

Haushaltegenerierung im Zensusstest

Dipl.-Volksw. Birgit Rimmelspacher, Dipl.-Volksw. Tilman von Roncador

Nach Empfehlungen der Bundesregierung soll eine künftige Volkszählung nicht mehr wie bisher mittels Befragung aller Personen und Haushalte durchgeführt werden, sondern sich auf die in den Verwaltungsregistern vorhandenen Daten stützen. – Die Statistischen Ämter des Bundes und der Länder wurden daher beauftragt, hierfür im Rahmen einer Testphase geeignete Modelle zu entwickeln und deren Ergebnisqualität im Vergleich zu derjenigen eines nach der traditionellen Methode durchgeführten Zensus zu prüfen. – Die Verwaltungsregister enthalten zwar Daten über Personen, nicht jedoch über Haushalts- und Wohnungszusammenhänge. Wohnungsstatistische Daten müssen daher auch in einem registergestützten Zensus primärstatistisch erhoben werden. – Zur Ermittlung der Haushaltsstrukturen wurde ein sehr komplexes Verfahren entwickelt, das diese aus dem Zusammenspiel der Informationen aus den Melderegistern und den primärstatistisch erhobenen Wohnungsdaten ableitet. Nachfolgend werden Grundzüge des entwickelten Verfahrens vorgestellt und die erzielten Ergebnisse diskutiert.

Vorgeschichte und Aufgaben des Tests

Registergestützter Zensus als Alternative zur herkömmlichen Vollerhebung

Im Zusammenhang mit der Empfehlung der Europäischen Union für Volks- und Wohnungszählungen in den Mitgliedsstaaten in den Jahren 2000/2001 hat sich die Bundesregierung sowohl in der 13. als auch in der 14. Wahlperiode aus Kosten- und Akzeptanzgründen gegen eine herkömmliche Vollerhebung nach dem Vorbild der Volkszählung von 1987 ausgesprochen. Daraufhin wurden die Statistischen Ämter des Bundes und der Länder beauftragt, Alternativmodelle zu entwickeln. Diese sollten sich statt auf eine primärstatistische Vollerhebung, also die Befragung aller Einwohner, auf die Nutzung von Daten aus vorhandenen Verwaltungsdateien, und zwar insbesondere der Melderegister (MR) der Gemeinden, stützen.

Mit diesem Umstieg auf einen registergestützten Zensus wird statistisches Neuland betreten. Daher mussten die neuen statistischen Verfahren umfassend getestet und weiter entwickelt werden, um eine ähnlich hohe Ergebnisqualität wie bei den bisherigen Volkszählungen zu garantieren. Hierzu diente der Zensusstest, der auf der rechtlichen Grundlage des Gesetzes zur Vorbereitung eines registergestützten Zensus (Zensusvorbereitungsgesetz) vom 27. Juli 2001 (BGBl I S.1882), zum Stichtag 5. Dezember 2001 durchgeführt wurde.

Die bisherigen Volkszählungen lieferten grundlegende personen-, haushalts- und wohnungsstatistische Daten. Die amtlichen Melderegister enthalten lediglich personenbezogene, nicht aber haushaltspezifische Daten. Register für Wohnungen und Haushalte bestehen in Deutschland nicht flächendeckend. Deshalb müssen wohnungsstatistische Daten auch in einem registergestützten Zen-

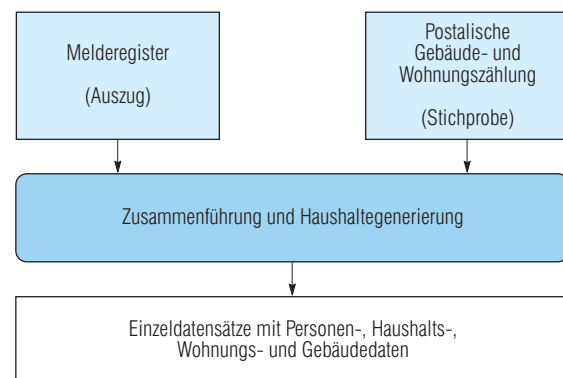
sus primärstatistisch erhoben werden. Um Ergebnisse für Haushalte nachweisen zu können, wurde im Zensusstest ein Verfahren zur Haushaltegenerierung entwickelt. Es leitet Haushaltszusammenhänge aus Informationen des Melderegisters in Zusammenspiel mit den erhobenen wohnungsstatistischen Daten ab.

Der Zensusstest wurde als Stichprobe durchgeführt. Er bestand u. a. aus einer Gebäude- und Wohnungszählung (GWZ), die postalisch bei den in der Stichprobe enthaltenen Gebäude- und Wohnungseigentümern durchgeführt wurde. Ferner wurden von den Gemeinden die erforderlichen Daten aus den Melderegistern angefordert. Auf Grundlage der in allen Bundesländern durchgeführten repräsentativen GWZ und der entsprechenden Auszüge aus den Melderegistern für die in der Stichprobe enthaltenen Adressen wurde im Bayerischen Landesamt für Statistik und Datenverarbeitung ein Verfahren zur Haushaltegenerierung entwickelt (vgl. Abb. 1). Gleich-

Datenquellen des Zensusstests

Ausschnitt aus dem Modell eines registergestützten Zensus

Abb. 1



Haushalte-
befragung
als Kontroll-
instrument

zeitig sollten Möglichkeiten, die Datenqualität der Melderegister zu verbessern, evaluiert werden. Um die Ergebnisse des Zensus-tests beurteilen zu können, wurde an den Stichprobenadressen in allen Bundesländern simultan eine Haushaltebefragung nach dem Muster der bisherigen Volkszählungen durchgeführt. Der Vergleich der Ergebnisse dieser Befragung mit den Ergebnissen des zur Generierung der Haushaltezusammenhänge entwickelten Verfahrens sollte zeigen, ob ein registergestützter Zensus eine vergleichbare Qualität wie die bisherigen Volkszählungen zu erreichen in der Lage ist.

Vor Beginn des Verfahrens der Zusammenführung und Haushaltegenerierung wurden die im Melderegister enthaltenen Personen mit denen der Haushaltebefragung verglichen. Dabei wurden Personen, die nur in den Melderegistern vorhanden waren, als „**Karteileichen**“ und Personen, die lediglich in der Haushaltebefragung vorkamen, als „**Fehlbestände**“ gekennzeichnet. Dadurch wurde es möglich, auch in einzelnen Verfahrensschritten zu erkennen, ob und in welchem Ausmaß Karteileichen zu Haushalten generiert wurden.

Nachfolgender Aufsatz skizziert das entwickelte Verfahren und stellt ausgewählte Ergebnisse dar. Mit Ausnahme von Tabelle 2 handelt es sich dabei stets um (für ganz Deutschland) hochgerechnete Zahlen.

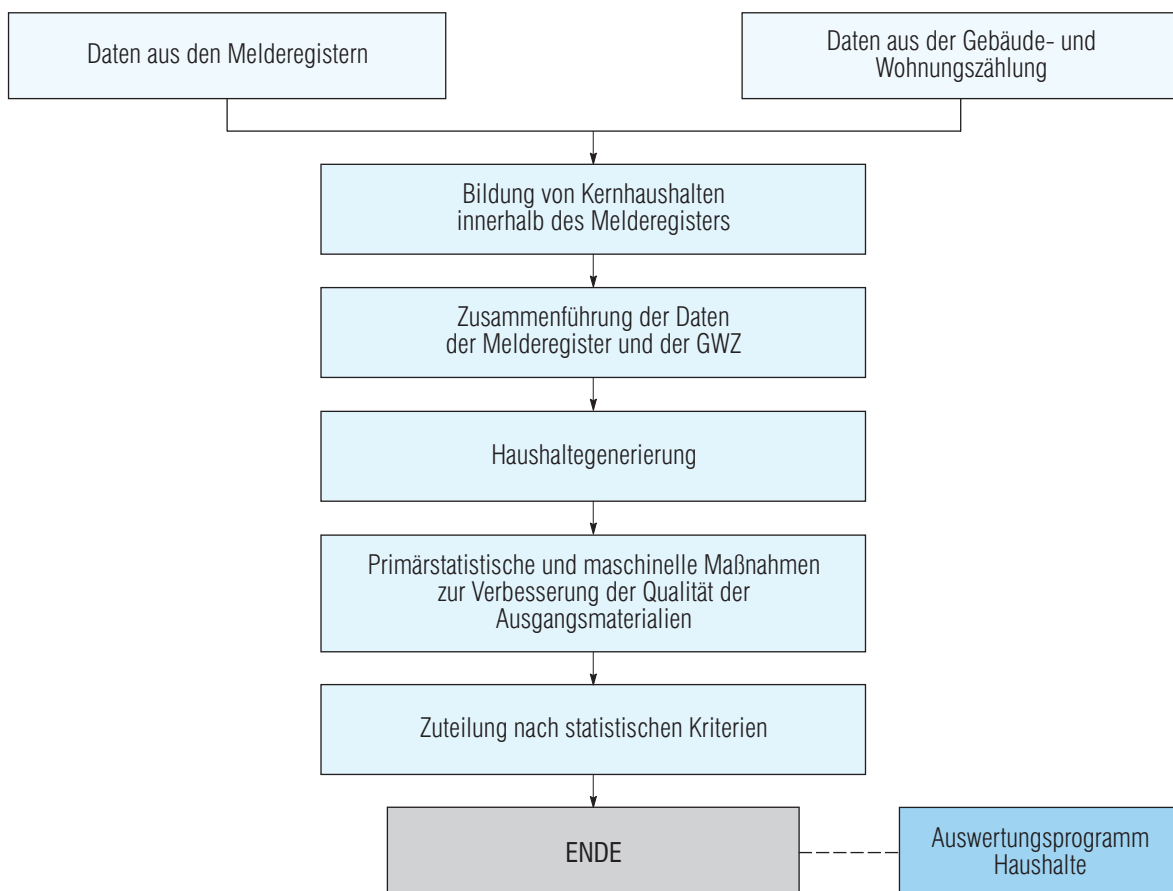
Grundstruktur des Verfahrens zur Haushaltegenerierung

