

Dipl.-Volkswirtin Ingeborg Vorndran

# Zensustest – Verfahren und Ergebnisse der Haushaltegenerierung

Daten über die Anzahl und Strukturen von privaten Haushalten sowie deren Wohnsituation sind sowohl für die öffentliche Planung als auch für die private Wirtschaft von großem Interesse. Angaben über Privathaushalte waren deshalb auch Kernbestandteil aller in der Vergangenheit durchgeführten Volkszählungen. Im Zuge der Entscheidung der Bundesregierung, künftig auf eine primärstatistische Vollerhebung zu verzichten, wurden die Statistischen Ämter des Bundes und der Länder beauftragt, ein alternatives Zensuskonzept zu entwickeln, das soweit möglich auf vorhandene Verwaltungsdaten zurückgreift. Informationen über Haushaltszusammenhänge können jedoch nur unvollständig, zu Wohnverhältnissen von Haushalten überhaupt nicht aus den vorhandenen Verwaltungsunterlagen gewonnen werden. Da es in Deutschland keine flächendeckenden Gebäude- und Wohnungsregister gibt, müssen gebäude- und wohnungsstatistische Daten auch in einem künftigen registergestützten Zensus primärstatistisch erhoben werden.<sup>1)</sup> Zur Bereitstellung von räumlich tief gegliederten Haushalts- und Wohnungsbelegungsdaten bei einem Methodenwechsel hin zu einem registergestützten Zensus wurde im Rahmen des Zensustests ein Verfahren entwickelt und getestet, das die Personendaten aus dem Melderegister mit den Gebäude- und Wohnungsdaten aus einer (postalischen) Gebäude- und Wohnungszählung (GWZ) adressenweise zusammenführt und in einem mehrstufigen Prozess die für diese Adresse gemeldeten Personen zu Haushalten zusammenführt (Haushaltegenerierung).

Der vorliegende Beitrag ist Teil der Berichtsreihe über die Ergebnisse des Zensustests<sup>2)</sup> und gibt zunächst einen Überblick über das im Rahmen des Zensustests entwickelte Verfahren der Zusammenführung/Haushaltegenerierung sowie die aufgrund der Testergebnisse vorgenommenen Modifikationen. Im Anschluss daran werden die Ergebnisse des Haushaltegenerierungsverfahrens den Ergebnissen der parallelen für die ausgewählten Adressen durchgeführten traditionellen Haushalbefragung durch Interviewer gegenübergestellt.<sup>3)</sup>

Dabei ist zu beachten, dass mit dem Verfahren der Haushaltegenerierung im Unterschied zu primärstatistischen Erhebungen nur Wohnhaushalte, im Sinne von Personen, die eine gemeinsame Wohnung bewohnen, ermittelt werden. Ein Nachweis des in der deutschen amtlichen Statistik bisher üblichen Haushaltsbegriffs als eine Personengruppe, die zusammen wohnt und wirtschaftet, ist nicht möglich, da es weder in den Melderegistern noch in einer Gebäude- und Wohnungszählung Informationen über die ökonomischen Beziehungen zwischen den Haushaltsmitgliedern gibt. Allerdings zeigen die Ergebnisse der Mikrozensus-Zusatzerhebung zur Wohnsituation aus dem Jahr 2002, dass 98,1% der Wohnungen nur mit einem Haushalt belegt sind. Zudem werden in den methodischen Empfehlungen des Statistischen Amtes der Europäischen Gemeinschaften (Eurostat) für die EU-weite Zensusrunde 2000/2001 beide Haushaltsdefinitionen akzeptiert.

1) Siehe hierzu Ginter, D.: „Zensustest – Ergebnisse der Gebäude- und Wohnungszählung“ in WiSta 11/2004, S. 1256 ff.

2) Siehe Statistische Ämter des Bundes und der Länder: „Ergebnisse des Zensustests“ in WiSta 8/2004, S. 813 ff.; Michel, B.: „Ergebnisse der Mehrfachfallprüfung im Rahmen des Zensustests“ in WiSta 9/2004, S. 979 ff. sowie Ginter, D., Fußnote 1.

3) Alle Ergebnisse basieren auf dem ausführlichen Abschlussbericht über das Verfahren der „Zusammenführung/Haushaltegenerierung“.

Die Ergebnisse in den folgenden Ausführungen beziehen sich nur auf private Haushalte (ohne Haushalte, die in Wohnheimen leben).

### 1 Zielsetzungen

Wichtigstes Ziel im Rahmen des Zensustests für die Arbeitsgruppe „Zusammenführung/Haushaltegenerierung“ war es, ein Verfahren zu entwickeln, mit dem in einem registergestützten Zensus kleinräumig zuverlässige Informationen über Zahl und Struktur von Haushalten sowie deren Wohnsituation bereitgestellt werden können. Darüber hinaus sollte die Eignung des Verfahrens der Zusammenführung/Haushaltegenerierung als ein Instrument zur Qualitätskontrolle der Melderegister untersucht werden. Zu diesem Zweck war zu prüfen, inwieweit mit diesem Verfahren Fehler des Melderegisters aufgedeckt werden können, die dann maschinell oder durch Rückfragen geklärt werden können. Fehler des Melderegisters liegen vor, wenn entweder Personen in den Melderegisterdaten einer Adresse enthalten sind, aber nicht in der Haushaltebefragung vorgefunden werden – so genannte Karteileichen – oder umgekehrt, wenn Personen nicht in den Melderegisterdaten enthalten sind, aber in der Haushaltebefragung als an dieser Adresse wohnhaft festgestellt werden, so genannte Fehlbestände.

### 2 Kurzdarstellung des Verfahrens der Zusammenführung/Haushaltegenerierung

Zur Bildung von Haushalten werden in einem ersten Schritt Personen über ihre in den Melderegistern gespeicherten Verzeigerungen, das heißt die zu jeder Person in den Melderegistern eingetragenen Namen vorhandener Ehegatten, Kinder und gesetzlicher Vertreter, zu so genannten Kernhaushalten zusammengeführt.

Nach diesem Schritt werden die bewohnten Wohnungen aus der Gebäude- und Wohnungszählung über die für jede Wohnung erfragten Namen der Wohnungsinhaber mit den Melderegister-Datensätzen verknüpft. Durch die zuvor durchgeführte Kernhaushaltbildung können nun, nachdem der Name des Wohnungsinhabers im Melderegister gefunden wurde, nicht nur der Wohnungsinhabers, sondern zugleich alle mit ihm verzeigerten Personen mit den Wohnungsangaben aus der GWZ verknüpft werden.

Im nächsten Schritt werden weitere aus dem Melderegister zur Verfügung stehende Informationen wie beispielsweise Namensgleichheiten, gleiches Einzugsdatum, gleiche frühere Wohnung genutzt, um Personen bzw. Kernhaushalte zu Personenverbänden (Wohnhaushalte) zusammenzuführen. Verbleibende Personen, die danach noch nicht mit einer Wohnung verknüpft werden konnten, werden mithilfe wohnungsstatistischer Daten wie zum Beispiel der Anzahl der Personen in einer Wohnung oder der Wohnfläche den Wohnungen zugeordnet, wobei die Obergrenze für die Zahl

der Wohnhaushalte die aus der GWZ genannte Anzahl der bewohnten Wohnungen darstellt.

### 3 Die zur Haushaltegenerierung genutzten Informationen

Bereits seit den 1980er-Jahren sind Bestrebungen vorhanden, die Datenbestände der Melderegister zu nutzen, um kleinräumig gegliederte aktuelle Daten über das Zusammenleben von Personen zu erhalten. In diesen beispielsweise vom KOSIS-Verbund<sup>4)</sup> entwickelten Verfahren werden ausschließlich die Informationen aus dem Melderegister zur Generierung von Haushalten herangezogen. Das im Rahmen des Zensustests entwickelte Verfahren verwendet zusätzlich zu den Angaben des Melderegisters auch gebäude- und wohnungsstatistische Informationen, um Personen zu Wohnhaushalten zusammenzuführen.

Im Einzelnen können folgende Informationen aus den Individualdaten des Melderegisters für die Zusammenführung von Personen zu Haushalten genutzt werden: Namen, Vornamen; gegenwärtige Anschriften; Geburtsdatum; Geschlecht; Staatsangehörigkeit; Familienstand; Wohnstatus; Datum des Beziehens der Wohnung; Datum des Auszugs aus der Wohnung; Datum des Wohnungsstatuswechsels; Ordnungsmerkmal der Meldebehörde für die gemeldete Person; Namen, Vornamen und Geburtsdatum des Ehegatten; Namen, Vornamen und Geburtsdatum der Kinder; bei Kindern: Namen, Vornamen und Geburtsdatum des gesetzlichen Vertreters; Ordnungsmerkmale der Meldebehörde für Ehegatten, Kinder und deren gesetzliche Vertreter; Datum der letzten Eheschließung; Datum der Beendigung der letzten Ehe; Anschrift und Status der Wohnung in der Gemeinde, aus der der Einwohner zugezogen ist; Datum des Zuzugs in die Gemeinde; Zuzug aus dem Ausland; Datum des Fortzugs ins Ausland; Anschrift der zuletzt bewohnten Wohnung in der Gemeinde; Name und Anschrift des Wohnungsgebers.

