

Möglichkeiten einer Haushaltegenerierung unter Verwendung von Indizien aus Einwohnermelderegistern und wohnungsstatistischen Daten

Einleitung

Die Struktur und Größe privater Haushalte ist sowohl für die öffentliche Planung als auch für die private Wirtschaft von großem Interesse. Angaben über Privathaushalte sind deshalb ständiger Inhalt primärstatistischer Erhebungen. Datenquellen waren hier bisher Volkszählungen sowie die 1 %ige Stichprobe des jährlichen Mikrozensus. Mit der Abkehr von einer primärstatistischen Vollerhebung stellt sich für die Zukunft die Frage, welche methodischen Ansätze zu Daten analog einem Zensus führen können.

Ausgangspunkt des hier vom Landesamt für Datenverarbeitung und Statistik Brandenburg (LDS BB) vorgestellten und weiterzuentwickelnden Haushaltegenerierungsverfahrens ist der Ansatz, dass Veränderungen im Generierungsablauf sowie die Nutzung weiter gehender Informationen (Indizien) aus den Melderegistern und die Kombination dieser Indizien mit wohnungsstatistischen Informationen die bei der Generierung vorhandenen Problembereiche abbauen und die Haushaltsstrukturen (Personenverbände) realitätsnäher abbilden können.

Zur Evaluation des Verfahrens werden die Ergebnisse der Generierung mit den erhobenen Einzeldaten des Mikrozensus verglichen, da gegenwärtig keine verlässlichere Datenbasis zur Kontrolle zur Verfügung steht.

Mit Testdaten einer kreisfreien Stadt wurden vier Szenarien mit unterschiedlicher Datenlage gewählt, um aufzuzeigen, welche Voraussetzungen erforderlich sind, das Haushaltegenerierungsverfahren sowohl in einem Zensus als auch als grundsätzliche Methode für die Abbildung realitätsnaher Haushaltsstrukturen einzubringen.

Lösungsansatz

Bereits **seit den 80er Jahren sind Bestrebungen** vorhanden, die **Datenbestände der Melderegister zu nutzen**, um aktuelle, kleinräumig gegliederte Angaben über das Zusammenleben der Personen in Haushalten zu erhalten. Mit Förderung des Bundesbauministeriums wurde im Auftrage der ehemaligen Bundesforschungsanstalt für Landeskunde und Raumordnung (BfLR) ein Verfahren entwickelt, das die unter der gleichen Adresse im Melderegister erfassten Personen in einem mehrstufigen iterativen Prozess zu Personenverbänden, die gemeinsam wohnen, gruppiert.

Der in der amtlichen Statistik bisher angewandte Haushaltsbegriff als eine Personengruppe, die gemeinsam wohnt und wirtschaftet, erhält damit einen neuen Inhalt.

Das **vom KOSIS-Verbund¹⁾ entwickelte Verfahren** wurde bereits in einigen deutschen Großstädten für kommunalstatistische Zwecke, u. a. zur Wohnungsmarktbeobachtung mit Erfolg eingesetzt. Dabei wurde anfangs auch der Gedanke verfolgt, neben Indizien aus dem Melderegister die Zahl der Wohnungen pro Gebäude aus einer Gebäudedatei zu verwenden. Dieser Ansatz wurde später nicht weiter verfolgt, weil die Vorhaltung von kommunalen Gebäudedateien durch das Volkszählungsgesetz 1987 enorm erschwert wurde²⁾.

Ausgangspunkt des hier vorgestellten Haushaltegenerierungsverfahrens ist der Ansatz, zusätzliche Indizien und Daten zu nutzen um den Generierungserfolg zu verbessern und die Haushaltsstrukturen (Personenverbände) so realitätsnah wie möglich abzubilden.

1) KOSIS-Verbund: Kommunales Statistisches Informationssystem

2) Haushaltsdaten aus dem Melderegister (Friedrich v. Klitzing/Hanno Ochsenburg in Stadtforschung und Statistik/95)

Mit den bisherigen Haushalgenerierungsverfahren, so auch dem uns zur Verfügung stehenden „**HHGEN 93**“, das inzwischen zum „**HHGEN 98**“ weiterentwickelt wurde, wurden bisher lediglich personenbezogene Einzeldaten des Einwohnermelderegisters (EMR) ohne Indizien zur Wohnung für die Darstellung von Zusammenhängen zwischen den unter einer Adresse gemeldeten Personen zur Bildung von Haushalten genutzt. Eine wesentliche Schwäche dieses Verfahrens ist, dass zu viele Haushalte gebildet werden, da hierbei „aus Personen, die nicht in einen durch die Algorithmen definierten Haushaltzusammenhang gebracht werden können, quasi automatisch Einpersonenhaushalte gemacht werden.“³⁾

Wir gingen in einem 1. Ansatz - das war Ende 1997 - davon aus, dass das **Verfahren auch für eine Siedlungsstruktur mit verschiedenen Meldebehörden zu testen** sei, bei der die für die Weiterentwicklung des Verfahrens zu lösenden Probleme auch mit Blick auf einen künftigen Zensus sichtbar würden. (Im Land BB gibt es Landkreise mit bis zu 18 Einwohnermeldeämtern.) Wir waren deshalb daran interessiert, einen Brandenburger Landkreis in die Tests einzubeziehen.

Die jedoch dabei **auf tretenden Hürden**, wie

- ❑ unterschiedliche Soft- und Hardware,
- ❑ Datenübermittlungsverordnung lt. Melderecht,
- ❑ Finanzierung der Gewinnung der Eingabedatensätze für das Generierungsverfahren

waren in dieser Phase nicht zu überwinden, so dass wir auf eine kreisfreie Stadt als Testpartner zurückgriffen. In Abstimmung mit dem Landesbeauftragten für den Datenschutz konnten so ca. 1 000 Datensätze des Einwohnermeldeamtes in den Test einbezogen werden.

Ausgewählt wurden die Datensätze der Personen, die im Mikrozensus erfasst werden, um somit, wie bereits zuvor im Statistisches Landesamt Hamburg praktiziert⁴⁾, eine Referenzdatei zur Bewertung der Generierungsergebnisse zur Verfügung zu haben.

Zur Ausgangssituation

Mit der 1996 entstandenen Situation,

- ❑ die Bundesregierung hatte sich in Vorbereitung eines Zensus 2001 **gegen eine erneute Totalerhebung als Volkszählung** ausgesprochen,
- ❑ vom Bund wurde eine **Variante eines Zensusmodells** favorisiert, mit der zwar der EU-Datenbedarf für 2001 gedeckt, aber nicht der Zensus-typische Datenkranz kleinräumig strukturiert auf der Basis eines kombinierfähigen Einzeldatensatzes zur Verfügung gestellt und auch das statistische Gesamtsystem nicht neu justiert werden konnte, wurden im Rahmen des Zensus-Ländermodells Überlegungen angestellt, wie zum einen
 - **Unplausibilitäten** zur Bereitstellung einer amtlichen Bevölkerungszahl **zu bereinigen** und
 - neben validen Einwohnerzahlen **auch kleinräumig strukturierte Haushalts(Personen)-zusammenhänge** zu erhalten wären.