Primäres
Ziel des
Verfahrens

Primäres Ziel des Verfahrens zur Haushaltegenerierung ist das Auffinden von Haushaltszusammenhängen der in den Melderegistern enthaltenen Personen. Da bei einem registergestützten Zensus grundsätzlich nur so genannte Wohnhaushalte gebildet werden können, ist dies gleichbedeutend damit, jede Person in die richtige Wohnung zu generieren.¹ Das Konzept der Wohnhaushalte, das nicht auf das gemeinsame Wirtschaften sondern auf das gemeinsame Wohnen abgestellt ist, findet international weite Verbreitung und soll nach Eurostat-Leitlinien auch Vorgabe in den Mitgliedstaaten werden.

¹ Im Gegensatz zum Konzept der Wirtschaftshaushalte, bei dem Mitglieder eines Haushalts nicht nur in einer gemeinsamen Wohnung wohnen, sondern auch gemeinsam wirtschaften, bilden im Konzept der Wohnhaushalte alle in einer Wohnung zusammen lebenden Personen immer auch einen gemeinsamen Haushalt.

Abb. 2 **Komponenten der Zusammenführung und Haushaltegenerierung**



Das im Rahmen des Zensus-tests entwickelte Verfahren zur Bildung von Wohnhaushalten untergliedert sich in mehrere Abschnitte (vgl. Abb. 2). Zunächst werden innerhalb der Melderegisterdaten die so genannten Kernhaushalte gebildet. Anschließend werden die Personen der Melderegister mit Wohnungen zusammengeführt. Auf der nächsten Stufe erfolgt die eigentliche Haushalgenerierung ehe die bis dato noch unverknüpften Personen über statistische Kriterien mit Wohnungen verknüpft werden. Am Ende des Verfahrens sind sämtliche in den Melderegistern enthaltenen Personendaten mit einer Wohnung verbunden.²

Bildung von Kernhaushalten innerhalb des Melderegisters

In diesem Verfahrensabschnitt werden unter Verwendung der „Verzeigerungen“ in den Personendatensätzen der Melderegister so genannte **Kernhaushalte** gebildet. Die Bezeichnung Kernhaushalte zeigt an, dass diese ausschließlich auf Grund von Verzeigerungen ihrer Mitglieder untereinander generiert wurden.

Verwendung der Verzeigerungen der Melderegister

Unter **Verzeigerungen** versteht man die zusätzlich zu den Meldedaten jeder Person in den Melderegistern eingetragenen Namen vorhandener Familienmitglieder wie Ehepartner, Eltern oder Kinder. Letztere sind, sofern unverheiratet, mit ihren Eltern in der Regel bis zum vollendeten 27. Lebensjahr verzeigert. Verzeigerungen lassen somit die Zusammengehörigkeit von Personen des Melderegisters innerhalb klassischer Familienstrukturen erkennen.

Über die Verzeigerungen wurden 20,3 Mill. Kernhaushalte mit mehr als einer Person gebildet (vgl. Tabelle 1). In diesen Kernhaushalten befanden sich 57,6 Mill. Personen.

Ergebnisse der Kernhaushaltsbildung

Tab. 1

Haushalte (ohne Wohnheime)	Kernhaushalte mit mehr als einer Person insgesamt	davon		
		Ehepaare mit Kind(ern)	Einelternteil mit Kind(ern)	Ehepaare ohne Kind(er)
	1 000			
	1	2	3	4
Kernhaushalte	20 267,0	8 151,1	2 737,2	9 378,7

Nach der Bildung der Kernhaushalte erfolgt im nächsten Abschnitt die Zusammenführung der Datensätze des Melderegisters mit den Wohnungsangaben aus der GWZ.

Zusammenführung der Daten der Melderegister und der GWZ

Durch die in diesem Abschnitt des Verfahrens durchgeführte Zusammenführung der beiden Datenquellen werden den gebildeten

Kernhaushalten aber auch Einzelpersonen Wohnungsdaten, wie etwa Größe oder Ausstattung, zugeordnet. Dadurch werden Wohnhaushalte gebildet.

Um die Daten zusammenführen zu können, wurden in der GWZ die Namen der Wohnungsinhaber erfragt. Unter Wohnungsinhaber ist der Nutzer der Wohnung, also der Hauptmieter oder der in seiner Wohnung lebende Eigentümer, zu verstehen. Weitere Merkmale wie Geburtsdatum und Geschlecht stehen für die Zusammenführung nicht zur Verfügung, da sie sich in der GWZ nicht erheben lassen.

Maschinell nicht identifizierbare Einheiten müssen in einem personal- und somit kostenintensiven Arbeitsgang manuell zusammengeführt werden. Es ist deshalb erforderlich, einen möglichst hohen Anteil von zusammengehörenden Datensätzen maschinell zu verknüpfen. Bei den hohen Datenvolumina von rund 37 Millionen Wohnungs- und ca. 90 Millionen Personendatensätzen (mit Haupt- bzw. Nebenwohnsitz), die in einem künftigen registergestützten Zensus voraussichtlich zu verarbeiten sein werden, führt bereits ein geringer Anteil maschinell nicht verknüpfter Datensätze zu erheblichen zusätzlichen Kosten. Gleichzeitig darf das maschinelle Verfahren jedoch nur zu einem äußerst geringen Anteil fehlerhafter Zusammenführungen von Wohnungs- mit Personendaten führen, da es sonst zu Verzerrungen der Wohnungs- und Haushaltszusammenhänge kommen kann.

Das im Bayerischen Landesamt für Statistik und Datenverarbeitung entwickelte Verfahren beruht auf einer Methode des Zeichenkettenvergleichs zur Quantifizierung von Namensähnlichkeiten sowie einem hierarchisch strukturierten Stufenmodell zur Bewertung von Datensatzübereinstimmungen.³

In einem mehrstufigen Abgleich wird adressweise der Name des Wohnungsinhabers in den Melderegisterdaten gesucht. Der Vergleich erfolgt hierbei mit den im Melderegister angegebenen Familien-, Geburts-, Ehe-, Ordens-, Künstler-, Vor- und Rufnamen. Im Anschluss an den Namensabgleich werden allen Kernhaushaltsmitgliedern von denjenigen Personen, die im Melderegister als Wohnungsinhaber identifiziert wurden, die Daten der entsprechenden Wohnung zugeordnet, so dass nun jeweils die gesamten zum Wohnungsinhaber gehörenden Kernhaushalte mit einer Wohnung verknüpft sind.

2 Ausführliche Darstellungen des Zensus-tests finden sich u. a. bei Füllrohr, M., Rimmelspacher, B.: Testuntersuchungen zur Vorbereitung eines registergestützten Zensus – Inhalt und Stand -, Bayern in Zahlen, Heft 1/2001, S.13 ff. sowie bei Lauer, T., Werner, J.: Der Zensus-test 2001 – Prüfung neuer Methoden als Alternative für eine Volkszählung, Baden-Württemberg in Wort und Zahl, Heft 11/2001, S.545 ff.
 3 Eine ausführliche Beschreibung des Verfahrens sowie der theoretischen Grundlagen findet sich bei: Füllrohr, M., Rimmelspacher, B., v. Roncador, T.: Zusammenführung von Datenbeständen ohne numerische Identifikatoren – ein Verfahren im Rahmen der Testuntersuchungen zu einem registergestützten Zensus, Bayern in Zahlen 7/2002, S. 308 ff.

Maschinelles
Verfahren mit
phonetischem
Abgleich

Da Namen derselben Personen in den beiden Datenquellen gelegentlich in unterschiedlicher Schreibweise enthalten sind, zielt der Namensabgleich nicht nur auf eine vollständige Übereinstimmung der Namen, sondern auch auf Namensähnlichkeiten („phonetischer“ Abgleich). Das aufwändige maschinelle Verfahren besteht aus insgesamt 47 Unterstufen.