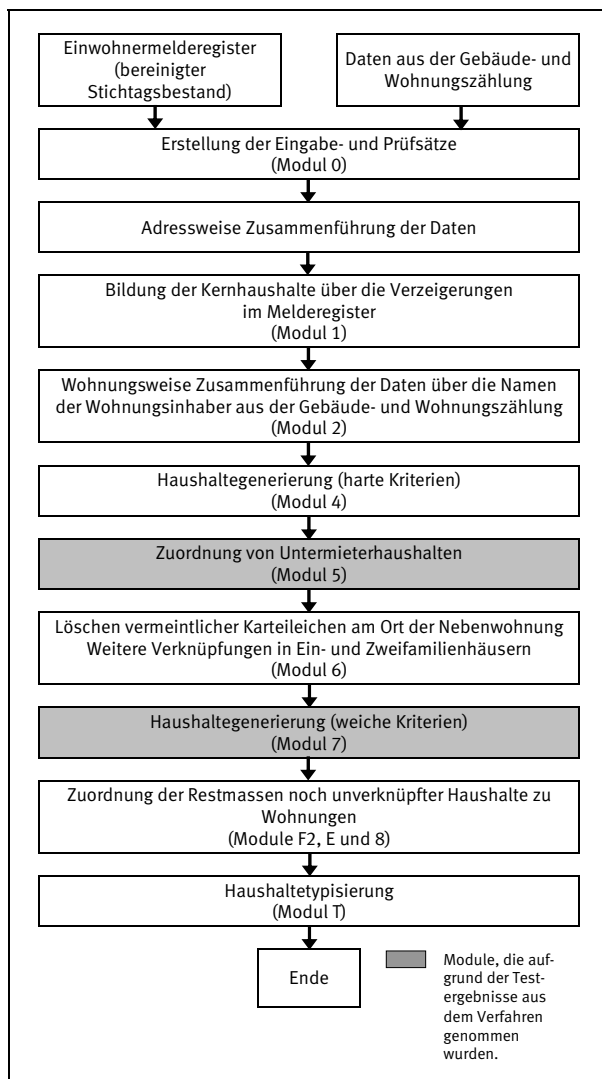
Aus der im Zensustest erprobten Gebäude- und Wohnungszählung können folgende Angaben für die Zusammenführung von Personen zu Haushalten verwendet werden: Anschrift des Gebäudes; Zahl der Wohnungen im Gebäude; Zahl der leerstehenden Wohnungen; Art der Nutzung der Wohnung; Zahl der Personen in der Wohnung; Namen und Vornamen der Wohnungsinhaber; Datum des Einzugs oder Beginn des Mietvertrags.

### 4 Darstellung des Verfahrens der Zusammenführung/Haushaltegenerierung

Das komplexe Verfahren der Zusammenführung/Haushaltegenerierung ist modular aufgebaut, die wichtigsten Programmteile enthält das Schaubild.

4) KOSIS-Verbund: Kommunales Statistisches Informationssystem.

## Das Verfahren der Zusammenführung/Haushalgenerierung



#### 4.1 Erstellung der Eingabesätze (Modul 0)

Zur Vorbereitung auf die einzelnen Arbeitsschritte werden die Eingangsmaterialien aus den Melderegistern und der Gebäude- und Wohnungszählung auf eindeutige und einheitliche Feldinhalte geprüft und um weitere für die Generierung notwendige Felder ergänzt (Modul 0).

In einem ersten Schritt werden die Daten dann aus dem Melderegister mit den Angaben aus der Gebäude- und Wohnungszählung auf Adressebene zusammengeführt. Eine Aufgabe, die bei einem künftigen Zensus – im Gegensatz zum Zensus, bei dem einheitliche numerische Schlüssel für Adressen vorhanden waren – den Aufbau eines Adressregisters erfordert.

#### 4.2 Bildung der Kernhaushalte (Modul 1)

Danach werden aus den Melderegisterdaten die so genannten Kernhaushalte gebildet (Modul 1). Alle Personen, die

gemäß Melderegister an der gleichen Adresse wohnen und miteinander verzeigert sind, und zwar als Ehegatten untereinander und/oder als Elternteil mit Nachkommen bzw. als Kind mit ihren gesetzlichen Vertretern (in der Regel Vater und Mutter) erhalten die gleiche Kernhaushaltsnummer und bilden somit einen Haushalt. Dadurch lassen sich bereits aus den Melderegisterdaten die klassischen Familienzusammenhänge erkennen. Alle Personen, die nicht einer anderen Person zugeordnet werden können, bekommen als Einpersonen-Kernhaushalte eine eigene Kernhaushaltsnummer.

#### 4.3 Verknüpfung der Wohnungsätze aus der GWZ mit den Personendaten des Melderegisters (Modul 2)

Im Anschluss an die Kernhaushaltbildung werden die Wohnungsdatensätze aus der Gebäude- und Wohnungszählung mit Personendatensätzen des Melderegisters verknüpft (Modul 2). Da hierfür kein Personenkennzeichen verfügbar ist, erfolgt diese Zusammenführung über die Namen der Wohnungsinhaber, die in der GWZ für jede Wohnung erfragt wurden (je Wohnung sollten zwei Wohnungsinhaber genannt werden). Hierzu wurde ein vom Bayerischen Landesamt für Statistik und Datenverarbeitung entwickeltes Namensabgleichprogramm verwendet, das zum Zweck der Zusammenführung von Personen auf Adressebene bei den Namen auch unterschiedliche Schreib- und Abkürzungsweisen sowie Doppelnamen berücksichtigt.<sup>5)</sup> Dieser maschinelle Namensabgleich erfolgt stufenweise und beginnt mit einer äußerst restriktiven Stufe, die nur dann einen GWZ-Wohnungsinhaber mit einem Personensatz des Melderegisters verknüpft, wenn der GWZ-Name und der Melderegister-Name exakt übereinstimmen. Die Suche nach allen weiteren Wohnungsinhabern eines Gebäudes wird Schritt für Schritt auf die Melderegistersätze beschränkt, die noch nicht mit einem Wohnungsinhaber verknüpft wurden (Prinzip der sukzessiven Massenreduktion). Mit jeder weiteren Stufe werden die Zuordnungskriterien gelockert und schließlich auch ein „phonetischer“ Abgleich durchgeführt, das heißt es ist ausreichend, wenn die Namen einen bestimmten Grad an Ähnlichkeiten aufweisen. Insgesamt wurden im Zensus mit diesem maschinellen Namensabgleichprogramm 92,4% der in der GWZ genannten Wohnungsinhaber im Melderegister identifiziert. Vermutlich wäre diese Trefferquote noch höher ausgefallen, wenn es bei der Durchführung des Zensus nicht zu Missverständnissen beim Begriff des in der Gebäude- und Wohnungszählung erfragten Wohnungsinhabers gekommen wäre. Statt des Namens eines Bewohners der entsprechenden Wohnung wurde teilweise der an einer anderen Adresse wohnende Besitzer (Eigentümer) der Wohnung genannt, der somit nicht im Melderegister gefunden werden konnte. Ein zweiter Grund, weshalb nicht alle aus der GWZ genannten Wohnungsinhaber in den Melderegisterdaten der entsprechenden Adresse gefunden werden konnten, war, dass es sich bei einem Teil der Wohnungsinhaber um Personen handelte, die nicht im Melderegister enthalten waren (Fehlbestände). Der sich

5) Ausführliche Beschreibung des Verfahrens siehe Fürnrohr, M./Rimmelspacher, B./von Roncador, T.: „Zusammenführung von Datenbeständen ohne numerische Identifikatoren“ in Bayern in Zahlen 7/2002, S. 308 ff.

daran anschließende manuelle Namensabgleich, der noch nicht verknüpfte Personen aus dem Melderegister und der GWZ über ein Dialogverfahren zuordnete, erhöhte die Trefferquote um 0,5%. Die Ergebnisse für den Namensabgleich zeigen auch, dass sich die Erfassung zweier Wohnungsinhaber bei der Zusammenführung von Wohnungen über den Namensabgleich bewährt hat. Rund 213 000 Wohnungen, bei denen der erste Wohnungsinhaber nicht im Melderegister gefunden wurde, konnten über den zweiten Wohnungsinhaber verknüpft werden.

Nach dieser Verknüpfung der Wohnungsinhaber mit den Melderegisterdaten erhalten nun aufgrund der Kernhaushaltbildung im vorangegangenen Programmteil nicht nur der Wohnungsinhaber, sondern alle mit ihm verbundenen Familienangehörigen die entsprechenden Wohnungsinformationen.

Mit diesem Modul konnten bereits 85,3% der Bevölkerung mit Hauptwohnsitz und 36,4% der Nebenwohnsitz-Bevölkerung mit einem Wohnungsdatensatz zusammengeführt werden (siehe Tabelle 2).

