

# Statistische Analysen und Studien Nordrhein-Westfalen

## Band 4

### Impressum

Herausgeber :  
Landesamt für Datenverarbeitung  
und Statistik Nordrhein-Westfalen

Redaktion:  
Bianca Klose, Hans Lohmann

Preis dieser Ausgabe: 4,50 EUR

Erscheinungsfolge: unregelmäßig

Bestellungen nehmen entgegen:

das Landesamt für Datenverarbeitung  
und Statistik NRW,  
Postfach 10 11 05,  
40002 Düsseldorf,  
Mauerstraße 51,  
40476 Düsseldorf  
Telefon: 0211 9449-2516/3516  
Telefax: 0211 442006  
Internet: <http://www.lds.nrw.de>  
E-Mail: [poststelle@lds.nrw.de](mailto:poststelle@lds.nrw.de)

sowie der Buchhandel.

Pressestelle:  
0211 9449-2521/2518

Zentraler Informationsdienst:  
0211 9449-2495/2525

© Landesamt für Datenverarbeitung  
und Statistik NRW, Düsseldorf, 2002

Für nicht gewerbliche Zwecke sind  
Vervielfältigung und unentgeltliche  
Verbreitung, auch auszugsweise, mit  
Quellenangabe gestattet. Die Verbrei-  
tung, auch auszugsweise, über elek-  
tronische Systeme/Datenträger bedarf  
der vorherigen Zustimmung. Alle üb-  
rigen Rechte bleiben vorbehalten.

Bestell-Nr. Z 08 1 2002 54

ISSN 1619-506X

### Inhalt

<b>editorial</b>	<b>3</b>
<b>Die Mehrfachfallprüfung im Zensustest 2001 – Konzeption und Ablaufplan</b> Dr. Marco Scharmer	<b>5</b>
<b>Die Statistikauglichkeit der Einwohnermelderegister Untersuchungen im Rahmen des Zensustests 2001</b> Dr. Michael Forster	<b>17</b>
<b>Postalische Gebäude- und Wohnungszählung im Zensustest 2001 – Erfahrungen in NRW</b> Dipl.-Sozialwirt Sebastian Scharnhorst und Dr. Henrik Egbert	<b>24</b>
<b>Die Haushaltegenerierung im Zensustest 2001 – Eine neue Methode in der amtlichen Statistik</b> Dr. Marco Scharmer	<b>31</b>
<b>Die Rolle der Haushaltebefragung im Zensustest</b> Dipl.-Statistiker Josef Schäfer	<b>41</b>
<b>Literatur</b>	<b>45</b>

## **Zeichenerklärung**

(nach DIN 55 301)

- 0 weniger als die Hälfte von 1 in der letzten besetzten Stelle, jedoch mehr als nichts
- nichts vorhanden (genau null)
- . Zahlenwert unbekannt oder geheim zu halten
- ... Zahlenwert lag bei Redaktionsschluss noch nicht vor
- ( ) Aussagewert eingeschränkt, da der Wert Fehler aufweisen kann
- / keine Angabe, da der Zahlenwert nicht sicher genug ist
- x Tabellenfach gesperrt, weil Aussage nicht sinnvoll
- p vorläufige Zahl
- r berichtigte Zahl

Abweichungen in den Summen erklären sich aus dem Runden der Einzelwerte.

## **editorial**

### **Deutschland auf dem Weg zum registergestützten Zensus**

Die Bundesregierung hatte sich 1996 gegen eine weitere Volkszählung nach dem herkömmlichen Muster entschieden. Vorausgegangen war eine Empfehlung der Europäischen Union, im Jahr 2001 in allen Mitgliedstaaten und Beitrittsländern einen flächendeckenden Zensus durchzuführen. Stattdessen wurde die amtliche Statistik in Bund und Ländern aufgefordert, alternative Erhebungswege zur direkten Befragung aller Einwohner zu entwickeln.

Das Ergebnis dieser Entwicklungen ist das Modell eines registergestützten Zensus. Seine wesentliche konzeptionelle Neuerung liegt im weitgehenden Verzicht auf den herkömmlichen Einsatz von Zählern. Dies wird durch ein Verfahren zur Zusammenführung von postalisch erhobenen Gebäude- und Wohnungsdaten mit Registerbeständen der Einwohnermeldeämter und der Bundesanstalt für Arbeit ermöglicht. Durch die Kombination dieser Angaben sollen Haushaltsdaten generiert werden, die qualitativ mit denen früherer Zensen vergleichbar sind – jedoch ohne die Einzeldaten direkt bei den Haushalten erheben zu müssen.

Der Ansatz, Registerangaben für Zensuserhebungen nutzbar zu machen, ist indes nicht neu. Neben anderen setzen insbesondere die skandinavischen Länder in ihren Volkszählungen bereits seit längerem auf Datenbestände der Verwaltungsregister – allerdings mit dem Unterschied, dass die dortigen Register zumeist statistiktauglicher gestaltet sind und die vorhandenen Angaben durch Personenkennziffern relativ problemlos miteinander kombiniert werden können. Daher liegt eine Besonderheit des deutschen Modells vor allem auch in dem Versuch, diese Angaben ohne eine solchermaßen eindeutige Kennzeichnung einander zuzuordnen.

Eine erstmalige Erprobung dieses registergestützten Zensusmodells, der Zensustest, wird derzeit von den Statistischen Ämtern des Bundes und der Länder in einer bundesweiten Stichprobe durchgeführt. In diesem Test sollen insbesondere die Rahmenbedingungen für einen erfolgreichen Einsatz des Modells und die Effektivität der dafür neu entwickelten statistischen Verfahren überprüft und beurteilt werden. Dabei werden, grob skizziert, folgende Punkte von maßgeblicher Bedeutung sein: die Statistikauglichkeit der Register von Einwohnermeldeämtern und Bundesanstalt für Arbeit; die Durchführbarkeit einer postalischen Erhebung von Gebäude- und Wohnungsdaten bei den Gebäudeeigentümern; die Möglichkeit, Mehrfachfälle sowie „Karteileichen“ und Fehlbestände zu identifizieren und zu bereinigen; die Zuverlässigkeit der Generierungsergebnisse und schließlich auch die Fähigkeit des Modells, eine der herkömmlichen Volkszählung vergleichbare Datenqualität zu erzielen. Trifft all dies zu, so könnte das registergestützte Verfahren beim nächsten Zensus in Deutschland flächendeckend zum Einsatz kommen.

Dabei weist das neue Modell auch im Rückblick auf bisherige Volkszählungen zwei wesentliche Vorteile auf. Zum einen ist dies der Kostenaspekt. Würde eine herkömmliche Zählung heute schätzungsweise mindestens 750 Mill. Euro kosten, so käme ein registergestützter Zensus bereits mit weitaus weniger als einem Drittel dieser Mittel aus. Zweitens würde sich mit dem Verzicht auf eine direkte Befragung der Haushalte auch die Belastung der Bevölkerung erheblich verringern.

Mit Erscheinen dieses Sonderbandes sind fast alle primärstatistischen Erhebungsteile des Zensustests abgeschlossen. In den kommenden Monaten können nun die Verfahren zur Zusammenführung der Einzeldaten und Generierung der Haushaltsdatensätze geprüft werden. Die Beiträge in diesem Heft widmen sich daher neben methodischen Beschreibungen der Erhebungsteile auch ersten Erfahrungen, die bei der Umsetzung und Erprobung einzelner Teilbereiche des neuen Modells in Nordrhein-Westfalen gesammelt wurden. Eine abschließende Auswertung des Tests mit Bundes- und Länderergebnissen und eine umfassende Beurteilung des registergestützten Zensusmodells werden voraussichtlich ab Mitte des Jahres 2003 vorliegen. Bereits zum jetzigen Zeitpunkt zeichnet sich aber ab, dass das Modell in Teilbereichen noch zu modifizieren sein wird.

**Jochen Kehlenbach**

Präsident



# Die Mehrfachfallprüfung im Zensusstest 2001 Konzeption und Ablaufplan

Dr. Marco Scharmer

*Volkszählungen oder auch Zensen bilden die zentrale Datenbasis der amtlichen Statistik in der Bundesrepublik Deutschland. Die letzte Volkszählung wurde in der Bundesrepublik Deutschland 1987, in der ehemaligen DDR im Jahr 1981 durchgeführt. Nicht zuletzt durch die Wiedervereinigung fanden in den Jahren seit diesen letzten Volkszählungen beträchtliche Wanderungsbewegungen statt, wodurch die zur Verfügung stehenden Daten immer weniger die tatsächlichen Verhältnisse abbilden. Die abnehmende Aktualität der Volkszählungsdaten gewinnt mehr und mehr an Bedeutung, da sich die amtliche Bevölkerungsstatistik auch weiterhin im Wesentlichen auf die Fortschreibung der in der Volkszählung 1987 ermittelten Bevölkerungszahlen und auf Schätzungen aus dem Mikrozensus stützt.*

## 1 Testerhebung für das Modell Registergestützter Zensus

Die Ergebnisse eines zukünftigen registergestützten Zensus müssen sich an der Qualität, im Sinne von Vollständigkeit, Vollzähligkeit und Stimmigkeit, der Ergebnisse früherer Volkszählungen messen lassen. Dieses gilt im Besonderen für die Ermittlung der amtlichen Einwohnerzahlen, welche die Grundlage für weitreichende finanzpolitische Entscheidungsprozesse bilden. Die Verwendung der Melderegister im Rahmen des registergestützten Zensus geht über die bisherige gesetzliche Zweckbestimmung dieser reinen Verwaltungsregister hinaus. Insofern sind umfangreiche Tests hinsichtlich der Qualität und Validität der aus den Registern gewonnenen Daten durchzuführen sowie Verfahren zu entwickeln, die u. a. zur Ermittlung korrekter und damit gerichtsfester amtlicher Einwohnerzahlen führen. Erst nach Auswertung dieser Tests könnte die amtliche Statistik in der Bundesrepublik Deutschland den Methodenwechsel von einer primärstatistischen Vollerhebung zu einem registergestützten Zensus vollziehen.

Hierzu haben Bundestag und Bundesrat das Gesetz zur Vorbereitung eines registergestützten Zensus (Zensusvorbereitungsgesetz) verabschiedet<sup>1)</sup>. Diese Gesetz sieht im Rahmen der Mehrfachfallprüfung vor, Auszüge aus den Einwohnermelderegistern für eine Geburtstagsauswahl auszuwerten. In einem weiteren Schritt werden diese Registerdaten anhand einer primärstatistischen Befragung unter den Personen der Geburtstagsstichprobe überprüft<sup>2)</sup>.

## 2 Zur Rolle der Mehrfachfallprüfung im Zensusstest

Für das im Zensusstest 2001 getestete Modell sind also die Einwohnermelderegister die zentrale Datenquelle. Die demographischen Grunddaten der Bevölkerung werden damit nicht mehr wie bisher durch Interviewer in Fragebögen erhoben, sondern durch eine Auswertung der bei den Gemeinden geführten Register gewonnen. Damit kommt der Aktualität und Vollständigkeit der Registereintragungen sowie der Qualität der demographischen Merkmale eine entscheidende Bedeutung zu. Bisherige Untersuchungen haben zumin-

dest Zweifel an der Vollständigkeit und Qualität der Melderegister aufgeworfen<sup>3)</sup>.

Bei den dezentral in jeder Gemeinde geführten Melderegistern ist jedoch nicht auszuschließen, dass Einwohnerinnen und Einwohner nicht oder in mehreren Gemeinden gleichzeitig mit alleiniger Wohnung oder Hauptwohnung gemeldet sind. Ebenso können Personen auch nur mit Nebenwohnung registriert sein. Diese Fälle, die zu so genannten Mehrfachfällen führen, können z. B. durch eine unterlassene oder nicht zeitgleich stattfindende An- bzw. Abmeldung und ihre verwaltungsmäßige Bearbeitung auftreten. Daneben könnte auch die Angabe zum Wohnungsstatus nicht mehr den aktuellen Wohnverhältnissen entsprechen, da keine oder keine zeitgleich erfolgte Meldung über diese Änderungen stattgefunden hat<sup>4)</sup>.

Die in jedem Zensus festgestellte amtliche Einwohnerzahl (Bevölkerung am Ort der Hauptwohnung) für Bund, Länder und Gemeinden, also die Zahl aller Personen, die innerhalb der jeweiligen Gebietseinheit ihre alleinige oder Hauptwohnung besitzen, dient als Basis für zahlreiche Entscheidungsprozesse, z. B. als Bemessungsgrundlage im vertikalen und horizontalen Finanzausgleich, zur Einteilung von Wahlkreisen oder für die Vergabe von Mitteln aus den EU-Strukturfonds. Aufgrund der herausragenden Bedeutung der amtlichen Einwohnerzahl sind die qualitativen Anforderungen an die Zensusergebnisse sehr hoch. Werden die Melderegisterangaben nun als zentrale Datenquelle herangezogen, wie

1) Vgl. Gesetz zur Vorbereitung eines registergestützten Zensus (Zensusvorbereitungsgesetz) vom 27. 7. 2001, in: Bundesgesetzblatt, Jg. 2001, Teil I Nr. 40, S. 1882 ff. Näheres siehe unter Abschnitt 3: „Rechtsgrundlagen und Erhebungsmerkmale“. – 2) Vgl. zu den Erhebungsteilen im Einzelnen: Forster, Michael, Die Zukunft der Volkszählung in Deutschland, Traditionelle Zählung oder registergestützter Zensus?, in: Statistische Analysen und Studien Nordrhein-Westfalen, 4/2001, S. 23 ff.; Scharmer, Marco, Haushaltstypisierung im Rahmen des registergestützten Zensusstests, in: Statistische Analysen und Studien Nordrhein-Westfalen, Bd. 1 (2002), S. 3 ff. sowie Egbert, Henrik/Scharmer, Marco, Die Ermittlung des Wohnungsinhabers im Rahmen der GWZ und deren Bedeutung für die Haushaltegenerierung, in: Statistische Analysen und Studien Nordrhein-Westfalen, 4/2001, S. 10 ff. – 3) Vgl. u. a. Eppmann, Helmut, Die Einwohnermelderegister als Basis der „amtlichen“ Einwohnerzahl, in: Statistische Rundschau Nordrhein-Westfalen, Heft 8/98, S. 407 ff. – 4) Der Status der Wohnung im Melderegister gibt an, ob die Wohnung in der Gemeinde als alleinige Wohnung, als Hauptwohnung, im Sinne des Mittelpunkts der Lebensführung, oder als Nebenwohnung genutzt wird. Eine Hauptwohnung setzt immer mindestens eine Nebenwohnung und einen Nebenwohnung immer eine Hauptwohnung unter einer anderen Adresse voraus.

es im registergestützten Zensus vorgesehen ist, besteht ohne eine Prüfung dieser Angaben die Gefahr, dass Personen nicht oder mehrfach, am falschen Ort oder mit falschem Wohnstatus gezählt und dadurch unzutreffende Einwohnerzahlen festgestellt würden.

Im Rahmen des Zensustests wurde daher von den statistischen Ämtern des Bundes und der Länder auch ein Verfahren entwickelt, mit dem die Melderegisterdaten auf Mehrfachfälle hin untersucht und um diese statistisch bereinigt werden können. Die Ziele der Mehrfachfallprüfung stellen sich damit wie folgt dar:

- Entwicklung und Test von Verfahren zur Beurteilung der Melderegisterqualität im Sinne von Vollständigkeit und Aktualität.
- Erstellung eines bundeseinheitlichen Datensatzes und dessen Überprüfung auf Mehrfachfälle.
- Primärstatistische Überprüfung der im Melderegister identifizierten Mehrfachfälle und Bestimmung des Aufwandes hierfür.
- Zuverlässige Schätzung der Anzahl von Mehrfachfällen und deren demographische und regionale Eigenheiten.
- Test der Datenlieferungen der Gemeinden sowie der Gebietsrechenzentren bzw. Software-Häuser an die statistischen Ämter.

Mit diesen Testergebnissen werden dann Aussagen über den Aufwand und die erreichbare Qualität bei der Ermittlung der amtlichen Einwohnerzahl im Rahmen eines registergestützten Zensus im Vergleich zu einer herkömmlichen Volkszählung möglich sein.

### **3 Rechtsgrundlage und Erhebungsmerkmale**

Rechtsgrundlage für die Mehrfachfallprüfung im Zensustest bildet das Artikelgesetz zur Vorbereitung eines

registergestützten Zensus (Zensusvorbereitungsgesetz) vom 27. Juli 2001. Hier wird in Artikel 1, dem Gesetz zur Erprobung eines registergestützten Zensus (Zensustestgesetz – ZensTeG), zur Überprüfung von Mehrfachmeldungen in den Einwohnermelderegistern eine bundesweite Stichprobenerhebung bei allen Meldebehörden angeordnet (§ 2 ZensTeG).

Bei dieser Stichprobe handelte es sich um eine Geburtstagsauswahl, welche alle gemeldeten Einwohnerinnen und Einwohner einbezog, die am 1. Januar, am 15. Mai oder am 1. September geboren wurden, sowie alle Einwohnerinnen und Einwohner mit unvollständig eingetragenen Geburtsdatum. In das Verfahren der Mehrfachfallprüfung (MFF-Prüfung) gingen damit alle gemeldeten Einwohnerinnen und Einwohner der genannten Geburtstage sowie diejenigen mit unvollständigen Geburtsangaben ein, die zum Stichtag 5. Dezember 2001 mindestens in einer Gemeinde eine alleinige Wohnung, Haupt- oder Nebenwohnung besaßen.

Um den Befragungsumfang der zu überprüfenden Mehrfachfälle möglichst gering zu halten und Erkenntnisse über das Meldeverhalten der Bevölkerung zu gewinnen, wurde für die Personen der Stichprobe ein zweiter Melderegisterauszug zum Stichtag 31. März 2002 angefordert. Durch den Abgleich beider Melderegisterauszüge können alle Fälle statistisch bereinigt werden, die nach dem Stichtag 5. Dezember 2001 rückwirkend Änderungen in der An- oder Abmeldung bzw. dem Wohnungstatus bei den Meldebehörden angezeigt haben.

Für die Personen der Stichprobe haben die Meldebehörden bzw. die mit der technischen Führung der Melderegister beauftragten Stellen (z. B. kommunale Rechenzentren) gemäß § 2 Abs. 2 ZensTeG, zu jedem der beiden Stichtage die folgenden Erhebungs- und Hilfsmerkmale an die jeweiligen statistischen Landesämter übermittelt:

#### **Erhebungsmerkmale**

- Geburtsmonat und -jahr
- Geschlecht
- Staatsangehörigkeiten
- bei im Ausland Geborenen: Geburtsstaat
- Familienstand
- Wohnort
- Status der Wohnung (alleinige Wohnung, Haupt- oder Nebenwohnung)

#### **Hilfsmerkmale**

- Namen, Vornamen
- gegenwärtige Anschriften
- Tag der Geburt
- Geburtsort
- Standesamt und Nummer des Geburtseintrags
- Anschrift und Status der künftigen Wohnung oder der Wohnung, in die die Einwohnerin/der Einwohner laut Rückmeldung verzogen ist
- Anschrift und Status der Wohnung in der Gemeinde, aus der die Einwohnerin/der Einwohner zugezogen ist
- Zuzug aus dem Ausland
- Anschrift der zuletzt bewohnten Wohnung in der Gemeinde
- Datum des Beziehens der Wohnung
- Datum des Auszugs aus der Wohnung
- Datum des Fortzugs ins Ausland
- Datum der Anmeldung bei der Meldebehörde
- Datum der Abmeldung bei der Meldebehörde
- Datum des Wohnungsstatuswechsels

#### **Nach § 2 Abs. 3 ZensTeG wurden von den Meldebehörden zusätzlich erhoben:**

- Bezeichnung und Anschrift der Meldebehörde
- Name und Telefonnummer der Person, die für Rückfragen zur Verfügung steht
- Bezeichnung der mit der Datenverarbeitung beauftragten Stelle
- technische Gegebenheiten der Führung des Melderegisters (Betriebssystem, Software, Möglichkeit des Datentransfers)

## 4 Konzept und Ablauf der Mehrfachfallprüfung

Die Mehrfachfallprüfung erstreckt sich zum einen auf die Suche von Dubletten in den Melderegisterdatensätzen und zum anderen auf die primärstatistische Überprüfung des Hauptwohnsitzes der als Dublette identifizierten Personen.

### 4.1 Datenanforderung bei den Gemeinden

Für die Mehrfachfallprüfung haben alle Gemeinden NRW – gemäß § 2 Abs. 2 ZensTeG – zu den Stichtagen 5. Dezember 2001 und 31. März 2002 für jede in ihrem Melderegister verzeichnete Person der Geburtstagsstichprobe einen Datensatz mit den genannten Erhebungs- und Hilfsmerkmalen an das LDS NRW übermittelt. Die Datenlieferungen der Geburtstagsauswahl zum ersten Stichtag gingen bis Anfang Februar 2002, die zweiten Datenlieferungen bis Anfang Mai ein. Die Gemeinden waren aufgefordert, die Melderegisterdaten möglichst auf elektronischem Weg und in der vorgegebenen Datensatzstruktur zu liefern. Da in allen nordrhein-westfälischen Gemeinden die Melderegister computergestützt geführt werden, erhielt das LDS NRW auch alle angeforderten Daten der Geburtstagsauswahl auf Datenträger. Allerdings entsprachen nicht alle Datenlieferungen der geforderten Datensatzstruktur und mussten daher vom LDS NRW umgesetzt werden.

Bundesweit wurden mit jedem Melderegisterauszug knapp 1,2 Prozent der Bevölkerung in die Geburtstagsstichprobe einbezogen. In Nordrhein-Westfalen gingen zum ersten Stichtag 234 889 und zum zweiten Stichtag 235 592 Personensätze in die Mehrfachfallprüfung ein. Hier von entfielen rund 41 Prozent auf den Geburtstag 1. Januar, jeweils gut 23 Prozent auf die Geburtstage 15. Mai und 1. September, und für knapp 13 Prozent der gemeldeten Einwohnerinnen und Einwohner war kein oder ein unvollständiges Ge-

burtsdatum verzeichnet. Insgesamt entsprach der Umfang der Geburtstagsstichprobe etwa 1,3 Prozent der nordrhein-westfälischen Bevölkerung. Dieser Anteil liegt über dem Bundesdurchschnitt, was auf den höheren Ausländeranteil in NRW und den in dieser Personengruppe größeren Anteil unvollständiger Geburtstageintragungen im Melderegister zurückzuführen ist. Gleiches gilt auch für Personen, die am 1. Januar geboren wurden, da unvollständige Geburtsdaten häufig durch die Meldebehörden mit dem 1. Januar vervollständigt werden.

Zur Vorbereitung der Datenlieferungen hat das LDS NRW bereits im Vorfeld des Gesetzgebungsverfahrens die 396 nordrhein-westfälischen Gemeinden, die kommunalen Statistikstellen, die Kreise sowie den Städte- und Gemeindebund über das Vorhaben Zensustest informiert. Da die Gemeinden in NRW häufig externe Stellen mit der Führung ihrer Melderegister beauftragt haben oder auf kommerzielle Software-Produkte zurückgreifen, sind zeitgleich auch die Hersteller von Software für die Melderegister sowie die bekannten Rechenzentren informiert worden. Da nach Gesetzesverabschiedung von den Gemeinden kurzfristig Daten zu liefern waren, hat das LDS NRW schon frühzeitig um die Benennung eines Ansprechpartners in den Meldebehörden sowie den betreuenden Rechenzentren gebeten.

Die Abfrage der technischen Gegebenheiten der Führung des Melderegisters gemäß § 2 Abs. 3 ZensTeG wurde, soweit die Informationen durch die erste Datenlieferung noch nicht vollständig vorlagen, im Laufe des Februars 2002 bei den genannten Ansprechpartnern in den Meldebehörden und Rechenzentren vorgenommen.

### 4.2 Mehrfachfallprüfung

Die Datenlieferungen der Gemeinden wurden im LDS NRW auf Vollständigkeit, Vollzähligkeit und Plausibilität kontrolliert. Da die Stichprobenziehung durch die Gemeinden

und Rechenzentren vorgenommen wurde, fehlten der amtlichen Statistik direkte Kontrollmöglichkeiten. Daher wurden die Gemeinden aufgefordert schriftlich mitzuteilen, wenn im Melderegister keine Einwohnerinnen und Einwohner mit einem der drei genannten Geburtstage oder mit unvollständigem Geburtsdatum gemeldet sind. Aufgrund der Gemeindegrößen in NRW waren in allen Lieferungen sowohl die drei Geburtstage als auch Datensätze mit unvollständigem Geburtsdatum enthalten. Eine weitere Kontrolle führte das LDS NRW über den Anteil der Geburtstagsauswahl an der jeweiligen Gemeindebevölkerung durch. Als plausibel galten Anteile von zwischen 0,8 und 1,5 Prozent. In 28 nordrhein-westfälischen Gemeinden lag dieser Anteil über und in acht Gemeinden unter dieser Bandbreite. Die Überschreitung zeigte sich vor allem in Großstädten, die einen überdurchschnittlich hohen Anteil von Personen mit ausländischem Geburtsort und unvollständigem Geburtsdatum in den Melderegistern aufwiesen. Die Unterschreitung trat ausschließlich in kleineren (ländlichen) Gemeinden NRW auf und war zufallsbedingt.

Die beiden Datenlieferungen zur Geburtstagsauswahl wurden nach der Vollständigkeits-, Vollzähligkeits- und Plausibilitätskontrolle sowie der Vereinheitlichung der Datensätze durch das LDS NRW gemäß § 3 ZensTeG an das Statistische Bundesamt geliefert. Dort wurden die einzelnen Lieferungen aller statistischer Landesämter zu jeweils einem bundeseinheitlichen Datenbestand (1. Stichtag und 2. Stichtag) zusammengeführt. Noch vor der Überprüfung des Datenbestandes auf Mehrfachfälle wurden die Anschriften durch die Deutsche Post (Post Direkt) standardisiert und auf die Richtigkeit von Straße, Hausnummer, Postleitzahl und Ort geprüft. Damit wurde sichergestellt, dass für jede später anzuschreibende Person postalisch korrekte Anschriften im Datensatz vorliegen.

Im Anschluss daran wurde die eigentliche Mehrfachfallprüfung durchgeführt. Grundsätzlich gilt das Geburts-

datum als unveränderliches Merkmal im Melderegisterdatensatz, es bildete daher bei der Suche nach Mehrfachfällen ein Ausschlusskriterium. Allerdings wies eine nicht unerhebliche Anzahl von Datensätzen (knapp 13 Prozent in NRW) kein oder nur ein unvollständiges Geburtsdatum auf<sup>5)</sup>. Aufgrund der Tatsache, dass diese unvollständigen Angaben im Zeitverlauf aufgefüllt werden können, die gleiche Person somit in einer Gemeinde mit unvollständigem und in einer anderen mit vollständigem Geburtsdatum gemeldet sein kann, mussten die jeweiligen Datensätze in der Mehrfachfallprüfung anhand bestimmter Merkmale, wie u. a. dem Geburtsdatum, unterschiedlich behandelt werden. Daher wurde der Gesamtdatenbestand in zwei Bestände aufgespalten:

- Datenbestand 1: Datensätze, die ein vollständiges Geburtsdatum an einem der drei Geburtstagsstichtage aufwiesen (Deutsche und im Ausland geborene Personen)
- Datenbestand 2: Datensätze, die ein unvollständiges Geburtsdatum besaßen sowie Datensätze von allen im Ausland geborenen Personen.

Bei den Personen mit unvollständig eingetragenen Geburtsdatum im Melderegister handelte es sich nahezu ausschließlich um im Ausland geborene Personen. Mit der Aufspaltung in zwei Datenbestände wurde erreicht, dass diese Personengruppe in beiden Beständen enthalten ist. So konnten Mehrfachfälle im zweiten Datenbestand auch dann aufgefunden werden, wenn es im Zeitverlauf möglicherweise zu einer Auffüllung des unvollständigen Geburtsdatums gekommen war. Weiterhin wurden die im ersten Datenbestand aufgefundenen Mehrfachfälle mit den im zweiten Datenbestand gefundenen Mehrfachfällen verglichen. Damit wird ausgeschlossen, dass eine Person mehrfach in die Mehrfachfallprüfung eingeht.

5) Der sehr hohe Anteil von Personen mit Geburtsdatum 1. Januar gegenüber der Anzahl von Personen der beiden übrigen Geburtstage (41 % zu jeweils 23 %) legt nahe, dass es im Zeitverlauf zu einer Auffüllung von unvollständigen Geburtsdaten auf den 1. Januar durch die Meldebehörden gekommen ist. Insofern ist der Anteil der Datensätze mit ursprünglich unvollständigem Geburtsdatum wesentlich höher als 13 %.