Der maschinelle Namensabgleich beginnt mit einer restriktiven Unterstufe, die nur dann einen GWZ-Wohnungsinhaber mit einem Personensatz des Melderegisters verknüpfte, wenn der GWZ-Name und der MR-Name, bestehend aus Familienname einschließlich aller Namensbestandteile sowie sämtlichen Vornamen, exakt übereinstimmt. Bereits hier wurde der mit Abstand größte Anteil der Wohnungsinhaber gefunden. Dies geht aus Tabelle 2 hervor. Sie enthält die Ergebnisse des maschinellen Namensabgleichs für sämtliche in der GWZ angegebenen Wohnungsinhaber aller Gebäude einschließlich sämtlicher Wohnheime **ohne Hochrechnung**. Von den insgesamt 127 652 Wohnungsinhaber der GWZ wurden 117 931 bzw. 92,4 % gefunden. Alle im maschinellen Namensabgleich nicht gefundenen Wohnungsinhaber wurden anschließend manuell überprüft. Hierbei wurden noch 681 weitere Wohnungsinhaber entdeckt.

Gefundene Wohnungsinhaber in allen Gebäuden einschließlich sämtlicher Wohnheime nach Stufen des maschinellen Namensabgleichs (ohne Hochrechnung)

Tab. 2

Stufen	Name des in der GWZ angegebenen Wohnungsinhabers im Melderegister gefunden	
	Anzahl	in % aller GWZ-Wohnungsinhaber
	1	2
Gefundene Wohnungsinhaber insgesamt	117 931	92,4
davon		
Nachnamen voll übereinstimmend	105 957	83,0
davon		
MR-Familienname und GWZ-Nachname ...	105 358	82,5
MR-Geburtsname und GWZ-Nachname	379	0,3
MR-Ehename und GWZ-Nachname	47	0,0
MR-Familienname vor Änderung und GWZ-Nachname	173	0,1
Nachnamen phonetisch übereinstimmend	7 258	5,7
davon		
MR-Familienname und GWZ-Nachname ...	7 159	5,6
MR-Geburtsname und GWZ-Nachname	72	0,1
MR-Ehename und GWZ-Nachname	7	0,0
MR-Familienname vor Änderung und GWZ-Nachname	20	0,0
Mehrfachfamilienname	701	0,5
Sonderformen	4 015	3,1
darunter		
Vertauschungen und Verknüpfungen von Vor- und Nachnamen	680	0,5
Künstlername und Ordensname	2	0,0

Im maschinellen Namensabgleich wurden die in Tabelle 2 angeführten Stufen durchgeführt. Dabei wurden aus Gründen der Übersichtlichkeit die Unterstufen zusammengefasst.

Für alle weiteren Unterstufen wurden die Zuordnungskriterien zunehmend „weicher“. Statt des Familiennamens wurde beispielsweise der Geburtsname mit dem in der GWZ angegebenen Nachnamen des Wohnungsinhabers verglichen. In späteren Unterstufen wurde auf eine vollständige Übereinstimmung der Namen verzichtet. Es war ausreichend, wenn die Namen einen bestimmten Ähnlichkeitsgrad aufwiesen.

In Gebäuden ohne Wohnheime wurden für das Bundesgebiet **hochgerechnet** insgesamt 48,1 Mill. Wohnungsinhaber angegeben. Davon wurden nach Durchführung aller Stufen 44,7 Mill. bzw. 92,9 % der Wohnungsinhaber im maschinellen Namensabgleich gefunden. Dies liegt unterhalb der ursprünglich in das Verfahren gesetzten Erwartungen. Die Ursachen hierfür lagen überwiegend jedoch nicht im Verfahren selbst begründet:

So wurde der Begriff Wohnungsinhaber, obgleich im Fragebogen der GWZ erläutert, im allgemeinen Sprachgebrauch häufig mit dem des Wohnungseigentümers verwechselt. Als Folge hiervon wurde statt des Namens der tatsächlich in der Wohnung lebenden Person der Eigentümer angegeben, der, falls andernorts wohnhaft, nicht im Melderegister gefunden werden konnte. Die entsprechenden Wohnungen konnten aufgrund dieses Fehlers nicht verknüpft werden. Bei einem künftigen Zensus sollte daher der Begriff des Wohnungsinhabers durch einen weniger missverständlichen ersetzt werden.

Eine weitere Ursache dafür, dass Wohnungsinhaber im Melderegister nicht identifiziert werden konnten, liegt im Melderegister selbst begründet. Mittels der Haushaltebefragungen wurden in den Meldedateien Untererfassungen (Fehlbestände) aufgedeckt. Ist ein in der GWZ angegebener und in den Haushaltebefragungen angegriffener Wohnungsinhaber nicht im Melderegister verzeichnet, so kann er verständlicherweise in diesem auch nicht identifiziert werden.

Ergebnisse der Zusammenführung der Daten der Melderegister und der GWZ

Insgesamt konnten 85,3 % aller Personen mit Hauptwohnsitz nach der Zusammenführung der Daten der Melderegister und der GWZ bereits mit einer Wohnung verknüpft werden (vgl. Tabelle 3 u. Titelschaubild). Es hat sich gezeigt, dass die in eine Wohnung generierten Personen mit einer äußerst geringen Wahrscheinlichkeit Karteileichen sind. Die Karteileichenrate der verknüpften Personen lag

Über 90 % der Wohnungsinhaber im maschinellen Namensabgleich gefunden

Bereits über 85 % der Hauptwohnsitzpersonen mit Wohnung verknüpft

Mit einer Wohnung verknüpfte und unverknüpfte Personen nach der Zusammenführung der Daten der Melderegister und der GWZ

Tab. 3

Gebäude lt. GWZ (ohne Wohnheime)	Personen lt. MR		darunter		davon (Sp. 1)						
	Karteileichen		mit einer Wohnung verknüpft		darunter		noch unverknüpft		darunter		
					Karteileichen				Karteileichen		
	1 000	in % von Sp.1	1 000	in % von Sp.1	1 000	in % von Sp.4	1 000	in % von Sp.1	1 000	in % von Sp.8	
1	2	3	4	5	6	7	8	9	10	11	
Bevölkerung am Ort der Hauptwohnung											
Gebäude insgesamt	77 964,4	2 179,7	2,8	66 524,8	85,3	651,7	1,0	11 439,7	14,7	1 528,0	13,4
Bevölkerung am Ort der Nebenwohnung											
Gebäude insgesamt	4 311,5	2 377,1	55,1	1 567,7	36,4	268,6	17,1	2 743,7	63,6	2 108,5	76,8

mit 1,0% deutlich unter den 13,4% der unverknüpften Personen. Diese Erkenntnis kann an späterer Stelle des Verfahrens als Indikator sowohl für eine primärstatistische als auch maschinelle Korrektur von Karteileichen verwendet werden (vgl. hierzu die entsprechenden Abschnitte des Aufsatzes).

Haushaltegenerierung

Bei der beschriebenen Zusammenführung der Daten der Melderegister und der GWZ bleiben diejenigen Haushalte ohne Wohnungszusammenhang, zu denen keine Wohnungsinhaber gehören. Solche Haushalte sind jedoch unter Umständen Bestandteil eines anderen Haushalts, der bereits mit einer Wohnung zusammengeführt werden konnte, sie wurden aber wegen fehlender Verzeigerungen nicht in den Kernhaushalt aufgenommen. Neben den Verzeigerungen enthält das Melderegister weitere Merkmale, die, wenn sie in bestimmten Kombinationen bei zwei Haushalten übereinstimmen, zwar nicht deterministisch jedoch mit hoher Wahrscheinlichkeit ein Zusammengehören dieser beiden Haushalte anzeigen. So wird beispielsweise ein Kernhaushalt A, bestehend aus Frau Müller und zwei mit ihr verzeigerten Kindern, wahrscheinlich zu Kernhaushalt B, bestehend aus Herrn Maier, gehören, wenn beide Kernhaushalte dasselbe Einzugsdatum und dieselbe frühere Adresse aufweisen. Diese Aussage gilt jedoch nur, wenn an derselben Adresse nicht noch weitere Kernhaushalte mit gleichem Einzugsdatum und gleicher früherer Adresse leben. In diesem Fall könnten an der Adresse Ersatzwohnungen für ein zu sanierendes Haus bereitgestellt worden sein, sodass dort zahlreiche Haushalte aus derselben früheren Adresse zeitgleich zugezogen sind.