Im Rahmen der Erarbeitung eines Berichtes für die Innenministerkonferenz zur Erarbeitung von Modellen für einen registergestützten Zensus im Zusammenhang mit dem **EU-Vorhaben „VZ 2001“** war beim StBA eine Projektgruppe tätig, die prüfen sollte, wie das Haushalgenerierungsverfahren in einen registergestützten Zensus einzubinden sei. Nachdem dieser Weg - mit Blick auf einen Zensus 2001 - vom Statistischen Bundesamt (StBA) nicht weiter verfolgt wurde, konzentrierte sich das LDS BB auf die Weiterentwicklung des eigentlichen Generierungsverfahrens als statistische Methode und seine Einbindung in einen registergestützten Zensus.

Unbestritten ist, dass gegenwärtig die **Bereinigung der Einwohnermelderegister** und die Validität der für einen Zensus aus dem Register zur Verfügung stehenden Datensätze einen Schwerpunkt darstellen. Zur Aufdeckung von Unplausibilitäten im Melderegister (Über- und/oder Untererfassungen) sind deshalb geeignete Verfahren zu entwickeln. Diesen Schritt soll, wie im Ländermodell beschrieben und vom Bayerischen Landesamt für Statistik und Datenverarbeitung jetzt publiziert, das als „**Integriertes Zusammenführungs- und Haushalgenerierungsverfahren**“ bezeichnete Modul vollziehen.⁵⁾

3) Bericht der Arbeitsgruppe des StBA „Gemeinschaftsweiter Zensus 2001“ (unveröffentlicht)

4) Bach, H.-J. (Haushaltsgenerierung: Eine neue Quelle zur Gewinnung von Haushaltsdaten, Hamburg in Zahlen 5/98)

5) Dr. Fürnrohr/König: Möglichkeiten einer Haushalgenerierung im Rahmen der Zusammenführung von Einzeldaten aus Melderegistern mit primärstatistisch gewonnenen Wohnungsdaten, Bayern in Zahlen 4/99

Nach diesen Überlegungen werden nach der Zusammenführung der Personendatensätze des Einwohnermelderegisters und der Daten einer Gebäude- und Wohnungszählung (GWZ) die Personen in Mehrfamilienhäusern um den Namen des Wohnungsinhabers unter Beibehaltung der Algorithmen des KOSIS-Verfahrens generiert. Wir gehen davon aus, dass die Zusammenführung der Personendatensätze mit den Datensätzen der GWZ einen unverzichtbaren notwendigen Schritt zur Bereinigung der Melderegisterdaten im Vorfeld der Haushalgenerierung darstellt und die Generierung der Personenhaushalte über den gesamten Datenbestand je Gebäude mit vorplausibilisierten Datensätzen durchzuführen ist.

Der **Schwerpunkt unserer Arbeit** liegt deshalb auf dem Abbau der Problembereiche der Generierungsstufen und auf der Programmierung entsprechender Algorithmen. Dem Verfahren gaben wir zur Verdeutlichung den Arbeitstitel „**Haushalgenerierung-Statistik**“ (HHGEN-Statistik).

Für einen ersten Generierungsversuch bedienten wir uns des vom KOSIS-Verbund für Testzwecke zur Verfügung gestellten HHGEN '93 und generierten damit ca. 1 000 Personendatensätze des Einwohnermeldeamtes Cottbus mit den bisher üblichen Indizien des Einwohnermelderegisters.

Um einen möglichst realitätsnahen Abgleich mit den Ergebnissen des Mikrozensus zu gewährleisten, wurde ein historischer Melderegisterauszug (Stichtag: 30.04.1997) verwendet.

Die **Auswertung** (siehe Tabelle) ergab:

- Einpersonenhaushalte wurden überhöht gebildet.
 - Bei Haushalten mit 2 bis 6 Personen ist die Haushaltsstruktur entsprechend verschoben.
 - Bei der Generierung werden auch falsche Personenzusammenhänge erzeugt.
- ca. 66 Prozent aller falsch gebildeten Haushalte waren fehlerhafte Einpersonenhaushalte, lediglich 72 Prozent der generierten Haushalte entsprachen in der personellen Zusammensetzung dem Ergebnis des Mikrozensus '97

- bei den fehlerhaft generierten Einpersonenhaushalten waren
- rund 23 Prozent nicht eheliche Lebensgemeinschaften und
 - rund 32 Prozent nicht zugeordnete erwachsene Kinder.

Vergleich der Haushalte im Mikrozensus 1997 mit den generierten Haushalten

Haushaltstyp	Mikrozensus 1997	HHGEN 1993	Abweichung HHGEN '93 zu MZ '97
	Anzahl		Prozent
Haushalte insgesamt	446	504	+ 13,0
davon mit			
1 Person	133	211	+ 58,6
2 Personen	162	163	+ 0,6
3 Personen	91	84	- 7,7
4 Personen	49	36	- 26,5
5 Personen	10	10	-
6 Personen	1	-	- 100,0

Diese ersten Ergebnisse führten zu Überlegungen, wie das Verfahren programmtechnisch weiter zu entwickeln sei und sowohl das Allein- als auch das Zusammenwohnen realitätsnäher abgebildet werden kann.

Dabei gingen wir von folgenden Prämissen⁶⁾ aus:

1. Die Haushalgenerierung ist eine praktikable Alternative zur Totalerhebung um Daten über kleinräumig strukturierte Personenhaushalte zu erhalten.
2. Mit zusätzlichen Indizien aus dem EMR und wohnungsstatistischen Daten ist die realitätsnahe Abbildung der zusammenwohnenden Personen möglich.
3. Die Generierung wird erfolgreicher, wenn alle Personendatensätze einer Adresse primär nach dem Generierungsgrund betrachtet werden.

6) An dieser Stelle soll nochmals verdeutlicht werden, dass wir die Haushalgenerierung aus Indizien der Melderegister und wohnungsstatistischen Daten zwar im Rahmen des Einstieges Deutschlands in einen registergestützten Zensus eingebettet sehen, das Verfahren selbst jedoch auch für die dann „normalen“ Zeiten eines registergestützten Zensus zu entwickeln ist, bei dem wohnungsstatistische Daten ggf. aus einem Gebäude- und Wohnungsregister zu entnehmen sind. Auf dem Weg dahin sind vielfältige Probleme im Verfahren selbst als auch bei den Rahmenbedingungen zu klären. Im gegenwärtig diskutierten Zensus-Ländermodell als Einstieg in eine Registerzählung wird, da Gebäude- und Wohnungsregister noch nicht existieren, u.a. auf eine GWZ als Datenquelle zurückgegriffen.