#### 4.4 Haushaltegenerierung nach „harten“ Kriterien (Modul 4)

Im nächsten Schritt des Verfahrens (Modul 4) werden bestimmte Merkmalsübereinstimmungen bzw. Merkmalskombinationen zwischen den Personensätzen aus den Melderegistern, teilweise ergänzt um Informationen aus der Gebäude- und Wohnungszählung, genutzt, um Personen zu Personenverbänden (Haushalten) zusammenzuführen. Die Reihenfolge des zehn Stufen umfassenden Algorithmus ist so angelegt, dass die weniger wahrscheinlichen Haushaltszusammenhänge von Personen erst auf einer späteren Stufe erfolgen.

Für die nachfolgend im Einzelnen beschriebenen Haushaltegenerierungsschritte gilt, dass sich diese stets auf Personen, die noch keinen gemeinsamen Haushalt bilden, also unterschiedliche Kernhaushaltsnummern aufweisen, beziehen. Wird in den Stufen ein Haushaltszusammenhang zwischen zwei Personen erkannt, so werden gleichzeitig die mit ihm verzeigten Kernhaushaltsmitglieder mit der entsprechenden Wohnung verknüpft. Des Weiteren werden keine Zuordnungen von Personen zugelassen, wenn diese Wohnungsinhaber einer eigenen Wohnung sind.

In der *ersten Stufe* werden Personen, die zunächst bei der Kernhaushaltbildung (Modul 1) wegen fehlender Verzeigerungen nicht zusammengeführt wurden, aber im darauffolgenden Schritt bei der Verknüpfung der Melderegistersätze mit den Wohnungssätzen (Modul 2) als Wohnungsinhaber derselben Wohnung erkannt wurden, zu einem Haushalt zusammengeführt. Insgesamt konnten in dieser Stufe etwa 2 Mill. Personen aufgrund identischer Wohnungsnummern, die zuvor noch keinen gemeinsamen Haushalt bildeten, zugeordnet werden (siehe Tabelle 1). Ein Vergleich mit den Ergebnissen der Haushaltebefragung zeigt, dass 91,5% dieser Zusammenführungen korrekt waren. Diese hohe Zahl an Zusammenführungen ist ein weiterer Beleg, dass sich die Abfrage zweier Wohnungsinhaber in der Gebäude-

Tabelle 1: Generierungserfolg der einzelnen Stufen des Moduls 4  
Hochgerechnetes Ergebnis des Zensustests vom 5. Dezember 2001

| Generierungsstufen | Beschreibung der Stufe   | In Modul 4 zu Haushalten mit Wohnungsnummern hinzugenerierte Personen |                              |
|--------------------|--|---|------------------------------|
|                    |  | insgesamt   | dar.: Wohnungsnummer richtig |
|                    |  | 1 000   | %                            |
| 1                  | Zusammenführung von Personen aufgrund gleicher Wohnungsnummer .....  | 1 976,9   | 91,5                         |
| 2                  | Zusammenführung von Personen aufgrund der Information zum Wohnungsgeber .....  | 422,0   | 45,6                         |
| 3                  | Zuordnung nicht- bzw. nicht vollständig verzeigter minderjähriger Kinder .....                                       | 382,0   | 92,4                         |
| 4                  | Bildung nichtehelicher Paare über Verzeigerungen zwischen Kind und nichtehelichem Vater .....                        | 352,5   | 93,2                         |
| 5                  | Zusammenführung von nicht verzeigten Ehepartnern .....   | 2,9   | 94,7                         |
| 6                  | Generierung von Nachkommen zu Eltern(teilen) bzw. Großeltern(teilen) .....   | 2 325,5   | 79,4                         |
| 7                  | Generierung älterer Personen zu Nachkommen .....   | 324,5   | 76,9                         |
| 8                  | Generierung erwachsener Personen über einen Geschwistersammenhang .....  | 92,8  | 65,6                         |
| 9                  | Generierung nichtehelicher Paare ..  | 356,6   | 77,2                         |
| 10                 | Generierung von Einzelpersonen zu Haushalten aufgrund der Identität von Einzugsdatum und/oder früherer Adresse ..... | 403,1   | 57,8                         |

und Wohnungszählung bewährt hat. Außerdem konnten durch die Angabe zweier Wohnungsinhaber auf dieser Stufe bereits nichteheliche Lebensgemeinschaften gebildet werden, die sonst nicht oder erst auf einer späteren Stufe mit weniger zuverlässigen Informationen hätten in einen Haushaltszusammenhang gebracht werden können.

In der *zweiten Stufe* werden Personen über das im Melderegisterdatensatz enthaltene Feld „Wohnungsgeber“ zusammengeführt. Ist im Melderegisterdatensatz der Person A ein Wohnungsgeber (Person B) eingetragen, so wird dieser über einen Namensabgleich in den Melderegisterdaten derselben Adresse gesucht. Wird Person B gefunden, so werden Person A und B zu einem Haushalt verknüpft. Fälle, für die im Melderegister das Feld „Wohnungsgeber“ belegt sein kann, sind beispielsweise später zugezogene Familienmitglieder, Lebenspartner oder auch Untermieter. Diese Generierungsstufe brachte jedoch noch keine ausreichenden Ergebnisse, wie der Anteil der richtigen Zusammenführungen von nur 45,6% zeigt. Die Ursache hierfür ist bei der Weiterentwicklung des Verfahrens zu suchen und zu beseitigen.

In der *dritten Stufe* werden nicht oder nicht vollständig verzeigte minderjährige Kinder mit Eltern(teilen), Großeltern(teilen), Geschwistern oder mit ihrem gesetzlichen Vertreter zusammengeführt. Bedingung für eine Zusammenführung ist, dass der Familienname des Kindes mit dem Familiennamen/Geburtsnamen/früheren Familiennamen der Eltern-/Großeltern(teile)/Geschwister übereinstimmt und die vorgegebenen Altersdifferenzen nicht über-

schrritten werden. Sind keine Familienmitglieder auffindbar, so erfolgt eine Prüfung, ob das Kind einen gesetzlichen Vertreter hat; sofern vorhanden, wird es diesem zugeordnet. Damit konnten insgesamt 382 000 minderjährige Kinder einem Haushalt zugeordnet werden; das Generierungsergebnis stimmt in 92,4% der Fälle mit dem Ergebnis der Haushalbefragung überein.

Über vorhandene Verzeigerungen eines Kindes sowohl zu seinem nichtehelichen Vater als auch zur Mutter werden in der *vierten Stufe* die beiden Elternteile zu einer nichtehelichen Lebensgemeinschaft zusammengeführt. Bundesweit wurden damit 352 500 Personen zugeordnet, davon 93,2% richtig.

*Stufe 5* führt nicht verzeigerte Ehegatten über einen Abgleich des Familiennamens unter Berücksichtigung des Familienstandes, des Geschlechts und bei gleichem Datum der Familienstandsänderung<sup>6)</sup> zusammen. Die geringe Fallzahl von nur 2 900 Fällen lässt vermuten, dass die Verzeigerungen im Melderegister für Ehepartner fast immer vorhanden sind, doch haben Untersuchungen mit dem Testmaterial ergeben, dass auf dieser Stufe eine größere Zahl von Generierungen hätte erfolgen müssen. Deshalb ist diese Stufe in Vorbereitung eines registergestützten Zensus weiter zu analysieren und zu verbessern.

In der *sechsten Stufe* werden über Namensabgleiche – analog der dritten Generierungsstufe – nicht verzeigerte erwachsene Nachkommen ihren Eltern(teilen) bzw. Großeltern(teilen) zugeordnet. Nach Vorliegen der ersten Ergebnisse musste diese Stufe überarbeitet werden, da mit diesem Zuordnungsschritt mehr Karteileichen als erwartet den Haushalten zugeführt wurden. Eine Erklärung hierfür ist, dass viele erwachsene Kinder (beispielsweise Studierende) noch bei ihren Eltern gemeldet sind, aber nicht mehr dort wohnen und somit in der Haushalbefragung nicht angegeben wurden. Um zu verhindern, dass zu viele Karteileichen bestehenden Haushalten zugeordnet werden,

wurde das Verfahren der Haushaltegenerierung ab Stufe 6 dahingehend modifiziert, dass – zuzüglich zu den jeweils vorgesehenen Prüfschritten – Personen nur dann zugeordnet werden, wenn die in der Gebäude- und Wohnungszählung erfragte „Anzahl der Personen in einer Wohnung“<sup>7)</sup> für diesen Wohnhaushalt nicht überschritten wird. Mit diesem so genannten „Blocker“ konnte die Karteileichenrate der zugeordneten Personen in der Stufe 6 um über ein Drittel gesenkt werden. Insgesamt ist diese Stufe mit 2,3 Mill. zugeordneten Personen, das sind rund 35% aller Generierungen, im Modul 4 am bedeutendsten, allerdings liegt hier die Trefferquote mit 79,4% – mit Ausnahme der zweiten – auch deutlich niedriger als auf den vorangegangenen Stufen.