Aus dem Beispiel wird deutlich, dass die Kombination der übereinstimmenden Merkmale sehr sorgfältig abgewogen werden muss. Dies berücksichtigend, wurde ein Verfahren zur Haushaltegenerierung mit zehn Haupt- und zahlreichen Unterstufen konzipiert. Jede Stufe prüft dabei eine bestimmte Kombination von Angaben aus den Melderegistern. Dabei wurde die Reihenfolge der Stufen so gewählt, dass die Angaben, die ein Zusammengehören zweier Haushalte mit

besonders hoher Wahrscheinlichkeit belegen, in den ersten Stufen abgearbeitet werden. Mit der Stufenzahl steigt die Irrtumswahrscheinlichkeit dann an. Im Einzelnen ergeben sich, in einem grob vereinfachten Schema dargestellt, die nachfolgenden Stufen:

Auf der ersten Stufe werden Personen, die dieselbe Wohnungsnummer wie ein Kernhaushalt haben, mit diesem zusammengeführt. Solche Fälle entstehen, wenn in der GWZ für dieselbe Wohnung zwei Wohnungsinhaber angegeben wurden und diese im Melderegister nicht miteinander verzeigert sind. Ein Beispiel hierfür sind zwei in der GWZ als Wohnungsinhaber angegebene Partner einer nichtehelichen Lebensgemeinschaft.

Auf der zweiten Stufe werden Personen, für die im Melderegister der Name eines Wohnungsgebers (z.B. bei später zugezogenen Familienmitgliedern, Lebenspartnern oder Untermietern) angegeben ist, über einen Namensabgleich mit diesem zusammengeführt, sofern er unter derselben Adresse gemeldet ist⁴. Dabei ist eine vollständige Übereinstimmung des Namens erforderlich. Phonetische Vergleiche wie bei der Zusammenführung der Daten von Melderegistern und GWZ erfolgen hier nicht.

Die dritte Stufe zielt auf fehlende oder unzureichende Verzeigerungen von Kindern mit ihren Eltern in den Melderegistern ab. Nicht oder nicht vollständig im Melderegister mit ihren Eltern verzeigerte minderjährige Kinder werden über Namensübereinstimmungen und unter Beachtung von Altersdifferenzen mit Eltern(teilen), Großeltern (teilen), Geschwistern oder – bei entsprechender Kennzeichnung – mit einem gesetzlichen Vertreter zu Haushalten generiert.

Nichteheliche Lebensgemeinschaften sind im Gegensatz zu verheirateten Paaren nicht im Melderegister miteinander verzeigert. Sie können daher nicht im Bereich Bildung von Kernhaushalten zusam-

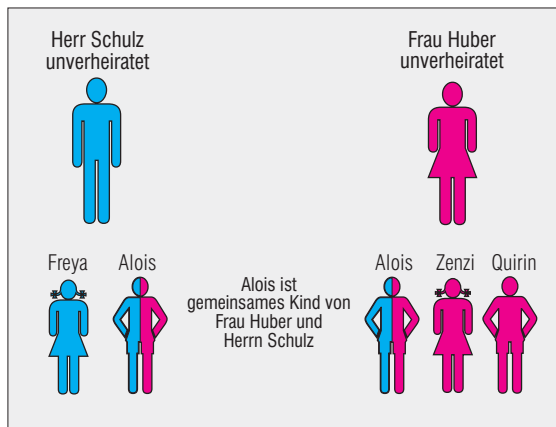
⁴ Weitere Kriterien wie etwa identische Stockwerke konnten nicht berücksichtigt werden, da diese Angaben aus den Melderegistern nicht vollständig vorliegen.

Zehnstufiges Verfahren zur Haushaltegenerierung

Generierung mit Merkmalen des Melderegisters

mengeführt werden. Eine Möglichkeit besteht jedoch, wenn die Partner mindestens ein gemeinsames Kind haben. Beide Partner sind dann jeweils mit demselben Kind verzeigert und können über dieses Indiz als nichteheliche Lebensgemeinschaft zu einem Haushalt zusammengefasst werden. Dies erfolgt auf der vierten Stufe (vgl. Abb. 3).

Abb. 3 **Bildung nichtehelicher Paare über Verzeigerungen* zwischen nichtehelichem Kind und seinen Eltern**



* Verzeigerungen: Zusätzlich zu den Meldedaten jeder Person in den Melderegistern eingetragene Namen vorhandener Familienmitglieder wie Ehepartner, Eltern oder Kinder (vgl. Kasten auf S. 342)

Ähnlich der dritten Stufe im Fall von nicht mit ihren Eltern verzeigerten Kindern führt Stufe fünf Ehepaare, deren Personendatensätze im Melderegister nicht über Verzeigerungen miteinander verknüpft sind, zu einem Haushalt zusammen. Dies erfolgt anhand von Namensübereinstimmungen der Partner sowie unter Berücksichtigung des Familienstands und des Geschlechts. Zusätzlich muss auch das Datum, an welchem sich der Familienstand geändert hat, bei beiden Partnern übereinstimmen.

Ab einem bestimmten Alter der Kinder werden deren Verzeigerungen mit den Eltern gelöscht, auch wenn sie weiterhin bei diesen wohnen. Spätestens erfolgt dies bei einem Alter von 27 Jahren. Für solche erwachsenen Kinder wurde die sechste Stufe konzipiert. Hier werden nicht (mehr) verzeigerte erwachsene Nachkommen mit ihren Eltern oder Elternteilen zu gemeinsamen Haushalten generiert. Als Kriterien dienen Namensübereinstimmungen, bestimmte Altersdifferenzen und gleiche Einzugsdaten. Auf dieser Stufe können auch erwachsene Kinder, die bei ihren Großeltern(-teilen) wohnen, mit diesen zu einem Haushalt zusammengefasst werden.

Ältere Personen, die bei ihren Nachkommen leben, sind nicht mit diesen im Melderegister verzeigert, werden also bei der Bildung von Kernhaushalten nicht zu diesen hinzugeneriert. Oftmals gibt es

hier jedoch Übereinstimmungen von Namen und Einzugsdaten. Hierauf greift Stufe sieben zurück und führt unter Beachtung von Altersdifferenzen ältere Einzelpersonen mit den Haushalten ihrer Nachkommen zusammen.

In einem Haushalt lebende Geschwister konnten, auch wenn sie dieselben Namen tragen, in den bisherigen Stufen nicht zusammengeführt werden, da ihre Altersdifferenz in der Regel unter der jeweils vorgeschriebenen liegt. In der achten Stufe werden Personen bei Übereinstimmung von Namen und Einzugsdatum unter Berücksichtigung entsprechender Altersdifferenzen als Geschwister zu einem Haushalt vereinigt.

Die beiden letzten Stufen des Verfahrens dienen der Bildung nichtehelicher Lebensgemeinschaften. Da hier eine Namensübereinstimmung kein Kriterium ist und vorhandene Altersdifferenzen eine verhältnismäßig geringe Aussagekraft haben, handelt es sich um die beiden Stufen mit der höchsten Wahrscheinlichkeit einer fehlerhaften Haushaltgenerierung. Gewisse Anhaltspunkte liefern noch die Übereinstimmung von Einzugsdaten und früheren Adressen.

Stufe neun generiert nichteheliche Paare (auch mit Nachkommen) bei gleichem Einzugsdatum oder gleicher früherer Adresse unter Berücksichtigung von Geschlecht und Altersdifferenzen. Stufe zehn führt schließlich Einzelpersonen bei Übereinstimmung des Einzugsdatums oder der früheren Adresse zusammen. Auf Altersdifferenzen wird hier verzichtet.

Da die Wahrscheinlichkeit einer fehlerhaften Generierung mit der Stufenhöhe zunimmt, wurde ab der sechsten Stufe eine zusätzliche Sicherung eingebaut: Zu einem Haushalt, der mit einer Wohnung zusammengeführt wurde, steht aus der GWZ eine Angabe bezüglich der Personenzahl in der Wohnung zur Verfügung. Entspricht diese Angabe der Zahl der bereits in dem Haushalt zusammenlebenden Personen, werden diesem keine weiteren mehr hinzugeneriert, auch wenn die Kriterien der jeweiligen Stufen dies erlauben würden. Dadurch soll verhindert werden, dass Karteileichen im Melderegister zu Haushalten hinzugefügt werden. Dieses Vorgehen hat sich im Testverlauf gut bewährt.

Sicherung gegen fehlerhafte Generierung

Ursprünglich war vorgesehen, weitere Stufen zu bilden, die im Vergleich zu den oben angeführten mit größeren Unsicherheiten verbunden gewesen wären. Nachdem sich im Testverlauf jedoch gezeigt hatte, dass diese Zusammenführungen zu sehr vielen Fehlern geführt hätten, wurden sie aus dem Verfahren herausgenommen.