Die **Zusammenführung der Daten einer Gebäude- und Wohnungszählung und der Melderegister** dient dabei nach unserem Verständnis

- der Aufdeckung von **Unplausibilitäten** in einem Melderegisterabzug zur Bereitstellung einer amtlichen Bevölkerungszahl sowie
- der Bereitstellung **vorplausibilisierter Datensätze** zur Haushalgenerierung.

Die Haushalgenerierung hat

- die **Ermittlung der** auf das Gebäude bezogenen **Personenhaushalte** (siehe Leitlinien zum EU-Vorhaben),⁷⁾
- die **Feststellung der Wohnungsbelegung** sowie
- die **Aufdeckung weiterer Unplausibilitäten** bei gebildeten Haushaltsstrukturen nach der Generierung (während der Testphase)

als Zielstellung.

Aus den Einzeldatensätzen der Melderegister werden folgende **Indizien als Tatbestand für die Zusammengehörigkeit von Personen herangezogen:**

1. Indizien, die bereits durch Datenübermittlungsverordnung (DÜVO) den Statistischen Landesämtern (StLÄ) übermittelt werden können:
 - Verzeigerung zwischen Ehegatten sowie zwischen Kindern und deren Elternteilen bzw. Erziehungsberechtigten (Kernhaushalt),
 - Personen mit Namensübereinstimmungen zwischen ihren Familiennamen, Geburtsnamen und früheren Namen in verschiedenster Kombination,
 - Personen mit übereinstimmender früherer Wohnadresse,
 - Personen mit übereinstimmendem Datum der Anmeldung an der gegenwärtigen Wohnadresse,
 - Personen mit übereinstimmendem Datum des Zuges in die Gemeinde,
 - Personen, deren demographische Merkmalskonstellationen bezüglich Alter, Geschlecht, Familienstand bestimmte familiäre Beziehungen nahe legen oder ausschließen

2. Indizien, die für den Generierungsverlauf vorteilhaft sind, aber derzeit noch nicht flächendeckend verfügbar sind:
 - Name des Wohnungsgebers (wohnt bei ...),
 - Etagen-Nr. der gemeldeten Person

Zu den **wohnungsstatistischen Daten** aus einer GWZ bzw. aus einem aktuellen Gebäude- und Wohnungsregister (GWR), die für die Generierung herangezogen werden, zählen:

1. Daten, die gegenwärtig Bestandteil der Generierung sind:
 - Anzahl der Wohnungen (WO) und lfd.-Nr. der Wohnung je Gebäude,
 - Nutzung der WO (bewohnt, leer stehend),
 - Name des Wohnungsinhabers (nur bei einer GWZ verfügbar)
2. Zur Plausibilisierung bzw. zur endgültigen Schaffung des Bezuges zwischen Haushalt und Wohnung (Algorithmen sind noch zu erarbeiten):
 - Anzahl der Räume je Wohnung,
 - Fläche der Wohnung,
 - Etagen-Nr. der Wohnung⁸⁾,
 - Einzugsdatum des Wohnungsinhabers,
 - Zahl der in der Wohnung lebenden Personen

Nicht betrachtet wird gegenwärtig die Einbeziehung einer Wohnungs-Nr. aus dem Einwohnermelderegister. Hier sind die erforderlichen rechtlichen Rahmenbedingungen in Deutschland gegenwärtig nicht gegeben. Darüber hinaus tragen Berichte über fehlgeschlagene Versuche in skandinavischen Ländern nicht dazu bei, diesen Schritt vorzubereiten.

Im Zuge ergebnisoffener Tests und dem damit verbundenen Wachsen eigener Erkenntnisse sollte dieser Weg jedoch nicht verschlossen bleiben.

Zum Generierungsverfahren

Wie bereits hervorgehoben, war der Anlass für die Weiterentwicklung eines Haushalgenerierungsverfahrens die vor einem politischen Hintergrund geführte Diskussion zur Erarbeitung eines alternativen Zensusmodells in Deutschland.

4) Leitlinien für das gemeinschaftliche Programm der Volks- und Wohnungszählungen im Jahr 2001 vom November 1997; Gemeinsames Wohnen: aggregierte Anzahl der eine Wohnungseinheit bewohnenden Personen
 5) Nur dann als Zuordnungsindiz verwendbar, wenn im Melderegister je Person eine Etagenangabe vorliegt.

Mit Blick auf eine über das Jahr 2001 hinausreichende Lösung wurde eine Neuprogrammierung in Angriff genommen, deren erste vorläufige Programmversion vorliegt. Es handelt sich dabei um ein **BATCH-Verfahren**, das in der Programmiersprache NATURAL als Basis für eine mögliche Verbundlösung realisiert wurde.

Im Gegensatz zum KOSIS-Verfahren HHGEN '93, bei dem der anliegende Datensatz nach möglichen Generierungen durch die Stufen geführt wird, erfolgt im **Verfahren HHGEN-Statistik** die Generierung primär nach dem Generierungsgrund. Dabei werden die Datensätze nach der Bearbeitung in einem Generierungsschritt nicht als beendet betrachtet, sondern nach weiteren Generierungsmöglichkeiten überprüft.

Für die Testphase des Verfahrens haben wir zur Einflussnahme auf den Generierungserfolg **Steuerparameter** zugelassen. Es handelt sich hier aus unserer bisherigen Erkenntnis um:

- den Erhebungstichtag,
- Altersdifferenzen,
- Einzugsspanne,
- Möglichkeit der Generierung mit/ohne Wohnungsangaben

Für die Anwender unseres Testprogrammes ist die Datensatz-Beschreibung Bestandteil der **Dokumentation**. Bei Bedarf bieten wir auch Mitwirkung bei der Erarbeitung des Vorprogrammes zur Erzeugung des Eingabedatensatzes an.

Zur Verdeutlichung der **Einflussgrößen** bestimmter Indizien und Zusatzinformationen auf den Generierungserfolg wurden **verschiedene Generierungsszenarien** getestet, wie:

- mit Zusatzinformationen aus dem Einwohnermelderegister (Name des Wohnungsgebers und Etagen-Nr.);
- mit Zusatzinformationen aus dem Einwohnermelderegister und wohnungsstatistischen Daten aus einer GWZ (Anzahl der Wohnungen, Name des Wohnungsinhabers);
- mit Zusatzinformationen aus dem Einwohnermelderegister und wohnungsstatistischen Daten aus einem Gebäude- und Wohnungsregister (GWR) (Anzahl der Wohnungen).

Im Zuge weiterer Tests, durch die Einbeziehung umfangreicherer Testdatenbestände, sollte sichtbar werden, unter welchen Bedingungen und mit welchen Ergebnissen dieses Verfahren als Alternative für eine Totalerhebung geeignet ist.

Ausgangspunkt des Verfahrens ist die adressweise Verknüpfung paariger Datensätze von Personen- und GWZ- bzw. GWR-Dateien über eine Straßenschlüsseldatei.