Im nächsten Generierungsschritt (*Stufe 7*) werden nun für allein stehende Elternteile oder Ehepaare über den Abgleich der Familien-, Geburts- bzw. früheren Familiennamen unter Berücksichtigung bestimmter Altersdifferenzen ihre erwachsenen Nachkommen gesucht und diesem Haushalt zugeordnet. Auf diese Weise konnten 324 500 Eltern(teile) den Haushalten ihrer Nachkommen zugeführt werden, allerdings mit einer Fehlerquote von 23,1%.

In der *achten Stufe* wird über die Identität von Familien- bzw. Geburtsnamen sowie entsprechender Altersdifferenzen geprüft, ob ein Geschwisterzusammenhang zwischen zwei Personen/Haushalten vorliegt. Über diese Kriterien wurde ein Geschwisterzusammenhang für 92 800 Personen erkannt, fast zwei Drittel davon richtig.

*Stufe 9* führt Personen (mit oder ohne Nachkommen) in Abhängigkeit vom Geschlecht, vorgegebenen Altersdifferenzen, gleichem Einzugsdatum oder gleicher früherer Adresse zu nichtehelichen Lebensgemeinschaften zusammen. Aufgrund dieser Bedingungen wurden 356 600 Personen einem Haushalt zugeordnet, die Übereinstimmungsquote mit der Haushalbefragung lag bei gut 77%.

Tabelle 2: Personen und Karteileichen nach Modulen und Wohnstatus  
Hochgerechnetes Ergebnis des Zensus-tests vom 5. Dezember 2001

| Modul                               | Personen laut Melderegister | Dar.: Karteileichen |                | Mit einer Wohnung verknüpft |                | Dar.: Karteileichen |                | Noch unverknüpft |                | Dar.: 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 |                             |                     |                |                             |                |                     |                |                  |                |                     |                |
| nach Modul 2 .....                  | 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           |
| nach Modul 4 .....                  | 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           |
| nach Modul 6 .....                  | 77 959,3                    | 2 176,8             | 2,8            | 72 637,5                    | 93,2           | 1 094,0             | 1,5            | 5 321,8          | 6,8            | 1 082,8             | 20,3           |
| Bevölkerung am Ort der Nebenwohnung |                             |                     |                |                             |                |                     |                |                  |                |                     |                |
| nach Modul 2 .....                  | 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           |
| nach Modul 4 .....                  | 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           |
| nach Modul 6 .....                  | 2 329,4                     | 796,4               | 34,2           | 1 978,4                     | 84,9           | 509,3               | 25,7           | 350,9            | 15,1           | 287,1               | 81,8           |

6) Das Feld Familienstandsänderung wird in Modul 0 belegt, es ist das jüngere Datum aus den Feldern „Datum der letzten Eheschließung“ und „Datum Beendigung letzte Ehe“ aus den Melderegistern.

7) Die Ergebnisse des Zensus-tests belegen, dass die Angabe aus der GWZ zur „Anzahl der Personen in der Wohnung“ – sofern sie erfolgte – häufig den Ergebnissen aus der Haushalbefragung entsprach.

In der zehnten und letzten Stufe von Modul 4 werden schließlich Personen aufgrund der Identität von Einzugsdatum und/oder gleicher früherer Adresse zusammengeführt. Die Zuordnungskriterien dieser letzten Stufe sind stärker fehleranfällig und weisen nach den Zuordnungen in der zweiten Stufe auch den höchsten Fehler auf. 42,2% von 403 100 Zuordnungen haben sich bei dem Vergleich mit den Ergebnissen der Haushalbefragung als nicht richtig erwiesen.

Insgesamt wurden mit der Haushalgenerierung nach diesen so genannten „harten Kriterien“ 6,6 Mill. Personen zugeordnet. Am Ende von Modul 4 waren 90,8% aller Personen mit Hauptwohnung mit einer Wohnung verknüpft.

Ein weiterer mit großem Aufwand entwickelter Verfahrensteil (Modul 7) zur Zusammenführung von Personen zu Haushalten nach so genannten „weichen“ Kriterien, etwa der Staatsangehörigkeit, wurde nach Vorliegen der Testergebnisse wegen zu hoher Fehlerraten aus dem Verfahren genommen. Ebenso erfolglos blieb der Versuch zur Abbildung von Untermieterhaushalten (Modul 5). Wegen mangelnder Übereinstimmung der in der Gebäude- und Wohnungszählung und der Haushalbefragung angegebenen Untermieterwohnungen war eine Überprüfung der erfolgten Zuordnungen nur sehr eingeschränkt möglich, sodass auch dieses Modul aus dem Generierungsverfahren entfernt wurde.

### 4.5 Weitere Verknüpfungen in Ein- und Zweifamilienhäusern (Modul 6)

Aufgrund der Ergebnisse des Zensustests neu entwickelt wurde hingegen das Modul 6. Zum einen werden in diesem Programmteil Karteileichen am Ort der Nebenwohnung maschinell gelöscht, da sich gezeigt hat, dass nach den Zuordnungen in Modul 4 noch nicht verknüpfte Nebenwohnsitz-Personen zu 80% Karteileichen sind (siehe hierzu auch Kapitel 6: Eignung des Haushalgenerierungsverfahrens zur Kontrolle der Melderegister).

Zum anderen wird die aus den Testergebnissen gewonnene Erkenntnis, dass die Angabe „Anzahl der Personen in der Wohnung“ gemäß GWZ im Ein- und Zweifamilienhausbereich von besonders guter Qualität ist, genutzt, um mit Hilfe dieser Angabe den Generierungsprozess für weitere Adressen abzuschließen.

Zu diesem Zweck werden in einem ersten Schritt noch nicht verknüpfte Wohnungen im Ein- und Zweifamilienhausbereich belegt, und zwar vorrangig mit unverknüpften Haushalten, bei denen die Personenzahl mit der Angabe zur Zahl der Personen aus der Gebäude- und Wohnungszählung übereinstimmt. In einem zweiten Schritt werden bereits mit einem Haushalt belegte Wohnungen mit – gemessen an der Personenzahl je Wohnung – noch freien Kapazitäten bis zur Personenzahl laut GWZ mit Hilfe der Informationen zum Wohnstatus und Geburtsdatum aus dem Melderegister aufgefüllt. Der Anteil der verknüpften Personen am Ort der Haupt- oder alleinigen Wohnung hat sich nach diesem Modul um 2,6% auf 93,2% erhöht.

### 4.6 Zuordnung der Restmassen (Module F2, E und 8)

Diese Programmschritte dienen der abschließenden Zuordnung der Restmassen noch unverknüpfter Personen/Haushalte zu Wohnungen.

Mit Modul F2 werden Adressen mit Wohnungen, die bisher noch nicht mit einem Personensatz aus dem Melderegister verbunden werden konnten, abgeschlossen, sofern die Anzahl der nicht verknüpften Wohnungen kleiner oder gleich der Anzahl der nicht verknüpften Haushalte ist. Die Verknüpfungen erfolgen nach der Größe in absteigender Reihenfolge, das heißt der größte Haushalt wird der größten Wohnung zugeordnet.

Mit Modul E werden alle noch nicht mit einer Wohnung zusammengeführten Personen/Haushalte in Einfamilienhäusern zu einem Wohnhaushalt zusammengefasst.

Anschließend werden im Modul 8 alle übrigen Gebäudegrößen abgeschlossen. Auch dieses Modul wurde aufgrund der Testergebnisse weiterentwickelt. Die ursprüngliche Fassung von Modul 8 sah vor, die auf dieser letzten Stufe noch vorhandenen geringen Restmassen ausschließlich nach wohnungsstatistischen Kriterien, etwa der Wohnfläche, auf die Wohnungen der jeweiligen Adresse zu verteilen. Die Testergebnisse haben aber gezeigt, dass mit dieser Verfahrensweise zu viele Fehler produziert werden, insbesondere auch deshalb, weil nach dem Wegfall von Modul 7 wesentlich mehr Personen als ursprünglich erwartet noch nicht mit einer Wohnung verknüpft werden konnten und deshalb in dieser letzten Generierungsstufe verarbeitet werden mussten.

Die Analyse der in diesem Programmteil erfolgten Zuordnungen mit den Testergebnissen ergab außerdem, dass viele vor diesem Modul richtige Haushaltszusammenhänge durch diese Verknüpfungen nach rein statistischen Kriterien zerstört wurden, da sich in dieser Restmasse der Personen ein hoher Karteileichenanteil befindet. Zudem stellte sich heraus, dass mit den bisherigen Verfahrensschritten – im Vergleich zur Haushalbefragung – einerseits zu wenige Haushalte vom Typ „nichteheliche Lebensgemeinschaften“ erzeugt und andererseits zu viele Elternteile mit Kindern sowie erwachsene Einzelpersonen noch nicht mit anderen Personen zu einem Haushalt verknüpft worden waren.