Mit einer Wohnung verknüpfte und unverknüpfte Personen nach der Zusammenführung von Haushalten

Tab. 4

Gebäude lt. GWZ (ohne Wohnheime)	Personen lt. MR		darunter		davon (Sp. 1)						
	Karteileichen		mit einer Wohnung verknüpft		darunter		noch unverknüpft		darunter		
	Karteileichen		Karteileichen		Karteileichen		Karteileichen		Karteileichen		
	1 000	in % von Sp.1	1 000	in % von Sp.1	1 000	in % von Sp.4	1 000	in % von Sp.1	1 000	In % von Sp.8	
1	2	3	4	5	6	7	8	9	10	11	
Bevölkerung am Ort der Hauptwohnung											
Gebäude insgesamt	77 964,4	2 179,7	2,8	70 777,8	90,8	992,6	1,4	7 186,6	9,2	1 187,1	16,5
Bevölkerung am Ort der Nebenwohnung											
Gebäude insgesamt	4 311,5	2 377,1	55,1	1 970,3	45,7	496,2	25,2	2 341,1	54,3	1 880,9	80,3

Ergebnisse des Verfahrens der Haushaltegenerierung

Sehr wenig
Karteileichen
bei verknüpften
Personen

Insgesamt wurden auf den zehn beschriebenen Stufen 5,5% aller Personen am Ort der Hauptwohnung mit einem anderen Haushalt zusammengeführt. Dadurch stieg der Anteil der mit einer Wohnung verknüpften Personen auf 90,8% (vgl. Tabelle 4). Die Karteileichenrate der verknüpften Personen mit Hauptwohnung lag mit 1,4% deutlich unter der Ausgangsrate des Melderegisters und war auch wesentlich geringer als die Karteileichenrate der noch nicht mit einer Wohnung verknüpften Personen. Ein Viertel aller verknüpften Personen mit Nebenwohnung erwies sich als Karteileiche, bei den unverknüpften Personen mit Nebenwohnung betrug die Karteileichenrate 80,3%.

Die einzelnen Stufen waren quantitativ von recht unterschiedlicher Bedeutung. Rund die Hälfte aller Verknüpfungen kam auf der sechsten Stufe zu Stande. Die weiteren Stufen blieben alle unterhalb eines Anteils von 10%.

Sowohl bei der Zusammenführung der Melderegisterdaten mit den Angaben aus der GWZ als auch in der Haushaltegenerierung zeigte sich, dass fehlerhafte Melderegisterdaten nur in deutlich unterproportionaler Anzahl zugeordnet wurden. Dadurch ist der Karteileichenanteil unter den mit einer Wohnung verknüpften Personen deutlich geringer als unter den unverknüpften. Die bisherigen Verfahrensabschnitte liefern somit Anhaltspunkte für mögliche Maßnahmen zur qualitativen Verbesserung der Melderegisterdaten.

Mit Hilfe eines maschinellen Korrekturverfahrens konnte die Karteileichenrate der Bevölkerung am Ort der Nebenwohnung deutlich gesenkt werden. Eine weitere Senkung dieser noch immer hohen Karteileichenrate lässt sich maschinell nur unter Inkaufnahme von Ergebnisverzerrungen erreichen. Deshalb bleiben hierfür nur primärstatistische Korrekturmaßnahmen, auf die später noch eingegangen wird.

Statistische Zuordnung der noch verbliebenen Personen

Nach Durchlaufen der Haushaltegenerierung sind noch 9,2% aller in den Melderegistern enthaltenen Personen mit Hauptwohnsitz und 54,3% aller Personen mit Nebenwohnsitz ohne Wohnung. Solche Personen können Karteileichen sein, die korrekterweise mit keinem anderen Haushalt in Verbindung gebracht wurden. Vor allem bei Personen mit Nebenwohnsitz war dies oftmals, wie oben beschrieben, der Fall. Sie können aber auch zu einer noch mit keinem Haushalt belegten Wohnung gehören, da in der Haushaltegenerierung nur Haushalte untereinander, nicht aber Haushalte und Wohnungen zusammengeführt wurden. Ein solcher Fall kann sich ergeben, wenn in der GWZ Wohnungsinhaber nicht oder falsch angegeben wurden.⁵ Als dritte Möglichkeit bleibt, dass die Personen zwar zu einem anderen Haushalt gehören, dies jedoch anhand der untersuchten Merkmale nicht festgestellt werden konnte.

Die nach der Haushaltegenerierung noch nicht mit einer Wohnung zusammengeführten Personen müssen nach statistischen Kriterien zugeordnet werden. Dies erfolgt in drei Etappen.

a) Zusammenführung mit noch unbelegten Wohnungen

Sofern an einer Adresse noch Wohnungen vorhanden sind, in die bislang noch kein Haushalt hineingeneriert wurde, werden diese zunächst mit jeweils einem noch unverknüpften Haushalt zusammengeführt. Da hier die Strukturen der einzelnen Haushalte unverändert bleiben, kann die Verknüpfung nach dem einfachen Muster größter Haushalt in die größte unbelegte Wohnung, zweitgrößter Haushalt in die zweitgrößte Wohnung usw. erfolgen.⁶

b) Bildung von nichtehelichen Lebensgemeinschaften

Im Test zeigte sich, dass die Stufen neun und zehn der Haushalte-

Zuordnung
von Personen
nach statisti-
schen Kriterien

⁵ Zu dieser Problematik vergleiche Abschnitt „Zusammenführung der Daten der Melderegister und der GWZ“.

⁶ Feinere Zuordnungsverfahren, etwa auf Basis eines Vergleichs der Struktur des zuzuordnenden Haushalts mit den Angaben aus der GWZ haben sich im Test nicht bewährt.

generierung viele der in der Haushaltebefragung vorhandenen nichtehelichen Lebensgemeinschaften nicht finden konnten, beispielsweise wenn ein Partner später zugezogen ist. Deshalb werden weitere nichteheliche Lebensgemeinschaften gebildet, bevor die Haushalte ohne Wohnungsnummern mit Wohnungen, in welchen sich bereits ein Haushalt befindet, zusammengeführt werden. Dabei wurde ähnlich wie in der neunten Stufe der Haushaltegenerierung vorgegangen, allerdings auf die Bedingung des gleichen Einzugsdatums bzw. gleicher früherer Adresse verzichtet.

c) Zusammenführung mit Wohnungen, in denen bereits ein Haushalt vorhanden ist

Gibt es unter der Adresse keine unverknüpfte Wohnung mehr, müssen Haushalte mit Wohnungen zusammengeführt werden, in denen sich bereits ein Haushalt befindet. Um dabei ein völlig willkürliches Verteilen einzelner Personen zu vermeiden, werden ausschließlich komplette Haushalte miteinander verbunden, ein Aufspalten bestehender Haushalte findet nicht statt. Im Prinzip ist ein analoges Vorgehen wie im oben beschriebenen zehnstufigen Verfahren erforderlich. Jedoch stehen jetzt keine weiteren Angaben aus den Melderegistern zur Verfügung, die auf ein Zusammengehören hinweisen. Die Zusammenführung kann nur noch nach statistischen Kriterien wie etwa der Haushalts- bzw. der Wohnungsgröße erfolgen.

In Gebäuden mit nur einer Wohnung ist die Unterteilung der beiden Zuordnungsarten nicht erforderlich. Da nur eine Wohnung zur Verfügung steht, müssen dieser sämtliche noch unverknüpfte Personen zugeordnet werden, unabhängig davon, ob sich in der einen Wohnung bereits ein Haushalt befindet. Eine Sperrung bestimmter Haushaltstypen ist hier nicht möglich.

Ergebnisse nach Durchführung der statistischen Zuordnungen

Insgesamt wurden in Gebäuden mit mehr als einer Wohnung rund 2,7 Mill. Personen am Ort der Hauptwohnung nach rein statistischen Kriterien mit einer Wohnung zusammengeführt (vgl. Tabelle 5). Hiervon waren 26,9% Karteileichen. Viele Personen wurden also vor der Zuteilung nach statistischen Kriterien korrekterweise nicht mit einer Wohnung zusammengebracht.

Nach statistischen Kriterien mit einer Wohnung zusammengeführte Personen am Ort der Hauptwohnung in Gebäuden mit mehr als einer Wohnung (ohne Wohnheime)

Gebäudegröße lt. GWZ	Personen am Ort der Hauptwohnung insgesamt	darunter			
		nach statistischen Kriterien mit einem Haushalt zusammengeführt		darunter Karteileichen am Ort der Hauptwohnung	
		1 000	%	1 000	%
		1	2	3	4
Gebäude mit zwei oder mehr Wohnungen	46 704,2	2 741,4	5,9	736,3	26,9
davon zusammengeführt mit Wohnung ohne Haushalt ...	X	965,8	2,1	210,0	21,7
Haushalt mit Wohnung	X	1 775,6	3,8	526,3	29,6

Tab. 5

Für Gebäude, in denen Haushalte nach statistischen Kriterien mit bestehenden Haushalten zusammengeführt werden mussten, wurde auffallend häufig in der GWZ eine im Vergleich zur Haushaltebefragung zu geringe Zahl von Wohnungen angegeben. Haushalte, die korrekt in den Daten des Melderegisters enthalten waren, konnten deshalb mangels einer Wohnung im Namensabgleich nicht zugeordnet werden.

Die Zuordnung nach statistischen Kriterien stellt ein verhältnismäßig grobes Instrumentarium dar, das in seiner gegenwärtigen Konzeption nicht in der Lage ist, Haushaltsstrukturen derselben Qualität zustande zu bringen wie die vorangegangenen Generierungen. Deshalb sollten, um ein gutes Generierungsergebnis zu erhalten, möglichst wenige Adressen erst mit der Zuordnung nach statistischen Kriterien abgeschlossen werden. Wie oben gezeigt wurde, trägt hierzu Datenmaterial von hoher Qualität maßgeblich bei, sei es über eine geringe Anzahl von Karteileichen im Melderegister oder die vollständige Angabe der Wohnungen einer Adresse in der GWZ.