Personendatei:

Stichtagsbezogener Melderegisterabzug lt. Eingabedatensatz (einschließlich Zusatzinformation: Etagen-Nr., Name des Wohnungsgebers)

Gebäude- und Wohnungsdatei:

aus GWZ oder kommunalen GWR

- Wohnungen je Gebäude
- Name des Wohnungsinhabers (nur bei einer GWZ vorhanden)

Die gegenwärtige Situation - Vorbereitung eines registergestützten Zensus ohne Vorhandensein dafür geeigneter Register in ausreichender Qualität - erfordert notwendige vorbereitende **Maßnahmen zur Prüfung und Klärung von Unplausibilitäten** wie:

- die Aufdeckung von Mehrfachfällen im Melderegister (Personen mit mehreren Hauptwohnsitzen) länderübergreifend,
- die Aufdeckung von Personen ohne Anmeldung mit Hauptwohnsitz sowie
- die Identifikation des Wohnungsinhabernamens im Melderegisterabzug der Adresse.

Diese Maßnahmen laufen nach unserem Verständnis im Vorfeld der eigentlichen Haushaltegenerierung, sind jedoch wegen der gegenwärtigen Datenlage Bestandteil eines Modells für den Einstieg in einen registergestützten Zensus.

In einer Haushaltegenerierungsvorstufe werden durch Nutzung der im Melderegister vorhandenen unmittelbaren Verwandtschaftsbeziehungen (**Verzeigerungen**) die so genannten „**Kernhaushalte**“ gebildet. Die Bezeichnung „Kernhaushalt“ hat sich als eigenständiger Begriff bei der Gewinnung von Haushaltsdaten aus Melderegistern herausgebildet.

Einen Kernhaushalt bilden alle Personen, die untereinander über Ehe- oder Elternteil-Kind-Beziehung verkettet (verzeigert) und an der gleichen Adresse gemeldet sind.⁹⁾

Bei über 50 Prozent der Mehrpersonenhaushalte in Deutschland handelt es sich bereits um Kernhaushalte. Weiterhin beinhaltet diese Vorstufe die Bestimmung der Stellung jeder Person im Kernhaushalt.

9) Haushaltsdaten aus dem Melderegister (Friedrich v. Klitzing/Hanno Ochsenberg, Stadtforschung und Statistik 1/95)

Das Haushalgenerierungsverfahren führt nun in **aufeinander folgenden Generierungsschritten** Kernhaushalte zu größeren Personenverbänden zusammen mit dem Ziel, sukzessive eine Annäherung an die tatsächliche Haushaltsstruktur zu erreichen. In das Verfahren gehen nur Daten von Personen ein, die zur Bevölkerung in Haushalten gehören. Das EMR führt jedoch auch Personen ohne eigene Haushaltsführung. Deshalb müssen vor der eigentlichen Generierung die Personen in Anstalten und Heimen einer besonderen Betrachtung unterzogen werden.

Die einzelnen **Generierungsschritte** sind auf haushaltsstrukturelle Inhalte ausgerichtet. Jeder Generierungsschritt ist einer bestimmten Art von Personen-Beziehungen gewidmet, die in Haushalten häufig vorkommen. In jedem Schritt werden also alle Personen einer Adresse systematisch nach der Bildung von Personenbeziehungen geprüft.

Schritt 1:

Dient der Zusammenführung von Personen aufgrund der eindeutigen, im Melderegister zur Verfügung stehenden Information zum Wohnungsgeber (DSMeld-Feld 1212). Über den Familiennamen des Wohnungsgebers, der nicht minderjährig sein darf, ist es möglich, Haushaltszusammenhänge von zusammenlebenden Personen zu bilden. Dabei bereits verzeigerte Kinder werden mit berücksichtigt.
(**Zuordnungsschritt: W**)¹⁰⁾

Schritt 2:

Alle minderjährigen Kinder einer Adresse werden überprüft, ob anhand von Verzeigerungen im Vorprogramm eine Zuordnung zur Mutter, oder, wenn an der Adresse nicht vorhanden, zum Vater erfolgte. Diese Prüfung macht sich notwendig, um fehlende Verzeigerungen minderjähriger Kinder vom Melderegister zu erkennen und zu beheben.

Wichtigstes Kriterium für eine mögliche Zusammenführung ist die Gleichheit des Familiennamens zwischen minderjährigem Kind und Elternteil bzw. auch Großelternteil, aber auch Geburtsnamen und frühere Familiennamen sowie Altersdifferenzen zwischen minderjährigem Kind und Elternteil werden beachtet. Weiterhin wird versucht, minderjährige Kinder über einen Geschwisterzusammenhang einem Personenverband zuzuordnen.

Nichtgenerierbare minderjährige Kinder werden extra ausgewiesen und deuten i. d. R. auf Fehler im Melderegister (Karteileichen) hin.

(**Zuordnungsschritt: M, V, G, S**)

Schritt 3:

Verzeigerungen zwischen minderjährigem Kind und nicht ehelichem Vater werden im Vorprogramm kenntlich gemacht. In diesem Schritt erfolgt die Zusammenführung beider Elternteile zu einem nicht ehelichen Paar. Steht aus einer Gebäude- und Wohnungszählung die Information zur Verfügung, dass beide Partner Wohnungsinhaber verschiedener Wohnungen an der gleichen Adresse sind, erfolgt keine Zusammenführung.
(**Zuordnungsschritt: U**)

Schritt 4:

Durch Fehler in den Einzeldatensätzen der Melderegister ist es möglich, dass Ehepaare nicht verzeigert sind. Solche Verzeigerungsdefizite werden über Namensabgleiche unter Berücksichtigung des Familienstandes, des Geschlechts und des Datums der Familienstandsänderung ausgeglichen und die Ehepartner zusammengeführt.

(**Zuordnungsschritt: E, L**)

Nach Durchlaufen dieser vier Generierungsschritte erfolgt ein **Vergleich der generierten Haushalte** mit der aus einer Gebäude- und Wohnungszählung bzw. einem Gebäude- und Wohnungsregister festgestellten Zahl der bewohnten Wohnungen je Gebäude:

- Entspricht die Zahl der bewohnten Wohnungen der Zahl der Haushalte wird davon ausgegangen, dass die Haushaltszusammenhänge bereits korrekt abgebildet wurden. Die Generierung wird an dieser Adresse beendet.
- Ist die Anzahl der Wohnungen größer als die bisher generierten Haushalte könnte dies ein Indiz für eine Untererfassung im Melderegister oder einen Fehler in den erhobenen Daten der Gebäude- und Wohnungszählung sein.
- Ist die Anzahl der gebildeten Haushalte größer als die Anzahl der bewohnten Wohnungen, wird in folgenden Generierungsschritten versucht, Personen zu Haushalten zusammenzuführen.

Nach jedem weiteren Generierungsschritt wird die **Anzahl der generierten Haushalte mit der Anzahl der bewohnten Wohnungen abgeglichen**. Bei erreichter Gleichheit wird die Bearbeitung der Adresse beendet. Generell gilt für jeden Generierungsschritt, dass keine Personenverbände gebildet werden, deren Personen Wohnungsinhaber verschiedener Wohnungen sind.

