bildungsregisters wäre, Daten aus dem Bildungsbereich, insbesondere Angaben zu Bildungsverläufen, zu erfassen und statistisch auswerten zu können. Vor diesem Hintergrund könnte ein Bildungsregister auch dazu genutzt werden, die oben genannten Zensusmerkmale zu ermitteln. Bildet ein Bildungsregister die Bildungsverläufe bereichsübergreifend von der Schulbildung bis zu den beruflichen und den Hochschulabschlüssen ab, ist es möglich, aus diesen Angaben den höchsten Bildungsabschluss zu ermitteln. Die erforderlichen Daten, die zum Aufbau eines solchen Bildungsregisters verwendet werden sollen, stammen aus bestehenden Bildungsstatistiken[8], die als Vollerhebungen unter Nutzung von Verwaltungsdaten durchgeführt werden. Allerdings müssten dafür Einzeldatensätze erfasst und an das Bildungsregister übermittelt werden. Die Schulstatistik als koordinierte Länderstatistik ermittelt jedoch bisher noch nicht flächendeckend Individualdaten. Dies müsste für den Aufbau eines Bildungsregisters perspektivisch geändert werden. Hierzu wäre vor allem die Schaffung der rechtlichen Grundlagen auf Bundes- und Länderebene erforderlich.

Die bisherigen Erläuterungen machen deutlich, dass im aufzubauenden Bildungsregister zunächst nur Personen erfasst werden könnten, die sich zu den jeweiligen Erhebungszeitpunkten in Bildungsgängen befinden. Das Bildungsregister würde also kontinuierlich befüllt werden. Personen, die ihren Bildungsweg bereits abgeschlossen haben, würden nicht erfasst. Daher ist zu untersuchen, welche weiteren Datenquellen für eine Ermittlung der Merkmale zum Bildungsstand in einem Registerzensus genutzt werden könnten, um übergangsweise den Bildungsstand von Personen, die sich nicht (mehr) in Schule, Studium oder Ausbildung befinden, mittels Hochrechnung zu ermitteln.

8 Das sind die Schulstatistik, die Berufsbildungsstatistik, die Statistik nach § 17 Berufsqualifikationsgesetz (zur Ermittlung von im Ausland erworbenen Bildungsabschlüssen) und die Hochschulstatistik.

Um das aufzubauende Bildungsregister pflegen und die Merkmale zum Bildungsstand ermitteln zu können, müssen die Angaben aus den unterschiedlichen Datenquellen in einem zuverlässigen, automatisierten Verfahren miteinander verknüpft werden. Dafür bestehende Möglichkeiten werden unter Berücksichtigung der fachlichen und technischen Anforderungen sowie vor dem Hintergrund der Wahrung des Datenschutzes und verfassungsrechtlicher Vorgaben derzeit geprüft (siehe Kapitel 4).

Erste vorbereitende Arbeiten zur Prüfung, ob und wie die Einführung eines Bildungsregisters möglich ist, haben unter Federführung des Bundesministeriums für Bildung und Forschung bereits begonnen.

6

Fazit

Sich wandelnde Anforderungen der Nutzerinnen und Nutzer sowie absehbare Lieferverpflichtungen aus der künftigen EU-Rahmenverordnung über Zensus und Bevölkerungsstatistiken erfordern nach dem Zensus 2021 eine grundlegende Neukonzeption des Zensusmodells in Deutschland. Nach den sich abzeichnenden Lieferverpflichtungen der geplanten EU-Rahmenverordnung sind Bevölkerungszahlen voraussichtlich ab dem Berichtsjahr 2024 jährlich geokodiert an die EU zu liefern. Geokodierte Ergebnisse zu den übrigen Zensusthemen sind voraussichtlich ab dem Berichtsjahr 2031 in teilweise kürzerer Periodizität als heute bereitzustellen. Langfristig werden die geänderten Anforderungen dabei nur mit einem registerbasierten Modell umzusetzen sein. Dieser Beitrag skizzierte eine Methode eines Registerzensus zur Gewinnung kleinräumiger Bevölkerungszahlen sowie der weiteren Zensusergebnisse zu den Themenbereichen Demografie, Gebäude und Wohnungen, Haushalte, Familien und deren Wohnsituation, Arbeitsmarktbeteiligung und Bildungsstand der Bevölkerung. Diese Methode ermöglicht es, die kommenden nationalen und internationalen Anforderungen zu bedienen.