Aus diesen Erkenntnissen heraus wurde das Modul 8 wie folgt überarbeitet: Die erste Stufe, in der sämtliche noch nicht mit einem Haushalt zusammengeführten Wohnungen mit jeweils einem – ebenfalls unverknüpften – Haushalt belegt werden, bleibt. In der zweiten Stufe wurde ein Programmteil zur Erzeugung weiterer Haushalte des Typs „nichteheliche Lebensgemeinschaften“ eingefügt, der analog der neunten Stufe im Modul 4, aber ohne die Bedingung eines gleichen Einzugsdatums bzw. gleicher früherer Adresse, Haushalte generiert. Des Weiteren wurden nach diesem Schritt die restlichen Personen nur noch auf Wohnungen mit „sonstigen Haushalten“ verteilt. Von dieser Regelung wurde nur dann abgewichen, wenn aufgrund dieser Bedingungen keine Wohnung mehr für eine Zuordnung

von Personen zur Verfügung gestanden hätte. Damit konnten die Ergebnisse dieses Programmteils wesentlich verbessert werden, trotzdem sind noch weitere Untersuchungen vorzunehmen, um die Zuordnungsregeln zu verfeinern. Nach Modul 8 sind alle Personen des Melderegisters einer Wohnung zugeordnet.

#### 4.7 Haushaltstypisierung (Modul T)

Für die Nachweisung von Ergebnissen nach Haushaltstypen werden im letzten Arbeitsschritt des Haushaltgenerierungsverfahrens (Modul T) die entsprechenden Zuordnungen (Typisierungen) vorgenommen. Zur Abgrenzung von Haushalten wurde in der Haushalthebefragung – analog der Vorgehensweise im Mikrozensus – eine Bezugsperson festgelegt. Es ist die Person, die im Erhebungsbogen als erste Person eingetragen ist, wobei hierfür eine Reihenfolge vorgegeben wurde (Ehegatten/Lebenspartner, Kinder, Verwandte, sonstige Personen). Angaben zur Festlegung einer Bezugsperson liegen im Melderegister nicht vor. Deshalb wurde zum Zweck der Haushaltstypisierung im Verfahren der Zusammenführung/Haushaltgenerierung folgende Vorgehensweise gewählt: Vor Durchführung der Haushaltgenerierung nach „harten“ Kriterien (Modul 4) wird für jeden Kernhaushalt eine Bezugsperson festgelegt. Bezugsperson bei einem Einpersonenhaushalt ist die Person selbst, bei einem Ehepaar ohne/mit Kind(ern) der Ehemann und bei einem allein lebenden Elternteil mit Kind(ern) der Elternteil. Wird in den darauf folgenden Generierungsschritten eine Person oder ein Kernhaushalt einem Haushalt zugeordnet, so überprüft das Verfahren den neu entstehenden Haushaltszusammenhang. Das Merkmal „Bezugsperson“ bei einem der zwei zusammengeführten Haushalte wird durch ein sich aus dem neuen Haushaltszusammenhang ergebendes Verhältnis zur Bezugsperson, zum Beispiel Lebenspartner, ersetzt.

Damit ist gewährleistet, dass für jeden Haushalt nur eine Bezugsperson zur Abgrenzung von Haushaltszusammenhängen vorhanden ist.<sup>8)</sup> Gleichzeitig wird bei jedem Zuordnungsschritt die Stellung der übrigen Personen des Haushalts zur Bezugsperson überprüft und – wenn erforderlich – korrigiert.

Mit diesen im Verfahren ermittelten Informationen zur Beziehung der Personen in einem Haushalt und den für jede Person ebenfalls im Generierungsprozess ermittelten Generationsziffern können dann unterschiedliche Haushaltstypen sowie auch Mehrgenerationenhaushalte abgebildet werden. Für den Ergebnisvergleich zwischen Haushaltgenerierung und Haushalthebefragung werden nur die folgenden Haushaltstypen betrachtet: Ehepaare ohne/mit Kind(ern); nichteheliche Lebensgemeinschaften ohne/mit Kind(ern), allein lebende Elternteile mit Kind(ern) sowie „sonstige“ Haushalte, in denen alle übrigen Haushaltszusammenhänge subsumiert werden.

Die Qualität der nun nach Abschluss des Verfahrens vorliegenden Haushaltsgrößen und -strukturen wird im nächsten Kapitel beurteilt.

## 5 Ergebnisse der Haushaltgenerierung im Vergleich zur Haushalthebefragung

Als Maßstab für die Beurteilung, wie gut das Verfahren der Zusammenführung/Haushaltgenerierung in der Lage ist, die Zahl und Struktur der Wohnhaushalte abzubilden, standen die Ergebnisse der Haushalthebefragung zur Verfügung. Wohnungen, für die in der Haushalthebefragung aufgrund von Antwortverweigerungen (vollständig oder nur partiell zu einzelnen Fragen) oder infolge des Nichtantretrens einzelner Haushalte kein (vollständiges) Befragungsergebnis erzielt werden konnte oder die offensichtliche Unplausibilitäten enthielten, können nicht als Vergleichsmaßstab herangezogen werden. Sie werden in den hier dargestellten Ergebnisvergleichen nicht berücksichtigt. Dadurch liegen die Haushaltszahlen in den folgenden Tabellen unter der Zahl der Haushalte für Deutschland insgesamt.

Weiterhin ist zur Bewertung der Ergebnisse des Verfahrens der Zusammenführung/Haushaltgenerierung zwischen zwei Fehlerursachen zu differenzieren. Abweichende Ergebnisse zur Vergleichsgröße „Haushalthebefragung“ können aus Fehlern in den verwendeten Datenquellen (Melderegister oder Gebäude- und Wohnungszählung) resultieren oder sind auf ein fehlerhaft arbeitendes Verfahren zurückzuführen. Um die Einflüsse der unterschiedlichen Fehlerquellen auf die Haushaltsergebnisse messen zu können, werden die Ergebnisse des Verfahrens sowohl unter Einbeziehung von Fehlern des Melderegisters (Karteileichen/Fehlbestände) und der GWZ als auch unter Ausschluss dieser Fehler in den verwendeten Datenquellen dargestellt. Letzteres dient zur Abschätzung der Leistungsfähigkeit des Verfahrens an sich.

### 5.1 Haushalte nach Haushaltsgröße

Bei einem Vergleich der Gesamthaushaltszahlen ist festzustellen, dass mit dem Generierungsverfahren insgesamt weniger Wohnhaushalte gebildet als in der Haushalthebefragung festgestellt wurden. Hauptursache hierfür ist, dass teilweise Wohnungen aus der Gebäude- und Wohnungszählung nicht mit Personendatensätzen aus dem Melderegister zusammengeführt werden konnten, weil die aus der GWZ angegebenen Wohnungsinhaber Fehlbestände des Melderegisters waren und somit nicht im Melderegister gefunden werden konnten.

Werden die Ergebnisse unter Einbeziehung der Register- und GWZ-Fehler betrachtet, fällt auf, dass die Anzahl der kleinen Haushalte unterschätzt und die der größeren Haushalte, vor allem der mit fünf oder mehr Personen, überschätzt wer-

<sup>8)</sup> Ausgenommen hiervon sind Wohngemeinschaften. Bei der Zuordnung von Personen zu Wohngemeinschaften wird das ursprünglich vergebene Merkmal „Bezugsperson“ nicht überschrieben, da keine familiären bzw. partnerschaftlichen Verhältnisse festgestellt werden können. Wohngemeinschaften werden als „sonstige Haushalte“ typisiert.

Tabelle 3: Haushalte laut Haushaltegenerierung und Haushaltebefragung nach Haushaltsgröße  
Hochgerechnetes Ergebnis des Zensustests vom 5. Dezember 2001

| Haushaltsgröße                                | Haushalte laut |                    | Abweichung |        |
|---|----------------|--------------------|------------|--------|
|   | Generierung    | Haushaltebefragung |            |        |
|   | 1 000          |                    | %          |        |
| Haushalte mit Register- und GWZ-Fehler .....  | 29 612,4       | 30 068,8           | - 456,4    | - 1,5  |
| davon mit ...                                 |                |                    |            |        |
| 1 Person .....                                | 9 006,5        | 9 517,1            | - 510,6    | - 5,4  |
| 2 Personen .....                              | 10 138,3       | 10 376,8           | - 238,5    | - 2,3  |
| 3 Personen .....                              | 4 760,3        | 4 690,3            | + 70,0     | + 1,5  |
| 4 Personen .....                              | 3 945,0        | 3 901,4            | + 43,6     | + 1,1  |
| 5 oder mehr Personen .....                    | 1 762,3        | 1 583,2            | + 179,1    | + 11,3 |
| Haushalte ohne Register- und GWZ-Fehler ..... | 25 398,8       | 25 499,4           | - 100,6    | - 0,4  |
| davon mit ...                                 |                |                    |            |        |
| 1 Person .....                                | 7 452,1        | 7 509,1            | - 57,1     | - 0,8  |
| 2 Personen .....                              | 8 965,4        | 8 910,5            | + 54,9     | + 0,6  |
| 3 Personen .....                              | 4 084,0        | 4 144,5            | - 60,5     | - 1,5  |
| 4 Personen .....                              | 3 481,7        | 3 520,7            | - 38,9     | - 1,1  |
| 5 oder mehr Personen .....                    | 1 415,6        | 1 414,6            | + 1,0      | + 0,1  |

GWZ = Gebäude- und Wohnungszählung.

den. Diese Verzerrung hin zu größeren Haushalten resultiert aus den Übererfassungen in den Melderegistern, wodurch im Generierungsverfahren mehr Personen den Wohnungen zugeordnet werden müssen. Durch Maßnahmen zur Bereinigung der Fehler im Melderegister können in einem registergestützten Zensus die Generierungsergebnisse entscheidend verbessert werden. Insgesamt zeigen die relativ geringen Abweichungen in den Ergebnissen ohne Fehler in den Datenquellen, dass das Verfahren zur Zusammenführung und Generierung von Haushalten geeignet ist, in einem künftigen registergestützten Zensus hinreichend zuverlässige Ergebnisse zur Zahl und Größe von Wohnhaushalten in Deutschland bereitzustellen.