10) Merkmal des jeweiligen Zuordnungsschrittes, in der die Person dem Haushalt zugeordnet wurde.

Schritt 5:

Da erwachsene Kinder im Melderegister nicht mit ihren Erziehungsberechtigten verzeigert sind, muß das Zusammenleben erwachsener Kinder (die auch bereits eigene Nachkommen haben können) mit der elterlichen Familie anhand anderer Indizien erkannt werden, wie

- Namensgleichheiten (Familiennamen oder Geburtsnamen oder frühere Familiennamen),
- Altersdifferenzen zwischen erwachsenem Kind und Elternteil bzw. Großelternanteil,
- das Einzugsdatums in die jetzige Wohnung. (Inwieweit die Einbeziehung des Einzugsdatums als Indiz des Melderegisters notwendige Zusammenführungen erschwert, müssen weitere Untersuchungen zeigen.)

(Zuordnungsschritt: K, O)

Schritt 6:

Dieser Schritt dient der Zusammenführung allein stehender Elternteile zum Haushalt ihrer Kinder. Über Namensgleichheiten, Beachtung von Altersdifferenzen und unter Berücksichtigung des Einzugsdatums in die gemeinsame Wohnung erfolgt die Haushaltsbildung.

(Zuordnungsschritt: A)

Schritt 7:

Für jede erwachsene Einzelperson einer Adresse wird geprüft, ob diese mit einer erwachsenen Geschwisterperson zu einem gemeinsamen Haushalt zusammengeführt werden kann. Über Namensgleichheiten und Beachtung von Altersdifferenzen erfolgt die Generierung. Eine Zusammenführung von Brüdern erfolgt nur, wenn eine gemeinsame Mobilität vorliegt.

(Zuordnungsschritt: B)

Schritt 8:

Dieser Schritt dient der Suche und Zusammenführung nicht ehelicher Paare. Bei Vorliegen bestimmter Altersdifferenzen in Abhängigkeit vom Geschlecht der Partner, gleichem Einzugsdatum oder gleicher früherer Herkunftsadresse werden Personen (mit oder ohne Nachkommen) zu einer nicht ehelichen Lebensgemeinschaft generiert.

(Zuordnungsschritt: N)

Schritt 9:

Alle Einzelpersonen jeder Adresse werden überprüft, ob sie mit anderen Personen am gleichen Tag an der jetzigen Adresse eingezogen sind und/oder eine gemeinsame Vorwohnung hatten. Bei Erfüllung dieser Bedingungen erfolgt die Zusammenführung zu einem gemeinsamen Haushalt. Um dadurch nicht übergroße Haushalte entstehen zu lassen wird die Personenzahl der so gebildeten Haushalte auf maximal 8 Personen begrenzt.

(Zuordnungsschritt: R)

Schritt 10:

Dieser Generierungsschritt kommt nur zur Anwendung, wenn aus einer Gebäude- und Wohnungszählung oder einem Gebäude- und Wohnungsregister eine Wohnungsanzahl je Gebäude als Vergleichsgröße herangezogen werden kann. Nach Ausschöpfen aller Möglichkeiten anhand von Indizien des EMR Personenbeziehungen zu bilden, erfolgt abschließend eine rein formale Zusammenführung von Einzelpersonen anhand von geschlechtsabhängigen Altersdifferenzen zu Lebensgemeinschaften.

(Zuordnungsschritt: Z, Y)

Für die **Evaluation des Verfahrens** sind nach der Generierung und Ergebnisdarstellung **Plausibilitätsprüfungen** des gebildeten Haushalt-Personenzusammenhangs über

- Haushaltstyp,
- Wohnungstyp und
- Personenzahl

durchzuführen.

Wir gehen davon aus, dass nach der Verarbeitung einer umfangreicheren Anzahl von Datensätzen weiterer Handlungsbedarf für die **Gestaltung der Generierungsalgorithmen** ersichtlich wird.

Zur Ergebnisinterpretation

Dem gegenwärtigen Stand der Diskussion folgend, wurden mit der vorliegenden Programmversion von HHGEN-Statistik **vier Szenarien getestet**. Damit soll - allerdings auf der Grundlage der zur Verfügung stehenden 1 000 Testdatensätze - der Einfluss einzelner Indizien bzw. wohnungsstatistischer Daten auf die Qualität der Ergebnisse der Haushaltegenerierung aufgezeigt werden.

Es soll nochmals verdeutlicht werden, dass mit den für die Tests zur Verfügung stehenden Daten eine weit gehende Optimierung des Verfahrens nicht möglich war und zur **Qualifizierung und weiteren Bewertung des Verfahrens** umfangreichere Tests erforderlich sind. Das gilt insbesondere für die Problembereiche des Haushalgenerierungsverfahrens wie

- die Zusammenführung nicht ehelicher Lebensgemeinschaften (bisher über vier Generierungsschritte gelöst) und
- die Zuordnung erwachsener Kinder zum elterlichen Haushalt.

Es ist zu erwarten, dass mit der Generierung größerer Personenzahlen, insbesondere in Ballungsgebieten, die gegenwärtigen Grenzen des Generierungsverfahrens deutlicher zu Tage treten und damit gleichzeitig den Handlungsbedarf für die weitere Gestaltung des Verfahrens aufzeigen.

Die **Ergebnisse der vorliegenden vier Testszenarien** beinhalten die Generierung mit unterschiedlicher Datenlage:

1. Ohne GWZ, ohne GWR, ohne Melderegister-Zusatz

- ohne Wohnungszahl
- ohne Wohnungsinhaber
- ohne zusätzliche Angaben aus dem Melderegister

2. Ohne GWZ, ohne GWR, mit Melderegister-Zusatz

- ohne Wohnungszahl
- ohne Wohnungsinhaber
- mit zusätzlichen Angaben aus dem Melderegister

3. Mit GWZ, mit Melderegister-Zusatz

- mit Wohnungszahl
- mit Wohnungsinhaber
- mit zusätzlichen Angaben aus dem Melderegister

4. Mit GWR, mit Melderegister-Zusatz

- mit Wohnungszahl
- ohne Wohnungsinhaber
- mit zusätzlichen Angaben aus dem Melderegister

Da gegenwärtig keine andere Datenbasis zur Kontrolle der tatsächlichen Haushaltszusammenhänge zur Verfügung steht, wurden, wie bereits angeführt, die Generierungsergebnisse mit den Datensätzen des aktuellen Mikrozensus verglichen und bewertet.