Statistische Zuordnung sollte möglichst vermieden werden

Problem der Änderung bestehender Haushaltsstrukturen

Wegen des im Zensustest strikt einzuhaltenden Konzeptes der Wohnhaushalte entsteht durch jede Zuordnung ein vergrößerter Haushalt. Die ursprüngliche Struktur der Haushalte wird dadurch geändert. Dies ist im Fall der statistischen Zuordnung problematisch, da die beiden zuvor vorhandenen Haushalte nach Kriterien gebildet wurden, die deutlich zuverlässiger sind als die statistischen. Somit würde durch eine fehlerhafte Zuordnung nach statistischen Kriterien nicht nur im Falle des zugeordneten Haushaltes ein Fehler begangen, sondern auch noch die Struktur eines höchstwahrscheinlich richtig generierten Haushaltes zerstört werden.

Aus diesem Grund werden bestimmte Haushaltskonstellationen in verknüpften Wohnungen von der Zuteilung der nicht verknüpften Personen ausgenommen. Dadurch soll verhindert werden, dass Haushaltstypen, wie beispielsweise Familien mit Kindern, bei denen das Vorhandensein von weiteren Personen weniger wahrscheinlich ist, durch ein Hinzugenerieren von Personen in den Typ „Sonstiger Haushalt“ umgewandelt werden. Dennoch ist, da die Zuordnung in diesem Modul ausschließlich unter Berücksichtigung von Haushaltsgröße und Wohnungsbelegung erfolgt, die Zuordnung nach statistischen Kriterien weitaus unsicherer als die Generierung von Haushalten.

Über 90% der Personen bereits vor Verfahrensschritt Statistische Zuordnung mit Wohnungen verknüpft

Zuordnungen in den einzelnen Verfahrensabschnitten

In Abb. 4 sind die Anteile der in den oben beschriebenen Verfahrensschritten jeweils den Wohnungen zugeordneten Personen dargestellt. Es wird deutlich, dass im Falle der Personen am Ort der Hauptwohnung, die 94,7% der Gesamtbevölkerung repräsentieren, eine große Mehrheit (90,8%) bereits vor dem Verfahrensabschnitt Statistische Zuordnung mit Wohnungen zusammengeführt ist. Dennoch wird auch ein beträchtlicher Anteil der Personen erst im Verfahrensschritt Statistische Zuordnung mit einer Wohnung zusammengeführt. Am Ort der Nebenwohnung werden knapp 85% aller nicht von der maschinellen Korrektur der Karteileichen betroffenen Personen vor dem Verfahrensabschnitt Statistische Zuordnung mit einer Wohnung zusammengeführt. Sowohl bei den Personen am Ort der Hauptwohnung als auch am Ort der Nebenwohnung lässt sich der Anteil der nach statistischen Kriterien zugeordneten Personen jedoch durch primärstatistische Maßnahmen zur Verbesserung der Qualität der Ausgangsmaterialien noch verringern.

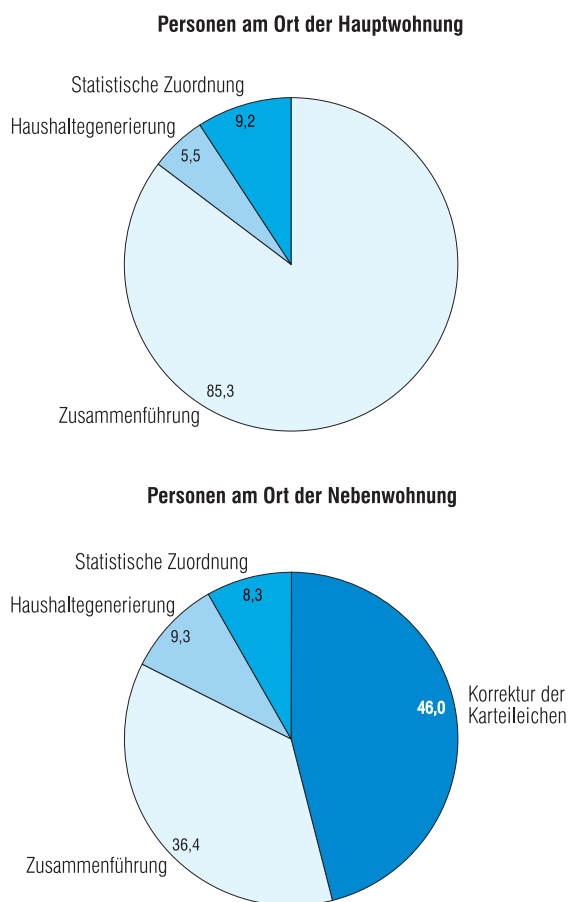
Möglichkeiten zur Aufdeckung und primärstatistischen Korrektur von Karteileichen

a) Möglichkeiten innerhalb des Verfahrens

Ziel des Zensustests war es, neben der Erstellung von Haushaltszusammenhängen auch herauszufinden, ob das Verfahren der Zusammenführung und Haushaltgenerierung auch Anhaltspunkte für eine primärstatistische Korrektur der Melderegisterdaten liefern kann.

Die Untersuchungen hierzu basieren auf der Annahme, dass die Wahrscheinlichkeit für ein Vorhandensein von Karteileichen vor allem an den Adressen hoch ist, an denen nach der Haushaltgenerierung die Zahl der Haushalte die Zahl der bewohnten Wohnungen übersteigt, also noch unverknüpfte Personen vorhanden sind. Diese Annahme stützt sich auf die Feststellung, dass sich unter den nach dem Verfahrensteil Haushaltgenerierung noch unverknüpften Personen ein deutlich höherer Anteil von Karteileichen befindet als unter den bereits verknüpften.

Abb. 4 **Anteile der Verfahrensabschnitte in Prozent**



Simulation primärstatistischer Korrekturen

In Gebäuden mit nach dem Verfahrensteil Haushaltgenerierung noch unverknüpften Haushalten wurden im Test primärstatistische Korrekturmaßnahmen simuliert. Hierfür wurden die Daten aus der Haushaltebefragung verwendet. Sie ersetzen in den entsprechenden Gebäuden die Generierungsergebnisse. Ein Vergleich der Ergebnisse vor und nach der Verwendung der Daten aus der Haushaltebefragung zeigt den Erfolg der primärstatistischen Korrekturen.

Für die Beurteilung der primärstatistischen Befragung von Haushalten ist es erforderlich, den eingesetzten Aufwand mit dem Ertrag zu vergleichen. Der für eine Befragung betriebene Aufwand lässt sich anhand der Zahl der befragten Haushalte messen. Der Ertrag besteht in der Anzahl der aufgedeckten Karteileichen und der damit erzielten Qualitätsverbesserung.

Im Test zeigte sich, dass nur in Gebäuden mit einer oder zwei bewohnten Wohnungen Aufwand und Ertrag einer primärstatistischen Befragung in einem angemessenen Verhältnis stehen. In größeren Gebäuden müssten hingegen für eine nennenswerte Aufdeckung von Karteileichen übermäßig viele Personen befragt werden.

In Ein- und Zweifamilienhäusern konnte durch simulierte primärstatistische Erhebungen bei etwa 1,1 Mill. Haushalten, das entspricht 6,6% der Haushalte in Ein- und Zweifamiliengebäuden, die Karteileichenrate der Hauptwohnsitzbevölkerung um mehr als die Hälfte reduziert werden. In dieser Gebäudekategorie leben mehr als 50% der gesamten Bevölkerung Deutschlands. Besonders deutlich sind die Auswirkungen bei den Gemeinden mit weniger als 10 000 Einwohnern (Gemeindegrößenklasse 1).

b) Ergänzende Stichprobenerhebung außerhalb des Verfahrens

Eine uneingeschränkte Verwendung der Melderegister als Datengrundlage für einen registergestützten Zensus würde nicht nur bezüglich der Haushaltszusammenhänge zu einer erheblichen Ergebnisverzerrung führen, sondern auch die Einwohnerzahlen unrichtig wiedergeben. Aus diesen Gründen ist es für die Ergebnisqualität eines registergestützten Zensus wichtig, die Qualität der aus den Melderegistern gewonnenen Daten durch statistische Verfahren vor ihrer Verwendung zu verbessern.