## 5.2 Haushalte nach Staatsangehörigkeit

Wie aus Tabelle 4 ersichtlich, haben die Testergebnisse eine starke Überzeichnung der Haushalte, in denen sowohl Personen mit deutscher als auch mit ausländischer Staatsbürgerschaft leben, durch das Haushaltegenerierungsverfahren ergeben. Werden alle Personen des Melderegisters in die Untersuchung einbezogen, liegt der Anteil dieser Haushalte in der Generierung um fast 40% über dem entsprechenden Wert aus der Haushaltebefragung. Aber auch ohne Unter- bzw. Übererfassungen des Melderegisters ist die Abweichung mit über 26% sehr hoch. Untersuchungen hierzu haben ergeben, dass hauptsächlich unterschiedliche Angaben zur Staatsangehörigkeit zwischen Melderegister und Haushaltebefragung zu dieser Differenz geführt haben. Für hochgerechnet 254 800 Personen wurde im Melderegister die deutsche und gleichzeitig in der Haushaltebefragung eine ausländische Staatsangehörigkeit angegeben. Umgekehrt besaßen laut Melderegister 467 000 Personen eine ausländische Staatsangehörigkeit, die sich in der Haushaltebefragung als Deutsche bezeichnet hatten.

Tabelle 4: Haushalte laut Haushaltegenerierung und Haushaltebefragung nach Staatsangehörigkeit  
Hochgerechnetes Ergebnis des Zensustests vom 5. Dezember 2001

| Haushaltstyp                                  | Haushalte laut |                    | Abweichung |        |
|---|----------------|--------------------|------------|--------|
|   | Generierung    | Haushaltebefragung |            |        |
|   | 1 000          |                    | %          |        |
| Haushalte mit Register- und GWZ-Fehler .....  | 29 612,4       | 30 068,8           | - 456,4    | - 1,5  |
| deutsche Haushalte .....                      | 27 204,4       | 27 967,0           | - 762,7    | - 2,7  |
| ausländische Haushalte .....                  | 1 107,2        | 1 163,7            | - 56,5     | - 4,9  |
| Mischhaushalte                                | 1 300,8        | 938,0              | + 362,8    | + 38,7 |
| Haushalte ohne Register- und GWZ-Fehler ..... | 25 398,8       | 25 499,4           | - 100,6    | - 0,4  |
| deutsche Haushalte .....                      | 23 717,2       | 23 996,8           | - 279,6    | - 1,2  |
| ausländische Haushalte .....                  | 751,1          | 765,0              | - 13,9     | - 1,8  |
| Mischhaushalte                                | 930,5          | 737,6              | + 192,9    | + 26,2 |

GWZ = Gebäude- und Wohnungszählung.

## 5.3 Ehepaare

Mit dem Generierungsverfahren wurde eine erheblich geringere Anzahl (- 7,2%) von Ehepaaren ohne Kind erzeugt als in der Haushaltebefragung ausgewiesen wurde. Deutlich niedriger liegt diese Differenz (- 1,4%), wenn es gelingt Fehler in den Datenquellen zu bereinigen. Das Gleiche gilt auch

Tabelle 5: Haushalte von Ehepaaren ohne und mit Kind(ern) laut Haushaltegenerierung und Haushaltebefragung  
Hochgerechnetes Ergebnis des Zensustests vom 5. Dezember 2001

| Haushaltstyp                                 | Haushalte laut |                    | Abweichung |       |
|--|----------------|--------------------|------------|-------|
|  | Generierung    | Haushaltebefragung |            |       |
|  | 1 000          |                    | %          |       |
| Haushalte mit Register- und GWZ-Fehler       |                |                    |            |       |
| Ehepaare ohne Kind .....                     | 6 904,4        | 7 437,0            | - 532,6    | - 7,2 |
| Ehepaare mit Kind(ern) unter 27 Jahren ..... | 6 734,8        | 7 030,5            | - 295,6    | - 4,2 |
| darunter mit ... unter 18 Jahren             |                |                    |            |       |
| 1 Kind .....                                 | 2 381,3        | 2 509,4            | - 128,1    | - 5,1 |
| 2 Kindern .....                              | 2 231,6        | 2 362,8            | - 131,1    | - 5,6 |
| 3 Kindern .....                              | 585,8          | 602,2              | - 16,5     | - 2,7 |
| 4 Kindern .....                              | 102,2          | 103,5              | - 1,3      | - 1,2 |
| 5 oder mehr Kindern .....                    | 49,5           | 53,4               | - 3,9      | - 7,3 |
| Haushalte ohne Register- und GWZ-Fehler      |                |                    |            |       |
| Ehepaare ohne Kind .....                     | 6 458,0        | 6 546,6            | - 88,6     | - 1,4 |
| Ehepaare mit Kind(ern) unter 27 Jahren ..... | 6 192,2        | 6 371,0            | - 178,8    | - 2,8 |
| darunter mit ... unter 18 Jahren             |                |                    |            |       |
| 1 Kind .....                                 | 2 187,3        | 2 256,1            | - 68,8     | - 3,0 |
| 2 Kindern .....                              | 2 093,2        | 2 156,6            | - 63,4     | - 2,9 |
| 3 Kindern .....                              | 543,9          | 541,4              | + 2,4      | + 0,5 |
| 4 Kindern .....                              | 86,0           | 93,0               | - 7,0      | - 7,5 |
| 5 oder mehr Kindern .....                    | 43,7           | 45,4               | - 1,7      | - 3,8 |

GWZ = Gebäude- und Wohnungszählung.

für Ehepaare mit Kind(ern) im Alter bis zu 27 Jahren, die durch das Haushaltgenerierungsverfahren um 2,8% unterschätzt werden. Hier ist davon auszugehen, dass durch eine fehlerhafte Zuordnung weiterer Personen (z.B. Elternteile, Geschwister) aus einigen Haushalten des Typs „Ehepaar ohne/mit Kind(ern)“ „sonstige Haushalte“ wurden.

### 5.4 Elternteile mit Kind(ern)

Bei dem Haushaltstyp „Allein lebender Elternteil mit Kind(ern)“ weisen die Abweichungen der Generierungsergebnisse zu den Ergebnissen der Haushaltebefragung ein unterschiedliches Vorzeichen auf, je nachdem ob die Ergebnisse mit bzw. ohne Registerfehler betrachtet werden. Ohne Unter-/Übererfassungen des Melderegisters wird der Haushaltstyp „Allein lebender Elternteil mit Kind(ern)“ durch die Generierung überschätzt (+ 2,5%), mit Registerfehler unterschätzt (- 3,4%). Auch hier wird durch die Zuordnung von Karteileichen, das heißt weiteren Personen, die Haushaltsstruktur verändert: Aus Haushalten des Typs „Allein lebender Elternteil mit Kind(ern)“ werden „nichteheliche Lebensgemeinschaften“ bzw. „sonstige Haushalte“, dies erklärt die Unterschätzung der Ergebnisse für diesen Haushaltstyp bei Einbeziehung der Registerfehler.

Tabelle 6: Haushalte von allein lebenden Elternteilen mit Kind(ern) laut Haushaltgenerierung und Haushaltebefragung  
Hochgerechnetes Ergebnis des Zensustests vom 5. Dezember 2001

| Haushaltstyp                                   | Haushalte laut |                    | Abweichung |      |
|--|----------------|--------------------|------------|------|
|  | Generierung    | Haushaltebefragung | 1 000      | %    |
| Haushalte mit Register- und GWZ-Fehler         |                |                    |            |      |
| Elternteil mit Kind(ern) unter 27 Jahren ..... | 1 367,5        | 1 415,1            | -47,6      | -3,4 |
| darunter mit ... unter 18 Jahren               |                |                    |            |      |
| 1 Kind .....                                   | 686,4          | 706,7              | -20,3      | -2,9 |
| 2 Kindern .....                                | 286,7          | 301,7              | -15,1      | -5,0 |
| 3 Kindern .....                                | 82,0           | 86,2               | -4,2       | -4,9 |
| Haushalte ohne Register- und GWZ-Fehler        |                |                    |            |      |
| Elternteil mit Kind(ern) unter 27 Jahren ..... | 1 172,4        | 1 144,0            | +28,4      | +2,5 |
| darunter mit ... unter 18 Jahren               |                |                    |            |      |
| 1 Kind .....                                   | 593,9          | 570,6              | +23,3      | +4,1 |
| 2 Kindern .....                                | 253,9          | 239,6              | +14,4      | +6,0 |
| 3 Kindern .....                                | 64,8           | 64,8               | +0,0       | +0,0 |

GWZ = Gebäude- und Wohnungszählung.