Neben dem Geburtsdatum wurden bei der Mehrfachfallprüfung noch weitere Melderegisterangaben als Schlüsselmerkmale einbezogen. Der erste Datenbestand wurde anhand der Merkmale Geschlecht, Geburtsjahr, Geburtsmonat, Geburtstag, Geburtsstaat (historisch)<sup>6)</sup>, Geburtsort, Geburtsname, Namensbestandteil des Geburtsnamens und Vorname auf Mehrfachfälle hin untersucht. Im zweiten Datenbestand wurde anhand der Merkmale Geschlecht, Geburtsjahr, Geburtsmonat, Staatsangehörigkeit, Geburtsstaat (historisch), Geburtsort, Geburtsname, Namensbestandteil des Geburtsnamens und Vorname geprüft, ob eine Person gleichzeitig in mehreren Gemeinden mit alleiniger Wohnung oder Hauptwohnung gemeldet oder nur mit Nebenwohnung registriert war.

Die Mehrfachfallprüfung in den aus dem Datenmaterial zum ersten Stichtag 5. Dezember 2001 gebildeten beiden Datenbeständen ist von fünf zuvor ausgewählten externen Dienstleistungsunternehmen durchgeführt worden. Die Prüfungen erfolgten nicht nur auf exakte, sondern auch auf teilweise Übereinstimmungen oder Namensähnlichkeiten. Im Datenmaterial zum zweiten Stichtag 31. März 2002 wurde nicht nach Mehrfachfällen gesucht.

Zusätzlich prüfte das Statistische Bundesamt in einem sogenannten 1:1-Abgleich (Vergleich auf vollständige Übereinstimmung von Geschlecht, Geburtsjahr, Geburtsmonat, Geburtstag, Geburtsstaat (historisch), Geburtsname, Namensbestandteil des Geburtsnamens, Vorname) die Melderegisterdatensätze aller Einwohnerinnen und Einwohner der Stichprobenauswahl zum ersten Stichtag auf widersprüchliche Angaben im Wohnungstatus. Im Anschluss daran hat das Statistische Bundesamt die Ergeb-

6) Der Geburtsstaat wird zuvor historisch umcodiert, damit den in den mittel- und osteuropäischen Ländern sowie den Nachfolgerepubliken der ehemaligen Sowjetunion seit Ende der 80er-Jahre stattgefundenen politischen Veränderungen Rechnung getragen werden kann. Somit wird verhindert, dass für eine Person unter einer Meldeadresse als Geburtsort z. B. UdSSR und unter einer anderen Adresse Estland eingetragen ist.

nisse der Mehrfachfallprüfung durch die fünf externen Firmen sowie des vom Statistischen Bundesamt durchgeführten 1:1-Abgleichs zusammengeführt.

Im Rahmen der Mehrfachfallprüfung werden zunächst auch Datensätze von Personen, die mit alleiniger bzw. einer Hauptwohnung und weiteren Nebenwohnungen gemeldet sind, als Dublette ausgewiesen. Wie die nachfolgende Tabelle zeigt, handelt es sich bei diesen Personen jedoch nicht um Mehrfachfälle im Sinne des Zensustests, weshalb die entsprechenden Datensätze für die weitere Bearbeitung unberücksichtigt bleiben.

Die externe Prüfung der Merkmale Geburtsort, Geburtsname, Namensbestandteil des Geburtsnamens und Vorname lässt eine gewisse Toleranz bei unterschiedlichen Schreibweisen zu. Durch die z. T. unterschiedlichen Prüfverfahren der ausgewählten Softwareanbieter haben sich bei der Zusammenführung nicht nur Dublettenpaare, sondern auch Dublettenketten ergeben. So identifiziert beispielweise Anbieter 1 die Datensätze A und B als Mehrfachfall, während Anbieter 2 die Datensätze A und C als Mehrfachfall identifiziert. Daraus ergibt sich bei der Zusammenführung der Ergebnisse die Dublettenkette Datensatz A – Datensatz B – Datensatz C. Durch die Zulassung einer gewissen Toleranz bei der Übereinstimmung der Prüfungsmerkmale im Rahmen der externen Prüfung wurden vermutlich nahezu alle Mehrfachfälle aufgedeckt. Gleichzeitig ist jedoch, in Abhängigkeit vom gewählten Verfahren, auch mit einer größeren Zahl von fälschlicherweise identifizierten Mehrfachfällen zu rechnen.

Das Konzept der Mehrfachfallprüfung sieht daher vor, die Dublettenketten entsprechend bestimmten Kriterien in Datensatzpaare zu zerlegen. Damit kann sichergestellt werden, dass den Vorgaben des Zensustestgesetzes gemäß die als Mehrfachfall identifizierten Personen möglichst nur einmal befragt werden müssen. Auch lassen sich Ab-



Klassifikation der Mehrfachfälle im Zensustest			
Alleinige Wohnung	Hauptwohnung	Nebenwohnung	Anmerkung
X			Keine Änderung der Einwohnerzahl
	X	X (X...)	
	X		
X		X (X...)	Änderung der Einwohnerzahl
		X (X...)	
XX			
	XX		
X	X		
XX		X (X...)	
	XX	X (X...)	
X	X	X (X...)	

stimmungsprobleme in den Fällen ausschließen, in denen die Wohnungen in verschiedenen Bundesländern liegen<sup>7)</sup>.

Nach der Zusammenführung der Ergebnisse der insgesamt sechs Mehrfachfallprüfungen im Datenmaterial zum ersten Stichtag wurden die Datensätze der so erstellten Mehrfachfallpaare mit der Datenlieferung zum zweiten Stichtag 31. März 2002 auf zutreffende Stichtagsrelevanz verglichen (Bildung einer Differenzdatei). Durch den Abgleich beider Melderegisterauszüge konnten alle Mehrfachfälle statistisch bereinigt werden, die nach dem Stichtag 5. Dezember 2001 rückwirkend Änderungen in der An- oder Abmeldung bzw. dem Wohnungsstatus bei den Meldebehörden angezeigt haben. Alle Fälle, die nur in der zweiten Datenlieferung enthalten und die nicht stichtagsrelevant waren (z. B. Zuzug oder Geburt) sowie alle Fälle, die nur in der ersten Datenlieferung enthalten waren (z. B. Fortzug oder Tod), werden für die weitere primärstatistische Überprüfung nicht mehr herangezogen. Damit wurde der Forderung im ZensTeG, den Befragungsumfang für die zu

7) Vgl.: Bundeslandsübergreifender Datenaustausch.

überprüfenden Mehrfachfälle so gering wie möglich zu halten, Rechnung getragen.

Für alle übrigen Mehrfachfälle ist nach der Zerlegung der Dublettenketten eine anschließende primärstatistische Überprüfung erforderlich. Sie wurden nach dem Bundesland, in dem die jüngste Zuzugsadresse liegt, sortiert und dem zuständigen statistischen Landesamt übergeben.

#### 4.3 Primärstatistische Überprüfung der identifizierten Mehrfachfälle

Personen in der Stichprobe mit mehr als einer alleinigen Wohnung oder Hauptwohnung oder nur mit Nebenwohnungen (die so genannten Prüffälle) und mit der jüngsten Zuzugsadresse in NRW werden ab Juni 2002 durch das LDS NRW zum tatsächlichen Hauptwohnsitz befragt. Bundesweit sind 9 159 als Mehrfachfall identifizierte Personen, in NRW 2 506 Personen zu befragen.

Einen Überblick über die Verteilung der Mehrfachfälle auf die 396 nordrhein-westfälischen Gemeinden gibt das nachstehende Kartogramm. Hier werden je Gemeinde die identi-

fizierten Mehrfachfälle im Verhältnis zur Anzahl der in die Geburtstagsstichprobe gelangten Personen dargestellt. Es zeigt sich, dass die meisten kreisfreien Städte höhere Mehrfachfallanteile als der Landesdurchschnitt NRW (1,1 Prozent) aufweisen. Sehr hohe Anteile von über 1,5 Prozent werden überwiegend in kleineren Gemeinden erreicht, was jedoch auf eine Überzeichnung aufgrund der geringen Auswahlgrundlage zurückzuführen ist.

Nach § 13 Abs. 3 und 1 ZensTeG sowie nach § 15 BStatG besteht eine Auskunftspflicht der betroffenen Einwohnerinnen und Einwohner. Kommen Personen ihrer Auskunftspflicht nicht nach, weil die Antwort nicht, nicht vollständig, nicht termingerecht oder offensichtlich nicht wahrheitsgemäß erteilt wurde, können verwaltungsrechtliche Zwangsmaßnahmen zur Durchsetzung der Auskunftserteilung durchgeführt werden.

Im Zuge der Vorbereitungsarbeiten hat das LDS NRW ein Konzept für den bundesweiten Ablauf der Befragung einschließlich des Mahnwesens erstellt (vgl. folgende Darstellung). Im Rahmen der Durchführung kommt grundsätzlich ein vierstufiges Mahnverfahren zum Einsatz.

Der Ablauf der Befragung bzw. des Mahnverfahrens geht von einem Termin 0 aus, an dem die Erhebungunterlagen an die Haushalte verschickt werden. Somit ist der Termin 0 kein festes Datum, sondern kann für jede befragte Person variieren (z. B. Datum des Poststempels).

#### Termin 0

Der Termin 0 ist der Tag des Versandes der Erhebungsvordrucke und der Erläuterungen. Zur Beantwortung wird den Auskunftspflichtigen in NRW eine Frist von einer Woche eingeräumt. Das LDS NRW rechnet damit, spätestens nach vier Wochen mit der folgenden Mahnstufe zu beginnen.

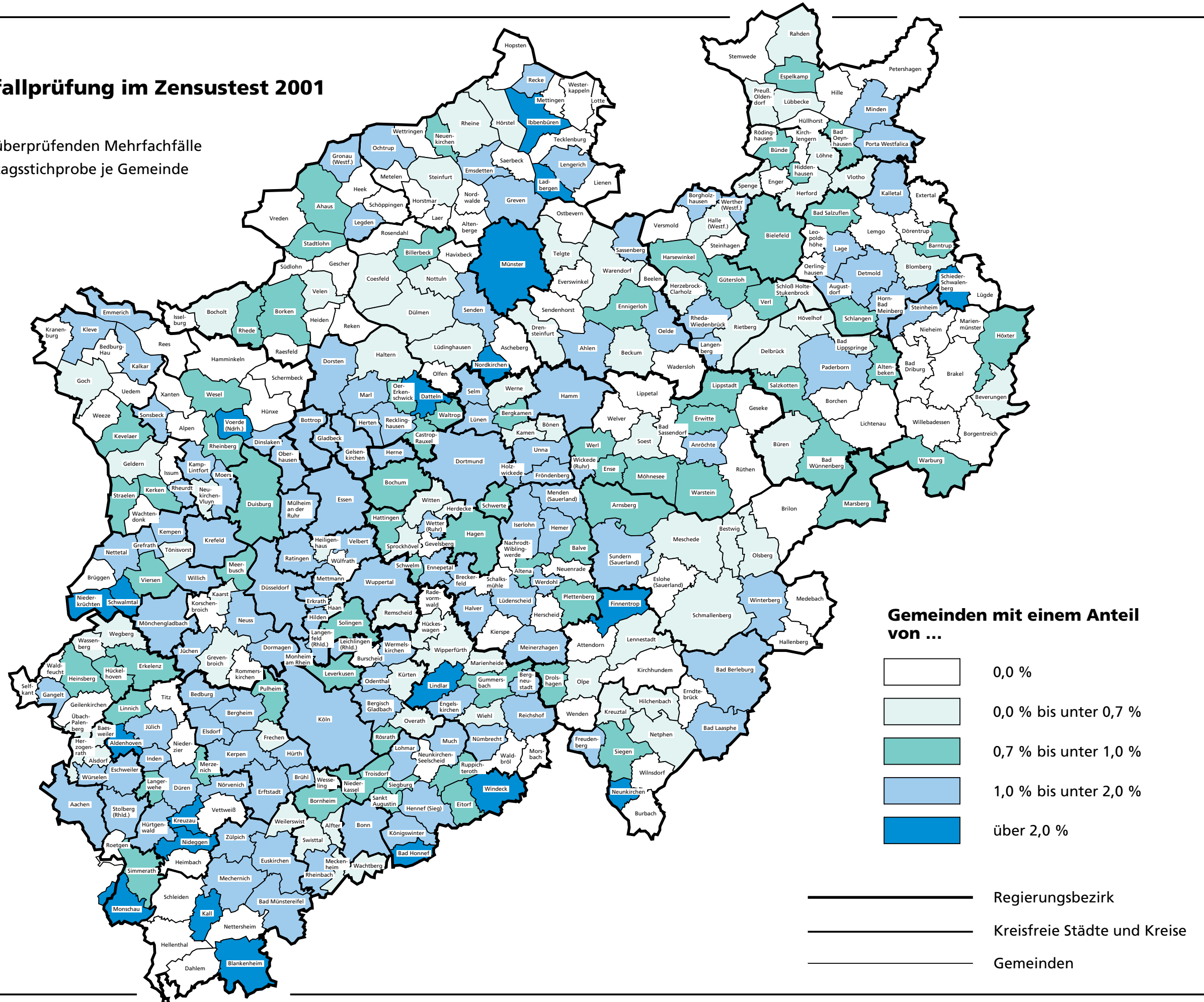
#### Termin 1

Zum Termin 1 erhalten alle Personen, die 28 Tage nach Termin 0 nicht



# Mehrfachfallprüfung im Zensusstest 2001

Anteile der zu überprüfenden Mehrfachfälle an der Geburtstagsstichprobe je Gemeinde





geantwortet haben, ein Mahnschreiben (1. Mahnstufe) in Form eines Erinnerungsschreibens mit den Erhebungspapieren als Anlagen. Zur Beantwortung haben die Auskunftspflichtigen eine einwöchige Frist. Für Personen, die eine Auskunft nach Termin 1 grundsätzlich verweigern, wird das Verfahren abgekürzt und direkt mit dem Verfahren nach Termin 3 fortgefahren. Auch hier wird davon ausgegangen, dass nach spätestens vier Wochen mit der folgenden Mahnstufe begonnen werden kann.

### Termin 2

Kann 28 Tage nach Termin 1 kein Rücklauf des Erinnerungsschreibens registriert werden, so wird ab Termin 2 eine telefonische Befragung (2. Mahnstufe) durchgeführt. Die Telefonbefragung soll sich in NRW über 61 Tage erstrecken, wobei nur an Werktagen und einigen Sonnabenden Interviews durchgeführt werden. Termin 2 ist auf 56 Tage nach Termin 0 festgelegt. Ergibt die telefonische Befragung, dass Sprach-, Verständnisprobleme oder sonstige Schwierigkeiten vorliegen, die am Telefon nicht gelöst werden können, sind grundsätzlich persönliche Befragung durch einen Interviewer vorgesehen.

### Termin 3

Personen, deren Telefonnummer nicht zu ermitteln ist, die in der telefonischen Befragung auch nach sechs Versuchen nicht erreicht werden oder die eine telefonische Befragung verweigern, erhalten zum Termin 3 (3. Mahnstufe) einen Heranziehungsbescheid. Diesem liegen wiederum die Erhebungspapiere als Anlagen bei. Es handelt sich um einen formalen Heranziehungsbescheid (HZB), in welchem den Auskunftspflichtigen eine Frist von 14 Tagen gesetzt und ein Zwangsgeld in Höhe von 100 € angedroht wird. Der Versand erfolgt gegen Postzustellungsurkunde. Termin 3 liegt 117 Tage nach Termin 0.

### Termin 4

Kann 14 Tage nach Termin 3 kein Rücklauf des Fragebogens registriert

werden, so ist in der letzten Mahnstufe (4. Mahnstufe) vorgesehen, ein Zwangsgeld festzusetzen. Diese Zwangsgeldfestsetzung sollte ebenfalls mit Postzustellungsurkunde verschickt werden. Dazu wäre vorher eine verwaltungsrechtliche Prüfung der Voraussetzungen für die Festsetzung von Zwangsgeld durchzuführen (hierzu gehört der Nachweis über die Zustellung der Androhung (HZB)). Neben der Festsetzung des Zwangsgelds wird dem Empfänger eine erneute Rücksendefrist der Unterlagen gesetzt.

Die Fristfestsetzung erfolgt in jedem Mahnschreiben individuell. Wie in der folgenden Tabelle zusammengefasst, ergäbe sich vom Erstversand bis zum Ende des Mahnverfahrens voraussichtlich ein Zeitaufwand von maximal 131 Tagen.

## 4.4 Erhebungsunterstützung

Im Rahmen der primärstatistischen Überprüfung von Mehrfachfällen kommen in NRW zwei Erhebungsunterstützungsprogramme zur Anwendung. Vom Statistischen Landesamt Sachsen wurde eine PC-gestützte Erhebungsorganisation erstellt, welche auf dem im Zuge der Gebäude- und Wohnungszählung (GWZ) eingesetzten Programm basiert. Hiermit wird die gesamte Verwaltung der Datensätze von Mehrfachfällen, die Rücklaufkontrolle und die Erfassung der Erhebungsunterlagen sowie die Durchführung des postalischen Mahnwesens gesteuert.

Das LDS NRW hat für die in der zweiten Mahnstufe (Termin 2) durchgeführten Telefoninterviews eine PC-gestützte Erhebungsorganisation auf

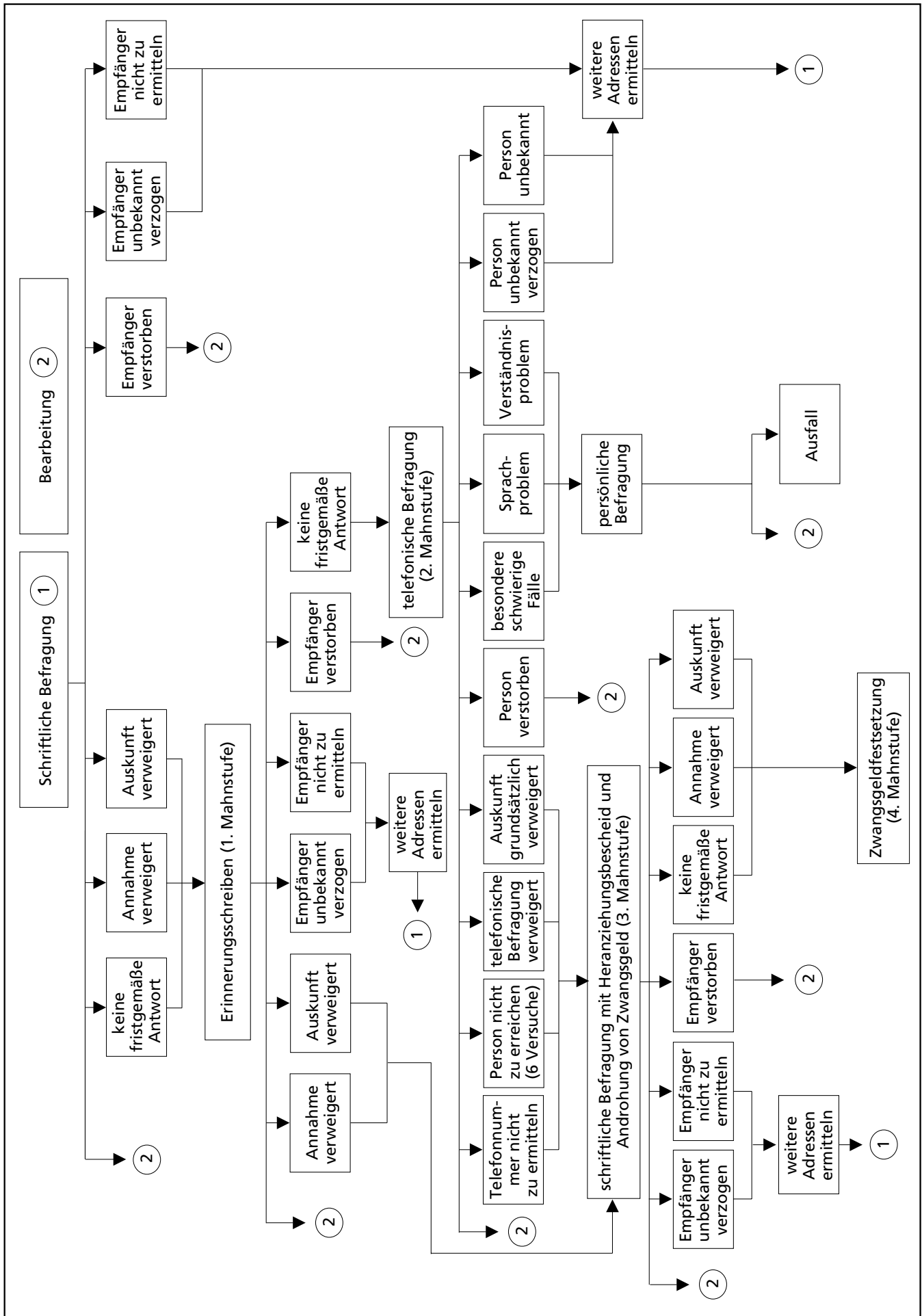
Basis der Programmiersprache Blaise entwickelt. Bei diesem CATI-Programm<sup>8)</sup> handelt es sich um eine Verbundlösung, d. h. das Programm kann von allen statistischen Landesämtern eingesetzt werden. Das LDS NRW hat eine Verknüpfung der Telefonnummernrecherche über eine Telefonauskunfts-CD-ROM hergestellt, so dass die Suchergebnisse automatisch in die CATI-Datenbank eingespielt werden können. Die eigentliche telefonische Befragung wird mit Hilfe der CATI-Module durchgeführt. Hierzu müssen die aus dem Erhebungsunterstützungsprogramm für die postalische Befragung übernommenen Datensätze verwaltet werden (z. B. Wiedervorlage zu bestimmten Tageszeiten und Wochentagen, Treffen von Verabredungen usw.). Weiterhin stehen zwei Module zur Spezifikation und Verwaltung zur Verfügung. Die programmierenden statistischen Landesämter in Sachsen und NRW haben die Schnittstellen zwischen beiden Erhebungsprogrammen angepasst, so dass ein Datenaustausch jederzeit möglich ist.

## 4.5 Bundeslandsübergreifender Datenaustausch

Die als Mehrfachfälle identifizierten Personen werden grundsätzlich nur durch das jeweilige statistische Landesamt zum Wohnungsstatus postalisch befragt, in dessen Zuständigkeitsbereich die jüngste Zuzugsadresse liegt. Da die Mehrfachfallprüfung auf Basis einer bundesweiten Stichprobe durchgeführt wird, können in jeder Befragungs-/Mahnstufe

8) CATI steht für Computer Assisted Telephone Interview.

Befragungsdauer und Befragungsumfang				
Befragungsdauer je Mahnstufe			Zum Zeitpunkt der Vorbereitung erwartete Fallzahlen in NRW	
T 0		0 Tage	Erstversand	2 506
T 1	T 0 + 28 Tage =	28 Tage	Erinnerungsschreiben	ca. 1 300
T 2	T 1 + 28 Tage =	56 Tage	Telefonische Befragung (davon Telefon-Nr. zu recherchieren)	ca. 700 (500)
T 3	T 2 + 61 Tage =	117 Tage	HZB	ca. 250
T 4	T 3 + 14 Tage =	131 Tage	Zwangsgeldfestsetzung (= Ausfälle in NRW)	ca. 125



bundeslandsübergreifende Mehrfachfälle auftreten. Liegt z. B. der Hauptwohnsitz einmal in Köln (NRW) und einmal in Frankfurt a. M. (Hessen) und kommen die Erhebungsunterlagen unter der jüngsten Zuzugsadresse als unzustellbar zurück, weil der Empfänger unbekannt verzogen ist, dann muss der Mehrfachfall unter der nächstälteren Adresse angeschrieben werden. Damit geht die Zuständigkeit für die Weiterbearbeitung dieses Datensatzes von einem statistischen Landesamt auf ein anderes über.

Damit eine reibungslose Übergabe dieser bundeslandsübergreifenden Mehrfachfälle zwischen den statistischen Landesämtern gewährleistet werden kann, ist folgendes vorgesehen: Die Datensätze werden – entsprechend ihres individuellen Rücklaufs und der Erfassung im erstversendenden statistischen Landesamt – elektronisch an das Statistische Bundesamt übermittelt. Hierzu ist es ausreichend, lediglich die Bogenummer, die Angabe der Adresse sowie das Versanddatum zu übermitteln und nicht den kompletten Datensatz. Dieser reduzierte Datensatz enthält keine personenbezogenen Angaben, wodurch im Rahmen des elektronischen Versands der Datenschutz gewährleistet bleibt.

Da der bundesweite Gesamtdatenbestand im Statistischen Bundesamt vorliegt, können dort täglich diese auszutauschenden Datensätze um alle Merkmale aufgefüllt werden. Im Anschluss daran findet über die jeweils aktuelle Adresse eine Sortierung nach Bundesländern statt. Die entsprechenden Datensätze werden den zuständigen statistischen Landesämtern in einem geschützten Bereich des Zentralen Datenservers der statistischen Ämter in Deutschland bereitgestellt. Die aktuellen Datensätze sind dann den statistischen Landesämtern täglich zugänglich, können in die jeweiligen Datenbanken der Erhebungsunterstützungsprogramme geladen werden und stehen für die primärstatistische Überprüfung zur Verfügung.

## 5 Weitere Schritte

Zu den Stichtagen 5. Dezember 2001 und 31. März 2002 haben alle Gemeinden in NRW für jede in ihrem Melderegister verzeichnete Person der Geburtstagsstichprobe einen Datensatz mit den genannten Erhebungs- und Hilfsmerkmalen übermittelt. Die Lieferungen gingen bis Ende Februar bzw. Anfang Mai beim LDS NRW ein und wurden dort, entsprechend den Vorgaben, in eine einheitliche Datensatzstruktur gebracht und auf Vollständigkeit, Vollzähligkeit sowie Plausibilität getestet. Seit Mitte März 2002 lag dem Statistischen Bundesamt aus allen Bundesländern das vereinheitlichte Datenmaterial vom ersten Stichtag vor. Dort wurde der bundesweite Gesamtdatenbestand für die eigentliche Mehrfachfallprüfung vorbereitet und an die externen Software-Anbieter übergeben. Im Mai 2002 wurden die Ergebnisse der insgesamt sechs Dublettenprüfungen durch das Statistische Bundesamt zusammengeführt und mit dem Datenmaterial vom zweiten Stichtag abgeglichen. Nach der Zerlegung von Dublettenketten wurden den statistischen Landesämtern im Juni die Datensätze zur primärstatistischen Überprüfung übergeben. Gegenwärtig führt das LDS NRW bei 2 506 Personen, deren Meldeadresse in NRW liegt, die postalische und telefonische Befragung zum Hauptwohnsitz durch. Die Erhebung wird voraussichtlich im November 2002 abgeschlossen sein, so dass im Anschluss daran ein bundesweites Ergebnis der primärstatistischen Überprüfung durch die statistischen Landesämter vorliegen könnte.

In jedem Zensus stellt die amtliche Einwohnerzahl für Bund, Länder und Gemeinden, also die Zahl der Personen, die mit alleinigem oder Hauptwohnsitz innerhalb der jeweiligen regionalen Gebietseinheit gemeldet sind, eine wichtige Basis für zahlreiche politische und wirtschaftliche Entscheidungsprozesse dar. Ziel der Mehrfachfallprüfung ist die Feststellung, inwieweit Personen mit mehreren alleinigen oder Hauptwohnun-

gen oder ausschließlich mit Nebenwohnungen in den Melderegistern verzeichnet sind. Auf diesen Ergebnissen aufbauend, müssen in einem registergestützten Zensus Verfahren entwickelt werden, mit denen Mehrfachfälle in den Melderegistern bereinigt und in Verbindung mit dem Verfahren der Haushaltgenerierung<sup>9)</sup> die amtliche Einwohnerzahl valide bestimmt werden kann.