Ergänzungsstichprobe zur Verbesserung der Ausgangsdaten

Hierfür bestehen Überlegungen, bei einem künftigen registergestützten Zensus neben einer bundesweit durchzuführenden Prüfung und Korrektur von Mehrfachfällen in den Melderegistern weitere korrigierende Maßnahmen in Form von ergänzenden Stichprobenerhebungen durchzuführen. Grundgedanke dieser Stichproben ist, in den Gemeinden – zusätzlich zur Auswertung der Melderegister – auf Stichprobenbasis eine Befragung von Personen durchzuführen mit dem Ziel, die Karteileichen- und Fehlbestandsrate für die einzelne Gemeinde zu ermitteln und die für die Gemeinde zunächst festgestellte Einwohnerzahl in dieser Größenordnung zu korrigieren. Unter Umständen ist es sinnvoll, solche Stichproben nur in Gemeinden mit 10 000 oder mehr Einwohnern (Gemeindegrößenklassen 2, 3 und 4) durchzuführen. Ob ein solches Vorgehen Akzeptanz findet, ist noch offen. Problematisch hierbei ist eine gewisse Ungleichbehandlung großer und kleiner Gemeinden, deren Einwohnerzahl nicht nach derselben Methode ermittelt werden würde.

Haushaltsstatistische Ergebnisse

Das zentrale Ziel der Zusammenführung/Haushaltgenerierung ist es, wie bereits ausgeführt, Haushaltsergebnisse von ähnlicher Qualität wie bei den bisherigen Volkszählungen vorzulegen. In welchem Umfang dieses Ziel erreicht wird, misst sich grundsätzlich an der Differenz der Haushaltsergebnisse laut Generierung zu den Ergebnissen der im Zensus durchgeführten Haushalbefragung. Abweichungen können allerdings neben den Schwächen des Generierungsverfahrens auch auf Datenlücken und Fehler in den Ergebnissen der Haushalbefragung sowie Fehler in den Datenquellen der Zusammenführung/Haushaltgenerierung, der GWZ und des Melderegisters, zurückzuführen sein. Deshalb muss vor dem Vergleich genau festgelegt werden, mittels welcher Daten er erfolgen soll.

Abgrenzung der Untersuchungsmassen für die haushalts- und wohnungsstatistischen Ergebnisse

Für eine sachgerechte Beurteilung der Haushaltgenerierung müssen die Abweichungen, die auf Fehler in der Haushalbefragung zurückzuführen sind, ausgeschaltet werden. Dies lässt sich nur durch Herausnahme der in der Haushalbefragung fehlerhaften Wohnungen und Adressen aus dem Vergleich erreichen. Hiervon

Ausschaltung von Fehlern, die den Vergleich Generierung mit Haushalbefragung beeinträchtigen

betroffen sind Wohnungen ohne oder mit fehlerhaftem Befragungsergebnis in der Haushalbefragung, sowie Adressen, für die dort eine falsche Anzahl von Wohnungen angegeben wurde. Die in den Vergleich eingehenden übrigen Wohnungen und Adressen wurden im Zensusstest als Untersuchungsmasse 1 bezeichnet (vgl. Tabelle 6).

Aber auch Untersuchungsmasse 1 enthält noch Fehler, die nicht auf das Generierungsverfahren per se zurückzuführen sind. Dies sind Über- und Untererfassungen von bewohnten Wohnungen in der GWZ aufgrund fehlerhafter Angaben der Eigentümer und Fehler im Melderegister in Form von Karteileichen und Fehlbeständen. In einem künftigen Zensus müssen diese Fehler für eine hohe Qualität der Ergebnisse vor der Generierung durch entsprechende Maßnahmen vermindert werden. Diese Fehler aus der Über- und Untererfassung der GWZ werden in der Untersuchungsmasse 2 ausgeklammert (vgl. Tabelle 6).

Abgrenzung der Untersuchungsmassen für die Haushaltauswertungen

Masse	Vor Hochrechnung	Nach Hochrechnung
	Anzahl	1 000
	1	2
Bewohnte Wohnungen lt. Haushalbefragung	94 517	35 305,1
- Bewohnte Wohnungen ohne vollständiges, plausibles und konsistentes Befragungsergebnis	14 038	5 236,3
= Untersuchungsmasse 1	80 479	30 068,8
- Wohnungen in Gebäuden mit Über- bzw. Untererfassungen in der GWZ	27 118	4 282,0
= Untersuchungsmasse 2	53 361	25 786,8

Tab. 6

Lässt man in die Auswertungen zu Untersuchungsmasse 2 nur sowohl in Melderegister als auch in Haushalbefragung vorkommende Personen (so genannte paarige Personen) einfließen, erhält man quasi idealtypische Bedingungen ohne Fehler der GWZ und des Melderegisters.

Vergleich der Haushaltsstrukturen

Um das Verfahren an sich beurteilen zu können, werden in den nachfolgenden Tabellen die Fehler der Melderegister und der GWZ ausgeschaltet. Dafür wird Untersuchungsmasse 2 verwendet, zusätzlich gehen nur paarige Personen in die Auswertung ein. Die Zahl der Haushalte laut Haushalbefragung insgesamt bleibt hierbei unterhalb der in Tabelle 6 ausgewiesenen Zahl der bewohnten Wohnungen in Untersuchungsmasse 2, da komplett aus Karteileichen bestehende Haushalte nicht in die Untersuchung einfließen. Aus Tabelle 7, die das Ergebnis vor dem Hintergrund eines weitgehend fehlerfreien Datenmaterials darstellt, wird deutlich, dass das Verfahren der Haushaltgenerierung zu beinahe identischen Haushaltsgrößen führt wie die Haushalbefragung. Ob dieses Ergebnis so

Generierung führt zu fast identischer Haushaltgröße

auch erreicht werden wird, hängt von dem in einem registergestützten Zensus betriebenen Aufwand zur Verminderung von Fehlern im Datenmaterial ab.

Größenvergleich der Haushalte in der Haushaltgenerierung und -befragung auf der Basis von Untersuchungsmasse 2 mit paarigen Personen

Tab. 7

Haushaltsgröße	Zahl der Haushalte laut		Anteile der Haushalte laut	
	Generierung	Haushaltebefragung	Generierung	Haushaltebefragung
	1 000		%	
	1	2	3	4
Haushalte insgesamt	25 398,8	25 499,4	100,0	100,0
davon mit ...				
1 Person	7 452,1	7 509,1	29,3	29,4
2 Personen	8 965,4	8 910,5	35,3	34,9
3 Personen	4 084,0	4 144,5	16,1	16,3
4 Personen	3 481,7	3 520,7	13,7	13,8
5 oder mehr Personen ..	1 415,6	1 414,6	5,6	5,5

Auf Ebene der vier Gemeindegrößenklassen ist die Abweichung bei den einzelnen Haushaltsgrößen vor allem in Gemeinden mit 800 000 Einwohnern oder mehr etwas größer als in der Gesamtauswertung (vg. Tabelle 8).

Die Staatsangehörigkeit der Personen in den Haushalten wird aus Gründen der besseren Übersichtlichkeit anhand eines vereinfachten Schemas dargestellt (vgl. Tabelle 9). Es werden lediglich drei Gruppen unterschiedlicher Haushalte betrachtet: Haushalte, die nur aus Personen deutscher Staatsangehörigkeit bestehen, Haushalte mit Personen von ausschließlich nicht deutscher Staatsangehörigkeit und solche mit Personen von deutscher und nichtdeutscher Staatsangehörigkeit.

Mehr deutsch- ausländische Haushalte in der Generierung

In der Generierung liegt der Anteil der Haushalte mit gemischten Staatsangehörigkeiten höher als in der Haushaltebefragung. Der wesentliche Grund für den hohen Anteil der Mischhaushalte dürfte in unterschiedlichen Angaben zur Staatsangehörigkeit in Melderegister und Haushaltebefragung bestehen. Hier ist davon auszugehen, dass die Angaben des Melderegisters zuverlässiger sind als die der Haushaltebefragung. So wäre es beispielsweise denkbar, dass noch nicht eingebürgerte Personen in der Haushaltebefragung bereits die deutsche Staatsangehörigkeit angeben. Die Zahl der Haushalte mit Personen unterschiedlicher Staatsangehörigkeit dürfte daher in der Haushaltebefragung als zu niedrig ausgewiesen sein.