Zu den Testszenarien

Auswertung der Ergebnisse HHGEN-Statistik im Vergleich zum Mikrozensus

Haushaltstyp	MZ '98	Szenario 1 und 2 ohne Wohnungszahl, ohne Wohnungsinhaber						Szenario 3 mit Wohnungszahl, mit Wohnungsinhaber			Szenario 4 mit Wohnungszahl, ohne Wohnungsinhaber		
		ohne zusätzliche Angaben aus EMR			mit zusätzlichen Angaben aus EMR			mit zusätzlichen Angaben aus EMR					
		HH	HH	richtig ¹⁾	Prozent	HH	richtig ¹⁾	Prozent	HH	richtig ¹⁾	Prozent	HH	richtig ¹⁾
Haushalte insgesamt	478	490	434	90,8	480	447	93,5	480	473	99,0	478	455	95,2
davon mit													
1 Person	138	163	126	91,3	148	127	92,0	140	136	98,6	139	128	92,8
2 Personen	175	162	153	87,4	164	159	90,9	177	174	99,4	175	167	95,4
3 Personen	95	95	89	93,7	96	91	95,8	93	93	97,9	92	90	94,7
4 Personen	58	57	55	94,8	59	58	100,0	58	58	100,0	60	58	100,0
5 Personen	9	10	8	88,9	10	9	100,0	9	9	100,0	9	9	100,0
6 Personen	3	3	3	100,0	3	3	100,0	3	3	100,0	3	3	100,0

1) gemessen am MZ-Ergebnis (MZ '98)

Hauptergebnisse der Testszenarien

Szenario 1: (ohne GWZ, ohne GWR, ohne Melderegister-Zusatzinfo)

Ohne Einbeziehung einer Wohnungszahl, ohne Information des Wohnungsinhabers und ohne zusätzliche Informationen aus dem Melderegister (Etagen-Angabe, Information des Wohnungsgebers) konnten 90,8 Prozent der im Mikrozensus '98 vorhandenen Haushalte in der richtigen personellen Zusammensetzung durch das Generierungsverfahren dargestellt werden. Nicht generiert werden konnten Personen, die lt. Mikrozensus '98 in nicht ehelicher Lebensgemeinschaft oder in Wohnungsgemeinschaft zusammenlebten, aber keinerlei Indizien im EMR für ein Zusammenleben aufwiesen (kein gemeinsames Einzugsdatum in die jetzige Wohnung, keine identische Angabe zur früheren Wohnung).

Ebenso wurden minderjährige Kinder, die in nicht elterlichen Haushalten (keine Verzeigerungen zu Mutter oder Vater vorhanden) leben und keinerlei Indizien zu dem entsprechenden Haushalt haben, nicht generiert. Diese Kinder bleiben als Klärungsfall als Einzelpersonen bestehen und werden nicht, wie im KOSIS-Verfahren, einem beliebigem Haushalt zugeordnet.

Ein weiteres Defizit in diesem Szenario besteht darin, dass Personen, die lt. Mikrozensus '98 nicht gemeinsam lebten, zu Haushalten zusammengeführt wurden. Hierbei handelt es sich um erwachsene Kinder, die an derselben Adresse mit eigener Wohnung lebten; um ältere Personen, die im Haus der Kinder mit eigener Wohnung lebten; um getrennt lebende Geschwister bzw. Ehepaare an der gleichen Adresse mit je einer eigenen Wohnung (Ohne Hinweis auf eine eigene Wohnung sind diese Zusammenführungen mit keinem Generierungsverfahren zu vermeiden.).

Szenario 2: (ohne GWZ, ohne GWR, mit Melderegister-Zusatzinfo)

Mit zusätzlichen Angaben des Melderegisters konnte das Generierungsergebnis verbessert werden. 93,5 Prozent der im Mikrozensus '98 erhobenen Haushalte wurden richtig generiert. Durch die Information des Wohnungsgebers war eine genaue Zuordnung der Person zum entsprechenden richtigen Haushalt möglich. Dies wirkte sich vor allem für die Erkennung nicht ehelicher Lebensgemeinschaften positiv aus, für die sonst keinerlei andere Indizien für ein gemeinsames Wohnen vorlagen.

Aufgrund des Fehlens der Wohnungsanzahl je Adresse und des Wohnungsinhabers wurden auch in diesem Szenario Personen zu Haushalten zusammengeführt, die in der Realität Wohnungsinhaber verschiedener Wohnungen sind (z. B. Erwachsene Kinder mit eigener Wohnung).

Szenario 3: (mit GWZ, mit Melderegister-Zusatzinfo)

Unter Einbeziehung einer Wohnungsanzahl, der Information zum Wohnungsinhaber einer Wohnung und zusätzlicher Informationen aus dem Melderegister konnten 99 Prozent der im Mikrozensus '98 vorhandenen Haushalte durch das Generierungsverfahren in der realen Haushaltsstruktur dargestellt werden. Dieses Ergebnis kann sicherlich aufgrund der geringen Datenmenge, die in die Generierung einbezogen wurde, noch nicht als repräsentativ bezeichnet werden. Es zeigt aber, dass unter Ausnutzung von Zusatzinformationen des Melderegisters und der Kombination mit wohnungsstatistischen Angaben ein fast realitätsnahes Generierungsergebnis erreicht werden kann. Ebenso ist mit diesen Informationen eine genaue Zuordnung der generierten Haushalte zu den tatsächlich bewohnten Wohnungen möglich, da bereits im Vorfeld die Wohnungsnummer aus einer GWZ (anhand des Namens des Wohnungsinhabers) in die zu generierende Melderegisterdatei übernommen wurde.

Lediglich drei nicht eheliche Lebensgemeinschaften bzw. Wohngemeinschaften wurden aufgrund fehlender Indizien nicht zusammengeführt.

Szenario 4: (mit Gebäude- und Wohnungsregister, mit Zusatzinfo aus Melderegister)

Eine fehlende Angabe zum Namen des Wohnungsinhabers beeinträchtigt die Qualität des Generierungsergebnisses negativ. Das Fehlen dieser Information bewirkt, wie auch bereits in den Szenarien 1 und 2, eine Verzerrung der Personenzusammengehörigkeiten in den 1- bis 3-Personenhaushalten.

Eine Zusammenführung der generierten Haushalte zur jeweils konkreten Wohnung kann direkt nicht erfolgen. Hier wäre es nur möglich, unter Ausnutzung wohnungsstatistischer Informationen (z. B. Wohnungsgröße, Raumzahl, Wohnungstyp) nach bestimmten festzulegenden Kriterien die generierten Haushalte entsprechend der Haushaltsgröße den Wohnungen zuzuordnen. Dabei ist zu erwarten, dass das Ergebnis weniger realitätsnah dargestellt wird.