Das im Rahmen des Zensus-tests angewandte Verfahren ist eine Stichprobenprüfung über eine Geburtstagsauswahl von bundesweit etwa 1,2 Prozent der Bevölkerung. Die Ergebnisse der Mehrfachfallprüfung dienen u. a. der Evaluation des Verfahrens Registergestützter Zensus. Da jedoch nicht alle Melderegisterdatensätze mit derselben Wahrscheinlichkeit in die Stichprobe gelangen, so sind die Sätze mit unvollständigem Geburtsdatum in der Stichprobe überrepräsentiert, ist eine Hochrechnung (d. h. eine Gewichtung) der Stichprobenergebnisse erforderlich. Weiterhin dürfen die Ergebnisse keine Verzerrungen im Sinne von systematischen Über- und Unterschätzungen einzelner Gruppen aufweisen, die durch das Antwortverhalten bestimmt sind, da trotz der im ZensTeG vorgeschriebenen Auskunftspflicht mit Antwortausfällen bei der primärstatistischen Überprüfung zu rechnen ist. Im Rahmen der Hochrechnung müssen diese Antwortausfälle berücksichtigt werden.

Unter den drei ausgewählten Geburtstagen nimmt der 1. Januar eine Sonderstellung ein, da häufig unvollständige Geburtsangaben von den Meldebehörden auf den 1. Januar vervollständigt werden. Damit unterscheidet sich dieser Geburtstag in seiner Struktur von den übrigen Geburtstagen, was auch in den deutlich höheren Fallzahlen zum Ausdruck kommt. Eine gleiche Gewichtung aller drei Stichprobengeburtstage bei der Hochrechnung

9) Vgl.: Scharmer, Marco, Haushaltstypisierung im Rahmen des registergestützten Zensus-tests, in: Statistische Analysen und Studien Nordrhein-Westfalen, Bd. 1 (2002), S. 3 ff.

würde somit die Ergebnisse verfälschen. Da die Anpassung an Eckwerte der Grundgesamtheit erfolgen wird und sowohl die Datensätze mit Geburtsdatum 1. Januar als auch Sätze mit unvollständigem Geburtsdatum für sich selbst stehen und damit als Totalerhebung in die Grundgesamtheit eingehen, ist hier keine Hochrechnung erforderlich. Im Rahmen der Evaluation werden überwiegend Befragungsergebnisse ausgewertet. Antwortausfälle, die

grundsätzlich bei allen Personen der Geburtstagsauswahl und auch bei solchen mit unvollständigen Geburtsdaten auftreten können, müssen hier also berücksichtigt werden. Zusätzlich ist angedacht, Antwortausfälle auf die Verteilung von demographischen Merkmalen hin zu analysieren und gegebenenfalls die entsprechenden Antwortausfälle mit einem Korrekturfaktor in die Hochrechnung einzubeziehen.

Die Auswertung des Erhebungsteils Mehrfachfallprüfung im Rahmen des Zensus 2001 wird sich bis in die erste Jahreshälfte 2003 erstrecken. Die Ergebnisse dieser Analyse werden entscheidend zur Evaluation und Weiterentwicklung des Modells registergestützter Zensus beitragen.

<b>Ablauf und Arbeitsschritte in der Mehrfachfallprüfung</b>		
Vorbereitende Arbeitsschritte/ Information der Gemeinden		ab Januar 2001
Datenanforderung zum 5. Dezember 2001: zur Mehrfachfallprüfung zur Zahl der gemeldeten Einwohner/-innen zur Führung der Melderegister		August – September 2001
Erste Datenlieferung: zur Mehrfachfallprüfung zur Zahl der gemeldeten Einwohner/-innen zur Führung der Melderegister		Dezember – Februar 2002
Eingangskontrolle erste Datenlieferung: Vereinheitlichung der Datensätze Vollständigkeit, Vollständigkeit, Plausibilität		Dezember – Februar 2002
Übermittlung der ersten Datenlieferung an das Statistisches Bundesamt	Datenanforderung zum 31. März 2002 zur Mehrfachfallprüfung	Februar 2002
Erstellung eines einheitlichen Datensatzes und Aufspaltung der Stichprobe in erster Datenlieferung	Zweite Datenlieferung: zur Mehrfachfallprüfung	April 2002
Mehrfachfallprüfung durch externe Softwareanbieter und 1:1-Abgleich durch Statistisches Bundesamt		April 2002
Zusammenführung aller identifizierten Mehrfachfälle	Eingangskontrolle zweite Datenlieferung: Vereinheitlichung der Datensätze Vollständigkeit, Vollständigkeit, Plausibilität	Mai 2002
Klärung von Mehrfachfällen anhand der Bildung von Dublettenketten	Übermittlung der zweiten Datenlieferung an das Statistisches Bundesamt	Mai 2002
	Erstellung eines einheitlichen Datensatzes mit zweiter Datenlieferung	Mai 2002
Abgleich erster mit zweiter Datenlieferung (Erstellung einer Differenzdatei) und Durchführung der Adressstandardisierung		Mai 2002
Zerlegung der Dublettenketten in Dublettenpaare		Mai 2002
Übermittlung identifizierter Mehrfachfälle an Statistische Landesämter		Juni 2002
Primärstatistische Überprüfung der Mehrfachfälle (postalisch und telefonisch)		Juni – November 2002
Lieferung an das Statistische Bundesamt und Auswertung der Ergebnisse		ab Dezember 2002



# Die Statistikauglichkeit der Einwohnermelderegister Untersuchungen im Rahmen des Zensus 2001

Dr. Michael Forster

*Bevor die Entscheidung fällt, zukünftig in Deutschland registergestützte Zensen durchzuführen, sollen im Rahmen des so genannten „Zensus 2001“ die Qualität und Statistikauglichkeit der zu verwendenden Register überprüft sowie die vorgesehenen Verfahren zur Zusammenführung und Haushaltegenerierung erprobt und verbessert werden. Im Folgenden wird ein Teilbereich des Zensus 2001, nämlich die Überprüfung der Statistikauglichkeit der Einwohnermelderegister behandelt. Aufgezeigt wird die Einbettung in den gesamten Zensus 2001, die methodischen Konzeption und Erfahrungen bei der Zusammenarbeit mit den Gemeinden.*

## 1 Die Erprobung eines registergestützten Zensus

Derzeit wird bundesweit ein neues Zensus-Verfahren erprobt. Ziel dieses Zensus 2001<sup>1)</sup> ist zum einen die Qualität der verwendeten Register (Melderegister und Register der Bundesanstalt für Arbeit) festzustellen sowie zum anderen die vorgesehenen Verfahren zu erproben und zu verbessern. Hierzu finden seit Ende des vergangenen Jahres verschiedene Testerhebungen statt, die sich alle auf den durch das Zensusgesetz<sup>2)</sup> vorgegebenen Stichtag – den 5. Dezember 2001 – beziehen. Es werden folgende Prüfungen durchgeführt:

- Test der Melderegister auf Über- und Untererfassungen (Register-test)
- Test der Verfahren eines registergestützten Zensus (Verfahrenstest)
- Untersuchung auf mehrfache Meldungen einzelner Personen (Mehrfachfallprüfung)

Beim **Register-test** soll in erster Linie festgestellt werden, wie hoch die Abweichungen zwischen den Eintragungen im Melderegister und den tatsächlichen Verhältnissen sind. Zu diesem Zweck werden bundesweit

ca. 450 000 Bewohner von ausgewählten Wohngebäuden in 555 zufällig bestimmten Gemeinden (4 % der Gemeinden in Deutschland) interviewt. In Nordrhein-Westfalen sind vom Register-test etwa 48 000 Personen in 4 138 Gebäuden<sup>3)</sup> verteilt auf 36 Städte und Gemeinden (9 % der Gemeinden in Nordrhein-Westfalen) betroffen. Die Ergebnisse der Befragungen werden in den statistischen Landesämtern mit den Daten verglichen, die die Gemeinden aus den Melderegistern für diese Adressen übermittelt haben.

Im Rahmen eines **Verfahrenstests** werden außerdem die bei einem registergestützten Zensus benötigten neuen statistischen Verfahren zur Zusammenführung von Daten und zur Generierung von Haushalten in 222 der bundesweit ausgewählten 555 Städte und Gemeinden – davon 18 in Nordrhein-Westfalen – überprüft. Betroffen sind davon bundesweit 16 000 zufällig ausgewählte Adressen mit ca. 230 000 Personen. Hierunter sind auch 1 650 Adressen mit etwa 20 000 Bewohnern in Nordrhein-Westfalen.

Im Laufe dieses Jahres werden darüber hinaus die Daten der Melderegister auf Mehrfachmeldungen untersucht (**Mehrfachfallprüfung**). Dafür liefern alle Gemeinden in Deutschland die Datensätze der Personen, die an drei zuvor festgelegten Tagen (1. 1., 15. 5. und 1. 9.) Geburtstag haben oder deren Geburtsdaten unvollständig im Melderegister gespeichert sind (sog. Geburtstagstichprobe). Diese Datensätze werden zentral im Statistischen Bundesamt geprüft. Soweit bei der Auswertung Doppelfälle (mehrfache alleinige oder Hauptwohnungen oder ausschließlich Nebenwohnungen) festgestellt werden, müssen die statistischen Landesämter durch Nachfragen bei den Betroffenen klären, ob sich die Angaben auf verschiedene Personen beziehen oder Registerfehler vorliegen.

Im Folgenden wird schwerpunktmäßig der Register-test behandelt. Zur Mehrfachfallprüfung und zu verschiedenen Elementen des Verfahrenstests sind weitere Arbeiten<sup>4)</sup> in der Reihe „Statistische Analysen und Studien“ erschienen.

## 2 Der Register-test

Die „Testerhebungen zur Untersuchung von Über- und Untererfassungen in Melderegistern bei Meldebehörden und Personen in ausgewählten Gebäuden“ – so die exakte Formulierung im Zensusgesetz – findet in 555 stellvertretend ausgewählten Gemeinden in Deutschland statt. Je nach Gemeindegröße wur-

1) Vgl. Forster, M.: Die Zukunft der Volkszählung in Deutschland – traditionelle Zählung oder registergestützter Zensus?; Statistische Analysen und Studien Nordrhein-Westfalen, 4/2001; S. 13 ff. – 2) Vgl. Gesetz zur Vorbereitung eines registergestützten Zensus (Zensusgesetz – ZensusTeG) vom 27. Juli 2001 BGBl. I. S. 1882. – 3) Bei der Stichprobenziehung wurden Adressen aus den Melderegistern ausgewählt. Erhoben werden die Bewohner aller Gebäude unter diesen Adressen. Da in den weitaus meisten Fällen je Adresse nur ein Gebäude vorhanden ist, werden die beiden Begriffe im Folgenden synonym verwendet. – 4) Vgl. weitere Aufsätze in diesem Heft sowie Egbert, H./Scharmer, M.: Die Ermittlung des Wohnungsinhabers im Rahmen der Gebäude- und Wohnungszählung (GWZ) und deren Bedeutung für die Haushaltegenerierung; Statistische Analysen und Studien Nordrhein-Westfalen, Bd. 1 (2002); S. 10 ff und Scharmer, M.: Haushaltstypisierung im Rahmen des registergestützten Zensus; Statistische Analysen und Studien Nordrhein-Westfalen, Bd. 1 (2002); S. 3 ff.

den 40, 80, 160 oder mehr<sup>5)</sup> bewohnte Adressen aus den Melderegistern hierfür ausgewählt. Zu allen unter diesen Adressen zum Stichtag gemeldeten Personen werden aus den Melderegistern folgende Erhebungsmerkmale abgefragt:

- a) Geburtsmonat und -jahr,
- b) Geschlecht,
- c) Staatsangehörigkeiten,
- d) Familienstand,
- e) Wohnort,
- f) Status der Wohnung (alleinige, Haupt- oder Nebenwohnung).

Neben diesen demographischen Grunddaten werden eine Reihe von Hilfsmerkmalen erhoben, die entweder zur zweifelsfreien Identifikation einer Person (Merkmale g – i) oder zur Feststellung der Stichtagsrelevanz (Merkmale j – o) benötigt werden:

- g) Namen, Vornamen,
- h) Tag der Geburt,
- i) gegenwärtige Anschriften,
- j) Datum des Beziehens der Wohnung,
- k) Datum des Auszugs aus der Wohnung,
- l) Datum der Anmeldung bei der Meldebehörde,
- m) Datum der Abmeldung bei der Meldebehörde,
- n) Datum des Wohnungsstatuswechsels,
- o) Ordnungsmerkmal der Meldebehörde für die gemeldete Person.

Die gleichen Merkmale – mit Ausnahme der Datumsangaben und des Ordnungsmerkmals (j – o) – werden parallel zur Abfrage in den Melderegistern auch durch Erhebungsbeauftragte bei den betroffenen Bürgern erhoben<sup>6)</sup> und später miteinander verglichen.

Für die Adressen, die auch für die Unterstichprobe Verfahrenstest ausgewählt wurden, mussten neben den hier genannten noch eine Reihe von weiteren Hilfsmerkmalen (u. a. die Verzeigerungen<sup>7)</sup>) aus den Melderegistern geliefert werden.

### 3 Kooperation mit den Gemeinden

In allen Erhebungsbereichen ist eine mehr oder weniger intensive Zusammenarbeit mit den Gemeinden erforderlich. Zu den Aufgaben der Gemeinden zählten insbesondere verschiedene Auswertungen und Datenerlieferungen aus den Melderegistern sowie die Übermittlung der Namen und Anschriften der Eigentümer oder Verwalter der ausgewählten Gebäude.

Bereits im Vorfeld des Gesetzgebungsverfahrens sind die statistischen Ämter der Gemeinden vom Landesamt für Datenverarbeitung und Statistik Nordrhein-Westfalen (LDS NRW) – im Rahmen der auf Landesebene turnusmäßig stattfindenden Arbeitsbesprechungen – über den Zensusstest informiert worden. In den auf Bundesebene arbeitenden Projektgruppen zur Vorbereitung der Erhebungen waren von Anfang an auch Vertreter des Deutschen Städtetages beteiligt. Nachdem das Zensusvorbereitungsgesetz – dessen erster Artikel das Zensusstestgesetz ist – am 18. Mai 2001 vom Deutschen Bundestag in dritter Lesung verabschiedet wurde, wurden alle Gemeinden des Landes über das Vorhaben informiert und um die Benennung von Ansprechpartnern (sowohl in den Gemeinden, als auch in den zuständigen Rechenzentren) gebeten. Die Hersteller von Software für das Meldewesen und die bereits bekannten Rechenzentren wurden zeitgleich zentral vom Statistischen Bundesamt informiert.

Die in Nordrhein-Westfalen für den Register- und Verfahrenstest ausgewählten 36 Gemeinden wurden zu einer Informationsveranstaltung am 31. Mai 2001 ins LDS NRW nach Düsseldorf eingeladen. An dieser Veranstaltung nahmen sowohl Vertreter der kommunalen Statistikstellen und der Meldebehörden teil als auch Vertreter bzw. Programmierer der kom-

munalen Rechenzentren. In weiteren Informationsschreiben wurde über den Fortgang des Gesetzgebungsverfahrens informiert.

Bis zum 22. August 2001 hatten alle 36 ausgewählten Gemeinden die erste Auswertung aus den Melderegistern (alle bewohnten Adressen mit der Anzahl der gemeldeten Personen) geliefert, so dass die Stichprobenziehung der Adressen im LDS NRW erfolgen konnte.

Daran schloss sich die Übermittlung der Gebäudeeigentümer bzw. -verwalter durch die Grundsteuerstellen der Gemeinden an. Danach wurden die Begehungslisten aus den Melderegistern (Namen und Geburtsjahr aller Personen, die unter den ausgewählten Adressen gemeldet sind) geliefert.

Die Lieferung der Datensätze zum Stichtag 5. Dezember 2001 erfolgte allerdings erst mit z. T. erheblichen Verzögerungen. Von den 36 ausgewählten Gemeinden lieferten nur 27 termingerecht bis zum 4. Januar 2002. 5 Wochen nach Termin konnte dieser Teilschritt abgeschlossen werden.

Daran schloss sich die als Hochrechnungsbasis benötigte Auswertung der Melderegister zur Zahl der gemeldeten Personen nach Wohnungsstatus und Staatsangehörigkeit (deutsch/nichtdeutsch) an.

Die im Rahmen der Mehrfachfallprüfung von allen Gemeinden angeforderten Datensätze (Geburtstagsauswahl) lagen vollständig mit 4 Wochen Terminverzögerung vor.

Positiv anzumerken ist, dass fast alle Gemeinden (in zwei Ausnahmefällen wurden EXCEL-Dateien geliefert) das vorgeschriebene Format (Text-File entsprechend einer vorgegebenen Datensatzbeschreibung) einhielten. Der Versandweg konnte von den Gemeinden frei gewählt werden, so dass verschiedene Datenträger (Dis-

5) Für 20 Gemeinden, die fest ausgewählt wurden (i. d. R. die größten im jeweiligen Bundesland) wurde der Stichprobenumfang jeweils derart festgelegt, dass sich für Gebäude gleicher Größe die gleichen Auswahlchancen innerhalb des Landes ergaben. In Nordrhein-Westfalen betrifft dies nur die Stadt Köln mit 218 Adressen. – 6) Vgl. Schäfer, J.: Die Rolle der Haushaltebefragung im Zensusstest 2001; Beitrag in diesem Heft, S. 41 ff. – 7) Mit Verzeigerungen werden die Angaben zu Ehegatten, Kindern oder gesetzlichen Vertretern bezeichnet, die im Melderegister i. d. R. zu jeder Person gespeichert sind.

Aufgaben der Kommunen in NRW			
Stichtag	Aufgabe	Solltermin für Lieferung	Letzter Eingang
7/2001	Auswertung der Zahl der gemeldeten Personen je Anschrift aus dem Melderegister	15. 8. 2001	22. 8. 2001
	Liste der „fiktiven Adressen“ (z. B. Meldeadressen Obdachloser) sowie nicht zu befragende Adressen (Übergangsheime, Frauenhäuser, Campingplätze u. Ä.)		
9/2001	Lieferung von Adressen der Gebäudeeigentümer oder -verwalter durch die Grundsteuerstellen	1. 10. 2001	15. 10. 2001
1. 10. 2001	Liste mit Namen und Geburtsjahr aller an ausgewählten Adressen gemeldeten Personen (sog. Begehungsliste)	24. 10. 2001	14. 11. 2001
5. 12. 2001	a) Auswertung des Melderegisters zur Zahl der mit alleiniger, Haupt- oder Nebenwohnung gemeldeten Personen nach Deutschen und Nichtdeutschen (Faxformular)	4. 1. 2002	25. 1. 2002
	b) Melderegisterauszüge aller Personen mit Geburtsdatum 1. 1., 15. 5. oder 1. 9. sowie Personen mit unvollständig gespeichertem Geburtsdatum für die Mehrfachfallprüfung		1. 2. 2002
	c) Melderegisterauszüge der gemeldeten Personen an allen ausgewählten Adressen für den Register- und den Verfahrenstest		12. 2. 2002
31. 3. 2002	a) Melderegisterauszüge aller Personen mit Geburtsdatum 1. 1., 15. 5. oder 1. 9. sowie Personen mit unvollständig gespeichertem Geburtsdatum für die Mehrfachfallprüfung	19. 4. 2002	2. 5. 2002
	b) Melderegisterauszüge der gemeldeten Personen an allen ausgewählten Adressen für den Register- und den Verfahrenstest		26. 4. 2002

ketten, CD-ROMs, Magnetband-Kassetten), aber auch E-Mails (i. d. R. verschlüsselte ZIP-Dateien) zum Einsatz kamen.

Technische Probleme bei den Datenlieferungen waren die Ausnahme. Lediglich von 4 Gemeinden mussten aus diesem Grund (defekte Disketten etc.) Neulieferungen angefordert werden. Bei 34 Gemeinden (28 für Geburtstagsauswahl, 6 für Adressauswahl) war eine Neulieferung aufgrund inhaltlicher Fehler (falsche Feldlängen, teilweise falsche Auswahl der Datensätze etc.) erforderlich. In einigen anderen Fällen wurden Fehler in den Datensätzen den Ansprechpartnern nur telefonisch mitgeteilt und direkt im LDS NRW behoben.

Die letzten Aufgaben der Kommunen – dies waren die Wiederholungen der Datenlieferungen für die Geburts-

tags- und die Adressauswahl, diesmal zum Stichtag 31. März 2002 – sind abgesehen von wenigen Ausnahmen weit gehend termingerecht sowie technisch und inhaltlich einwandfrei erfolgt. Damit steht der weiteren Bearbeitung im Rahmen des Zensus-tests nichts entgegen.

Die Kooperation mit den Ansprechpartnern in den Gemeinden und den beauftragten Rechenzentren war i. d. R. problemlos. Die angeforderten Neulieferungen erfolgten innerhalb kurzer Zeit. Die Fragen, die sich im LDS NRW wegen einiger Unstimmigkeiten ergaben, konnten in sehr kurzer Zeit (meistens direkt telefonisch) beantwortet werden.

Insgesamt kann ein aus Sicht des LDS NRW erfreulich positiver Verlauf dieser komplexen und für alle Seiten neuen Arbeitsschritte festgestellt werden.

## 4 Zusammenführung der beiden Melderegister-Auszüge

### 4.1 Stichtagsrelevanz

Für die Ermittlung von Über- und Untererfassungen sowie für die Überprüfung der demographischen Merkmale in den Melderegistern (MR) werden die Registerdaten mit den Ergebnissen der Haushaltebefragung satzweise abgeglichen. Hierzu müssen zunächst die beiden Lieferungen (zum Stichtag 5. 12. 2001 und vier Monate später zum 31. 3. 2002) aus den Melderegistern zusammengeführt und anschließend die Stichtagsrelevanz jeder einzelnen Person festgestellt werden und zwar unabhängig davon, ob sie nur in einer oder in beiden MR-Lieferungen enthalten ist. Mit Stichtagsrelevanz wird die Tatsache bezeichnet, dass eine Person am Stichtag noch (oder schon) unter der Adresse gewohnt hat (bzw. gemeldet war). Des Weiteren müssen auch alle veränderlichen Merkmale soweit möglich auf ihre Ausprägung am Stichtag überprüft werden.

Es wurde beschlossen, dass Veränderungen in den Melderegistern am Stichtag noch berücksichtigt werden, d. h. der für die Auswertungen relevante Zeitpunkt wurde auf 24.00 Uhr am 5. Dezember 2001 festgelegt<sup>8)</sup>. Wer am Stichtag auszieht, wird also nicht mehr als relevant angesehen, wer noch am Stichtag einzieht, wird dagegen bereits mitgezählt. Ausschlaggebend sind in jedem Fall die Wohnverhältnisse um 24.00 Uhr.

Für die stichtagsrelevanten Personen muss dann für alle Erhebungsmerkmale (soweit möglich) die Situation am Stichtag wiedergegeben werden. Für die Hilfsmerkmale (z. B. Datum der Anmeldung) ist eine Prüfung i. d. R. nicht notwendig. Eine Ausnahme bilden hier die Verzeigerungen von Ehegatten und Kindern, die für Zwecke der Haushaltegenerie-

<sup>8)</sup> Dies entspricht nicht dem Vorgehen bei der letzten Volkszählung. Damals galten die Verhältnisse am Stichtag um 0.00 Uhr. Diese Änderung schränkt die Vergleichbarkeit der Ergebnisse mit früheren Zählungen aber in keiner Weise ein.

rung auch wie die Erhebungsmerkmale stichtagsrelevant ermittelt werden sollen.

#### 4.2 Zusammenführung der Personensätze aus den Melderegister-Lieferungen

Die Personensätze aus der ersten und der zweiten MR-Lieferung (MR1 und MR2) werden adressweise anhand des (unveränderlichen) Ordnungsmerkmals der Person (OMP) zusammengeführt. Wenn in einem Datensatz das OMP fehlt (z. B. weil die Meldebehörde kein derartiges Kennzeichen führt oder es aufgrund landesrechtlicher Bestimmungen nicht liefern darf), wird ein künstliches OMP erzeugt. Hierzu werden die ersten Buchstaben von Geburtsname und Vorname, das Geburtsdatum und das Geschlecht benutzt. Da der Familienname sich im Zeitablauf durchaus ändern kann, wird er hier nicht einbezogen.

Es wird zunächst in den beiden MR-Lieferungen je Adresse eine eindeutige laufende Nummer der Person vergeben. Innerhalb einer Adresse erhalten die jeweils gefundenen paarigen Datensätze aus MR1 und MR2 die gleiche laufende Nummer. Die unpaarigen Sätze aus MR1 oder

MR2 erhalten jeweils die weiteren laufenden Nummern. Die laufende Nummer der Person wird für jede Adresse bei 1 beginnend vergeben. Zusätzlich zur laufenden Nummer der Person wird ein Wert für den Status vergeben, der das Ergebnis der Zusammenführung anzeigt. Es wird hier im ersten Schritt nur die Paarigkeit bzw. Unpaarigkeit festgestellt. Daher erhalten alle Sätze zunächst für den Status entsprechend ihrem Vorkommen in MR1 und/oder MR2 einen der Werte 1 bis 3. Bei einem Wert von 1 kommt der Datensatz in beiden, bei einem Werte von 2 oder 3 jeweils nur in einer MR-Lieferung vor.

In den ausgewählten Gebäuden in Nordrhein-Westfalen waren rd. 48 000 Personen gemeldet, wovon nur 45 700 (95 %) maschinell als paarig in MR1 und in MR2 gefunden wurden. Alle maschinell nicht paarig gesetzten Datensätze wurden nochmals manuell (anhand von Drucklisten) überprüft. Da einzelne der genannten Merkmale sich zwischenzeitlich geändert haben können (durch tatsächliche Änderungen und/oder durch Korrekturen im Melderegister), muss für die Feststellung der Paarigkeit keine 100%ige Übereinstimmung aller Merkmale vorliegen. Nur wenige Fälle wurden im LDS NRW noch manuell zusammengeführt, wodurch

der Anteil von 5 % der tatsächlich unpaarigen Fälle sich nicht veränderte. Größtenteils hatte die maschinelle Zusammenführung dann versagt, wenn sich das OMP entgegen den vorherigen Erwartungen geändert hatte (z. B. bei Namensänderungen oder bei Korrekturen des Geburtsdatums im Melderegister).

Die Stichtagsrelevanz einer (in einer oder in beiden Lieferungen enthaltenen) Person wird anhand des Geburts-, Einzugs- und Auszugsdatums überprüft (das Sterbedatum gehört nicht zum Lieferumfang der Meldebehörden und steht daher hier nicht zur Verfügung). Das Datum der Geburt und des Einzugs (bzw. der Anmeldung von Amts wegen) darf nicht nach dem Stichtag liegen und das Datum des Auszugs (bzw. der Abmeldung von Amts wegen) muss – falls überhaupt im Datensatz vorhanden – nach dem Stichtag liegen. In diesem Fall gehört eine Person zum stichtagsrelevanten Bestand. Alle nicht stichtagsrelevanten Personen werden besonders gekennzeichnet und ebenfalls in die Datei aufgenommen, da sie für Kontrollzwecke im Rahmen der Haushalgenerierung benötigt werden. Für sie finden aber keine weiteren Prüfungen zur Stichtagsrelevanz der einzelnen Merkmale statt.

MR1 (ZT250)				
AGS	Adressnummer	lfd. Nr. der Person	Status	Merkmale 1. Datenlieferung (MR1)
05111000	00001	0001	1	OMP; xxxxxxxxxxx
05111000	00001	0002	1	OMP; xxxxxxxxxxx
05111000	00001	0003	2	OMP; xxxxxxxxxxx
05354016	00232	0001	1	OMP; xxxxxxxxxxx

MR2 (ZT255)				
AGS	Adressnummer	lfd. Nr. der Person	Status	Merkmale 2. Datenlieferung (MR2)
05111000	00001	0001	1	OMP; xxxxxxxxxxx
05111000	00001	0002	1	OMP; xxxxxxxxxxx
05111000	00001	0004	3	OMP; xxxxxxxxxxx
05354016	00232	0001	1	OMP; xxxxxxxxxxx
05354016	00232	0002	3	OMP; xxxxxxxxxxx

Nach der Zusammenführung und Prüfung auf Paarigkeit wird in einem zweiten Schritt die Stichtagsrelevanz der einzelnen Personen anhand von Geburts-, Einzugs- und Auszugsdaten überprüft. Falls die Person nicht stichtagsrelevant ist, werden für den Status statt der Werte 1 bis 3 entsprechend die Werte 4 bis 6 vergeben.

Der Wert zum Status kann also abhängig vom Zusammenführungsergebnis und der Prüfung auf Stichtagsrelevanz folgendermaßen aussehen:

1. paarig und stichtagsrelevant
2. nur in MR1 und stichtagsrelevant
3. nur in MR2 und stichtagsrelevant
4. paarig und nicht stichtagsrelevant
5. nur in MR1 und nicht stichtagsrelevant
6. nur in MR2 und nicht stichtagsrelevant

Im Rahmen des Registertests wird der Status in der Kombination mit dem Ergebnis der Haushaltebefragung (z. B. wohnhaft und befragt; zweifelsfrei wohnhaft, aber nicht befragt; usw.) für eine Klassifikation bezüglich Karteileichen<sup>9)</sup> und Fehlbeständen verwendet.