### 5.5 Nichteeliche Lebensgemeinschaften

Wie vorstehend erläutert zeigte sich, dass mit der Herausnahme von Modul 7, das neben Modul 4 weitere Generierungsschritte zur Bildung von „nichtehelichen Lebensgemeinschaften“ enthielt, zunächst zu wenige Haushalte dieses Typs erzeugt wurden. Der daraufhin in die letzte Stufe des Verfahrens integrierte Generierungsschritt hat dazu geführt, dass nun im Vergleich zur Haushaltebefragung zu viele nichteheliche Lebensgemeinschaften nachgewiesen werden. In beiden Untersuchungsmassen ist die

Tabelle 7: Haushalte von nichtehelichen Lebensgemeinschaften laut Haushaltgenerierung und Haushaltebefragung  
Hochgerechnetes Ergebnis des Zensustests vom 5. Dezember 2001

| Haushaltstyp  | Haushalte laut |                    | Abweichung |       |
|---|----------------|--------------------|------------|-------|
|   | Generierung    | Haushaltebefragung | 1 000      | %     |
| Haushalte mit Register- und GWZ-Fehler                                |                |                    |            |       |
| Nichteeliche Lebensgemeinschaften ohne Kind .....                     | 1 271,2        | 1 161,3            | +109,9     | +9,5  |
| Nichteeliche Lebensgemeinschaften mit Kind(ern) unter 27 Jahren ..... | 640,0          | 549,0              | +91,0      | +16,6 |
| darunter mit ... unter 18 Jahren                                      |                |                    |            |       |
| 1 Kind .....  | 359,5          | 324,6              | +34,9      | +10,7 |
| 2 Kindern .....   | 164,9          | 133,9              | +31,0      | +23,2 |
| 3 Kindern .....   | 44,7           | 33,7               | +11,0      | +32,7 |
| Haushalte ohne Register- und GWZ-Fehler                               |                |                    |            |       |
| Nichteeliche Lebensgemeinschaften ohne Kind .....                     | 951,2          | 903,6              | +47,6      | +5,3  |
| Nichteeliche Lebensgemeinschaften mit Kind(ern) unter 27 Jahren ..... | 516,4          | 448,3              | +68,1      | +15,2 |
| darunter mit ... unter 18 Jahren                                      |                |                    |            |       |
| 1 Kind .....  | 294,0          | 262,0              | +32,0      | +12,2 |
| 2 Kindern .....   | 138,5          | 111,2              | +27,3      | +24,5 |
| 3 Kindern .....   | 34,5           | 26,1               | +8,3       | +31,9 |

GWZ = Gebäude- und Wohnungszählung.

Anzahl der nichtehelichen Lebensgemeinschaften für die Haushalte ohne Kind um 9,5 bzw. 5,3%, für die Haushalte mit Kind(ern) sogar um 16,6 bzw. 15,2% höher als die der Haushaltebefragung. Daher wird eine Weiterentwicklung des Generierungsverfahrens angestrebt, mit dem Ziel die Zuordnungskriterien für die nichtehelichen Lebensgemeinschaften zu verfeinern.

### 5.6 Seniorenhaushalte

Die durch die demografische Entwicklung künftig immer bedeutsamer werdende Gruppe der Seniorenhaushalte, hier in der Darstellung für die Kategorien „Ehepaare, beide

Tabelle 8: Seniorenhaushalte laut Haushaltgenerierung und Haushaltebefragung  
Hochgerechnetes Ergebnis des Zensustests vom 5. Dezember 2001

| Haushaltstyp                                    | Haushalte laut |                    | Abweichung |      |
|---|----------------|--------------------|------------|------|
|   | Generierung    | Haushaltebefragung | 1 000      | %    |
| Haushalte mit Register- und GWZ-Fehler          |                |                    |            |      |
| Ehepaare, beide 65 Jahre oder älter .....       | 2 244,8        | 2 372,8            | -128,0     | -5,4 |
| Ehepaare, ein Partner 65 Jahre oder älter ..... | 906,5          | 991,3              | -84,9      | -8,6 |
| Allein Lebende(r), 65 Jahre oder älter .....    | 3 604,6        | 3 706,8            | -102,1     | -2,8 |
| Haushalte ohne Register- und GWZ-Fehler         |                |                    |            |      |
| Ehepaare, beide 65 Jahre oder älter .....       | 2 082,3        | 2 129,5            | -47,2      | -2,2 |
| Ehepaare, ein Partner 65 Jahre oder älter ..... | 857,3          | 874,5              | -17,3      | -2,0 |
| Allein Lebende(r), 65 Jahre oder älter .....    | 3 100,3        | 3 087,8            | +12,5      | +0,4 |

GWZ = Gebäude- und Wohnungszählung.

65 Jahre oder älter“, „Ehepaare, ein Partner 65 Jahre oder älter“ und „Allein Lebende(r), 65 Jahre oder älter“ wird mit dem Generierungsverfahren, bezogen auf die Ergebnisse ohne Berücksichtigung der Registerfehler, etwas unterschätzt. Die stärkere Unterschätzung durch das Generierungsverfahren in den Ergebnissen mit Berücksichtigung der Registerfehler ist wiederum darauf zurückzuführen, dass durch die Zuordnung von Karteileichen aus Seniorenhaushalten „sonstige Haushalte“ wurden. Die in das Modul 8 eingebaute Sperre für diesen Haushaltstyp hat eine noch größere Differenz zu den Ergebnissen der Haushaltebefragung verhindert.

## 6 Eignung des Haushaltegenerierungsverfahrens zur Kontrolle der Melderegister

Nach den Ergebnissen des Zensustests enthalten die Melderegister in Deutschland bei der Bevölkerung am Ort der Hauptwohnung 2,8% Karteileichen und 1,5% Fehlbestände. Bei Personen am Ort der Nebenwohnung beträgt die Karteileichenrate sogar 55,1% und die Fehlbestandsrate 5,6%. Diese Fehler des Melderegisters sind zu reduzieren, um bei der Durchführung eines registergestützten Zensus valide Einwohnerzahlen zu erhalten.

Deshalb war ein weiteres Untersuchungsziel des Zensustests, festzustellen, ob durch das Zusammenführen der Daten aus den Melderegistern mit den Daten der Gebäude- und Wohnungszählung Unplausibilitäten, also Karteileichen und Fehlbestände, des Melderegisters aufgedeckt und mittels nachträglicher Befragungen bzw. maschineller Verfahren korrigiert werden können.

Um Fehlbestände aufdecken zu können, wurden die Adressen, die nach dem Namensabgleich (Modul 2) noch nicht mit einer Wohnung zusammengeführt werden konnten, in die Untersuchungen mit den Testergebnissen einbezogen. Die dieser Vorgehensweise zugrunde liegende Annahme war, dass Wohnungen, deren Wohnungsinhaber nicht im Melderegister gefunden werden konnte, auf einen Fehlbestand hindeuten. Die Ergebnisse dieser Untersuchungen ergaben jedoch, dass die Wohnungen mit Fehlbeständen stark streuen und somit zur Aufdeckung eines angemessenen Anteils an Fehlbeständen ein enormer Befragungsaufwand nötig wäre. Deshalb sind weitere Untersuchungen des Testmaterials durchzuführen, mit dem Ziel effizientere Verfahren zur Korrektur von Fehlbeständen zu finden.

Hinsichtlich der Möglichkeiten zur Bereinigung von Karteileichen im Melderegister konzentrierten sich die Untersuchungen zunächst auf das Auffinden von Personengruppen mit hohen Registerfehlern. Entsprechende Teilpopulationen wie zum Beispiel erwachsene Kinder, Ledige, Nichtdeutsche usw. mit deutlich höheren Karteileichenraten als andere Bevölkerungsgruppen wurden auch gefunden, doch war die Streuung der Karteileichen für eine primärstatistische

Reinigung zu groß und die Höhe der Karteileichenrate nicht ausreichend, um maschinell zu bereinigen. Daraufhin fokussierten sich die Analysen auf einzelne Module des Generierungsprozesses und dort vor allem auf Personen, die bis zu dieser Stufe noch keiner Wohnung zugeordnet werden konnten. Es zeigte sich, dass die Personen am Ort der Nebenwohnung, die auch nach den Generierungsstufen von Modul 4 noch keinem Haushalt zugeordnet werden konnten, eine Karteileichenrate von 80,3% aufwiesen. Aufgrund dieser hohen Karteileichenrate war eine maschinelle Lösung, ohne nennenswerte Verzerrung der demografischen Strukturen, möglich. Es wurde ein Verfahren entwickelt, das vier von fünf unverknüpften Personen mit Nebenwohnsitz unter Berücksichtigung haushaltsstruktureller Aspekte löscht. Dieses Konzept ist für einen registergestützten Zensus noch zu verfeinern.