Die Voraussetzungen für den Umstieg auf einen Registerzensus sind derzeit noch nicht in allen Bereichen geschaffen. Für einige Themengebiete muss die für eine registerbasierte Datengewinnung erforderliche Registerinfrastruktur aufgebaut werden. Teilweise sind bestehende Register zu ertüchtigen. Zudem ist ein datenschutzkonformes Verfahren zur Registerverknüpfung zu entwickeln und einzuführen. Der hierfür notwendige politische Willensbildungsprozess ist noch nicht abgeschlossen. Allerdings kann der Registerzensus hier von der geplanten Modernisierung der Registerlandschaft im Kontext des Ausbaus digitaler Verwaltungsleistungen profitieren. Gleichzeitig kann die amtliche Statistik ihre Erfahrungen in diesem Bereich einbringen. Der Aufbau und die Ertüchtigung von Registern sowie die Schaffung der Voraussetzungen für eine Verknüpfung von Registern sind komplex und erfordern in der operativen Umsetzung das Zusammenspiel verschiedener Akteure in der deutschen Verwaltungslandschaft.

Angesichts der Komplexität der notwendigen fachlich-technischen Grundlagen, des erforderlichen zeitlichen Vorlaufs für die Schaffung der rechtlichen Voraussetzungen sowie der im Rahmen der Einführung durchzuführenden Tests sollte mit den notwendigen Arbeiten sofort begonnen werden. Das Statistische Bundesamt hat daher bereits im Frühjahr 2018 den Stabsbereich Projekt Registerzensus eingerichtet, um mit den Vorbereitungen zu beginnen. Die Amtsleiterkonferenz der Statistischen Ämter des Bundes und der Länder hat im Herbst 2018 eine gemeinsame Arbeitsgruppe gegründet und diese beauftragt, einen Registerzensus zu konzipieren, zu prüfen und gegebenenfalls vorzubereiten. Für das Gelingen eines derart komplexen Projekts ist darüber hinaus eine frühzeitige Einbindung der Kommunen und der registerführenden Behörden sowie der Datennutzerinnen und Datennutzer, des Datenschutzes und der Wissenschaft erforderlich.

Der Umstieg auf einen Registerzensus kann nicht nur die Belastung der Bürgerinnen und Bürger sowie die Kosten reduzieren, sondern auch dazu beitragen, das Informationsangebot nachhaltig auszubauen: Bevölkerungszahlen und Zensusergebnisse stehen häufiger, aktueller und regional tiefer gegliedert für Entscheidungen und als Planungsgrundlage zur Verfügung. Die Geokodierung der Daten bietet darüber hinaus die Chance, regionale Daten passgenau bereitzustellen und mit vorhandenen Regionalinformationen zu kombinieren.

LITERATURVERZEICHNIS

Bierschenk, Michaela/Leischner, Sonja. *Zur Verfassungsmäßigkeit der Vorschriften über den Zensus 2011*. In: WISTA Wirtschaft und Statistik. Ausgabe 1/2019, Seite 11 ff.

Dittrich, Stefan. *Der registergestützte Zensus 2021*. In: WISTA Wirtschaft und Statistik. Sonderheft Zensus 2021. 2019, Seite 5 ff.

Eurostat. *Budapest Memorandum „Population Movements and Integration Issues - Migration Statistics"*. 103. Konferenz der Leiterinnen und Leiter der Nationalen Statistischen Ämter. Budapest 2017. [Zugriff am 13. Juni 2019]. Verfügbar unter: https://ec.europa.eu

Massih-Tehrani, Nilgun/Reichert, Adrian. *Die Haushaltegenerierung im Zensus 2021*. In: WISTA Wirtschaft und Statistik. Sonderheft Zensus 2021. 2019, Seite 36 ff.

Nationaler Normenkontrollrat. *Stellungnahme des Nationalen Normenkontrollrates gemäß § 6 Absatz 1 NKRG – Entwurf eines Gesetzes zur Vorbereitung eines registergestützten Zensus einschließlich einer Gebäude- und Wohnungszählung 2021 (Zensusvorbereitungsgesetz 2021 – ZensVorbG 2021)*. NKR-Nr. 3821. 2016.