#### **4.3 Stichtagsrelevanz der Erhebungsmerkmale**

Grundsätzlich sind die Angaben aus MR1 maßgeblich für die Prüfung der Statistikauglichkeit der Melderegister. Fehlen diese (weil die Person vollständig fehlte, oder weil nur ein Merkmal im Datensatz fehlte) werden die Werte aus MR2 übernommen. Nur bei einigen wenigen Merkmalen sind Prüfungen, ob der Wert aus MR1 oder MR2 übernommen wird, notwendig und möglich.

Da Geburtsdatum und Geschlecht als unveränderlich gelten, kann eine Abweichung i. d. R. nur aufgrund einer Korrektur des MR zustande kommen. Bei Abweichungen ist daher der Wert aus MR2 zu übernehmen.

Die Staatsangehörigkeiten können sich im Zeitverlauf (durch Einbürgerungen) durchaus ändern. Da aber keine Merkmale bzgl. des Zeitpunktes der Änderung vorhanden sind bzw. erhoben werden, gilt der Wert aus MR1 als stichtagsrelevant.

Das Merkmal Familienstand wird in Registertest und Verfahrenstest unterschiedlich behandelt. Im Registertest muss es (mangels Datumsangaben zur Ehe) ungeprüft aus MR1 übernommen werden, während es im Verfahrenstest bei Abweichungen zwischen MR1 und MR2 anhand der Hilfsmerkmale „Datum der letzten Eheschließung“ und „Datum der Beendigung der letzten Ehe“ geprüft werden kann. Letzteres ist notwendig, um für die Haushaltegenerierung den stichtagsrelevanten Familienstand bestimmen zu können.

<sup>9)</sup> Mit Karteileichen werden Übererfassungen des Melderegisters – also Personen, die nicht mehr unter der angegebenen Adresse wohnen – bezeichnet.

Dieses Vorgehen könnte zu unterschiedlichen Ergebnissen der Registerqualität im Vergleich zur Haushaltebefragung in den beiden Teilstichproben führen. Da nur in sehr wenigen Fällen damit zu rechnen ist, dass sich der Familienstand zwischen MR1 und MR2 ändern wird (und wiederum aufgrund der Prüfung im Verfahrenstest nur in einem Teil dieser Fälle nicht der Wert aus MR1 übernommen wird), wird es zwischen den beiden Teilstichproben vermutlich nicht zu signifikanten Unterschieden kommen. Auf eine getrennte Darstellung der beiden Teilstichproben in den Ergebnissen kann daher voraussichtlich verzichtet werden.

Das Merkmal Wohnungsstatus kann mit Hilfe des „Datums des letzten Wohnungsstatuswechsels“ bzw. „Datum des Wohnungsstatuswechsels von Amts wegen“ geprüft werden.

#### **4.4 Stichtagsrelevanz der Hilfsmerkmale**

Namen und Vornamen können verändert werden (durch Heirat, Adoption, Namensänderung). Der genaue Name zum Stichtag ist aber für die statistische Auswertung irrelevant, so dass keine exakte Überprüfung stattfinden muss. Bei Unterschieden zwischen den beiden Lieferungen werden die Namen aus MR2 übernommen, da die Kenntnis zwischenzeitlicher Namensänderungen für die Haushaltegenerierung hilfreich sein kann.

Die gegenwärtige Anschrift wird wie der Wohnort durch die Stichprobenauswahl vorgegeben und ist daher kein überprüfbares Merkmal. Der Name des Wohnungsgebers ist Teil der Anschrift und wird daher auch aus MR1 übernommen.

Das Datum des Beziehens der Wohnung und das Datum der Anmeldung bei der Meldebehörde werden, da i. d. R. unveränderlich, aus MR1 übernommen.

Datum des Auszugs aus der Wohnung und Datum der Abmeldung bei der Meldebehörde sowie Datum des Wohnungsstatuswechsels werden immer aus MR2 übernommen. Das Datum der letzten Eheschließung und das Datum der Beendigung der letzten Ehe werden (weil der aktuellere Stand) ebenfalls aus MR2 übernommen. Nach dem Stichtag liegende Daten können Indizien dafür sein, dass die Ehepartner auch am Stichtag schon bzw. nicht mehr zusammenlebten.

Für die Zwecke der Haushaltegenerierung müssen die Verzeigerungen von Ehegatten auf Stichtagsrelevanz überprüft werden. Diese Prüfung kann relativ einfach mit den Prüfungen des Familienstandes kombiniert werden. Ist der stichtagsrelevante Familienstand „Verheiratet“, so wird die Verzeigerung des Ehegatten aus MR1 beibehalten bzw. aus MR2 aufgenommen. Bei allen anderen Ausprägungen des Familienstands wird eine evtl. vorhandene Verzeigerung zum Ehegatten gelöscht.

Verzeigerungen für Kinder werden aus MR1 beibehalten. In MR2 neu hinzukommende Verzeigerungen werden nur aufgenommen, wenn das Geburtsdatum vor oder am Stichtag liegt. Verzeigerungen für gesetzliche Vertreter können nicht geprüft werden (in MR2 neu hinzukommende werden soweit möglich in MR1 aufgenommen).

Sowohl bei den Verzeigerungen als auch bei allen anderen hier angesprochenen Prüfungen findet keine datensatzübergreifende Bearbeitung statt, d. h. es werden keine korrespondierenden Merkmale bei anderen Personen des Haushalts untersucht, insbesondere findet an dieser Stelle keine Generierung von Haushalten statt.

Anschrift und Status der vorherigen Wohnung in dieser oder einer anderen Gemeinde wird ungeprüft aus MR1 übernommen, da nicht veränderlich. Gleiches gilt für die Angabe des Zuzugs aus dem Ausland.

<b>Übersicht zur Prüfung der Stichtagsrelevanz der einzelnen Merkmale</b>	
<b>Erhebungsmerkmale</b>	
Geburtsdatum Geschlecht	ungeprüft aus MR2, da dies i. d. R. ein korrigierter bzw. ergänzter Stand
Staatsangehörigkeiten	ungeprüft aus MR1
Familienstand	Registertest: ungeprüft aus MR1 Verfahrenstest: Prüfung anhand Datum des Beginns bzw. der Beendigung der letzten Ehe
Wohnungsstatus	Prüfung anhand Datum des Statuswechsels
Wohnort	wird durch Auswahl vorgegeben, nicht zu prüfen
<b>Hilfsmerkmale</b>	
Ordnungsmerkmal der Meldebehörde	kann sich i. d. R. nicht ändern, nicht zu prüfen
Namen, Geburtsnamen, Vornamen	ungeprüft aus MR2
gegenwärtige Anschrift	wird durch Auswahl vorgegeben, nicht zu prüfen
Datum des Beziehens der Wohnung	ungeprüft aus MR1
Datum der Anmeldung bei der Meldebehörde	ungeprüft aus MR2
Datum des Auszugs aus der Wohnung	
Datum der Abmeldung bei der Meldebehörde	
Datum des Wohnungsstatuswechsels	
Datum der letzten Eheschließung	
Datum der Beendigung der letzten Ehe	
Verzeigerungen für Ehegatten, Kinder und gesetzliche Vertreter	<b>Prüfung</b> anhand des Familienstands, bzw. der Geburtsdaten der Kinder
Anschrift und Status der Wohnung in der Gemeinde, aus der der Einwohner zugezogen ist	ungeprüft aus MR1, da nicht veränderbar
Zuzug aus dem Ausland	
Anschrift der zuletzt bewohnten Wohnung in der Gemeinde	
Name des Wohnungsgebers	ungeprüft aus MR1

## 5 Abgleich mit der Haushaltebefragung

Nach der Zusammenführung der beiden Lieferungen aus dem Melderegister und der Feststellung der Stichtagsrelevanz erfolgt anschließend der Abgleich mit der Haushaltebefragung. Hierzu werden unter jeder ausgewählten Adresse die Namen der in der Haushaltebefragung erhobenen Personen im Bestand des Melderegisters gesucht. Für diesen Namensabgleich wurde ein Programm, das zu diesem Zweck im Bayerischen Landesamt für Statistik und Datenverarbeitung entwickelt wurde, eingesetzt.

Zuvor wird in allen Datenbeständen eine Vereinheitlichung der Namensschreibweisen durchgeführt. Alle Namen werden auf Großbuchstaben umgesetzt und alle Umlaute werden aufgespalten (Ä zu AE etc.). Ebenso werden „PH“ durch „F“ und „TH“ durch „T“ ersetzt. Sonstige Zeichen,

die in den Namensfeldern des Melderegisters vorkommen können (Bindestrich, Punkt, Plus), werden durch Leerzeichen ersetzt.

Die Suche der Personen erfolgt unter Berücksichtigung der Angaben zu Geburtsdaten und Geschlecht. Während für das Geschlecht immer eine exakte Übereinstimmung vorliegen muss, wird die Prüfung des Geburtsdatums schrittweise gelockert. Im ersten Durchlauf wird innerhalb der Adresse nach einer Person gleichen Geschlechts und exakt gleichen Geburtsdatums gesucht. Für diese wird geprüft, ob der Name ganz oder (bei Doppelnamen) teilweise übereinstimmt oder zumindest große Ähnlichkeit aufweist (sog. phonetischer Abgleich). Für alle derart identifizierten Personen wird die Personnummer aus dem Melderegisterbestand in die Haushaltebefragungsdaten übernommen. Im zweiten Durchlauf wird für die restlichen Personen geprüft, ob bei Gleichheit von Ge-

schlecht und Geburtsjahr eine Namensübereinstimmung gefunden werden kann. Schließlich wird im dritten Durchlauf auf eine Prüfung des Geburtsdatums gänzlich verzichtet. Dafür wird aber von Namensähnlichkeiten abgesehen und nur exakt übereinstimmende Namen noch als gleiche Person identifiziert.

Alle in einer derartigen maschinellen Kontrolle nicht gefundenen Personen werden nochmals manuell überprüft. So können auch noch Vertauschungen von Vor- und Nachnamen (ein häufiger Fehler bei fremdsprachigen Namen) oder Erfassungsfehler bei Geburtsdatum oder Geschlecht erkannt und die entsprechende Zusammenführung der Personen vorgenommen werden.

Für alle nicht direkt befragten Haushalte (Verweigerer, nicht erreichte Haushalte etc.) werden anschließend Ersatzvornahmen aus dem Melderegister durchgeführt<sup>10</sup>. Die verbleibenden unpaarigen Fälle in den Melderegisterdaten und in den Haushaltebefragungsdaten ergeben nach der Hochrechnung Aufschlüsse über die Qualität der Melderegister.

Im Rahmen des Verfahrenstests wird im Herbst 2002 die Zusammenführung mit der Gebäude- und Wohnungszählung mit der anschließenden Haushaltegenerierung zentral im Bayerischen Landesamt für Statistik und Datenverarbeitung durchgeführt.

## 6 Hoch- und Fehlerrechnung

Für den Registertest wurde eine zweistufige Stichprobe gezogen, bei der auf der ersten Stufe innerhalb der Bundesländer geschichtet nach Größenklassen einzelne Gemeinden ausgewählt wurden. Auf der zweiten Stufe wurden dann aus den laut Melderegister bewohnten Adressen eine bestimmte Anzahl – je nach Gemeindegrößenklasse 40, 80, 160 oder bis zu 320 – selektiert. Auf beiden

<sup>10</sup>) Vgl. Schäfer, J.: Die Rolle der Haushaltebefragung ...; a. a. O.

Stufen erfolgte die Auswahl größenproportional, d. h. die Auswahlwahrscheinlichkeit war jeweils proportional zur Zahl der Einwohner bzw. der gemeldeten Bewohner. Adressen, unter denen auch im Rahmen des Mikrozensus 2001/2002 eine Befragung stattfand, sowie Anstalten, Botschaften, Kasernen, fiktive Adressen etc. wurden vor der Stichprobenziehung gekennzeichnet und von der Auswahl ausgeschlossen.

Vor der Interpretation der Ergebnisse der Testerhebung ist eine Hochrechnung erforderlich, da Feststellungen bei verschiedenen Adressen wegen der unterschiedlichen Auswahlwahrscheinlichkeiten auch mit unterschiedlichem Gewicht in die Gesamtergebnisse einfließen müssen. Es wird zunächst eine Hochrechnung mit den Kehrwerten der Auswahlwahrscheinlichkeiten auf Gemeinde- und Adressebene vorgenommen. Anschließend erfolgt eine Anpassung für Antwortausfälle (Non-Response der Haushaltebefragung) und eine weitere Anpassung für fehlende Mikrozensusadressen und Anstaltsgebäude. Schließlich werden die Ergebnisse je Gemeinde-

größtenklasse an Rahmendaten aus den Melderegistern (Zahl der gemeldeten deutschen/nichtdeutschen Personen zum 5. 12. 2001) angepasst. Je nachdem, ob die Gesamtstichprobe für den Registertest oder nur die Unterstichprobe für den Verfahrenstest ausgewertet wird, und je nachdem, ob Personen, Haushalte oder Wohnungen bzw. Gebäude betrachtet werden, ergeben sich so insgesamt 12 verschiedene Hochrechnungsfaktoren.

## **7 Ausblick**

Erste Ergebnisse des Registertests werden voraussichtlich Anfang nächsten Jahres vorliegen. Der Bericht zu allen Teilen des Zensusstests soll bis Mitte 2003 fertig gestellt werden. Nach der Diskussion der darin enthaltenen Ergebnisse auch in Politik und Wissenschaft können die Entscheidungen über die Ausgestaltung und den Termin des nächsten Zensus in Deutschland getroffen werden. Ausschlaggebend hierbei wird nicht so sehr die Qualität der Melderegister an sich, sondern die Qualität der mit dem Modell eines

registergestützten Zensus erreichbaren Ergebnisse – also nach der statistischen Bereinigung von Über- und Untererfassungen – sein.

Das Zensusstestgesetz sieht für einen Teil der Hilfsmerkmale eine Aufbewahrungsfrist von fünf Jahren vor (danach sind alle Hilfsmerkmale sowohl auf den Erhebungsbogen als auch in den Dateien zu löschen), damit weitere Untersuchungen und die detaillierte Ausarbeitung des zukünftigen Zensusmodells über einen längeren Zeitraum möglich sind. Angesichts der Tatsache, dass die amtliche Statistik mit den Modellen für einen registergestützten Zensus methodisch vollkommenes Neuland beschreitet, ist es unumgänglich, die zur Verfügung stehende Zeit für weitere Untersuchungen zu nutzen. Da die wesentlichen Daten nach diesem Zeitraum nicht mehr zur Verfügung stehen werden, müssen die Arbeiten zur Beantwortung der noch offenen Fragen unmittelbar in Angriff genommen werden. Voraussichtlich werden diese Arbeiten übergangslos in die Vorbereitungen für den nächsten flächendeckenden Zensus in Deutschland münden.

# Postalische Gebäude- und Wohnungszählung im Zensusstest 2001 Erfahrungen in NRW

Dipl.-Sozialwirt Sebastian Scharnhorst  
Dr. Henrik Egbert

*Durch den registergestützten Zensus soll zukünftig auf die flächendeckende Befragung einzelner Haushalte verzichtet werden. Die postalische Gebäude- und Wohnungszählung (GWZ) soll dabei die notwendigen Daten zu Gebäuden und Wohnungen liefern. Da diese mangels geeigneter Quellen nicht aus Beständen der Verwaltungsregister entnommen werden können, werden sie im Modell des registergestützten Zensus auf postalischem Wege bei den Eigentümern bzw. Verwaltern der Gebäude erfragt. Im Idealfall könnte auf diese Weise die direkte Befragung einzelner Haushalte vermieden werden.*

*Die bisherigen Erkenntnisse aus der Erprobung der postalischen Gebäude- und Wohnungszählung im Zensusstest lassen jedoch Zweifel daran aufkommen, dass der Verzicht auf Haushaltebefragungen vollständig umgesetzt werden kann. So erscheint das Verfahren im Allgemeinen zwar durchaus geeignet, bei der Behandlung bestimmter Sonderfälle ergeben sich aber Problembereiche, welche durch eine rein postalische Befragung wahrscheinlich kaum gelöst werden können.*

*Diese problematischen Bereiche einer postalischen Gebäude- und Wohnungszählung werden im Folgenden anhand der Vorbereitung und Durchführung der Erhebung in Nordrhein-Westfalen aufgezeigt und mit Blick auf die Bedeutung der GWZ im Modell des registergestützten Zensus bewertet.*

## **Die Rolle der Gebäude- und Wohnungszählung im Modell eines registergestützten Zensus und im Zensusstest**

Um Volkszählungsergebnisse in herkömmlicher Qualität bereitstellen zu können, ist ein breites Datenspektrum notwendige Voraussetzung. Da einige dieser Daten bereits in Verwaltungsregistern erfasst sind, wird im registergestützten Zensusmodell auf diese Datenbestände zurückgegriffen. Allein mittels der Registerbestände kann der Datenbedarf jedoch nicht gedeckt werden. Zwar enthalten die Melderegister die benötigten personenbezogenen Daten, nicht aber Angaben zur Haushaltsstruktur. Diese sollen im Zensusmodell durch die Haushaltgenerierung bereitgestellt werden. Hierfür benötigt das Generierungsverfahren jedoch Informationen über die Wohnsituation dieser Personen.

Der postalischen Gebäude- und Wohnungszählung kommt deshalb die Aufgabe zu, gebäude- und wohnungstatistische Daten bereitzustellen, mittels derer in der anschließenden Haushaltgenerierung auf Haushaltszusammenhänge zwischen den gemeldeten Personen geschlossen werden kann. Erst mit den durch die GWZ erhobenen Angaben zum Namen der Wohnungsinhaber, zur Anzahl der in der Wohnung lebenden Personen und zu deren Einzugsdatum ist es möglich, die laut Melderegister unter einer Adresse wohnhaften Personen den einzelnen Wohnungen entsprechend der Wohnungsgröße und -belegung zuzuordnen und sie dadurch zu Haushaltseinheiten zusammenzufassen. Weiterhin lässt sich erst auf Basis der GWZ-Daten im Zuge des Generierungsverfahrens auch eine Beurteilung der Melderegister vornehmen, etwa hinsichtlich bestehender Über- oder Untererfassungen.

Die Erfahrungen und Ergebnisse aus dem Zensusstest werden ganz wesentlich zur Beurteilung und Verbesserung des Modells beitragen. Da eine postalische GWZ in dieser Form bisher nicht durchgeführt wurde, bildet ihre Erprobung im Rahmen des Zensusstests eine wichtige Entscheidungsgrundlage im Hinblick auf das methodische Vorgehen. So muss geklärt werden, ob die Eigentümer und Verwalter der Gebäude tatsächlich in der Lage sind, die Angaben zu Wohnungen und Bewohnern zu machen und ob dieses Vorgehen im Ergebnis mit einer direkten Befragung der Bewohner (also der Haushalte) vergleichbar ist. Die Bereitstellung einer solchen Vergleichsmöglichkeit ist daher einer der Gründe für die im Zensusstest parallel zur GWZ stattfindende Interviewerbefragung bei den Haushalten (Haushaltebefragung). Da die mittels der GWZ erhobenen Daten eine wichtige Grundlage der weiteren Verfahrensteile bilden, sollen zudem verschiedene methodische Ansätze zur Vervollständigung der aufgrund von Befragungsausfällen entstehenden Erhebungslücken erprobt werden.

## **Die GWZ in NRW**

Die postalische Gebäude- und Wohnungszählung im Rahmen des Zensusstests umfasst in Nordrhein-Westfalen 1 650 Adressen in 18 Gemeinden, bei denen zum Stichtag 5. Dezember 2001 die Daten zu Gebäuden und Wohnungen sowie zu Wohnungsinhabern und zur Wohnungsbelegung erhoben wurden. Abhängig von der Gemeindegrößenklasse entfallen dabei auf jede Gemeinde 40, 60, 110 oder 150 Adressen (siehe Grafik S. 26). Diese „Stichprobe Verfahrenstest“, mit der vorwiegend die neue Methodik des Zensusmodells



geprüft werden soll, ist Teil einer 36 Gemeinden umfassenden landesweiten Stichprobe zum Zensustest in NRW.

Auswahlgrundlage der Stichprobe bildeten dabei nicht die Gebäude selbst, sondern einzelne Adressen in Form der Anschrifteneinträge in den Melderegistern der Städte und Gemeinden. Die Auswahlwahrscheinlichkeit einer Adresse lag daher umso höher, je mehr Personen unter ihr gemeldet waren. Adressen mit einer hohen Anzahl gemeldeter Personen sind somit absichtlich in der Stichprobe überrepräsentiert, weil diese als potentiell problematisch für das neue Erhebungsverfahren gelten und im Zensustest verstärkt untersucht werden sollen.

### Arbeitsschritte der postalischen Gebäude- und Wohnungszählung

Grundsätzlich bestünde der Ablauf einer postalischen Gebäude- und Wohnungszählung in einem registergestützten Zensus aus drei Phasen: der Vorbereitung, in der die Eigentümeranschriften ermittelt und ggf. vervollständigt werden; der Durchführungsphase, welche die postalische Erhebung umfasst, und schließlich der Verarbeitung der gewonnenen Daten einschließlich Erfassung, Plausibilitätsprüfung und einer statistischen Korrektur von Befragungsausfällen.

Auch im Zensustest entsprach die Vorgehensweise grundsätzlich diesem Schema. Darüber hinaus wurden hier jedoch noch weitere Schritte durchgeführt (siehe Ablaufplan). So wurden die nach dem Zensustestgesetz<sup>1)</sup> derzeit in Frage kommenden Quellen für Eigentümer- und Verwalteranschriften im Zuge der Vorbereitung zunächst einmal auf ihre Eignung zur Lieferung der benötigten Daten hin überprüft und beurteilt. Während der Durchführung der postalischen GWZ ergaben sich zudem weitere Ergänzungen im Ab-

1) „Gesetz zur Vorbereitung eines registergestützten Zensus (Zensusvorbereitungsgesetz)“ vom 27. 7. 2001

Ablaufphasen und Arbeitsschritte der postalischen GWZ im Zensustest	
Vorbereitung	Befragung der potentiellen Quellen zu vorhandenen Anschriften von Gebäudeeigentümern Lieferung der Anschriften durch die Quelle(n) Information und Vorbefragung der Auskunftspflichtigen auf Basis der gelieferten Anschriften <sup>1)</sup> Aktualisierung und Abgleich der übermittelten Anschriften durch den Rücklauf der Vorbefragung
Erhebung	Hauptbefragung bei Eigentümern bzw. Verwaltern Zusatzerhebung bei Wohnungseigentümern <sup>1)</sup>
Verarbeitung	Erfassung Plausibilisierung Ersatzvornahmen aus Stichprobe und <i>Haushalbefragung</i> . <sup>2)</sup>
Anschließend	Haushaltgenerierung aus den Daten der GWZ und der Melderegister

*Kursiv* gesetzte Schritte finden zunächst ausschließlich im Zensustest statt.  
1) Diese Arbeitsschritte waren ursprünglich nicht Bestandteil der Ablaufplanung, könnten aber auch im registergestützten Zensus Anwendung finden. – 2) An die Stelle der Daten aus der Haushalbefragung des Zensustests würde in einem späteren Zensus die Begehung der entsprechenden Adressen oder Gebäude treten.

laufschemata, welche auch bei einer Durchführung im registergestützten Zensus Anwendung finden könnten. So wurden die Eigentümer und Verwalter bereits im Vorfeld der eigentlichen Erhebung angeschrieben und über die bevorstehende Befragung informiert. Mit diesem Schritt sollte vor allem die Qualität der gelieferten Anschriftendaten festgestellt werden, um ggf. im Vorfeld der Hauptbefragung noch Korrekturen oder Aktualisierungen vornehmen zu können. Ein weiterer Zusatzschritt wurde nötig, weil in manchen Gebäuden mit Eigentumswohnungen die Angaben zu Wohnungen und Wohnungsinhabern nicht durch die zuständigen Verwalter gemacht werden konnten. In diesen Fällen wurden die einzelnen Wohnungseigentümer gesondert befragt, was bei der Ablaufplanung zunächst nicht vorgesehen war.

### Registerdaten als Anschriftenquellen

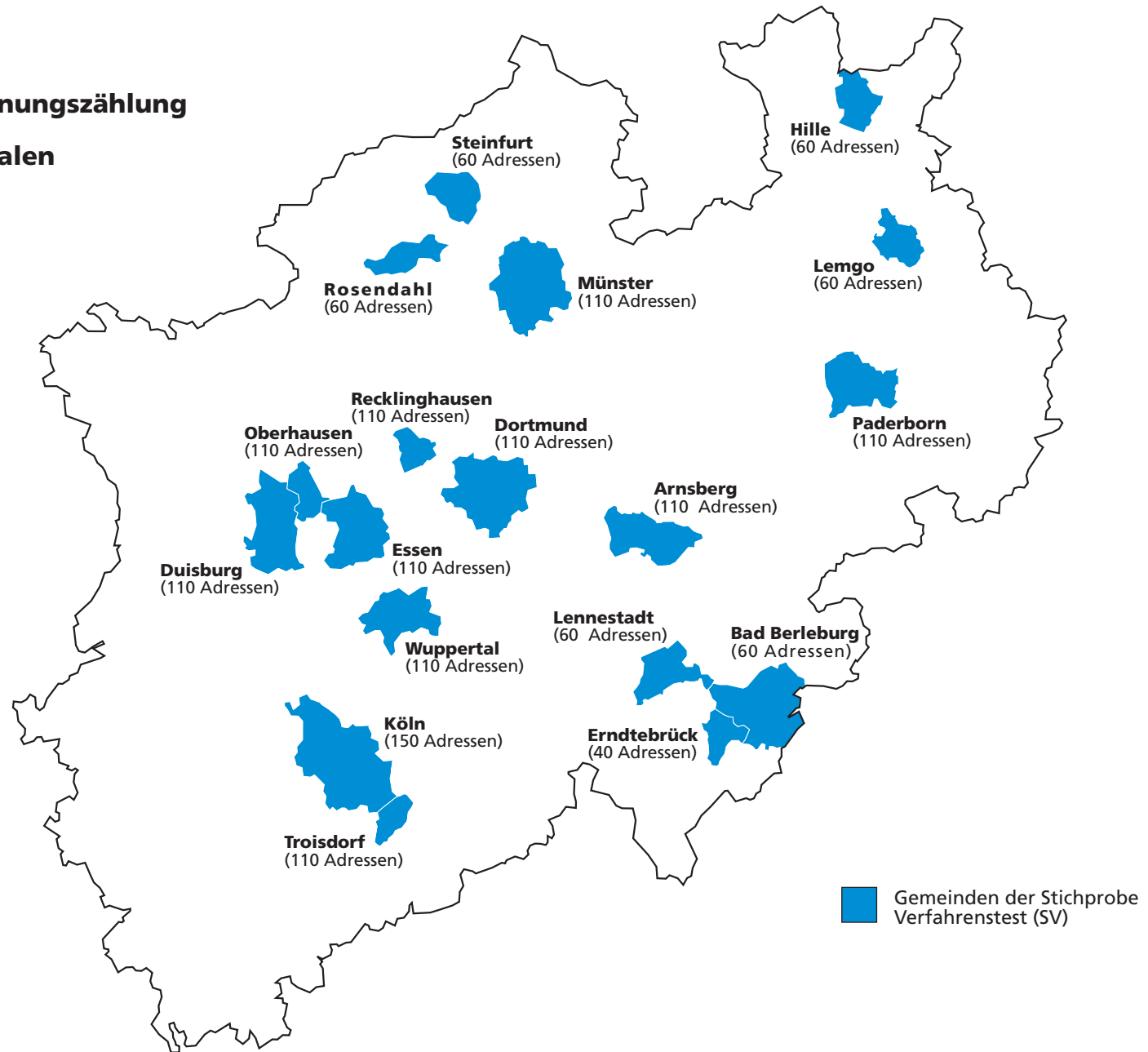
Voraussetzung der postalischen Befragung ist die Bereitstellung von Anschriften der Gebäudeeigentümer oder -verwalter. In der Bundesrepublik gibt es jedoch keine ausgewiesenen Register, in denen derartige Daten vollständig und aktualisiert geführt werden. Das Zensustestgesetz sieht für deren Ermittlung deshalb

mehrere potentielle Quellen vor, bei denen entsprechende Datenbestände in anderen Zusammenhängen vermutet wurden. Dies sind die Gemeinden, die für die Führung der Grundbücher und der Liegenschaftskataster zuständigen Stellen, die Finanzbehörden, die für die Gebäudebrandversicherung zuständigen juristischen Personen des öffentlichen Rechts sowie die Ver- und Entsorgungsbetriebe.