Auch für unverknüpfte Personen am Ort der Hauptwohnung belegen die Zwischenergebnisse nach den Modulen 2, 4 und 6 eine hohe Karteileichenrate. So ist zum Beispiel die Karteileichenrate bei unverknüpften Personen mit Hauptwohnsitz nach Modul 4 mit 16,5% zwölfmal so hoch wie die der verknüpften Personen (siehe Tabelle 2). Um diese Karteileichen aufzufinden, müsste jedoch das Fünffache an Personen befragt werden. Dieser Aufwand für die Bereinigung des Melderegisters ist im Verhältnis zum Ertrag, also der Aufdeckung von Karteileichen, zu hoch. Weitere Simulationsrechnungen mit den Daten des Zensustests nach Gebäudegrößenklassen haben ergeben, dass nur für den Ein- und Zweifamilienhausbereich die Bereinigung von Hauptwohnsitzkarteileichen durch retrospektive Befragungen einen akzeptablen Wirkungsgrad aufweist. Durch eine Befragung von 6,6% der im Ein- und Zweifamilienhausbereich wohnenden Haushalte könnten rund 459 000 Karteileichen, das sind 53,5% aller festgestellten Karteileichen in diesen Gebäuden aufgedeckt werden. Dieser Bereinigungsprozess bewirkt für die kleineren Gemeinden eine deutliche Absenkung der Karteileichenrate für die Bevölkerung am Ort der Hauptwohnung, da dort ein großer Teil der Bevölkerung in Ein- oder Zweifamilienhäusern wohnt. Vor allem aber führt er zu einer starken Angleichung der Qualität der amtlichen Einwohnerzahl für die Gemeinden mit weniger als 10 000 Einwohnern.<sup>9)</sup>

Für die Mehrfamiliengebäude würde die Klärung unplausibler Fälle in der Haushaltegenerierung ein wesentlich schlechteres Kosten-Nutzen-Verhältnis aufweisen. So müssten bei den Gebäuden mit drei bis sechs Wohnungen 15% aller Gebäude befragt werden, um etwa 42% der Karteileichen auffinden zu können und bei den Gebäuden mit sieben und mehr Wohnungen müssten für eine Halbierung der Karteileichenrate rund 25% der Gebäude dieser Gebäudegrößenklasse primärstatistisch erhoben werden. Aufgrund dieser Ergebnisse wird vorgeschlagen, das Verfahren Zusammenführung/Haushaltegenerierung in einem registergestützten Zensus nur für Gemeinden unter 10 000 Einwohnern, neben der Mehrfachfallprüfung, zur Bereinigung der Melderegister einzusetzen.

<sup>9)</sup> Siehe Statistische Ämter des Bundes und der Länder: „Ergebnisse des Zensustests“ in WiSta 8/2004, S. 818.

## 7 Fazit

Zusammenfassend betrachtet haben die Ergebnisse des Zensustests hinsichtlich der Gewinnung von Haushalts- und Wohnungsbelegungsdaten sowie der Eignung als Instrument zur Bereinigung der Fehler im Melderegister mit dem Verfahren der Zusammenführung/Haushaltegenerierung folgende Erkenntnisse gebracht: Eine maschinelle Generierung von Wohnhaushalten durch Zusammenführung von Melderegisterdaten mit Angaben aus der Gebäude- und Wohnungszählung ist technisch-organisatorisch machbar. Das zur Zusammenführung der beiden Datenquellen entwickelte Namensabgleichprogramm erreicht hohe Trefferquoten. Die Erfassung zweier Wohnungsinhaber in der GWZ hat sich bewährt.

Die Zahl der Haushalte insgesamt wird durch das Verfahren leicht unterschätzt. Ursachen hierfür sind zum einen Fehlbestände des Melderegisters sowie in geringerem Umfang Übergenerierungen, das heißt die fehlerhafte Zuordnung von Personen zu einer bestimmten Wohnung, die tatsächlich in verschiedenen Wohnungen leben.

Die Haushaltsgrößen werden mit dem Verfahren als solches realistisch abgebildet, liegen Übererfassungen in den Daten der Melderegister vor, kommt es zu einer Überschätzung der Zahl größerer Haushalte, vor allem der Haushalte mit fünf oder mehr Personen.

Insgesamt können mit dem Verfahren der Zusammenführung/Haushaltegenerierung belastbare Ergebnisse zur Zahl und Struktur der Haushalte in einem registergestützten Zensus bereitgestellt werden. Die Ergebnisse zeigen aber auch die starke Abhängigkeit des Verfahrens von der Qualität der zugrunde liegenden Datenquellen. Zur statistischen Beseitigung der Fehler in den Melderegistern sollen in einem registergestützten Zensus eine Mehrfachfallprüfung sowie ergänzende primärstatistische Bereinigungsmaßnahmen durchgeführt werden.<sup>10)</sup> Zur Verbesserung der Qualität der GWZ-Daten sind erhebungsteilübergreifende Plausibilitätskontrollen zwischen den beiden Datenquellen zu entwickeln. Aber auch der Algorithmus zur Generierung von Haushalten muss für einen künftigen registergestützten Zensus noch weiter optimiert werden. Dies gilt besonders für die Generierungsschritte zur Bildung nichtehelicher Lebensgemeinschaften.

Hinsichtlich der Eignung des Verfahrens als Instrument zur Bereinigung von Unplausibilitäten im Melderegister belegen die vorgestellten Ergebnisse einen wesentlich höheren Karteileichenanteil bei Personen, die durch den Generierungsprozess bis nach Modul 4 nicht mit einer Wohnung verknüpft werden konnten, als bei verknüpften Personen. Für nicht verknüpfte Personen am Ort der Nebenwohnung war der Karteileichenanteil so hoch, dass hierfür ein Verfahren zur maschinellen Bereinigung ohne nennenswerte Verzerrungen der demografischen Strukturen entwickelt werden konnte.

Die Analysen für Personen am Ort der Hauptwohnung ergaben, dass nur im Ein- und Zweifamilienhausbereich mit einem vertretbaren Aufwand ein hinreichend hoher Anteil (53,5%) an Karteileichen entdeckt und durch retrospektive Befragungen geklärt werden kann. Da in Gemeinden unter 10 000 Einwohnern sehr viele Ein- und Zweifamilienhäuser zu finden sind, würde dieser Bereinigungsprozess, ergänzend zur Mehrfachfallprüfung, zu einer deutlichen Absenkung der Karteileichenrate in diesen Gemeinden führen, sodass hier bei einem registergestützten Zensus auf eine ergänzende Stichprobe zur weiteren Korrektur von Registerfehlern verzichtet werden könnte. [u](#)

<sup>10)</sup> Art und Umfang der nachträglichen Befragungen sind aufgrund von Aufwand-Nutzen-Überlegungen politisch zu entscheiden. Die statistischen Ämter haben hierzu mehrere Modellvarianten entwickelt. Siehe Fußnote 9, S. 813 ff.

## Auszug aus Wirtschaft und Statistik

© Statistisches Bundesamt, Wiesbaden 2004

Für nichtgewerbliche Zwecke sind Vervielfältigung und unentgeltliche Verbreitung, auch auszugsweise, mit Quellenangabe gestattet. Die Verbreitung, auch auszugsweise, über elektronische Systeme/Datenträger bedarf der vorherigen Zustimmung. Alle übrigen Rechte bleiben vorbehalten.

Herausgeber: Statistisches Bundesamt, Wiesbaden

Schriftleitung: Johann Hahlen  
Präsident des Statistischen Bundesamtes  
Verantwortlich für den Inhalt:  
Brigitte Reimann,  
65180 Wiesbaden

- Telefon: +49 (0) 6 11/75 20 86
- E-Mail: [wirtschaft-und-statistik@destatis.de](mailto:wirtschaft-und-statistik@destatis.de)

Vertriebspartner: SFG Servicecenter Fachverlage  
Part of the Elsevier Group  
Postfach 43 43  
72774 Reutlingen  
Telefon: +49 (0) 70 71/93 53 50  
Telefax: +49 (0) 70 71/93 53 35  
E-Mail: [destatis@s-f-g.com](mailto:destatis@s-f-g.com)

Erscheinungsfolge: monatlich



Allgemeine Informationen über das Statistische Bundesamt und sein Datenangebot erhalten Sie:

- im Internet: [www.destatis.de](http://www.destatis.de)

oder bei unserem Informationsservice  
65180 Wiesbaden

- Telefon: +49 (0) 6 11/75 24 05
- Telefax: +49 (0) 6 11/75 33 30
- E-Mail: [info@destatis.de](mailto:info@destatis.de)