Auswertung der Ergebnisse mit HHGEN-Statistik nach Zuordnung der tatsächlichen Personen

Haushaltstyp	Szenario 1 und 2 ohne Wohnungszahl, ohne Wohnungsinhaber						Szenario 3 mit Wohnungszahl, mit Wohnungsinhaber			Szenario 4 mit Wohnungszahl, ohne Wohnungsinhaber		
	ohne zusätzliche Angaben aus EMR			mit zusätzlichen Angaben aus EMR			mit zusätzlichen Angaben aus EMR					
	HH	richtig ¹⁾	Prozent	HH	richtig ¹⁾	Prozent	HH	richtig ¹⁾	Prozent	HH	richtig ¹⁾	Prozent
Haushalte insgesamt	490	434	88,6	480	447	93,1	480	473	98,5	478	455	95,2
davon mit												
1 Person	163	126	77,3	148	127	85,8	140	136	97,1	139	128	92,1
2 Personen	162	153	94,4	164	159	97,0	177	174	98,3	175	167	95,4
3 Personen	95	89	93,7	96	91	94,8	93	93	100,0	92	90	97,8
4 Personen	57	55	96,5	59	58	98,3	58	58	100,0	60	58	96,7
5 Personen	10	8	80,0	10	9	90,0	9	9	100,0	9	9	100,0
6 Personen	3	3	100,0	3	3	100,0	3	3	100,0	3	3	100,0

1) gemessen am MZ-Ergebnis (MZ' 98)

Eine zahlenmäßige Gegenüberstellung der Generierungsergebnisse mit der Haushaltszahl des Mikrozensus'98 sagt nichts über die Korrektheit des Generierungsergebnisses, unterteilt nach Haushaltstypen, aus.

Die Darstellung in obiger Tabelle dokumentiert die Anzahl der in der personellen Zusammensetzung richtig generierten Haushalte gemessen am Generierungsergebnis des jeweiligen Szenario insgesamt und ermöglicht dadurch eine **Bewertung der Generierungsergebnisse**.

- Mit der Generierung nach Szenario 1 wurden 490 Haushalte gebildet. Nur 88,6 Prozent aller mit diesem Szenario generierten Haushalte bilden die richtige personelle Zusammensetzung der Haushalte des Mikrozensus'98 ab.

- Je mehr Informationen für eine Generierung zur Verfügung stehen, einmal aus dem Melderegister als auch wohnungsstatistische Informationen, umso höher wird der Anteil der nach der personellen Zusammensetzung richtig generierten Haushalte. Mit den Testdaten wurden unter Nutzung aller Informationsmöglichkeiten (Szenario 3) 480 Haushalte gebildet, von denen 98,5 Prozent die richtige personelle Zusammensetzung aufwiesen. Alle nach der Generierung vorhandenen 3-, 4-, 5- und 6-Personenhaushalte entsprachen in der personellen Zusammensetzung der Realität (dem Mikrozensus'98).

Zuordnungsschritte HHGEN - Statistik (bezogen auf Daten des Mikrozensus '98)

Zuordnungsschritte	Szenario 1 und 2 ohne Wohnungsanzahl, ohne Wohnungsinhaber				Szenario 3 mit Wohnungsanzahl, mit Wohnungsinhaber		Szenario 4 mit Wohnungsanzahl, ohne Wohnungsinhaber	
	ohne zusätzliche Angaben aus EMR		mit zusätzlichen Angaben aus EMR		mit zusätzlichen Angaben aus EMR			
	insgesamt	darunter: falsch ¹⁾	insgesamt	darunter: falsch ¹⁾	insgesamt	darunter: falsch ¹⁾	insgesamt	darunter: falsch ¹⁾
	generierte Personen							
ohne Zuordnungsschritt (Verzeig.)	976	-	969	-	973	-	969	-
W über Wohnungsgeber	0	x	14	-	14	-	14	-
M minderjährige Kinder zur Mutter	5	-	5	-	5	-	5	-
V minderjährige Kinder zum Vater	0	x	0	x	0	x	0	x
G minderjährige Kinder zu Großeltern	0	x	0	x	0	x	0	x
S minderj. Kinder über Geschwister	0	x	0	x	0	x	0	x
U nichtehel. LG über verzeig. Kinder	12	1	12	1	11	-	12	1
E Zuordnung Ehepartner	5	-	5	-	5	-	5	-
L Lösung von vermeintl. ehel. Verbind.	0	x	0	x	0	x	0	x
K erwachsene Kinder zu den Eltern	48	8	48	8	36	-	43	4
O erwachsene Kinder zu Großeltern	1	-	1	-	1	-	1	-
A ältere Personen zu Nachkommen	4	3	3	2	3	-	1	-
B Personen über Geschwister	3	-	3	-	3	-	3	-
N nicht eheliche Lebensgemeinschaft. (LG)	14	-	10	-	10	-	8	-
R gleich: Einzugsdat./u.o. frühere WO	2	-	1	-	1	-	1	-
Z LG anhand Wohnungsanzahl	0	x	0	x	9	-	9	3
Y Lösung von LG anhand WO-zahl	0	x	0	x	0	x	0	x
doppelt generierte Personen	3	-	3	-	3	-	3	-
nichtgenerierbares minderj. Kind	1	1	0	x	0	x	0	x

1) Im Vergleich zum Mikrozensus 98 falscher Haushaltszusammenhang entstanden

Eine **analytische Betrachtung der Zuordnungsschritte** bringt nachfolgende Ergebnisse:

- Im Unterschied zu Szenario 1: „Generierung ohne Wohnungsanzahl, ohne Wohnungsinhaber und ohne zusätzliche Angaben aus dem Melderegister“ ist ersichtlich, dass im **Schritt W** in den anderen Szenarien 14 Haushalte über den Wohnungsgeber generiert werden konnten. Dabei handelt es sich um 13 nicht eheliche Lebensgemeinschaften und ein minderjähriges Kind im fremden Haushalt lebend.
- Im Schritt M konnten 5 nicht verzeigerte minderjährige Kinder der Mutter zugeordnet werden. Ein Verzeigerungsdefizit minderjähriger Kinder zum Vater, eine Zuordnung von minderjährigen Kindern zu Großeltern, bzw. über einen Geschwisterzusammenhang war bei den Testdaten nicht feststellbar.
- Nicht eheliche Lebensgemeinschaften wurden im **Schritt U** über eine Verzeigerung minderjähriger Kinder zu ihrem Vater gebildet, wobei die Genauigkeit mit Angabe des Wohnungsinhabers (Szenario 3) verbessert werden konnte. In einem Fall lebt der mit dem Kind verzeigerte Vater an gleicher Adresse in einer eigenen Wohnung.
- Über das Programm war es auch möglich 5 Verzeigerungsdefizite von Ehepaaren aufzudecken (**Schritt E**).
- Der **Generierungsschritt K** [Zuordnung von erwachsenen Kindern zu Eltern(-teilen)] ist am stärksten besetzt. Ohne Informationen aus einer GWZ oder einem GWR wurden 18 Kinder im Alter von 18 bis 27 Jahren und 26 noch ältere erwachsene Kinder mit 4 mitgeführten minderjährigen Kindern ihren Eltern(-teilen) zugeordnet. Ein Vergleich mit den direkten Ermittlungsergebnissen des Mikrozensus vor Ort zeigt jedoch, dass ein Teil der erwachse-

nen Kinder unter der gleichen Adresse in einer eigenen Wohnung leben. Ein korrektes Ergebnis wird erst mit Vorliegen der Information zum Wohnungsinhaber und einer Wohnungszahl je Gebäude erreicht (Szenario 3).