Es erschien bereits im Vorfeld der GWZ als relativ gesichert, dass mit den Grundsteuerstellen der Gemeinden bereits eine Quelle zur Verfügung steht, welche über hinreichend vollständige Adressdaten von Gebäudeeigentümern verfügt. Um zu klären, ob neben diesen Grundsteuerstellen auch andere der genannten Quellen für eine postalische Befragung der Gebäudeeigentümer herangezogen werden könnten, wurden von den im Zensustestgesetz genannten Quellen die Amtsgerichte für die Grundbücher, die Kreise und das Landesvermessungsamt für die Liegenschaftskataster und die in den Gemeinden tätigen Ver- und Entsorgungsbetriebe in Form einer schriftlichen Voranfrage auf ihre Eignung zur Lieferung dieser Anschriften überprüft. Dies diente gleichzeitig auch zur Information über den bevorstehenden Zensustest und zur Feststellung

## Zensusstest 2001

### Gebäude- und Wohnungszählung im Verfahrenstest in Nordrhein-Westfalen



möglicher Schwierigkeiten im Zusammenhang mit der Datenübermittlung.

Nach der Auswertung dieser Voranfrage lässt sich die Eignung der genannten Quellen wie folgt resümieren:

- Bei den Amtsgerichten liegen die Anschriften der Eigentümer, soweit sie bei Erwerb eines Grundstücks registriert wurden, zwar vor; sie werden danach aber i. d. R. nicht mehr aktualisiert, da die Eigentümer auch nicht verpflichtet sind, den Grundbuchstellen eine Änderung ihrer Wohnadresse zu melden.
- Ver- und Entsorgungsbetriebe verfügen aufgrund ihrer Rechnungslegung über hinreichend aktuelle Anschriftendateien, diese Kundenadressen sind jedoch nicht immer nur die Adressen der Gebäudeeigentümer oder -verwalter. Zu den registrierten Kunden gehören ebenso und teilweise stattdessen auch einzelne Mieter. Eine Differenzierung der Anschriften nach Eigentümern/Verwaltern und Mietern ist aufgrund der Datenlage nicht möglich.
- Zum Zeitpunkt der Voranfrage waren auch die vorliegenden Daten der Liegenschaftskataster für die Gebäude- und Wohnungszählung nicht geeignet. Ein Vorteil besteht hier gegenüber anderen Quellen jedoch darin, dass die Daten von allen Katasterämtern einheitlich, nämlich als Auszug aus dem automatisierten Liegenschaftsbuch (ALB), geliefert werden könnten. Dies würde ihre weitere Verarbeitung wesentlich erleichtern. Gegenwärtig werden im ALB zwar noch keine Anschriften der Gebäudeeigentümer geführt. Da es aber in Zukunft um diese Daten erweitert werden soll, stünde zu einem späteren Zeitpunkt möglicherweise auch das ALB als geeignete Anschriftenquelle zur Verfügung.

Demzufolge sind neben den Grundsteuerstellen derzeit keine weiteren Quellen zur vollständigen Lieferung der Eigentümeranschriften in der Lage. Allenfalls kämen manche von

ihnen in Einzelfällen zur Nachrecherche oder Ergänzung der Anschriften in Betracht.

Die Grundsteuerstellen, die im Zuge der Voranfrage ebenfalls angeschrieben wurden, verfügen dagegen über relativ vollständige Adressdateien. Allerdings wurden bei der Auswertung der im September und Oktober 2001 gelieferten Daten auch hier einige Schwierigkeiten offenbar. So erhalten die Grundsteuerstellen nicht automatisch beim Umzug von Gebäudeeigentümern eine Änderungsmitteilung. Weiterhin sind den Grundsteuerstellen bei größeren Gebäuden teilweise nur die Eigentümer von Gebäudeteilen oder Einzelwohnungen bekannt. Da sich die GWZ in letzteren Fällen zur Minimierung des Befragungsumfangs zunächst an die zuständigen Verwalter wenden sollte (möglichst nur ein Ansprechpartner pro Gebäude), mussten die entsprechenden Anschriften nachträglich ermittelt werden, etwa durch telefonische Befragung der Wohnungseigentümer. In anderen Fällen waren den Grundsteuerstellen zu den ausgewählten Adressen auch gar keine Ansprechpartner bekannt, so dass diese anderweitig ermittelt werden mussten.

Insgesamt zeigt sich, dass dieser Weg der Anschriftenermittlung zwar in weiten Teilen zufriedenstellende Resultate liefert, jedoch nicht ohne selektive Einzelfallrecherche zu beschreiten ist. Da zum gegenwärtigen Zeitpunkt zudem nur eine einzige hinreichend aktuelle Anschriftenquelle verfügbar ist, könnten bestehende Lücken nur bedingt durch Abgleich oder Zusammenführung der verschiedenen Quellen geschlossen werden.

### **Simulation der Begehung von Problemgebäuden**

Bei einigen der Adressen, für die von den Grundsteuerstellen keine Eigentümer- oder Verwalteranschriften übermittelt wurden, konnte auch durch die Heranziehung anderer Quellen kein Eigentümer oder zuständiger Verwalter ermittelt werden. In diesen Fällen wurde versucht,

die jeweiligen Ansprechpartner durch die Interviewer der Haushaltebefragung feststellen zu lassen. Damit wurde quasi eine Begehung dieser Problemfälle simuliert. Mit dieser Befragung der (Gebäude-)Bewohner konnte schließlich für alle bis dahin noch offenen Adressen, rund 1% der Stichprobe, tatsächlich eine entsprechende Auskunftsperson gefunden werden.

Hierbei ist aber zu beachten, dass die eingesetzten Interviewer nur im Rahmen des Zensusstests durch die parallel stattfindende Haushaltebefragung zur Verfügung standen. Die Inanspruchnahme eigens hierfür einzusetzender Begehungsbeauftragter würde in einem registergestützten Zensus einen wesentlichen Zusatzaufwand bedeuten. Angesichts der positiven Erfahrung aus der Simulation wäre jedoch zu überlegen, ob durch solche Begehungen auch Teile der teilweise extrem aufwändigen Einzelfallrecherche vermieden werden könnten.

### **Information und Vorbefragung der Auskunftspflichtigen**

Um die Richtigkeit der von den Grundsteuerstellen übermittelten und teilweise selbst recherchierten Anschriften zu prüfen und diese ggf. korrigieren zu können, wurden im Rahmen einer gesonderten Vorbefragung im Oktober und November 2001 alle Auskunftspflichtigen schriftlich über die bevorstehende Befragung informiert. Diese sollten dazu mitteilen, ob sie im Sinne des Zensusstestgesetzes zu den Angaben über das jeweilige Gebäude, den darin enthaltenen Wohnungen und deren Inhabern auskunftsfähig sind oder andernfalls den richtigen Ansprechpartner angeben, soweit ihnen dieser bekannt war.

Etwa zwei Drittel der Schreiben wurden ausgefüllt zurückgesandt. Durch den Rücklauf unzustellbarer Sendungen konnte zudem für die Gesamtheit der vorliegenden Daten weitgehend festgestellt werden, welche Eigentümernamen und Anschriften

nicht korrekt waren und in welchem Umfang somit weiterer Recherchebedarf bestand.

Im Zuge dessen wurde deutlich, dass etwa 10% der Daten der Grunderstellung korrigiert werden mussten. Im Wesentlichen handelte es sich jedoch um kleinere Korrekturen oder Ergänzungen der Namen oder Anschriften – nur rund 1% der angegebenen Gebäudeeigentümer oder -verwalter waren nicht bzw. nicht unter dieser Anschrift bekannt. In den letztgenannten Fällen wurden wiederum Bewohner der entsprechenden Gebäude, welche unter dieser Anschrift in elektronischen Telefonverzeichnissen registriert waren, telefonisch um eine Auskunft zum entsprechenden Eigentümer oder Verwalter gebeten. Letztlich verblieb so nur eine geringe Anzahl an Gebäuden, die wegen fehlender Ansprechpartner ausfielen.

Generell ist davon auszugehen, dass jede im Vorfeld der postalischen Erhebung korrigierte Anschrift die Kosten und vor allem auch den Arbeits- und Zeitaufwand in späteren Erhebungsphasen erheblich reduziert. Insofern sind der Einsatz und das Ergebnis der Vorbefragung durchweg positiv zu bewerten. Ein weiterer Vorteil dieses Vorgehens besteht darin, dass damit in weiten Teilen bereits vorzeitig festgestellt werden kann, in welchen Fällen die angeschriebenen Gebäudeverwalter oder Eigentümer selbst keine entsprechenden Angaben machen können.

### **Auskunfts-fähigkeit und Auskunfts-bereitschaft der Eigentümer und Verwalter**

Ein konzeptioneller Vorteil der postalischen Gebäude- und Wohnungszählung im Modell des registergestützten Zensus ist es, die erforderlichen Daten für jedes Gebäude zentral bei den jeweiligen Eigentümern und Verwaltern erheben zu können. Dadurch lassen sich Umfang und Aufwand der Befragung erheblich reduzieren.

Grundsätzlich scheint dieses Konzept aufzugehen. Bis auf einige noch zu

schildernde Sonderfälle verfügen die Gebäudeverwalter und/oder -eigentümer über ausreichende Informationen, um die erforderlichen Angaben zu Wohnungen und Wohnungsinhabern zu machen. Einschränkungen bestehen hier fast ausschließlich aufgrund von Alter oder Krankheit der befragten Personen.

Im Hinblick auf die Auskunftsbereitschaft bietet sich dagegen ein differenzierteres Bild. Seitens der Auskunftspflichtigen wurden verschiedene Argumente gegen diese Form der Erhebung angeführt, vor allem hinsichtlich der Pflicht zur Erteilung von Auskünften über Dritte. Sowohl private Eigentümer wie auch Verwaltungsfirmen äußerten hier Bedenken hinsichtlich des Datenschutzes. Zuständigkeitsbezogene Schwierigkeiten traten vorwiegend bei einem Wechsel der Eigentümer oder Verwalter zwischen Stichtag bzw. Erhebung und anschließenden Mahnschritten auf. Dabei stellte sich oftmals ein logistisches Problem, da ehemalige Verwalter oder Eigentümerangaben, nicht mehr über die notwendigen Gebäudeunterlagen zu verfügen. Das Zensusgesetz gibt für derartige Fälle im Wortlaut keine eindeutige Regelung zur Auskunftspflicht vor. So ist zumindest fraglich, ob auch diejenigen Verfügungsberechtigten auskunftspflichtig sind, welche diese Aufgabe erst nach dem Erhebungstichtag übernommen haben. Ein Großteil dieser Fälle konnte schließlich jedoch im Dialog mit den Beteiligten zufriedenstellend gelöst werden.

### **Grenzbereiche: Zusatzerhebung bei Wohnungseigentümern**

Ein Sonderproblem stellen Gebäude mit Eigentumswohnungen dar. Hier waren die zuständigen Verwalter oftmals nicht in der Lage, die Angaben zu Wohnungen und Wohnungsinhabern zu machen. Dieses Problem betraf vermietete Eigentumswohnungen und von den Eigentümern selbst bewohnte Wohnungen gleichermaßen – auch bei letzteren richtet sich die postalische Gebäude- und Wohnungs-

zählung zunächst an die zuständigen Verwalter. Zur Lösung dieses Problems wurden die einzelnen Wohnungseigentümer gesondert angeschrieben, um die Angaben zu Wohnungen und Wohnungsinhabern zu erhalten. Die hierzu erforderlichen Anschriftendaten konnten von den Gebäudeverwaltern mitgeteilt werden.

In NRW betraf dies gut 2 % der Stichprobengebäude mit insgesamt rund 500 Wohnungen. Durch die Befragung der einzelnen Wohnungseigentümer konnten die Daten für einen überwiegenden Teil (94 %) dieser Gebäude vollständig erhoben werden. Dennoch führte die Zusatzbefragung neben dem erheblichen Mehraufwand für Recherche und Befragung, welcher einer parallelen Einzelerhebung gleichkommt, auch zu einem Zeitverzug im Verfahrensablauf, da der Versand der Erhebungsunterlagen hier erst wesentlich später als bei der Hauptbefragung erfolgen konnte.

Diese Sonderfälle waren bei der Planung der postalischen Gebäude- und Wohnungszählung ursprünglich nicht berücksichtigt worden. Trotz der in NRW vergleichsweise geringen Dimension dieses Problems im Zensus muss im Hinblick auf einen registergestützten Zensus davon ausgegangen werden, dass bei einer nicht unerheblichen Anzahl von Adressen mit einem zusätzlichen Aufwand für die Ermittlung und Befragung einzelner Wohnungseigentümer zu rechnen ist.

### **Ursachen und Behandlung von Gebäudeausfällen**

Für rund 5 % der 1 650 Stichprobenadressen im Verfahrenstest lagen nach Abschluss der Befragung keine Daten zu Wohnungen und Wohnungsinhabern vor. Diese teilen sich entsprechend ihrer Ausfallursache in zwei Gruppen, die im weiteren Verfahrensablauf der GWZ jeweils unterschiedlich behandelt werden müssen. Die erste Gruppe umfasst die so genannten stichprobenneutralen Ausfälle. Hierzu zählen all jene Adressen, die zwar in der Stichprobe enthalten sind, jedoch für die postalische Ge-

bäude- und Wohnungszählung nicht herangezogen werden können und sollen. Neben vollständig gewerblich genutzten sowie leer stehenden Gebäuden sind dies insbesondere auch so genannte Anstalten. Der Begriff „Anstalt“ umfasst in diesem Zusammenhang alle Fälle, bei denen es sich nicht um Haushalte im Sinne einer eigenständigen Wohn- und Wirtschaftsgemeinschaft handelt. Dazu gehören etwa Krankenhäuser, Wohn- und Betreuungsheime oder militärische Einrichtungen. Zur Gruppe der stichprobenneutralen Ausfälle zählen weiterhin auch fiktive Adressen, die nur ersatzweise im Melderegister geführt werden. Um die Anstalten von der Stichprobenziehung ausschließen zu können, wurden die dem statistischen Landesamt und den Gemeinden bekannten Anstaltsadressen bereits im Vorfeld entsprechend berücksichtigt. Dennoch verblieben einige Anstalten unerkannt in der Stichprobe, so dass diese erst im Verlauf der Erhebung auffielen. Das verspätete Auftreten stichprobenneutraler Ausfälle (hier besonders solche aufgrund gewerblicher Nutzung und Leerstand von Gebäuden) könnte aber reduziert, zumindest jedoch früher festgestellt werden, wenn im Zuge der genann-

ten Vorbefragung unter Hinweis auf die Auskunftspflicht ausdrücklich um entsprechende Rückmeldung gebeten würde. Es sollte auch im Interesse der Auskunftspflichtigen liegen, eine nochmalige Inanspruchnahme ihrer Person in späteren Erhebungsphasen zu vermeiden, wenn die entsprechende Befragungsgrundlage nicht vorliegt.

In die zweite Gruppe der Befragungsausfälle sind dagegen alle übrigen Ausfälle einzuordnen. Diese Gruppe ist damit wesentlich unbestimmter. Neben den eigentlichen Befragungsausfällen – d. h. es wurde ein entsprechender Ansprechpartner für ein erhebungsrelevantes (!) Gebäude festgestellt, der jedoch die Auskunft nicht erteilt oder explizit verweigert hat – schließt diese Gruppe möglicherweise auch Adressen unentdeckter stichprobenneutraler Ausfälle ein, da diese letztlich nur aufgrund einer entsprechenden Benachrichtigung durch den Postzusteller oder die angeschriebene Person festgestellt werden können. Erfolgt dagegen eine solche Rückmeldung nicht, so sind diese Ausfälle nicht von Auskunftsverweigerungen zu unterscheiden. Deswegen ist nicht auszuschlie-

Ben, dass sich in der Gruppe der Befragungsausfälle noch Adressen stichprobenneutraler Ausfälle befinden. Dieser Umstand ist jedoch eine potenzielle Fehlerquelle für den nachfolgenden Arbeitsschritt, in dem die GWZ-Daten für die anschließende Haushaltegenerierung aufbereitet werden.

Da das Generierungsverfahren für Adressen ohne Wohnungsangaben nicht oder nur unzureichend durchgeführt werden kann, wird im Rahmen der GWZ versucht, solche Erhebungslücken durch so genannte Ersatzvornahmen aufzufüllen. Für diese kommen im Zensus zwei unterschiedliche Vorgehensweisen zum Einsatz, die im Ergebnis miteinander verglichen werden sollen. Zum einen ist dies die Einsetzung künstlicher Ersatzwohnungen und -gebäude. Dabei werden nach der Zahl der Wohnungen ähnliche Gebäude aus der Stichprobe samt der dazugehörigen Angaben gedoppelt. Eine zweite Möglichkeit besteht in der Übernahme der Befragungsergebnisse aus der Haushaltebefragung. Im Hinblick auf den registergestützten Zensus würde letztere Variante eine Begehung dieser Gebäude durch Interviewer be-

<b>Gebäudeausfälle der GWZ in NRW (1 650 Gebäude in der Stichprobe)</b>				
Nach erfolgtem Arbeitsschritt	Verbleibende offene Fälle		Davon Ausfälle	
	Anzahl	davon	Grund	Anzahl
Adresslieferung der Grundsteuerstellen	36	36 (keine Angabe vorhanden)		
		(24) Telefonrecherche im LDS NRW (12) Vor-Ort-Recherche durch Interviewer		
Vorbefragung der Gebäudeeigentümer	182	(12) Anstalt/gewerbliche Nutzung (23) nicht zustellbar (147) Empfänger nicht Eigentümer	<b>Anstalten</b> <b>Gewerbliche Nutzung</b>	<b>5</b> <b>7</b>
		Telefonrecherche im LDS NRW		
Erstversand Erhebungsbogen	111	(59) unbekannt/nicht zustellbar (16) Anstalt/gewerbliche Nutzung (36) Wohnungseigentümerbefragung (WE)	<b>Anstalten</b> <b>Gewerbliche Nutzung</b>	<b>8</b> <b>8</b>
		Telefonrecherche im LDS NRW		
Mahnverfahren Gebäude (3 Stufen)	55	(29) keine Rückmeldung/Verweigerung (26) unbekannt/unbekannt verzogen (02) leer stehend	<b>Keine Rückmeldung/Verweigerung</b> <b>Unbekannt/unbekannt verzogen</b> <b>Leer stehend</b>	<b>29</b> <b>26</b> <b>2</b>
			<b>Summe der Gebäudeausfälle</b>	<b>85</b>
			davon	
			Stichprobenneutrale	30
			Befragungsausfälle	55

deuten, da die Haushaltebefragung nur im Rahmen des Zensusstests stattfindet und nicht Bestandteil des registergestützten Verfahrens ist.

Eine Ersatzvornahme darf jedoch nur für die Adressen aus der Gruppe der Befragungsausfälle durchgeführt werden, also nur für tatsächlich bewohnte Gebäude und Wohnungen. Da diese Gruppe, wie oben gezeigt, jedoch durchaus noch unentdeckte stichprobenneutrale Ausfälle enthalten kann, führt eine Ersatzvornahme in diesen Fällen dazu, dass in diese Gebäude unberechtigterweise Haushalte generiert werden, sofern entsprechende Einträge im Melderegister vorliegen. Erfahrungsgemäß kann letzteres durchaus auch bei Anstalten, Leerständen oder einer gewerblichen Nutzung der Gebäude der Fall sein.

Das Problem berührt hierbei jedoch nur die Ersatzvornahmen der ersten Variante, da hierbei die betreffenden Gebäude nicht in Augenschein genommen werden. Dagegen würde beim zweiten Verfahren durch die Interviewer sehr wahrscheinlich festgestellt werden, dass es sich bei der Adresse um einen stichprobenneutralen Ausfall handelt, so dass diese Adresse entsprechend umgruppiert werden könnte. Im Zensusstest finden Ersatzvornahmen daher nur dann statt, wenn in der Haushaltebefragung tatsächlich ein bewohntes Gebäude festgestellt wurde. Diese generelle Heranziehung der Begehungsergebnisse entspricht zwar nicht den authentischen Bedingungen eines registergestützten Zensus, ist für einen methodischen 1:1-Vergleich der beiden Verfahrensweisen jedoch sinnvoll.

Sollte das Dopplungsverfahren im registergestützten Zensus jedoch ohne den Abgleich mit Begehungsergebnissen eingesetzt werden – was im Grunde genommen seine einzig zweckmäßige Anwendung wäre, da bei einer ohnehin erfolgenden Begehung ja auch die Daten erhoben werden könnten – verursacht eine fälschliche Zuordnung zu den Befragungsausfällen unter Umständen Verzerrungen

im weiteren Verfahren. Da durch die Haushaltegenerierung auch Fehlbestände und Karteileichen in den Melderegistern nachgewiesen werden sollen, führen unberechtigt vorgenommene Ersatzvornahmen möglicherweise zu fehlerhaften Schätzungen der Melderegisterfehler.

### **Perspektiven einer postalischen Gebäude- und Wohnungszählung im registergestützten Zensus**

Die bisherigen Erfahrungen mit der postalischen Gebäude- und Wohnungszählung im Zensusstest haben gezeigt, dass Daten zu Gebäuden und insbesondere zu Wohnungen und Wohnungsinhabern größtenteils durchaus bei den Eigentümern und Verwaltern der einzelnen Gebäude erhoben werden können. Ebenso wurde an verschiedenen Stellen dieses Beitrags aber auch deutlich, dass diese Form der Befragung allein nicht ausreichend sein wird, um möglichst vollständige Ergebnisse zu erzielen.

Es zeichnet sich vielmehr ab, dass bei manchen Gebäuden oder Adressen eine Begehung auch in einem registergestützten Zensus notwendig werden dürfte. So hätten etwa in NRW ohne das Mittel der Vor-Ort-Befragung allein für rund 1% der Stichprobengebäude erst gar keine Ansprechpartner ermittelt werden können. Da von Problemen mit der rein postalischen Befragung besonders Adressen mit einer höheren Anzahl an Wohnungen und hoher Fluktuation der Bewohner betroffen waren, dürfte die Zahl der auf die Wohnbevölkerung bezogenen Ausfälle hier vermutlich noch deutlich höher liegen. Verschärfend kommt hinzu, dass sich solche Fälle auch im Zuge der Haushaltegenerierung kaum klären lassen werden.

Angesichts der Höhe zu erwartender Befragungsausfälle und des unverhältnismäßig hohen Zusatzaufwandes durch Sonderfälle im Rahmen einer rein postalischen Erhebung

scheint die Begehung in bestimmten Fällen kaum verzichtbar zu sein, auch wenn diese durch das neue Zensusmodell eigentlich möglichst vermieden werden soll.

Zum einen kommt hierfür eine nachträgliche Begehung, etwa bei Befragungsausfällen oder zur Klärung verbleibender Unplausibilitäten im Zuge des Generierungsverfahrens in Betracht. Diese ist jedoch mit dem Problem behaftet, dass sich der zeitliche Abstand zwischen dem Erhebungstichtag und der Begehung möglicherweise negativ auf die Befragungserfolge auswirkt. Zusätzlich wäre deshalb zu überlegen, die als besonders problematisch geltenden Gebäudetypen bereits im Vorfeld der Erhebung zu begehen. Auch in denjenigen Fällen, in denen die postalische Befragung erfahrungsgemäß nur umständlich und mit großem zeitlichen Aufwand durchgeführt werden kann, könnte so eine stichtagsnahe Erhebung sichergestellt werden. Voraussetzung wäre hierfür allerdings, dass sich diese Adressen bereits im Vorfeld anhand bestimmter Kriterien zuverlässig identifizieren lassen.

So bleibt abzuwägen, in welchen Fällen der Kosten- und Zeitaufwand anderer Recherchewege sowie die hinzunehmenden Befragungsausfälle noch im Verhältnis zu den aufgezeigten Vorteilen und Sicherheiten einer lokalen Begehung bestimmter Adressen stehen. Ausdrücklich ist dabei aber zu betonen, dass eine solche Begehung nur als Ergänzung der postalischen Befragung herangezogen werden soll, welche ansonsten nach bisherigem Stand äußerst zufriedenstellende Resultate erbringt. Endgültige Ergebnisse der postalischen Gebäude- und Wohnungszählung im Zensusstest werden voraussichtlich ab Anfang 2003 vorliegen. Dann wird auch abzusehen sein, anhand welcher Kriterien, in welchem Umfang und in welcher Form eine solche ersatzweise Begehung in einem registergestützten Zensus erfolgen könnte.

# Die Haushaltegenerierung im Zensus 2001

## Eine neue Methode in der amtlichen Statistik

Dr. Marco Scharmer

*Im Zuge der Entscheidung der Bundesregierung gegen eine herkömmlich Volkszählung, wurden die statistischen Ämter des Bundes und der Länder beauftragt, ein alternatives Zensusverfahren zu entwickeln, mit dem ohne eine direkte Befragung aller Bürger dennoch fachlich und räumlich tief gegliederte Ergebnisse zur Zahl, Struktur und Wohnsituation von privaten Haushalten geliefert werden kann. Der Test des neuen Modells für einen registergestützten Zensus, der so genannte Zensus 2001, wird mit Stichtag 5. Dezember 2001 durchgeführt<sup>1)</sup>.*

*Im Mittelpunkt des registergestützten Zensus mit seinen einzelnen Erhebungs- und Verfahrensteilen steht die Generierung von Haushalten. Als Haushaltegenerierung werden Verfahren bezeichnet, mit denen die Personendatensätze der Meldebehörden systematisch ausgewertet und gegebenenfalls zusätzlich gebäude- und wohnungsstatistische Informationen einbezogen werden, um Haushaltszahlen und -strukturen zu ermitteln. Der vorliegende Beitrag zeigt zunächst in kurzer Form verschiedene Ansätze von Haushaltegenerierungsverfahren auf, die in der Bundesrepublik Deutschland entstanden sind. Im Mittelpunkt der Betrachtung steht die Darstellung des für den Zensus 2001 entwickelten und getesteten Haushaltegenerierungsverfahrens.*

1) Vgl. Gesetz zur Vorbereitung eines registergestützten Zensus (Zensusvorbereitungsgesetz) vom 27. 7. 2001, in: Bundesgesetzblatt, Teil 1, Nr. 40, S. 1882 ff.

### 1 Die Haushaltegenerierung als statistische Methode

Sowohl für öffentliche wie auch für privatwirtschaftliche Institutionen sind die sozialen und ökonomischen Verhältnisse sowie die Wohnsituation der Bevölkerung von hohem Interesse. Regional gegliederte Haushaltsdaten waren somit immer ein zentraler Bestandteil der amtlichen Statistik im Rahmen der Volkszählungen und des jährlich durchgeführten Mikrozensus. Aktuelle Volkszählungsergebnisse standen jedoch nur in langen Zeitabständen zur Verfügung. Beim Mikrozensus handelt es sich um eine 1 %-Stichprobe, dessen Ergebnisse regionalisierte Aussagen nur eingeschränkt zulassen<sup>2)</sup>. Da jedoch insbesondere im kommunalen Bereich und in der Marktforschung

eine große Nachfrage nach aktuellen und kleinräumig gegliederten Haushaltsdaten besteht, wurde bereits Ende der 70er-, Anfang der 80er-Jahre die Idee geboren, mit Hilfe von Informationen aus den Melderegistern haushaltsstatistische Ergebnisse zu erzeugen<sup>3)</sup>.

Eine wichtige Voraussetzung für verlässliche haushaltsstatistische Ergebnisse ist zum einen das Vorhandensein eines ausgereiften Verfahrens zur Haushaltegenerierung und zum anderen eine qualitativ hochwertige Datengrundlage. Für die in der Bundesrepublik Deutschland entwickelten Haushaltegenerierungsverfahren bilden die Personendatensätze<sup>4)</sup> aus den Einwohnermelderegistern diese Datenbasis. Ziel der Haushaltegenerierung ist es, haushaltstypische Per-

sonenbeziehungen zu bilden und damit Erkenntnisse über Anzahl, Größe und Struktur der Haushalte zu erlangen.