Der Anteil der Ehepaare mit und ohne Kinder liegt in der Generierung niedriger als in der Haushaltebefragung (vgl. Tabelle 10). Der Haushaltstyp „Ehepaare“ wird im Bereich „Bildung der Kernhaushalte“ unter Verwendung der Verzeigerungen des Melderegisters

Größenvergleich der Haushalte in der Haushaltgenerierung und -befragung auf der Basis von Untersuchungsmasse 2 mit paarigen Personen nach Gemeindegrößenklassen (Aufgliederung der Tabelle 7)

Tab. 8

Haushaltsgröße	Zahl der Haushalte laut		Anteile der Haushalte laut	
	Generierung	Haushaltebefragung	Generierung	Haushaltebefragung
	1 000		%	
	1	2	3	4
Gemeinden mit unter 10 000 Einwohnern				
Haushalte insgesamt	7 192,6	7 223,1	100,0	100,0
davon mit ...				
1 Person	1 576,4	1 609,1	21,9	22,3
2 Personen	2 545,5	2 518,3	35,4	34,9
3 Personen	1 314,1	1 331,9	18,3	18,4
4 oder mehr Personen	1 756,5	1 763,8	24,4	24,4
Gemeinden mit 10 000 bis unter 50 000 Einwohnern				
Haushalte insgesamt	8 191,8	8 235,4	100,0	100,0
davon mit ...				
1 Person	2 201,6	2 226,2	26,9	27,0
2 Personen	2 829,1	2 822,6	34,5	34,3
3 Personen	1 360,4	1 371,6	16,6	16,7
4 oder mehr Personen	1 800,8	1 815,1	22,0	22,0
Gemeinden mit 50 000 bis unter 800 000 Einwohnern				
Haushalte insgesamt	7 775,5	7 801,6	100,0	100,0
davon mit ...				
1 Person	2 736,6	2 734,0	35,2	35,0
2 Personen	2 810,2	2 805,4	36,1	36,0
3 Personen	1 132,2	1 150,7	14,6	14,7
4 oder mehr Personen	1 096,5	1 111,4	14,1	14,2
Gemeinden mit 800 000 oder mehr Einwohnern				
Haushalte insgesamt	2 238,9	2 239,3	100,0	100,0
davon mit ...				
1 Person	937,5	939,8	41,9	42,0
2 Personen	780,6	764,2	34,9	34,1
3 Personen	277,3	290,3	12,4	13,0
4 oder mehr Personen	243,6	245,0	10,9	10,9

generiert. Auf der dritten und sechsten Stufe der Haushaltgenerierung kommen unter Umständen noch weitere Personen dazu. Abgesehen von diesen beiden Stufen verändert jede Person, die im gesamten sonstigen Generierungsprozess noch zu Ehepaaren hinzugeneriert wird, den Typ „Ehepaare“ in „Sonstigen Haushalt“. Die Zahl der Haushalte des Typs „Ehepaar“ sinkt also mit jeder solchen

Ausgewählte Haushaltstypen in der Haushaltgenerierung und -befragung auf der Basis von Untersuchungsmasse 2 mit paarigen Personen (Staatsangehörigkeit)

Tab. 9

Haushalte	Zahl der Haushalte laut		Anteil der Haushalte laut	
	Generierung	Haushaltebefragung	Generierung	Haushaltebefragung
	1000		%	
	1	2	3	4
Haushalte insgesamt	25 398,8	25 499,4	100,0	100,0
davon				
deutsche Haushalte	23 717,2	23 996,8	93,4	94,1
ausländische Haushalte ...	751,1	765,0	3,0	3,0
Mischhaushalte	930,5	737,6	3,7	2,9

Generierung. Die im Vergleich zur Haushaltebefragung auf Seiten der Haushaltegenerierung zu niedrige Anzahl der Ehepaare dürfte daher in den meisten Fällen auf ein fehlerhaftes Hinzugenerieren einer Person zu einem Ehepaar entstanden sein. Hier ist noch eine weitere Verfeinerung des Verfahrens erforderlich.

Ausgewählte Haushaltstypen in der Haushaltegenerierung und -befragung auf der Basis von Untersuchungsmasse 2 mit paarigen Personen (Ehepaare mit und ohne Kinder)

Tab. 10

Haushalte	Zahl der Haushalte laut		Anteil der Haushalte laut	
	Generierung	Haushaltebefragung	Generierung	Haushaltebefragung
	1000		%	
	1	2	3	4
Haushalte insgesamt	25 398,8	25 499,4	100,0	100,0
darunter				
Ehepaare ohne Kind	6 458,0	6 546,6	25,4	25,7
Ehepaare mit Kindern				
27 Jahren	6 192,2	6 371,0	24,4	25,0
darunter mit ... unter				
18 Jahren				
1 Kind	2 187,3	2 256,1	8,6	8,8
2 Kindern	2 093,2	2 156,6	8,2	8,5
3 oder mehr Kindern	673,6	679,9	2,7	2,7

Höherer Anteil des Haushaltstyps Elternteil mit Kindern

Der Haushaltstyp „Elternteil mit Kindern“ kommt in der Generierung auf einen höheren Anteil als in der Haushaltebefragung (vgl. Tabelle 11). Das Generierungsverfahren bringt zu viele Haushalte dieses Typs hervor. Gleiches gilt auch für die nichtehelichen Lebensgemeinschaften. Die korrekte Bildung dieses sehr schwierig einzugrenzenden Haushaltstyps stellt eines der zentralen Probleme des Generierungsverfahrens dar. Die Auswertung zeigt, dass die bisher angewendeten Kriterien offenbar über das Ziel hinausschießen. Im Bereich der Bildung von nichtehelichen Lebensgemeinschaften besteht daher noch Forschungsbedarf, um das Verfahren gegebenenfalls weiter zu verfeinern.

Die Zahl der Seniorenhaushalte wurde vereinfachend in die drei Kategorien „Ehepaar, beide 65 Jahre oder älter“, „Ehepaar, ein Partner 65 Jahre oder älter“ und „allein lebend, 65 Jahre oder älter“ untergliedert (vgl. Tabelle 12). Die Abweichungen zwischen Generierung und Haushaltebefragung sind verhältnismäßig gering, bei der zahlenmäßig größten Gruppe der allein lebenden Senioren sogar nur minimal.

Fazit

Ein wesentliches Ziel des Zensustests bestand in Entwicklung und Test eines Verfahrens, das es erlaubt, Haushaltsstrukturen in einer registergestützten Volkszählung zu ermitteln. Das hierbei entstandene Verfahren gestaltet sich äußerst komplex und aufwändig. An einigen Stellen sind auch noch Weiterentwicklungen erforderlich. Dennoch lässt es sich technisch und organisatorisch prinzipiell

Ausgewählte Haushaltstypen in der Haushaltegenerierung und -befragung auf der Basis von Untersuchungsmasse 2 mit paarigen Personen (Elternteile mit Kindern und Nichteheleiche Lebensgemeinschaften)

Tab. 11

Haushalte	Zahl der Haushalte laut		Anteil der Haushalte laut	
	Generierung	Haushaltebefragung	Generierung	Haushaltebefragung
	1000		%	
	1	2	3	4
Haushalte insgesamt	25 398,8	25 499,4	100,0	100,0
darunter				
Elternteil mit Kindern unter				
27 Jahren	1 172,4	1 144,0	4,6	4,5
darunter mit Kindern unter				
18 Jahren	912,7	875,0	3,6	3,4
Nichteheleiche Lebensgemeinschaften ohne Kinder	951,2	903,6	3,7	3,5
Nichteheleiche Lebensgemeinschaften mit Kindern unter				
27 Jahren	516,4	448,3	2,0	1,8
darunter mit Kindern unter				
18 Jahren	467,0	399,3	1,8	1,6

Ausgewählte Haushaltstypen in der Haushaltegenerierung und -befragung auf der Basis von Untersuchungsmasse 2 mit paarigen Personen (Seniorenhaushalte)

Tab. 12

Haushalte	Zahl der Haushalte laut		Anteil der Haushalte laut	
	Generierung	Haushaltebefragung	Generierung	Haushaltebefragung
	1000		%	
	1	2	3	4
Haushalte insgesamt	25 398,8	25 499,4	100,0	100,0
darunter				
Ehepaare, beide 65 Jahre oder älter	2 082,3	2 129,5	8,2	8,4
Ehepaare, ein Partner 65 Jahre oder älter	857,3	874,5	3,4	3,4
Alleinlebender 65 Jahre oder älter	3 100,3	3 087,8	12,2	12,1

realisieren. Es führt zu tragfähigen Ergebnissen und hat sich somit für einen künftigen Zensus als tauglich erwiesen.

Ein wesentlicher Ausgangspunkt für dieses Verfahren sind die Melderegister. Sie enthalten zahlreiche Merkmale, die sich für die Bestimmung von Haushaltstypen verwenden lassen. Am wichtigsten sind hier sicherlich die Verzeigerungen, die eine sehr genaue Zuordnung von Ehepartnern und Kindern erlauben. Daneben stehen Angaben zu Geschlecht, Familienstand, Alter, Staatsangehörigkeit und Wohnstatus aus den Melderegistern zur Verfügung, aber auch Zugzugsdatum und frühere Adresse. Damit bieten die Melderegister ein breiteres Spektrum an Möglichkeiten, Haushaltstypen zu bestimmen, als die Angaben aus dem Personenbogen der Volkszählung 1987.

Die Qualität der generierten Haushaltsstrukturen reagiert jedoch sehr empfindlich auf Fehler in den verwendeten Ausgangsmaterialien, also GWZ und Melderegister. Dadurch wird die Qualität der verwendeten Ausgangsmaterialien zum zentralen Problem eines künftigen registergestützten Zensus.

Für den Erfolg eines künftigen registergestützten Zensus und die Möglichkeit, mittels der Haushaltegenerierung gute Haushaltsergebnisse hervorzubringen, sind daher eine hinreichende Korrektur der Fehler im Melderegister sowie eine optimale Durchführung der GWZ zur Vermeidung der dort im Zensusstest aufgetretenen Fehler unumgänglich.