Nationaler Normenkontrollrat. *Stellungnahme des Nationalen Normenkontrollrates gemäß § 6 Absatz 1 NKRG – Entwurf eines Gesetzes zur Durchführung des Zensus im Jahr 2021*. NKR-Nr. 4684. 2019.

Rat für Sozial- und Wirtschaftsdaten (RatSWD). *Empfehlungen des RatSWD zum Zensus 2021 und zu späteren Volkszählungen*. Berlin 2016. [Zugriff am 13. Juni 2019]. Verfügbar unter: www.ratswd.de

UNECE (United Nations Economic Commission for Europe – Wirtschaftskommission für Europa der Vereinten Nationen). *Measuring Population and Housing. Practices in the UNECE countries in the 2010 round of censuses*. Genf, New York 2014.

UNECE. *Conference of European Statisticians Recommendations for the 2020 Censuses of Population and Housing*. Genf, New York 2015.

UNECE. *Guidelines on the use of registers and administrative data for population and housing censuses*. Beschlossen von der 66. Konferenz Europäischer Statistiker (CES). Genf, 18. bis 20. Juni 2018. [Zugriff am 9. April 2019]. Verfügbar unter: www.unece.org

Valente, Paolo. *Censuses: Current Approaches and Methods*. In: Wright, James D. (Herausgeber). International Encyclopedia of the Social and Behavioral Sciences. 2. Auflage. Band 3. Oxford 2015.

Valente, Paolo. *Survey Results on Future Censuses Beyond 2020*. Vortrag gehalten bei der Sitzung der UNECE Expert Group on Population and Housing Censuses. Genf 2018.

RECHTSGRUNDLAGEN

Durchführungsverordnung (EU) 2018/1799 der Kommission vom 21. November 2018 über die Einführung einer zeitlich begrenzten statistischen Direktmaßnahme für die Verbreitung ausgewählter Themen der Volks- und Wohnungszählung 2021 geokodiert auf ein 1-km^2-Gitter (Amtsblatt der EU Nr. L 296, Seite 19).

Verordnung (EG) Nr. 763/2008 des Europäischen Parlaments und des Rates vom 9. Juli 2008 über Volks- und Wohnungszählungen (Amtsblatt der EU Nr. L 218, Seite 14).

Verordnung (EU) Nr. 1260/2013 des Europäischen Parlaments und des Rates vom 20. November 2013 über europäische Statistiken (Amtsblatt der EU Nr. L 330, Seite 39).

Verordnung (EG) Nr. 862/2007 des Europäischen Parlaments und des Rates vom 11. Jui 2007 zu Gemeinschaftsstatistiken über Wanderung und internationalen Schutz und zur Aufhebung der Verordnung (EWG) Nr. 311/76 des Rates über die Erstellung von Statistiken über ausländische Arbeitnehmer (Amtsblatt der EU Nr. L 199, Seite 23).

Herausgeber
Statistisches Bundesamt (Destatis), Wiesbaden

Schriftleitung
Dr. Daniel Vorgrimler
Redaktionsleitung: Juliane Gude
Redaktion: Ellen Römer

Ihr Kontakt zu uns
www.destatis.de/kontakt

Erschienen im August 2019
Das Archiv älterer Ausgaben finden Sie unter www.destatis.de

Print
Einzelpreis: EUR 19,– (zzgl. Versand)
Jahresbezugspreis: EUR 114,– (zzgl. Versand)
Bestellnummer: 1010200-19907-1
ISSN 0043-6143
ISBN 978-3-8246-1090-7

Download (PDF)
Artikelnummer: 1010200-19907-4, ISSN 1619-2907

Vertriebspartner
IBRo Versandservice GmbH
Bereich Statistisches Bundesamt
Kastanienweg 1
D-18184 Roggentin
Telefon: +49 (0) 382 04 / 6 65 43
Telefax: +49 (0) 382 04 / 6 69 19
destatis@ibro.de

Papier: Design Offset, FSC-zertifiziert

© Statistisches Bundesamt (Destatis), 2019
Vervielfältigung und Verbreitung, auch auszugsweise, mit Quellenangabe gestattet.