Vor allem in den skandinavischen Ländern werden seit langem Registerauswertungen als Ersatz für einen traditionellen Zensus durchgeführt. In diesen Ländern wurden in der Vergangenheit Statistikregister zu verschiedenen Themenbereichen – wie Bevölkerung, Erwerbstätigkeit und Wohnen – aufgebaut, deren Einzeldatensätze über einheitliche Nummern (Personenkennzeichen) eindeutig zusammengeführt werden können. Hier für bestehen in der Bundesrepublik Deutschland weder rechtlich noch technisch die Voraussetzungen<sup>5)</sup>, da die vorhandenen Register ausschließlich Verwaltungszwecken dienen. So enthalten die Register nicht alle gewünschten Merkmale und auch eine einheitliche Kennzeichnung der Personen ist nicht gegeben. In diesem Sinne sind die Register nur sehr eingeschränkt statistiktauglich. Dennoch wurden in der Bundesrepublik Deutschland Verfahren entwickelt, mit denen aus vorhandenen Registern Haushaltszusammenhänge hergestellt werden können.

Mit dem Urteil des Bundesverfassungsgerichts vom 15. Dezember 1983<sup>6)</sup> und der Entscheidung der Bundesregierung von 1996, künftig keine herkömmliche Vollerhebung nach dem Vorbild der Volkszählung 1987 mehr durchzuführen, gewann die Weiterentwicklung dieser Haushaltegenerierungsverfahren als Ersatz für primärstatistische Vollerhebungen zunehmend an Bedeutung.

2) Aussagen sind lediglich auf der Ebene von Anpassungsschichten möglich. Unter Anpassungsschichten sind i. d. R. benachbarte Kreise zu verstehen. Da auf Kreisebene die Ergebnisse des Mikrozensus teilweise so dünn besetzt sind, dass der Stichprobenfehler in dieser tiefen regionalen Gliederung zu groß sein würde, sind alle Kreise bzw. kreisfreien Städte, die weniger als 300 000 Einwohner haben, zu Anpassungsschichten zusammengefasst. – 3) Vgl. von Klitzing, Friedrich/Osenberg, Hanno, Haushaltsdaten aus dem Melderegister, in: Beiträge zur Stadtforschung und Statistik, Köln 1995, Heft 1, S. 15 ff. – 4) Als Personendatensatz wird der Einzeldatensatz bezeichnet, der sich aus den bei der Meldebehörde gespeicherten Daten (z. B. Vor-, Geburts-, Ehe- und Familienname, Geburtsdatum, Anschrift, Einzugsdatum, usw.) zusammensetzt. – 5) Vgl. Urteil des Bundesverfassungsgerichts vom 15. 12. 1983 zum Volkszählungsgesetz 1983, in: Bundesanzeiger, hg. vom Bundesminister für Justiz, Nr. 241a, Jg. 35, 24. 12. 1983. – 6) Vgl. Urteil des Bundesverfassungsgerichts, a. a. O., 1983.

## 2 Verfahren der Haushaltegenerierung

Das Bemühen insbesondere der kommunalen Statistikstellen, Verfahren zur Haushaltegenerierung zu entwickeln, reicht über 20 Jahre zurück. Bereits Ende der 70er-Jahre konstruierte die Stadt München mit der so genannten Familienstatistik ein Verfahren zur Gewinnung von Informationen über Haushaltszusammenhänge aus den Angaben des Melderegisters. In den 80er-Jahren folgte im Rahmen eines Forschungsvorhabens des Bundesministeriums für Forschung und Technologie das Verfahren SIHAGEN. Hier wurde in einem mehrstufigen Verfahren versucht, die im Einwohnermelderegister unter einer Adresse gemeldete Personen zu Personenverbänden zusammenzuführen. Aufgrund von Erfahrungen in der praktischen Anwendung wurde daraus das Verfahren HHGEN93 entwickelt. In den 90er-Jahren entstanden im statistischen Landesamt Bayern und Landesbetrieb Brandenburg zwei weitere Erfolg versprechende Verfahren.

Grundsätzlich kann unterschieden werden zwischen:

- Verfahren, die ausschließlich auf einer Auswertung von Informationen aus den Melderegistern beruhen.
- Verfahren, die zusätzlich gebäude- und wohnungsstatistische Informationen einbeziehen.

Die Konzepte dieser Verfahren sowie die Erfahrungen in Praxistests stellen die wichtigste Grundlage zur Weiterentwicklung der Methode „Haushaltegenerierung“ im Rahmen des Zensus tests dar.

### 2.1 Modell des KOSIS-Verbundes (HHGEN93/HHGEN98)

Das Modell HHGEN93 wurde 1993 vom Verbund der Städtestatistiker (KOSIS = Kommunales statistisches In-

formationssystem) zur Gewinnung von Informationen über die Anzahl und Zusammensetzung von Haushalten in Großstädten entwickelt. Eine Erweiterung dieses Verfahrens wird als HHGEN98 eingesetzt.

HHGEN93/98 zieht ausschließlich die in den Melderegistern enthaltenen Informationen zur Generierung von Haushalten heran. Die Meldebehörden liefern hierzu einen Datenabzug aus den Einwohnermelderegistern in Form einer standardisierten Basisdatei. Hierbei akzeptiert das Verfahren sowohl einen stichtagsbezogenen Einwohnerbestand als auch einen zeitraumbezogenen Einwohnerstamm. Der Verfahrensaufbau sieht vor, durch Merkmalsübereinstimmungen bzw. -kombinationen eine sukzessive Annäherung an die tatsächlich vorliegenden Haushaltsstrukturen zu erreichen. Hierbei ist jeder Schritt darauf ausgerichtet, unter Verwendung geeigneter Merkmale des Melderegisters, eine bestimmte haushaltstypische Personenbeziehung zwischen den unter einer Adresse gemeldeten Personen herzustellen<sup>7)</sup>.

Mit dem mehrstufigen Vorgehen in HHGEN93/98 wird versucht, den Kernhaushalten über Namensübereinstimmungen, gemeinsame Mobilität, Geschlecht und plausible Altersdifferenzen sukzessive weitere Personen zuzuordnen. Als Kernhaushalt werden Personenverbände bezeichnet, deren Datensätze im Melderegister über Verzeigerungen miteinander verbunden sind<sup>8)</sup>.

Ein Nachteil dieses Verfahrens ist, dass bereits erfolgte Zuordnungen von Personen nicht mehr überprüft bzw. geändert werden können. Weitaus schwerwiegender ist jedoch, dass mit diesem Verfahren niemals zwei Kernhaushalte zusammengeführt<sup>9)</sup> werden und ebenso die Bildung von gleichgeschlechtlichen Lebensgemeinschaften oder Wohnge-

meinschaften unterbleibt. In Kombination mit der Tatsache, dass gebäude- und wohnungsstatistische Daten bezüglich der Anzahl von im Gebäude vorhandenen und bewohnten Wohnungen nicht Bestandteil des Verfahrens sind, kommt es mit HHGEN93/98 zur Bildung von zu vielen Einpersonenhaushalten. Praxistests haben gezeigt, dass einem großen Überhang an Einpersonenhaushalten ein Defizit an Zweipersonenhaushalten gegenübersteht. Bei Drei- und Vierpersonenhaushalten erreicht HHGEN93/98 eine gute Übereinstimmung mit der Realität. Insgesamt zeigten die Tests, dass mit dem Verfahren bei jüngeren Personen zu wenige Haushaltszusammenhänge gebildet werden und bei älteren Personen (über 60 Jahre) zu häufig generiert wird.

### 2.2 Modell des Landesamtes für Statistik und Datenverarbeitung Bayern

Im Rahmen der Methodendiskussion für einen zukünftig registergestützten Zensus hat das Landesamt für Statistik und Datenverarbeitung Bayern ein Alternativverfahren entwickelt. Das Verfahren zur Zuführung von Einzeldaten aus Melderegistern mit primärstatistisch gewonnenen Wohnungsdaten sollte die grundsätzliche Tauglichkeit von Haushaltegenerierungsverfahren sowie die Qualität der Melderegister im Hinblick auf einen registergestützten Zensus testen<sup>10)</sup>.

Das bayerische Haushaltegenerierungsmodell bezog neben den Personendatensätzen der Einwohnermelderegister auch gebäude- und wohnungsstatistische Informationen ein, welche über eine Zusatzbefragung im Rahmen des Mikrozensus 1998 zur Verfügung standen. Untersuchungsgegenstand waren die im Mikrozensus einbezogenen Gebäude mit bis zu zehn Wohnungen sowie

7) Vgl. von Klitzing, Friedrich/Osenberg, Hanno, a. a. O., 1995, S. 17 ff. sowie von Klitzing, Friedrich, Haushaltegenerierungsverfahren HHGEN98, Verfahrensdokumentation im Auftrag des KOSIS-Verbunds vom 1. 1. 1998. – 8) Verzeigerungen sind Merkmale in einem Personendatensatz, die auf den Datensatz einer anderen Person hinweisen. Dieses ist der Fall bei Ehepartnern, minderjährigen Kindern und deren Eltern(-teilen) sowie minderjährigen Kindern und deren gesetzlichen Vertretern (einschließlich Stief- und Pflegeeltern). Teilweise enthalten die Melderegister in Deutschland auch Verzeigerungen zwischen Eltern und deren erwachsenen Kindern bis 27 Jahren. – 9) Es ist demnach nicht möglich, eine nichteheliche Lebensgemeinschaft, bestehend aus zwei Elternteilen mit ihren minderjährigen Kindern, zu einem Haushalt zusammenzuführen. – 10) Vgl. Fürnrohr, Michael/König, Markus, Möglichkeiten einer Haushaltegenerierung im Rahmen der Zuführung von Einzeldaten aus Melderegistern mit primärstatistisch gewonnenen Wohnungsdaten, in: Bayern in Zahlen, 4/99.



die Melderegistereinzeldatensätze der unter diesen ausgewählten Adressen gemeldeten Personen in den Städten Ingolstadt und Fürth. Da im Mikrozensus Gebäude mit mehr als zehn Wohnungen lediglich unvollständig erhoben werden, blieben größere Gebäude unberücksichtigt.

Die Einbeziehung von gebäude- und wohnungsstatistischen Daten, wie beispielsweise Wohnungsgröße oder Wohnungsanzahl je Gebäude, ermöglichte es, die Zahl der zu bildenden Haushalte durch die Anzahl von bewohnten Wohnungen je Gebäude von vorne herein zu determinieren. Eine Bildung zu vieler Haushalte wie im Verfahren HHGEN93/98 konnte dadurch vermieden werden. Wurden dennoch mehr Haushalte generiert als bewohnte Wohnungen im Gebäude vorhanden waren, so war dieses als Hinweis auf Schwächen im Melderegister zu deuten.

In dem mehrstufigen Verfahren wurde versucht, über die hierarchisch angeordneten Verknüpfungskriterien Haushaltszusammenhänge zu bilden<sup>11)</sup>. Zunächst zeigte die Überprüfung des Verfahrens anhand der Mikrozensusergebnisse, dass die Einbeziehung von gebäude- und wohnungsstatistischen Daten eine deutliche Qualitätsverbesserung gegenüber dem Verfahren HHGEN93/98 erbrachte. Das angewendete Verfahren wurde durch Auswertung des Zuordnungserfolgs je Generierungsstufe kontrolliert. Auf fast allen Stufen wurden akzeptable Ergebnisse erzielt.

Die Untersuchung wies jedoch einige Schwächen auf. So gelangten nur Gebäude des Mikrozensus in die Auswahl, wodurch nur Gebäude mit maximal zehn Wohnungen berücksichtigt wurden. Inwieweit das Verfahren auf die Gesamtheit der Gebäude in der Bundesrepublik Deutschland, vor allem auf Großgebäude, anwendbar wäre, blieb offen. Weiterhin wurden auch hier die

einmal erfolgten Zuordnungen der Personen nicht mehr überprüft bzw. geändert. Ebenfalls ist die Zusammenführung von zwei Kernhaushalten nicht möglich. Auch der Namensvergleich führte aufgrund teilweise unterschiedlicher Schreibweisen bzw. Schreibfehler im Melderegister und in den Angaben des Mikrozensus nicht immer zu eindeutigen Ergebnissen. Das Verfahren verfügte zudem nicht über eine Plausibilitätsüberprüfung hinsichtlich der maximalen Wohnungsbelegung. Eine Übergenerierung von Haushalten mit potenziellen Karteileichen<sup>12)</sup> aus dem Melderegister war somit nicht zu verhindern<sup>13)</sup>.

Weitere wichtige Erkenntnisse der bayerischen Untersuchungen waren hinsichtlich der Nutzung von Daten aus den Einwohnermelderegistern im Rahmen eines registergestützten Zensus, dass insbesondere die Erfassung von Personen mit Nebenwohnsitz in den Melderegistern unzureichend ist sowie teilweise große Divergenzen zwischen dem Einzugsdatum laut Melderegister und dem Einzugsdatum gemäß der Auskünfte im Mikrozensus vorliegen<sup>14)</sup>.

### **2.3 Modell des Landesbetriebes für Datenverarbeitung und Statistik Land Brandenburg (HHGEN-Statistik)**

HHGEN-Statistik ist ein weiteres vom Landesbetrieb für Datenverarbeitung und Statistik Land Brandenburg Ende der 90er-Jahre entwickeltes Verfahren zur Haushaltgenerierung. Ebenso wie bei dem bayerischen Verfahren werden Melderegisterdaten mit gebäude- und wohnungsstatistischen Informationen aus einer Zusatzbefragung des Mikrozensus kombiniert. Untersuchungsgegenstand waren alle im Mikrozensus einbezogenen Gebäude mit bis zu zehn Wohnungen sowie Melderegistereinzeldaten der ausgewählten Adressen in der Stadt Cottbus<sup>15)</sup>.

Die Tests mit HHGEN-Statistik zielten neben einer korrekten Ermittlung der Größe und Struktur von Haushalten darauf ab, Qualitätsunterschiede des Verfahrens durch Variation der zur Generierung herangezogenen Informationen zu untersuchen. Zu diesem Zweck wurde die Generierung derselben Personendatensätze in vier verschiedenen Testszenerarien durchgeführt und das jeweilige Ergebnis mit den entsprechenden Angaben aus dem Mikrozensus verglichen. In einem ersten Szenario wurden weder gebäude- oder wohnungsstatistische Informationen noch Zusatzinformationen aus dem Melderegistern verwendet. Dieses Vorgehen entsprach weitgehend dem Verfahren HHGEN93/98. In einer zweiten Variante wurden der Name des Vermieters (Wohnungsgeber) und, soweit vorhanden, die Etagennummer als Zusatzinformationen aus dem Melderegister herangezogen. Ein drittes Szenario berücksichtigte darüber hinaus die Anzahl der Wohnungen je Gebäude aus dem Mikrozensus. In einer vierten Variante wurden neben allen einbezogenen Zusatzinformationen der vorherigen Szenarien außerdem die Namen der Wohnungsinhaber aus dem Mikrozensus genutzt.

Auch HHGEN-Statistik zielte darauf ab, haushaltstypische Personenbeziehungen aufzudecken. Im vierten Szenario wurden Personendatensätze sowie Gebäude- und Wohnungsdatensätzen adressweise über eine Straßenschlüsseldatei verknüpft. Zur Bildung von Kernhaushalten nutzte das Verfahren die im Einwohnermelderegister enthaltenen Informationen über Verzeigerungen. An diese Vorstufe schlossen sich zehn aufeinander folgende Generierungsschritte an, welche jeweils auf einen bestimmten haushaltsstrukturellen Zusammenhang abstellten.

Im Gegensatz zu den oben beschriebenen Generierungsverfahren ist HHGEN-Statistik nicht streng hierarchisch aufgebaut. Auch bereits zuge-

11) Als Kriterien dienen: Verzeigerungen, Namensübereinstimmungen und identische Einzugsdaten. – 12) Karteileichen: Personen, die noch gemeldet sind, inzwischen aber nicht mehr unter der Meldeadresse wohnhaft sind (= Registerüberhang oder Übererfassung). – 13) Als Übergenerierung wird bezeichnet, wenn Personen zu Personenverbänden zusammen geführt werden, obwohl sie eigenständige Haushalte bilden. – 14) Diese Divergenzen müssen nicht unbedingt aus Fehlern in den Melderegistern resultieren. Es ist durchaus denkbar, dass der Auskunftgebende sein Einzugsdatum in der Mikrozensusbefragung falsch angegeben hat. – 15) Vgl. Fest, Manfred/Herzog, Petra, Möglichkeiten einer Haushaltgenerierung unter Verwendung von Indizien aus Einwohnermelderegistern und wohnungsstatistischen Daten, in: Daten und Analysen, 5/99.

ordnete Datensätze wurden auf weitere Generierungsmöglichkeiten hin überprüft, wodurch auf einer früheren Generierungsstufe geknüpfte Verbindungen wieder gelöst werden konnten. Nicht überraschend zeigte sich, dass die mit HHGEN-Statistik erzielten Ergebnisse mit einer steigenden Anzahl von Informationen (viertes Szenario: Gebäude- und Wohnungsdaten sowie Melderegister-Zusatzinformationen) – gemessen an den Ergebnissen des Mikrozensus – an Qualität zunahm. Die auf den verschiedenen Stufen erfolgten Zuordnungen erwiesen sich weitgehend als verlässlich.

Auch diese Untersuchung offenbarte Schwächen. Wie auch im bayerischen Modell gelangten nur Gebäude des Mikrozensus mit maximal zehn Wohnungen in die Auswahl. Die Frage nach einer Eignung des Verfahrens auf die Gesamtheit der Gebäude in der Bundesrepublik Deutschland bleibt auch hier unbeantwortet. Weiterhin führt der Namensvergleich bei unterschiedlichen Schreibweisen bzw. Schreibfehlern im Melderegister und in den Angaben des Mikrozensus nicht immer zu eindeutigen Ergebnissen. Ebenso traten Fehler bei der Generierung nichtehelicher Lebensgemeinschaften und bei Wohngemeinschaften auf. Probleme bereitete ebenfalls die Zuordnung erwachsener Kinder zu ihren Eltern, dort kam es relativ häufig zu einer Übergenerierung. Das Verfahren verfügt nicht über eine hinreichende Plausibilitätsüberprüfung hinsichtlich der Wohnungsbelegung unter Beachtung der Wohnfläche und Raumzahl. Eine Übergenerierung von Haushalten mit potenziellen Karteileichen aus dem Melderegister war somit nicht ausgeschlossen. Weiterhin war das Verteilen der zum Verfahrensende noch nicht einem Haushalt zugeordneten Personen nach Wohnungsgröße mit deutlichen Fehlern behaftet. Ohne eine Weiterentwicklung des Verfahrens, könnte im Rahmen eines registergestützten Zensus an dieser Stelle nicht auf eine primärstatistische Überprüfung verzichtet werden.

### 3 Vorbedingungen für ein Haushaltegenerierungsverfahren

Mit dem Übergang von einer traditionellen Vollerhebung zu einem registergestützten Zensus stellte sich die Frage, wie künftig Haushaltsergebnisse erhoben werden können. In einem herkömmlichen Zensus geschah dieses durch Primärbefragungen. Diese Möglichkeit ist in einem registergestützten Zensus nur in Einzelfällen vorgesehen. Eine Alternative bietet die Haushaltegenerierung. Die Erfahrungen in Praxistests mit den beschriebenen drei Haushaltegenerierungsverfahren haben gezeigt, dass es grundsätzlich möglich ist, die Zahl, Größe und Struktur von Haushalten realitätsnah abzubilden. Die Untersuchungen haben aber auch deutlich gemacht, dass die Qualität der Ergebnisse mit einer steigenden Anzahl von Informationen, insbesondere der Einbeziehung von gebäude- und wohnungsstatistischen Angaben, entscheidend zunimmt.

Werden dagegen lediglich Angaben aus den Einwohnermelderegistern herangezogen, so kommt es tendenziell zur Bildung von zu vielen Haushalten, da alle Personen, die nicht in einen Haushalt zusammenhang zu bringen waren, als Einpersonenhaushalte ausgewiesen wurden. Der adressenweise Vergleich von Melderegisterangaben mit Ergebnissen des Mikrozensus zeigte zudem gravierende Divergenzen in der Personenzahl. Damit ist es ohne die Einbeziehung weiterer Informationen, wie gebäude- und wohnungsstatistischer Angaben, nicht möglich, verlässliche und damit gerichtsfeste Einwohnerzahlen zu ermitteln. Durch die Zusammenführung von Melderegisterangaben mit gebäude- und wohnungssta-

tistischen Daten können in vielen Fällen die in den Praxistests aufgetretenen Fehler des Melderegisters aufgedeckt und statistisch bereinigt werden. Dieses ist von entscheidender Bedeutung bei der Generierung und damit bei der validen Ermittlung von Zahl, Größe und Struktur der Haushalte sowie der Wohnungsbelegung<sup>16</sup>.

Angesichts der Bedeutung der Einwohnermelderegister als zentrale Datengrundlage in einem registergestützten Zensus ist es notwendig, die Qualität der Melderegister und damit auch die erreichbare Qualität eines registergestützten Zensus zu überprüfen. Nach derzeitigem Kenntnisstand weisen die Melderegister in der Bundesrepublik Deutschland flächendeckend nicht eine Qualität auf, die eine ungeprüfte und unbereinigte Nutzung für einen registergestützten Zensus zulässt<sup>17</sup>. Im Rahmen des Zensusstests wurden deshalb Verfahren entwickelt, mittels derer die Qualität der Melderegister hinsichtlich der Anzahl von so genannten Mehrfachfällen<sup>18</sup>, der Über-<sup>19</sup> und Untererfassung<sup>20</sup> sowie hinsichtlich der Vollständigkeit und des Zutreffens der zu einer Person gespeicherten Merkmale überprüft und gegebenenfalls statistisch bereinigt werden kann<sup>21</sup>.

Daneben wird im Rahmen des Zensusstests in den ausgewählten Gebäuden eine Befragung aller Haushalte zu demographischen sowie erwerbs- und wohnungsstatistischen Merkmalen durchgeführt. Diese Befragung entspricht dem Vorgehen in einer herkömmlichen Volkszählung und soll als Maßstab für die Überprüfung der Qualität der Melderegister sowie der im Verfahren der Haushaltegenerierung gebildeten Haushaltssammenhänge dienen<sup>22</sup>. Der Ver-

16) Vgl. Fürnrohr, Michael/Rimmelspacher, Birgit, Testuntersuchungen zur Vorbereitung eines registergestützten Zensus, in: Bayern in Zahlen, 1/01. – 17) Eppmann, Helmut, Die Einwohnermelderegister als Basis der „amtlichen“ Einwohnerzahl, in: Grohmann, Heinz/Sahner, Heinz/Wiegert, Rolf (Hrsg.), Volkszählung 2001: Von der traditionellen Volkszählung zum Registertest, Allgemeines Statistisches Archiv/Sonderheft 33, 1999, S. 65 ff. – 18) Ein Mehrfachfall liegt vor, wenn eine Person in mehreren Gemeinden gleichzeitig mit alleinigem Wohnsitz, mit Hauptwohnsitz oder nur mit Nebenwohnsitz gemeldet ist. – 19) Übererfassung: Gemeldete Personen, die inzwischen nicht mehr unter der Meldeadresse wohnhaft sind (= Karteileichen oder Registerüberhang). – 20) Als Untererfassung (Fehlbestände) werden Personen bezeichnet, die an dem Ort ihrer Wohnung nicht gemeldet sind. Es handelt sich damit um eine Untererfassung im Melderegister. – 21) Zum Test der Melderegisterangaben vgl. Forster, Michael, Die Statistikauglichkeit der Einwohnermelderegister – Untersuchungen im Rahmen des Zensusstests 2001, Beitrag in diesem Heft. – 22) Vgl. Scharmer, Marco, Haushaltstypisierung im Rahmen des registergestützten Zensusstests, in: Statistische Analysen und Studien Nordrhein-Westfalen, Bd. 1 (2002).

gleich der Ergebnisse der Haushaltegenerierung mit denen der Haushaltebefragung stellt damit eine Kontrolle der Testergebnisse dar und dient als Maßstab für die Beurteilung des Modells registergestützter Zensus<sup>23</sup>). Zudem werden die Befragungsergebnisse in solchen Fällen herangezogen, in denen die Haushaltegenerierung nicht zu plausiblen Ergebnissen führt. In einem echten registergestützten Zensus würde hier eine primärstatistische Überprüfung notwendig werden.

#### 4 Das Haushaltegenerierungsverfahren im Zensus

Vor einem Methodenwechsel der amtlichen Statistik in der Bundesrepublik Deutschland, von einer primärstatistischen Vollerhebung zu einem registergestützten Zensus, sind somit umfangreiche Tests hinsichtlich der Qualität und Validität der aus den Registern gewonnenen Daten sowie die Weiterentwicklung der Haushaltegenerierungsverfahren durchzuführen. Für die Realisierung und Überprüfung der Haushaltegenerierung als zentralem Bestandteil eines registergestützten Zensus wurden im Rahmen des derzeit realisierten Zensus in einer Zufallsstichprobe bundesweit rund 16 000 Adressen in 222 Gemeinden ausgewählt (Verfahrenstest). In Nordrhein-Westfalen wurden hierzu in 18 Gemeinden 1 650 Adressen mit etwa 20 000 Einwohnern einbezogen. Für diese ausgewählten Adressen<sup>24</sup>) wurden von den statistischen Landesämtern Melderegisterauszüge aller dort gemeldeten Einwohner angefordert<sup>25</sup>).

##### 4.1 Umfangreiche Weiterentwicklung des Verfahrens

Die Erfahrungen in Praxistests mit den von dem statistischen Landesamt Bayern und Landesbetrieb Bran-

denburg entwickelten Generierungsverfahren haben gezeigt, dass die Qualität der Ergebnisse mit einer steigenden Anzahl von Informationen, vor allem mit der Einbeziehung von gebäude- und wohnungsstatistischen Daten, deutlich zunimmt. Daher wird auch das Generierungsverfahren im Zensus eine Zusammenführung von Daten aus vorhandenen Melderegistern und zusätzlichen primärstatistischen Erhebungen vornehmen. Aufgrund von fehlenden aktuellen und flächendeckenden gebäude- und wohnungsstatistischen Daten wurde eine postalische Gebäude- und Wohnungszählung (GWZ) bei den Eigentümern oder Verwaltern der Auswahlgebäude durchgeführt<sup>26</sup>). Beide Erhebungsteile sind weitgehend abgeschlossen, so dass nach Aufbereitung der Daten ab September 2002 mit der Haushaltegenerierung, also der Zusammenführung von Melderegister- und GWZ-Daten begonnen wird.

Das im Zensus eingesetzte Verfahren zur Haushaltegenerierung wurde vor allem durch die statistischen Landesämter Bayern und Nordrhein-Westfalen sowie den Landesbetrieb Brandenburg entwickelt. Hierbei wurden drei Zielsetzungen verfolgt:

- Gewinnung von Erkenntnissen über Zahl und Struktur von privaten Haushalten
- Gewinnung von Erkenntnissen über die Wohnungsbelegungen und damit über die Wohnsituation von privaten Haushalten
- Kontrolle und statistische Bereinigung der Melderegisterangaben auf Adressebene

In das Generierungsverfahren im Zensus flossen die Erfahrungen mit den bisher entwickelten Modellen ein. Die zutage getretenen Schwächen der bereits entwickelten Verfahren machte für den Zensus die Entwicklung eines erheblich erweiterten Programms notwendig.

Die entscheidende Änderung gegenüber den bisherigen Modellen besteht darin, dass im Zensus Personendatensätze aus den Melderegistern systematisch ausgewertet und auf Ebene der Einzeldatensätze adressenweise mit den in der GWZ erhobenen Wohnungsinhabern und deren Wohnungssätze zusammengeführt werden. Anschließend werden über die Algorithmen der Haushaltegenerierung die übrigen Personendatensätze im Melderegister mit den Wohnungsinhaberdatensätzen zu Haushalten verknüpft<sup>27</sup>). Dabei gibt die in der GWZ ebenfalls erfragte Zahl der bewohnten Wohnungen die Gesamtzahl der (Wohn-)Haushalte vor<sup>28</sup>).

##### 4.2 Der modulare Aufbau

Die einzelnen Verfahrensteile (Stufen) der Haushaltegenerierung sind ihrem Inhalt nach modularisiert. Das vorgeschaltete **Modul 0** dient zur Vorbereitung der Erstellung des Eingangs- und Prüfdatenmaterials und beinhaltet alle ausgewählten Personendatensätze der Melderegister sowie alle Wohnungsdatensätze aus der GWZ.

Das eigentliche Verfahren zur Zusammenführung und Haushaltegenerierung sieht in **Modul 1** vor, aus den Verzeigerungen in den Personendatensätzen der Melderegister Kernhaushalte zu bilden. Soweit in den Personendatensätzen der Melderegister Verzeigerungen enthalten sind, ist eine Zusammenführung von Personen zu Haushalten unkompliziert. Bei diesen Personenpaaren wird jeweils im Datensatz des Einen auf den Datensatz des Anderen verwiesen und vice versa. Die über Verzeigerungen zusammengeführten Personen bilden den so genannten Kernhaushalt, welcher ein vollständiger Haushalt sein kann, im weiteren Verlauf des Generierungsverfahrens aber auch zu einem Teil eines Haushalts mit weiteren nicht verzeigerten Personen werden kann<sup>29</sup>).

23) Zur Haushaltebefragung vgl.: Schäfer, Josef, Die Rolle der Haushaltebefragung im Zensus 2001, Beitrag in diesem Heft. – 24) Da das Haushaltegenerierungsverfahren im Zensus nicht zwischen Adresse und Gebäude unterscheidet, werden im weiteren Verlauf die beiden Begriffe synonym verwendet. – 25) Vgl. Forster, Michael, Die Statistikauglichkeit der Einwohnermelderegister (...), a. a. O., 2002, Beitrag in diesem Heft, S. 17 ff. – 26) Vgl. zur Durchführung der postalischen GWZ: Scharnhorst, Sebastian/Egbert, Henrik, Postalische Gebäude- und Wohnungszählung im Zensus 2001 – Erfahrungen in NRW, Beitrag in diesem Heft, S. 24 ff. – 27) Im Rahmen des beschriebenen Haushaltegenerierungsverfahrens werden als Haushalte auch Einzelpersonen i. S. v. Einzelpersonen verstanden. In den folgenden Ausführungen sind deshalb unter Haushalten auch immer Einpersonenhaushalte zu verstehen. – 28) Zum Begriff Wohnhaushalt vgl. Punkt „Der modulare Aufbau“, insbesondere unter Modul T Haushaltetypisierung. – 29) Ein Beispiel hierfür wäre, wenn im Haushalt einer Kernfamilie (Mutter, Vater, Kind) der Elternteil eines Ehepartners lebt.

In **Modul 2** wird versucht, für alle Gebäude mit mindestens einer bewohnten Wohnung über einen Namensabgleich die Wohnungsinhaber laut GWZ in den Melderegisterdatensätzen der entsprechenden Adresse zu identifizieren und mit der jeweiligen Wohnungsnummer zu verknüpfen. Hierzu hat das Landesamt für Statistik und Datenverarbeitung Bayern einen mehrstufigen maschinellen Abgleich zwischen den Vor- und Zunamen sowie dem Einzugsdatum der Wohnungsinhaber aus der GWZ mit den Familien-, Geburts-, Ehe-, Künstler-, Vor- und Rufnamen sowie den Einzugsdaten aus den Melderegisterdatensätzen programmiert. Aus dem Melderegister werden nur Datensätze von Personen herangezogen, die mindestens 15 Jahre alt sind, da jüngere Personen nicht als Wohnungsinhaber in Frage kommen. Der Wohnungsinhaber wird auch in den Verzeigerungen der gemeldeten Personen gesucht, selbst wenn er selber dort nicht gemeldet sein sollte<sup>30</sup>. Durch eine Variation der Parameter im Namensabgleich ist es möglich, auch dann eine maschinelle Identifikation der Wohnungsinhaber in den Melderegisterdatensätzen vorzunehmen, wenn die Namensschreibweisen nicht völlig identisch sind. Damit werden häufig auftretende Unterschiede in der Namensschreibweise oder Schreibfehler berücksichtigt. Nicht im Melderegister zu identifizierende Wohnungsinhaber sind potenzielle Fehlbestände.

Wie in konventionellen Erhebungen stellen Wohnungen von Angehörigen ausländischer Streitkräfte, diplomatischer oder konsularischer Vertretungen sowie Ferien- und Freizeitwohnungen auch im Zuge der Haushaltgenerierung einen Sonderfall dar. In aller Regel ist der angeführte Personenkreis nur selten gemeldet, weshalb diese Adressen gesondert behandelt werden müssen. Hierzu dient das **Modul 5**, welches im Anschluss an Modul 2 zum Einsatz kommt. In Modul 5 können entsprechende Woh-

nungen und gegebenenfalls zugeordnete Personendatensätze aus dem weiteren Generierungsverfahren herausgenommen werden.

An die bisherigen Schritte schließt sich das **Modul F1** an, welches hier erstmals gestartet wird und nach jeder Stufe bzw. Generierung der Module 3, 4, 5, 6, und 7 zur Anwendung kommt. Modul F1 gewährleistet, dass Gebäude in denen alle Haushalte und Wohnungen verknüpft sind, nicht mehr für das weitere Generierungsverfahren zur Verfügung stehen. Für diese Adressen ist die Haushaltgenerierung abgeschlossen.

Für Fälle, in denen es nicht gelungen ist, einen Wohnungsdatensatz aus der GWZ mit einem Personendatensatz aus dem Melderegister zu verknüpfen, hält **Modul 3** zwei Optionen bereit. So kann an dieser Stelle entschieden werden, die aufgedeckten Unpaarigkeiten nicht zu berücksichtigen. Die Datensätze gehen damit in das weitere Generierungsverfahren ein. Alternativ wird die Möglichkeit vorgesehen, aufgedeckte Unpaarigkeiten über die Gewinnung zusätzlicher Informationen (Rückfragen) zu korrigieren. Im Zensustest kann auf eine Rückfrage verzichtet werden, da die benötigten Informationen über die als Teil des Zensustests durchgeführte Haushaltebefragung unter den ausgewählten Adressen zur Verfügung stehen (primärstatistische Simulation). Die Entscheidung darüber, welche der zwei Optionen wahrgenommen wird, hängt vor allem von den im Modul 2 erzielten Ergebnissen ab, d. h. wie sich das Verhältnis von unverknüpften Wohnungen zur Zahl aller bewohnten Wohnungen im Gebäude darstellt. Überschreitet dieses Verhältnis den zuvor gesetzten Grenzwert (Parameter), so wird das Gebäude primärstatistisch simuliert.

Die Generierung von Haushalten aufgrund der im Melderegister enthaltenen Verzeigerungen (Modul 1) führt in der Regel zu einer validen Haus-

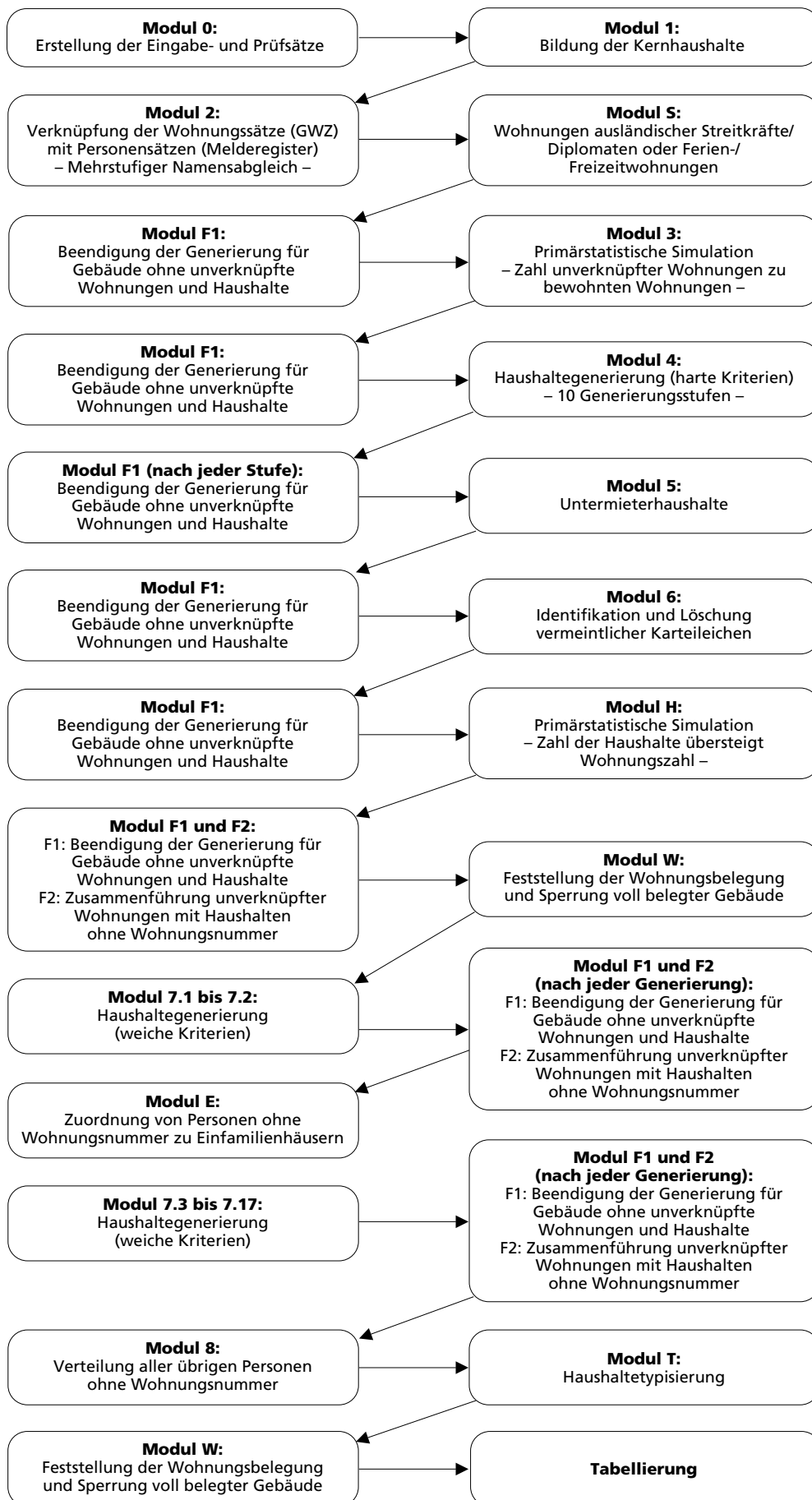
haltsbildung. Wesentlich schwieriger gestaltet sich hingegen eine Zusammenführung von Personen, deren Melderegisterdatensätze keine Verzeigerungen aufweisen. Häufig bestehen mehrere Möglichkeiten, Personen zu verschiedenen Haushalten zusammenzuführen<sup>31</sup>). In den weiteren Modulen des Haushaltgenerierungsverfahrens wird versucht, jeder Personenbeziehung, die einen Haushaltszusammenhang bedingen kann, eine bestimmte Übereinstimmung oder Kombination der im Personendatensatz enthaltenen Merkmale zu unterstellen.

Ab **Modul 4** wird für jede potenzielle Personenbeziehung eine Anzahl von Bedingungen in Form von Merkmalsübereinstimmungen bzw. Merkmalskombinationen zwischen den Personendatensätzen aus den Melderegistern und der GWZ definiert. Sind diese erfüllt, wird die Zugehörigkeit der betreffenden Personen zum entsprechenden Haushalt unterstellt.

Modul 4 beinhaltet zehn Generierungsstufen, welche auf haushaltstypische Personenbeziehungen fokussiert sind. Wie in den zuvor beschriebenen Modellen aus Bayern und Brandenburg, werden hier unter Heranziehung von Daten aus den Melderegistern und der GWZ für die unter einer Adresse wohnhaften Personen Haushalte generiert. In der ersten Stufe von Modul 4 werden alle Personen der verschiedenen zuvor in Modul 1 gebildeten Kernhaushalte zusammengeführt, welche in Modul 2 eine gleiche Wohnungsnummer erhielten. Damit ist es möglich, verschiedene Kernhaushalte zu einem Wohnhaushalt zu generieren. In der zweiten Stufe werden Personen, für die im Melderegister der Name des Wohnungsgebers gespeichert ist, über einen Namensabgleich mit diesem zusammengeführt. In der dritten Stufe werden nicht oder nicht vollständig verzeigerte minderjährige Kinder über Namensübereinstimmungen und unter Beachtung von Altersdifferenzen mit erwachsenen

30) Mit diesem Schritt soll einer Konstellation Rechnung getragen werden, in der eine Person einen Mietvertrag für ein Familienmitglied abgeschlossen hat, damit also Wohnungsinhaber ist, selber aber nicht unter der ausgewählten Adresse wohnt und gemeldet ist (z. B. Eltern schließen Mietvertrag für minderjährige Kinder). – 31) So können mit dem Verfahren der Haushaltgenerierung Haushalte mit nichtehelichen Lebensgemeinschaften oder Wohngemeinschaften nur sehr schwer gebildet werden, da hier Merkmalsübereinstimmungen, die auf Namensgleichheiten oder gemeinsamer Mobilität aufbauen, nur selten zutreffen.

### Modularer Ablauf der Zusammenführung und Haushaltgenerierung



Grafik: LDS NRW

Personen zu Haushalten generiert. Die vierte Stufe bildet nichteheliche Paare über im Melderegister enthaltene Verzeigerungen zwischen minderjährigen Kindern und dem jeweiligen nichtehelichen Vater. Stufe fünf führt Ehepaare, deren Personendatensätze nicht über Verzeigerungen miteinander verknüpft sind, mittels Namensübereinstimmungen und unter Berücksichtigung des Familienstandes, des Geschlechtes sowie des Familienstandsänderungsdatums zusammen. In der sechsten Stufe werden nicht verzeigerte erwachsene Nachkommen mit Eltern(-teilen) oder Großeltern(-teilen) über Namensübereinstimmungen, Altersdifferenzen und Einzugsdaten zu Haushalten generiert. Stufe sieben generiert ebenfalls über Namensübereinstimmungen und unter Beachtung von Altersdifferenzen und Einzugsdaten ältere Einzelpersonen zu deren Nachkommen. In der achten Stufe werden Geschwisterzusammenhänge aufgrund von Namensübereinstimmungen, Altersdifferenzen und Einzugsdaten gebildet. Stufe neun generiert nichteheliche Paare (auch mit Nachkommen) unter Berücksichtigung von Altersdifferenzen und Einzugsdaten. Stufe zehn führt schließlich alle Einzelpersonen über Einzugsdaten und/oder gleiche Herkunftsadressen unter Beachtung einer maximalen Wohnungsbelegung zu Haushalten zusammen.

In **Modul 5** kann den Personendatensätzen von Einzelpersonen mit einem Mindestalter von 15 Jahren, die bislang noch nicht mit einem Wohnungsdatensatz verknüpft waren, ein Wohnungsdatensatz mit Untermieterkennzeichnung zugewiesen werden, wenn diese Kennzeichnung vorhanden ist. Diese Verknüpfung wird unabhängig davon durchgeführt, ob der jeweilige Wohnungsinhaber im Melderegister identifiziert wurde oder nicht.

Erste Auswertungen der adressweisen Zusammenführung von Melderegisterangaben mit den Haushalbefragungsergebnissen haben gezeigt, dass sich Registerüberhänge, die so genannten Karteileichen, vor allem auf erwachsene Nachkommen und

geschiedene Personen sowie auf Personen mit Nebenwohnsitz unter der Auswahladresse konzentrieren. In einem gegenwärtig entwickelten **Modul 6** sollen anhand verschiedener Indizien aus Melderegister- und GWZ-Angaben diese vermeintlichen Karteileichen möglichst maschinell erkannt und statistisch bereinigt werden. Ziel von Modul 6 ist es, alle Personen, die tatsächlich nicht mehr unter der Meldeadresse wohnhaft sind, als vermeintliche Karteileichen zu kennzeichnen und nicht mehr in das weitere Generierungsverfahren einzubeziehen. Gelingt dieser Schritt, so kann mit Modul 6 die Qualität der Einwohnermelderegister hinsichtlich eines Zutreffens des Tatbestandes Karteileiche beurteilt werden.

Übersteigt in einem Gebäude die Anzahl der bisher gebildeten Haushalte die Zahl der Wohnungen um einen zuvor festgelegten Grenzwert, so ist im sich anschließenden **Modul H** (analog Modul 3) eine primärstatistische Simulation vorgesehen. Anders als in Modul 3 ist hier keine gezielte Simulation einzelner Haushalte möglich, da nicht für alle Haushalte bekannt ist, zu welchen Wohnungen diese tatsächlich gehören. Somit werden in Modul H alle Personen- und Wohnungsätze unter der Adresse aus der Haushalbefragung übernommen.

Hieran schließen sich die Module F1 und F2 an. **Modul F2** kann nach jeder Stufe der Module 7 und 8 zur Anwendung kommen und wird immer nur nach einem Durchlauf von Modul F1 gestartet. In Modul F2 werden noch nicht mit einer Wohnung verknüpfte Haushalte mit Wohnungen, deren Wohnungsinhaber nicht in den Melderegisterdatensätzen identifiziert werden konnten, unter Beachtung von Einzugsdatum, Personenzahl, Wohnungsgröße, Alterskonstellationen und Staatsangehörigkeiten zusammengeführt. Voraussetzung für den Start von Modul F2 ist, dass die Anzahl der noch nicht mit einer Wohnung verknüpften Haushalte kleiner oder gleich der Zahl der unverknüpften Wohnungen innerhalb des Gebäudes ist. Für entsprechende

Gebäude ist damit die Generierung abgeschlossen.

Im Anschluss kommt **Modul W** zum Einsatz. Hier wird überprüft, ob eine Wohnung bereits voll belegt ist. Eine Wohnung gilt als voll belegt, wenn das Verhältnis von Personenzahl zur Raum- bzw. Quadratmeterzahl der Wohnung einen zuvor festgelegten plausiblen Grenzwert überschreitet. In diesem Fall kann keine weitere Person zu den bereits mit der Wohnung verknüpften Haushaltsmitgliedern hinzugeneriert werden. Die betreffende Wohnung steht für das weitere Generierungsverfahren nicht mehr zur Verfügung.

**Modul 7** stellt auf die Bildung von weiteren typischen Haushaltsformen ab. Im Vergleich zu Modul 4 sind die hier gebildeten Haushaltszusammenhänge mit größeren Unsicherheiten verbunden. Modul 7 versucht alle Einzelpersonen, die zuvor noch nicht mit einer Wohnung verknüpft und die in Modul 6 nicht als vermeintliche Karteileiche aus dem Generierungsverfahren herausgenommen wurden, zu Haushalten zusammenzufassen. In den Stufen eins bis neun werden zunächst (Einpersonen-)Haushalte ohne Wohnungsinhaber mit Haushalten eines Wohnungsinhabers, dessen Wohnung für weitere Generierungsschritte nicht in Modul W gesperrt wurde, zusammengeführt. Da bei diesen Generierungsstufen die jeweilige Wohnungsgröße bekannt ist, erfolgt eine Generierung weiterer Personen zu einem Wohnungsinhaberhaushalt nur dann, wenn der zuvor festgelegte Schwellenwert der maximalen Wohnungsbelegung (analog Modul W) nicht überschritten wird. Sollten nach Stufe neun noch nicht mit einer Wohnung verknüpfte Personen unter der Adresse vorhanden sein, so werden mit Ausnahme der zweiten alle Stufen nochmals durchlaufen. In diesem zweiten Lauf sollen – unter den gleichen Bedingungen des ersten Laufs – noch nicht mit einer Wohnung verknüpfte Personen zu Haushalten zusammengeführt werden. Eine Überprüfung der maximalen Wohnungsbelegung kann hier somit nicht stattfinden.

In der ersten Stufe von Modul 7 werden Einzelpersonen aufgrund gleichen Einzugsdatums und unter Beachtung von Nachkommeneigenschaften zu Haushalten zusammengeführt. Ein abweichendes Einzugsdatum innerhalb einer gewissen Zeitspanne ist hierbei zulässig. In der zweiten Stufe werden erwachsene Einzelpersonen ohne Wohnungsnummer<sup>32)</sup> unter Beachtung von Geschlechts- und Alterskonstellationen mit Einzelpersonen oder Elternteilen zu Lebensgemeinschaften generiert.

Nach der zweiten Stufe von Modul 7 wird für alle Gebäude mit genau einer bewohnten und nicht gesperrten Wohnung **Modul E** gestartet. Hier werden sämtliche unter der Adresse im Melderegister genannten und nicht deaktivierten Personen ohne Wohnungsnummer zu einem gemeinsamen Wohnhaushalt zusammengeführt. Eine Plausibilitätsprüfung hinsichtlich der maximalen Wohnungsbelegung findet nicht statt. Für entsprechende Gebäude ist der Generierungsprozess nach Modul E abgeschlossen.

Nach Modul E wird mit der dritten Stufe von **Modul 7** fortgefahren. In der dritten Stufe werden alle nicht-deutschen Einzelpersonen, die kein Wohnungsinhaber sind, mit Mehrpersonenhaushalten zusammengeführt, bei denen sämtliche Haushaltsmitglieder die gleiche Staatsangehörigkeit wie die Einzelperson besitzen. Stufe vier führt alle übrigen nichtdeutschen Einzelpersonen, die keine Wohnungsinhaber sind, mit Mehrpersonenhaushalten zusammen, in denen mindestens ein Haushaltsmitglied die gleiche Staatsangehörigkeit wie die Einzelperson besitzt. Stufe fünf führt die übrigen nichtdeutschen Einzelpersonen gleicher Staatsangehörigkeit zusammen, sofern die betreffenden Personen nicht Inhaber verschiedener Wohnungen sind. Befindet sich ein Wohnungsinhaber unter den hier generierten Personen, wird unter Beachtung der maximalen Wohnungsbelegung zusammengeführt. In Stufe

32) Hierunter sind Personen zu verstehen, deren Personendatensatz noch nicht mit einem Wohnungsdatensatz verknüpft wurde.

sechs findet eine Zusammenführung von jugendlichen Einzelpersonen zu Wohngemeinschaften statt. Die Generierung erfolgt hierbei zunächst für Personen im Alter zwischen 18 und 35 Jahren und anschließend für minderjährige Personen im Alter von 15 bis 17 Jahren. In Stufe sieben werden gleichgeschlechtliche Einzelpersonen zu Haushalten generiert, wenn diese nicht Inhaber unterschiedlicher Wohnungen sind. Minderjährige ausländische Kinder im Alter zwischen 14 und 17 Jahren werden in Stufe 8 einem beliebigen Haushalt mit Kindern im schulfähigen Alter zugeordnet. Schließlich werden in der neunten Stufe minderjährige Kinder unter Beachtung von Alterskonstellationen mit einem beliebigen Haushalt, dem eine erwachsene weibliche Einzelperson angehört, zusammengeführt.

Da es auch nach den vorangegangenen Generierungsstufen bei einer geringen Anzahl von Fällen möglich ist, dass noch nicht alle gebildeten Haushalte mit einer Wohnung verknüpft sind, werden in **Modul 8** zunächst die verbleibenden Haushalte ohne Wohnungsnummer mit unverknüpften Wohnungen zusammengeführt. Hierbei kommen die Bedingungen von Modul F2 zur Anwendung, ohne dass hierbei aber die Eingangsbedingung „Anzahl von Haushalten ohne Wohnungsnummer darf nicht größer als die Zahl der unverknüpften Wohnungen innerhalb des Gebäudes sein“ erfüllt sein muss. Alle übrigen unverknüpften Haushalte werden entsprechend der Wohnfläche auf die unverknüpften bewohnten Wohnungen des Gebäudes verteilt. Spätestens mit Modul 8 ist die Haushaltegenerierung für die jeweilige Adresse abgeschlossen.

Durch die Haushaltegenerierung werden Haushaltszusammenhänge von Personen aufgrund bestimmter Übereinstimmungen oder Kombinationen der im Personendatensatz der Melderegister enthaltenen Merkmale gebildet. In Modul T werden die generierten Haushalte entsprechend der Definition Wohnhaushalte typisiert, um Haushaltsstrukturen darstellen zu können. An dieser Stelle ver-

lässt die amtliche Statistik ihre bisher übliche Haushaltsdefinition<sup>33)</sup> – ein Haushalt ist eine Personengruppe, die gemeinsam wohnt und wirtschaftet, d. h. insbesondere ihren Lebensunterhalt gemeinsam finanziert. Im Rahmen der Haushaltegenerierung kann nur auf das gemeinsame Wohnen der haushaltszugehörigen Personen abgestellt werden (Wohnhaushalte), da die Melderegister als zentrale Datengrundlage keine Informationen über die wirtschaftlichen Verhältnisse einzelner Haushaltsmitglieder liefern können<sup>34)</sup>.

## 5 Fazit und Perspektiven des Haushaltegenerierungsverfahrens in einem registergestützten Zensus

Grundsätzlich können Haushaltegenerierungsverfahren eine praktikable Alternative zu Totalerhebungen darstellen, um aktuelle und kleinräumig gegliederte Daten über Anzahl, Größe und Struktur von Haushalten sowie deren soziale und wirtschaftliche Verhältnisse zu erhalten. Die bisher entwickelten Haushaltegenerierungsverfahren zeigten zum Teil jedoch deutliche Schwächen, wodurch deren Eignung als Ersatz einer primärstatistischen Vollerhebung eingeschränkt ist.

Das im Rahmen des Zensusstests entwickelte Verfahren wird voraussichtlich über die zusätzlich zu Melderegistereinzeldaten einbezogenen gebäude- und wohnungsstatistischen Informationen wesentliche Verbesserungen hinsichtlich der Qualität von Generierungsergebnissen erzielen. Durch die Verwendung von Gebäude- und Wohnungsdaten ist die Zahl der Wohnungen je Gebäude bekannt und somit die Anzahl von Haushalten je Gebäude von vornherein bestimmt. Damit wird eine Untergenerierung wie im KOSIS-Verfahren vermieden. Weiterhin liegen Wohnungsgröße und Raumzahl der Wohnungen im Gebäude vor. Hier-

33) beispielsweise in der Volkszählung 1987 oder im Mirozensus – 34) Zum Übergang vom Wirtschafts- zum Wohnhaushalt und der damit verbundenen Problematik vgl.: Scharmer, Marco, Haushaltstypisierung im Rahmen des registergestützten Zensusstests, a. a. O., 2002.

durch kann eine Plausibilitätsprüfung hinsichtlich einer maximalen Wohnungsbelegung durchgeführt werden. Eine mögliche Übergenerierung von Haushalten mit potenziellen Karteileichen, wie in den bayerischen und brandenburger Verfahren, wird ausgeschlossen.

Die Durchführung einer GWZ ist im Vorfeld einer Haushaltegenerierung somit unverzichtbare Voraussetzung, um Haushaltsstrukturen in zensustypischer Qualität abbilden zu können. Da zum gegenwärtigen Zeitpunkt in der Bundesrepublik Deutschland weder in der amtlichen noch in der nicht amtlichen Statistik Quellen vorhanden sind, denen aktuelle und flächendeckende gebäude- und wohnungsstatistische Daten entnommen werden können, müssen diese Information auch in einem flächendeckenden registergestützten Zensus primärstatistisch erhoben werden<sup>35</sup>.

35) Vgl. in diesem Zusammenhang: Egbert, Henrik/Scharmer, Marco, Die Ermittlung des Wohnungsinhabers im Rahmen der Gebäude- und Wohnungszählung (GWZ) und deren Bedeutung für die Haushaltegenerierung, in: Statistische Analysen und Studien Nordrhein-Westfalen, Bd. 1 (2002).

Die Qualität der Ergebnisse des Haushaltegenerierungsverfahrens hängt somit entscheidend von der Ausgestaltung des Verfahrensablaufs und von der Qualität der Eingangsdaten, insbesondere der Melderegisterangaben, ab. So ist im Rahmen des Zensusstests das Generierungsprogramm hinsichtlich des Aufdeckens und der maschinellen Bereinigung von Karteileichen im Melderegister zu optimieren. Daneben steht die Haushaltegenerierung vor der Aufgabe, geeignete Merkmalsübereinstimmungen bzw. Merkmalskombinationen zu finden und diese in eine Reihenfolge zu bringen, mit denen die in der Realität vorliegenden Haushaltszusammenhänge verlässlich abgebildet werden können. Gleichzeitig werden die Melderegister flächendeckend keine optimale Qualität aufweisen. Unstimmigkeiten der Melderegister (Überhänge, Fehlbestände) könnten eine valide Haushaltegenerierung unter bestimmten Adressen ohne zusätzliche primärstatistische Informationen (Haushaltebefragung)

unmöglich machen. Daher wird in einem künftigen registergestützten Zensus eine Haushaltebefragung unter bestimmten Adressen notwendig sein.

Im Zensusstest stehen in Fällen einer unplausiblen Generierung die Ergebnisse der Haushaltebefragung zur Verfügung, welche in einem maschinellen Verfahren als Ersatzvornahme übernommen werden können. In einem späteren registergestützten Zensus muss bei unplausiblen Ergebnissen unter entsprechenden Adressen eine Haushaltebefragung erst organisatorisch aufwändig im Nachhinein durchgeführt werden. Der Zensusstest wird zeigen, ob vermutlich problematische Adressen bereits vorab aus den Melderegisterangaben identifiziert und gegebenenfalls primärstatistisch erhoben werden können. Die Analyse der Testergebnisse wird somit entscheidend zur Evaluation und Weiterentwicklung des Modells registergestützter Zensus beitragen.



# Die Rolle der Haushaltebefragung im Zensustest

Dipl.-Statistiker Josef Schäfer

*Wie lässt sich beurteilen, ob das registergestützte Zensusmodell geeignet ist, eine herkömmliche Volkszählung zu ersetzen? Die Antwort auf diese Frage liegt auf der Hand: das Modell muss unter den gegebenen Umständen in der Lage sein, eine vergleichbare Datenqualität zu erzielen!*

*Um einen solchen Vergleich zu ermöglichen, fand im Rahmen des Zensustests auch eine primärstatistische Befragung einzelner Haushalte statt, wie sie auch in früheren Volkszählungen angewandt wurde. Sie soll einen Vergleichsmaßstab für die im Test zu erprobenden neuen Zensusverfahren liefern und Aussagen über die Qualität der Melderegister ermöglichen.*

## 1 Aufgaben

Eine Haushaltebefragung ist nicht Bestandteil des registergestützten Zensusmodells. Sie findet lediglich im Zensustest statt und dient dazu, einen Vergleich des registergestützten Modells mit der herkömmlichen Volkszählung zu ermöglichen.

Verglichen wird dabei zweierlei. Im Rahmen des so genannten Verfahrenstest soll sich zeigen, ob die für den registergestützten Zensus entwickelten Verfahren und Methoden ähnliche Ergebnisse liefern, wie eine Zählung auf dem herkömmlichen Wege der Haushaltebefragung durch Interviewer. Um diese beiden Erhebungsvarianten am konkreten Fall zu vergleichen ist es also notwendig, beide parallel durchzuführen und die jeweiligen Ergebnisse gegenüberzustellen. Des Weiteren soll mittels der Haushaltebefragung geprüft werden, ob in der Bundesrepublik die Voraussetzungen gegeben sind, ein registergestütztes Verfahren in hinreichender Qualität durchführen zu können. Sollen zukünftige Zensen auf Basis der Verwaltungsregister erfolgen, so muss zunächst die Eignung dieser Register für ein solches Verfahren bewertet werden. Ein Vergleich mit der Haushaltebefragung soll Aufschluss darüber geben, inwieweit die Einwohnermelderegister im Hinblick auf Karteileichen und Fehlbestände, also Über- bzw. Untererfassungen, die für die Anwendung des registergestützten Zensus notwendige Aktualität und Validität aufweisen.

## 2 Stichproben und Auswahlverfahren

Entsprechend ihrer Aufgaben teilt sich die Haushaltebefragung in zwei Bereiche, denen jeweils unterschiedliche Stichproben zugrunde liegen.

Die Überprüfung der Verwaltungsregister, der so genannte Registertest, basiert auf einer Zufallsstichprobe (Hauptstichprobe) von 38 000 Adressen in bundesweit 555 Gemeinden, wovon 36 Gemeinden mit 4 138 Adressen und etwa 55 000 Personen in Nordrhein-Westfalen liegen. Die Überprüfung einzelner Verfahren des Zensusmodells, der Verfahrenstest, findet in einer Unterstichprobe dieser Hauptstichprobe statt. Von den darin bundesweit enthaltenen 222 Gemeinden mit rund 16 000 Adressen entfallen 18 Gemeinden mit 1 650 Adressen auf Nordrhein-Westfalen.

Die Auswahl dieser Stichproben erfolgte anhand einer komplexen mehrstufigen Zufallsstichprobe mit unterschiedlichen Auswahlwahrscheinlichkeiten auf mehreren Ebenen. Die Aufteilung der bundesweiten Stichprobe auf die Bundesländer erfolgte nicht größenproportional, sondern aufgrund von Vorgaben zur Erreichung eines relativen Standardfehlers von 15 % bzgl. der Überhänge und Fehlbestände der Melderegister in jedem einzelnen Bundesland, unter Berücksichtigung der jeweiligen gemeindegrößenabhängigen Siedlungsstruktur. Nordrhein-

Westfalen erhält dabei mit 4 138 Adressen einen geringeren Anteil als es seiner Einwohnerzahl entspräche. Innerhalb der Bundesländer erfolgte die Auswahl zweistufig. Auf der ersten Stufe wurden Gemeinden nach Größenklassen geschichtet und innerhalb der Schichten mit einer Wahrscheinlichkeit proportional zur Einwohnerzahl mittels einer systematischen Auswahl gezogen. Innerhalb der Gemeinden (zweite Stufe) war die Anzahl ausgewählter Adressen schicht-, d. h. größenklassenspezifisch, vorgegeben. Die Wahrscheinlichkeit der Auswahl einzelner Adressen war innerhalb der Gemeinden proportional zur Zahl der darin gemeldeten Personen. Größere Gebäude sind daher (gewollt) stark überrepräsentiert, was durch das Hochrechnungsverfahren wieder ausgeglichen wird. Der Grund für die ungleichen Auswahlwahrscheinlichkeiten liegt zum einen in der zu erwartenden größeren Streuung bzgl. des Meldeverhaltens bei größeren Gebäuden, die zur Optimierung der Gesamtstichprobe einen erhöhten Anteil größerer Gebäude verlangt. Zum anderen wird bei größeren Gebäuden die Haushaltegenerierung i. d. R. schwieriger durchzuführen sein, und im Test sind verstärkt die potenziellen Problemfälle zu eruieren.

Für die Unterstichprobe erfolgte zunächst eine Vorauswahl von Gemeinden aus der Hauptstichprobe. Eine Schichtung nach Größenklassen und eine innerhalb der Schichten erfolgte systematische Auswahl bei einer Sortierung nach der Einwohnerzahl stellte dabei eine gleichmäßige Aufteilung der Stichprobenteile bzgl. der Einwohnerzahlen der Gemeinden sicher. Die Aufteilung der Stichprobenteile innerhalb der vorausgewählten Gemeinden war schichtspezifisch vorgegeben.

Sowohl bzgl. der Haupt- als auch der Unterstichprobe sind Ergebnisse aufgrund unterschiedlicher Auswahl-

wahrscheinlichkeiten grundsätzlich hochzurechnen. Bei ungewichteten Auszählungen ist mit erheblichen statistischen Verzerrungen rechnen.

### 3 Die Haushalbefragung als Maßstab zur Bewertung der Registerqualität

Aufgabe des Registertests ist die Überprüfung der Qualität der einem künftigen registergestützten Zensus zu Grunde liegenden Einwohnermelderegister<sup>1)</sup> sowie der Statistikauglichkeit der Register der Bundesanstalt für Arbeit. Die Haushalbefragung als Stichproben-Simulation einer klassischen Volkszählung, d. h. eines bewährten und allgemein akzeptierten Verfahrens, stellt dabei den Bewertungsmaßstab dar. Überhänge und Fehlbestände des Melderegisters werden als Differenzen zu den Feststellungen der Haushalbefragung ermittelt. Wird ein 4-Personen-Haushalt befragt, bei dem nur 3 Personen im Melderegister enthalten sind, ergibt sich ein Fehlbestand des Melderegisters. Enthält das Melderegister dagegen fünf Personen gleichen Namens, von denen keine Person einem anderen Haushalt der Adresse zugeordnet werden kann, ergibt sich ein Melderegister-Überhang. Entsprechendes gilt auch für Überhänge und Fehlbestände vollständiger Haushalte.

Eindeutige Aussagen ergeben sich bei Primärerhebungen von Haushalten i. d. R. nur als Resultat einer tatsächlich durchgeführten Befragung oder aufgrund der eindeutigen Feststellung eines Leerstands oder einer gewerblichen Nutzung der Wohnung durch Erhebungsbeauftragte. Informationen über Haushalte, von denen auch nach Durchführung des Mahnverfahrens keine Auskunft zu erhalten war oder die zwischen Stichtag und Begehung fortgezogen sind, können nicht die Sicherheit ei-

ner vor Ort erfolgten Befragung aufweisen. Zur Dokumentation der Sicherheit der Feststellung der Existenz oder Nicht-Existenz von Personen im Rahmen der Befragung wurde das Befragungsergebnis in der sog. Verteilungsliste vom/von der Interviewer(in) wie folgt verschlüsselt:

- 01 Angaben vollständig erhalten
- 02 Angaben unvollständig (d. h. Angaben liegen nur für einzelne Personen, nicht für den vollständigen Haushalt vor)
- 03 Selbstausfüllerwunsch (d. h. der Haushalt wünschte eine schriftliche Beantwortung ohne Beteiligung des/der Interviewer(s/in))
- 04 Auskunftsverweigerung
- 05 Name und Zahl der Bewohner ermittelt, Wohnung augenscheinlich bewohnt
- 06 Nur Name ermittelt, Zahl der Bewohner unbekannt, Wohnung augenscheinlich bewohnt
- 07 Kein Name ermittelt, aber Wohnung augenscheinlich bewohnt
- 08 Von Angehörigen ausländischer Streitkräfte/Diplomaten genutzt
- 09 Ausschließlich gewerblich genutzte Wohnung
- 10 Wohnung leer stehend
- 11 Bereits im Gebäude erfasster Haushalt

Befragungsergebnis „01“ bedeutet eine durchgeführte Befragung und somit eine sichere Feststellung. Die übrigen Befragungsergebnisse einer von Auskunftspflichtigen bewohnten Wohnung („02“ – „07“) werden zur Dokumentation der Befragung zunächst als solche erfasst und gespeichert. Nach Rücklauf eines Erhebungsbogens werden die ursprünglich nicht unmittelbar befragten Haushalte ebenfalls wie Befragungsergebnis „01“ bewertet. Eine entsprechende Umschlüsselung erfolgt in einem anderen Feld des Datensatzes, so dass auch das ursprüngliche Befragungsergebnis für die Auswertung erhalten bleibt. Auch die (negativen) Befragungsergebnisse „08“<sup>2)</sup> – „11“ werden als sichere

2) Personen in Haushalten von Angehörigen ausländischer Streitkräfte oder Diplomaten können im Einzelfall im Melderegister enthalten sein. Ein Nachweis von Melderegisterüberhängen ist daher bei Befragungsergebnis 08 nicht möglich.

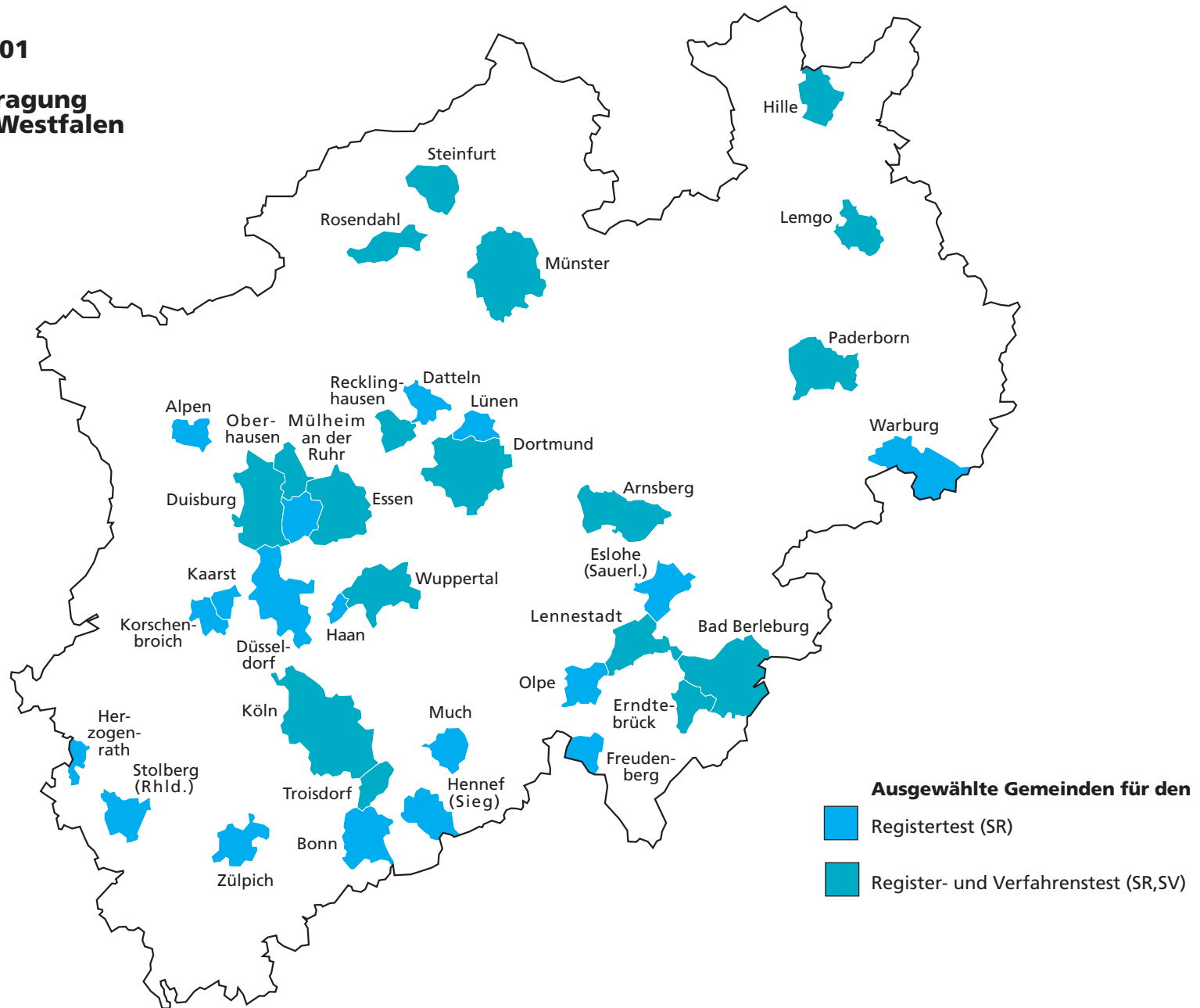
Feststellungen angesehen. Befragungsergebnisse „02“ und „03“ werden in Bezug auf die endgültige Bewertung entweder zu „01“ (Rücklauf eines ausgefüllten Erhebungsbogens) oder zu „04“ (Verweigerung) und kommen daher letztendlich nicht mehr vor. Es verbleiben damit die Befragungsergebnisse „04“ – „07“ als unsichere Feststellungen.

Sind der Name des Haushalts und die Zahl der Personen bekannt („05“, teilweise auch „04“), so ist es i. d. R. möglich, Personen des Melderegisters eindeutig dem jeweiligen Haushalt zuzuordnen und überzählige oder fehlende Personen auf Seiten des Melderegisters, zumindest deren Anzahl nach, zu ermitteln. Ist nur der Name bekannt („06“, teilweise auch „04“) wird ggf. die Anzahl der Personen gleichen Namens aus dem Melderegister akzeptiert. Es wird hierbei im Zweifel für das zu prüfende Melderegister entschieden, wodurch die Schätzung der Registerfehler konservativ erfolgt, d. h. im Zweifelsfall wird die Rate der Überhänge und Fehlbestände der Melderegister unterschätzt. Dies gilt insbesondere auch für den Fall, dass keine Person mit dem vom Interviewer festgestellten Namen im Melderegister enthalten ist. In diesem Fall kann ein Fehlbestand des Melderegisters nur für eine Person nachgewiesen werden, obwohl es sich auch um mehrere Personen handeln kann. Noch unsicherer wird die Feststellung bei Haushalten, für die kein Name ermittelt werden konnte. Ein Überhang des Melderegisters ist in diesem Fall nur nachzuweisen, wenn mehrere Indizien (Namen, Einzugsdatum, Zusammensetzung der verbleibenden Personen nach Geschlecht, Alter und Familienstand) dafür sprechen, dass einem Haushalt mit Befragungsergebnis „07“ mehrere Haushalte des Melderegisters gegenüberstehen. Der Nachweis eines Fehlbestands kann bei einem Befragungsergebnis „07“ nur für jeweils eine Person erfolgen, und auch nur dann, wenn das Melderegister keine noch nicht einem (anderen) Haushalt zugeordnete Person mehr enthält.

1) Zur Bewertung der Qualität von Einwohnermelderegistern vgl. Eppmann, Helmut: Die Einwohnermelderegister als Basis der amtlichen Einwohnerzahl, in: Statistische Rundschau Nordrhein-Westfalen Heft 8/1998, S. 407 ff. sowie Joggerst, Klaus: Bevölkerungszahlen aus Melderegistern und aus der Fortschreibung – ein Vergleich, in: Statistische Rundschau Nordrhein-Westfalen, Heft 12/1998, S. 623 ff.

# Zensusstest 2001

## Haushaltsbefragung in Nordrhein-Westfalen



Eine fehlende Befragung wird durch sog. Ersatzvornahmen aus dem Melderegister (fehlende Angaben aus der Befragung werden durch Melderegisterangaben der jeweils dem Verteilungslisteneintrag zugeordneten Personen ersetzt) ergänzt. Nach der Devise „Im Zweifel für das Melderegister“ erfolgt die Zuordnung bei unsicheren Befragungsergebnissen „großzügig“, d. h. es erfolgen eher zu viele als zu wenige Ersatzvornahmen, um dem Melderegister keine nicht nachweisbaren Fehler anzulasten. Dies führt zu einer konservativen Schätzung, d. h. zu einer Unterschätzung der Überhänge und Fehlbestände der Melderegister. Die resultierende Aussage wäre demnach „Mindestens xx % des Registerbestands sind Überhänge bzw. Fehlbestände“.

Neben der Ermittlung von Überhängen und Fehlbeständen der Melderegister ermöglicht die Haushaltebefragung auch einen Vergleich der Merkmale „Alter“, „Geschlecht“, „Familienstand“, „Staatsangehörigkeit“ sowie der Frage nach „alleiniger Wohnung“, „Haupt- oder Nebenwohnung“ mit den jeweiligen Angaben des Einwohnermelderegisters sowie der Merkmale „Beteiligung am Erwerbsleben“, „Stellung im Beruf“ und „Arbeitsort“ mit den entsprechenden Angaben der Register der Bundesanstalt für Arbeit<sup>3)</sup>.

#### **4 Die Haushaltebefragung als Maßstab zur Bewertung des Zensusmodells**

Neben der Überprüfung der Qualität bzw. der Statistikauglichkeit der für einen künftigen registergestützten Zensus eingesetzten Verwaltungsregister (Registertest) dient die Haushaltebefragung vor allem der Bewertung des im Rahmen des Zensus-tests erstmals eingesetzten neuen Verfahrens der Haushaltegenerierung. Aus Angaben der Einwohnermeldeämter

und einer schriftlich bei den Gebäudeeigentümern durchgeführten Gebäude- und Wohnungszählung werden Haushalte plausibel zusammengesetzt (Verfahrenstest). Wichtige Fragen in diesem Zusammenhang sind:

- Inwieweit erkennt das Verfahren der Haushaltegenerierung auf Adressebene Überhänge und Fehlbestände des Einwohnermelderegisters?
- Bei welchem Umfang wird beim künftigen Zensus eine Begehung vor Ort (Haushaltebefragung) erforderlich sein?
- Gibt die Haushaltegenerierung die Anzahl und die Struktur der Haushalte richtig wieder?
- Welche Merkmalsresultate differieren zwischen einer schriftlichen Befragung der Gebäudeeigentümer und einer Interviewerbefragung der Haushalte?

Bei den genannten Fragen bildet die Haushaltebefragung den Vergleichs- und Bewertungsmaßstab. Die Parameter der Haushaltegenerierung sind dabei so einzustellen, dass sowohl die Differenzen zur Haushaltebefragung (Qualität) als auch der Umfang der erforderlichen Begehungen vor Ort (Kosten) im vertretbaren Rahmen bleiben. Diese Aufgabe einer gleichzeitigen Qualitäts- und Kostenoptimierung kommt einer Gratwanderung gleich. Die Qualität eines registergestützten Zensus, gemessen an den Abweichungen zwischen Haushaltegenerierung und -befragung beim repräsentativ erhobenen Zensus-test soll sich an den Ergebnissen früherer Zensus, insbesondere an der Volkszählung 1987, messen lassen. Nimmt man die obere Grenze des 95 %-Konfidenzintervalls der 1987 in Nordrhein-Westfalen durchgeführten Hoch- und Fehlerrechnung des Vergleichs zwischen Volkszählung und Wiederholungs-befragung<sup>4)</sup> als Orientierungswert, so ergibt sich daraus die Forderung so-

wohl einer Überhangs- als auch einer Fehlbestandsrate von maximal 0,6 %. Weiterhin wird die Streuung von Überhangs- und Fehlbestandsraten zwischen den Gemeinden im Hinblick auf die Sicherstellung einheitlicher Ergebnisse auf Gemeindeebene zu analysieren sein. Ob dies unter vertretbarem Aufwand, insbesondere im Hinblick auf den Umfang der dann notwendigen Begehung vor Ort (Haushaltebefragung), zu realisieren ist, wird ein zentrales Ergebnis des Zensus-tests sein.

Die Wiedergabe der richtigen Haushaltsstrukturen stellte bei dem vom KOSIS-Verbund entwickelten und allein auf den Einwohnermelderegistern basierenden Haushaltegenerierungsverfahren<sup>5)</sup> ein Problem dar, das im Rahmen des Zensus-tests durch die Einbeziehung der Merkmale „Zahl der Wohnungen“ und „Name des Wohnungsinhabers“ gelöst werden soll<sup>6)</sup>. Inwieweit dieses Ziel – wiederum gemessen an den Ergebnissen der Haushaltebefragung – erreicht wird, wird ebenfalls eine wichtige Erkenntnis aus dem Zensus-test sein.

Individualvergleiche von Gebäude- und Wohnungsangaben zwischen einer schriftlichen Eigentümerbefragung und einer mittels Erhebungsbeauftragten durchgeführten Haushaltebefragung sind – zumindest für die amtliche Statistik – neu. Die sich daraus ergebenden methodischen Erkenntnisse werden noch zu analysieren sein, insbesondere im Hinblick auf die Vergleichbarkeit früherer mit künftigen Zensus und Gebäude- und Wohnungszählungen.

3) Die genannten Merkmale zur Erwerbssituation wurden nur in der Unterstichprobe für den sog. „Verfahrenstest“ erhoben. Gleiches gilt für das Merkmal „Art des überwiegenden Lebensunterhalts“. Bei Letzterem ist anhand der Ergebnisse des Zensus-tests zu prüfen, ob auch dieses Merkmal aus vorhandenen Registern weitgehend „konstruiert“ werden kann. – 4) Vgl. Eppmann, Helmut, Schäfer, Josef: Volkszählung 87 – Eine Untersuchung zur Qualität der Ergebnisse, in: Statistische Rundschau Nordrhein-Westfalen, Heft 4/1991. – 5) Vgl. von Klitzing, Friedrich: Haushaltegenerierungsverfahren 93 – Verfahren zur Generierung von Haushalten aus Einwohnerregisterdaten, KOSIS-Verbund, Nürnberg 1993. – 6) Vgl. Egbert, Henrik, Scharmer, Marco: Die Ermittlung des Wohnungsinhabers im Rahmen der Gebäude- und Wohnungszählung (GWZ) und deren Bedeutung für die Haushaltegenerierung, in: Statistische Analysen und Studien Nordrhein-Westfalen, Bd. 1 (2002), S. 10 ff. Einschränkend ist anzumerken, dass eine Haushaltegenerierung auf Wohnungsebene nur Wohnhaushalte und keine Wirtschaftshaushalte (wie dies bei früheren Volkszählungen und beim Mikrozensus der Fall war und ist) abbilden kann.

# Literatur zum Thema

## Allgemeine Aufsätze

- [1] Krug, Walter (2002): Strategien zur deutschen Volkszählung unter Beachtung internationaler Erfahrungen, in: Jahrbücher für Nationalökonomie und Statistik, Bd. 222/3, S. 337 – 345
- [2] Gabka, Dieter (2002): Volkszählungen – gestern und morgen, in: Statistische Monatshefte Mecklenburg-Vorpommern, Heft 04/2002, S. 96 – 115
- [3] Grohmann, Heinz (2000): Geschichte und Zukunft der Volkszählung in Deutschland, in: Berliner Statistik, Monatsschrift 7 – 12/00, S. 216 – 223
- [4] Grohmann, H./Shaner, H./Wiegert, R. (Hrsg.) (1999): Volkszählung 2001 – Von der traditionellen Volkszählung zum Registerzensus, Sonderheft 33 zum Allgemeinen Statistischen Archiv, Göttingen
- [5] Wiegert, Rolf (1999): Der Zensus – Ein Bericht zu Tradition und Moderne; Jahrbücher für Nationalökonomie und Statistik 1999, Bd. 219/5+6, S. 712 ff.
- [6] Statistisches Bundesamt (1998): Bericht der AG „Gemeinschaftsweiter Zensus 2001“. Wiesbaden

## Methodische und konzeptionelle Diskussion

- [7] Scharmer, Marco (2002): Haushaltstypisierung im Rahmen des registergestützten Zensus, in: Statistische Analysen und Studien Nordrhein-Westfalen, Heft 01/2002, S. 3 – 9
- [8] Werner, Joachim (2001): Der Zensus 2001 – Prüfung neuer Methoden als Alternative für eine Volkszählung, in: Baden-Württemberg in Wort und Zahl, Heft 11/2001, S. 545 – 561
- [9] Rompel, Heinz-Kurt (2001): Testerhebungen für einen registergestützten Zensus, in: Staat und Wirtschaft in Hessen, Heft 10/01, S. 322 – 330
- [10] Forster, Michael (2001): Die Zukunft der Volkszählung in Deutschland. Traditionelle Zählung oder registergestützter Zensus?, in: Statistische Analysen und Studien Nordrhein-Westfalen, Heft 04/2001, S. 12 – 24
- [11] Fürnrohr, Michael/Rimmelspacher, Birgit (2001): Testuntersuchungen zur Vorbereitung eines registergestützten Zensus, in: Bayern in Zahlen, Heft 01/2001, S. 13 – 18
- [12] Eichler, Ulrich (2000): Probleme der Verknüpfung personenbezogener Einzeldaten aus verschiedenen Registern, in: Allgemeines Statistisches Archiv, Heft 84, S. 83 – 93
- [13] Werner, Joachim (1999): Zensus 2001 – Modellentwürfe, in: Baden-Württemberg in Wort und Zahl, Heft 2/99, S. 103 ff.
- [14] Eppmann, Helmut, Köster, Gabriele und Rost, Reinhard (1998): Zensus 2001 – Stand der Methodendiskussion, in: Statistische Rundschau Nordrhein-Westfalen, Ausgabe 8/1998, S. 389 ff.
- [15] Joggerst, Klaus (1998): Bevölkerungszahlen aus Melderegistern und aus der Fortschreibung – ein Vergleich; Statistische Rundschau Nordrhein-Westfalen, Heft 12/98, S. 623 ff.
- [16] Eppmann, Helmut (1998): Die Einwohnerzahl als Basis der „amtlichen“ Einwohnerzahl, in: Statistische Rundschau Nordrhein-Westfalen, Ausgabe August 1998, S. 407 – 419
- [17] Statistisches Bundesamt (1995): Künftige Zensen in Deutschland. Ergebnisse der Untersuchungen der Arbeitsgruppe „Künftige Zensen“ zu Inhalt und Methode, Wiesbaden

## Zu einzelnen Bereichen des Zensusmodells

- [18] Egbert, Hendrik/Scharmer, Marco (2002): Die Ermittlung des Wohnungsinhabers im Rahmen der Gebäude- und Wohnungszählung (GWZ) und deren Bedeutung für die Haushaltegenerierung, in: Statistische Analysen und Studien Nordrhein-Westfalen, Heft 01/2002, S. 10 – 14
- [19] Fest, Manfred/Herzog, Petra (1999): Möglichkeiten einer Haushaltegenerierung unter Verwendung von Indizien aus Einwohnermelderegistern und wohnungsstatistischen Daten, in: Daten + Analysen, hrsg. vom LDS Brandenburg, 5/1999, S. 1 – 12
- [20] Fürnrohr, Michael/König, Markus (1999): Möglichkeiten einer Haushaltegenerierung im Rahmen der Zusammenführung von Einzeldaten aus Melderegistern mit primärstatistisch gewonnenen Wohnungsdaten, in: Bayern in Zahlen, hrsg. vom LfStaD BY, 4/1999, S. 161 – 166