- Im **Generierungsschritt Z** (Bildung von Lebensgemeinschaften) wird ersichtlich, dass bei Vorliegen einer Wohnungszahl und der Angabe zum Wohnungsinhaber einer Wohnung nach der Durchführung von Vergleichen mit bestimmten Bedingungen noch nicht generierte Personen zu Lebensgemeinschaften zusammengeführt werden können. Dennoch gibt es Personen die aufgrund fehlender Indizien (unterschiedliche Einzugsdaten in die jetzige Wohnung) nicht zu einer Lebens- bzw. Wohngemeinschaft generiert werden können (im Test bestand für 3 Lebensgemeinschaften keine Generierungsmöglichkeit).

Ein **Abbau von Generierungsdefiziten** ist durch die Möglichkeit der Mehrfachgenerierung von Personen gegeben. In den vorliegenden Testszenarien wurden z. B. 3 Personen doppelt generiert: im Generierungsschritt 2 wurden minderjährige Kinder der Mutter zugeführt und diese Mutter (mit Kind) später ihrem nichtverzeigertem Ehepartner zugeordnet.

Auffällig ist die unterschiedlich hohe Besetzung der Generierungsschritte in den vier Generierungsszenarien. Die Ursache dafür liegt in den zur Verfügung stehenden und die Generierungsabfolge beeinflussenden Informationen.

Resümee

Ausgangspunkt des vorgestellten Verfahrens war, **im Rahmen des Ländermodells zensus-relevante Ergebnisse zur Struktur der Privathaushalte zu erhalten**. Im Grundmodul des Ländermodells integriert, ging es darum, durch die **Zusammenführung von Melderegisterdaten und wohnungsstatistische Daten**:

- die **Qualität der aus dem Melderegister zu gewinnenden Bevölkerungsdaten zu verbessern** (Unter- und Übererfassung) sowie
- kleinräumig strukturierte Haushaltsdaten je Gebäude zu erhalten**.

Wir konzentrierten uns im LDS BB auf die Erarbeitung und Qualifizierung des hier vorgestellten Verfahrens HHGEN-Statistik zur Generierung von Haushalten (Personen, die zusammen bzw. alleine wohnen), das auf zusammengeführten plausiblen Datensätzen aufsetzt.

Zielstellung des Verfahrens ist es, in den Generierungsschritten und einer nachfolgenden zusätzlichen Plausibilisierung (noch zu erarbeiten) die real zusammenwohnenden Personen je Wohnung abzubilden.

Dabei soll aufgezeigt werden, durch welche (gegenwärtig) zusätzlichen Indizien aus dem Einwohnermelderegister und wohnungsstatistischen Daten, z. B. aus noch nicht vorhandenen Gebäude- und Wohnungsregistern, der Generierungserfolg verbessert werden kann.

Es wird davon ausgegangen, dass eine Wohnungsnummer in Deutschland gegenwärtig nicht einzuführen ist. Für die Generierung werden Indizien der Einwohnermelderegister herangezogen, wie:

- Etagen-Nr. (DSMeld-Feld 1210)
- Name des Wohnungsgebers (wohnt bei ...) (DSMeld-Feld 1212)

die teilweise bereits genutzt werden, für deren Erfassung es im EMR jedoch keine verbindlichen Regelungen gibt.

Weiterhin bedarf es eines gesetzlichen Rahmens um zukünftig die für die Generierung benötigten wohnungsstatistischen Merkmale in kommunalen GWR aufnehmen und fortschreiben zu können. Für den Namen des Wohnungsinhabers, der den Generierungserfolg positiv beeinflusst, gibt es - außer bei einer GWZ - keine weitere Quelle. Hier könnte als langfristige Alternative die Einführung einer Wohnungs-Nr. in Betracht gezogen werden. Anderenfalls ist das ohne diese Zusatzinformation zu erhaltene Ergebnis mit dem noch auszuweisenden Fehlerbereich zu akzeptieren (siehe Szenario 4).

Aufgrund fehlender rechtlicher Rahmenbedingungen, die wir uns in Zukunft mit der Vorbereitung für einen registergestützten Zensus erhoffen, beschränkten sich die bisherigen Tests auf die Datensätze des Einwohnermelderegisters einer kreisfreien Stadt im Land Brandenburg.

Wo ordnet sich das **Haushaltgenerierungsverfahren** gegenwärtig ein?

1. In ein alternatives Zensusmodell

Hier wäre es Aufgabe der amtlichen Statistik, das Verfahren weiterzuentwickeln und die Ergebnisse den Nutzern im Rahmen eines Volkszählungsgesetzes zur Verfügung zu stellen.

2. Als ein auf die Zukunft gerichtetes Verfahren

Über einen Zensus hinaus wäre das Verfahren als statistische Methode weiterzuentwickeln und den Anwendern als eine Alternative zum bisherigen vom KOSIS-Verbund entwickelten Verfahren anzubieten.

Die gewählten Testszenarien machen deutlich, dass durch die Kombination von Informationen des Datensatzes eines Melderegisters mit wohnungsstatistischen Daten über einen mehrstufigen Generierungsalgorithmus die realitätsnahe Abbildung der in einem Gebäude wohnenden Personen im Haushaltszusammenhang möglich ist.

Wie aus dem Szenario 3 ersichtlich, kann die **Zielstellung** erreicht werden, wenn

1. **durch verbindliche Regelungen** flächendeckend im Einwohnermeldewesen die **Möglichkeiten des Meldedatensatzes ausgeschöpft** werden sowie
2. **Merkmale eingeführt** werden, **mit denen die Zuordnung von Personen zu ihrer Wohnung zweifelsfrei** - wie hier über den Namen des Wohnungsinhabers - **möglich ist**.

In der Umsetzung dieser Erkenntnisse sollte der Schwerpunkt der Arbeiten zur Etablierung des HHGEN-Verfahrens als Methode der amtlichen Statistik liegen.

Unbestritten sollte sein, dass, wenn das HHGEN-Verfahren als alternatives Verfahren zur Ermittlung kleinräumiger Haushaltsstrukturdaten analog einem Zensus-Bestand haben soll, die Voraussetzungen zur Bereitstellung der entsprechenden Datenbasis zu schaffen sind.

Dringend notwendig ist es weiterhin, den Rahmen für die Prüfung der generierten Ergebnisse zu schaffen.

Aufgabenstellungen für weiter gehende Tests wären somit nach gegenwärtigem Kenntnisstand:

- Einbeziehung kompletter Gebäude mit mehr als 10 Wohnungen,
- Bedeutung des Einzugsdatums im Melderegister im Vergleich mit einer im Rahmen einer GWZ erhobenen Angabe für die Generierung,
- Auswirkungen einer Wohnungs-Nr. aus dem Melderegister auf die Generierung,
- maschinelle Prüfungen bzw. ein Methodenmix aus maschineller Prüfung bzw. Befragung zur Klärung von Unplausibilitäten bei der Feststellung der Wohnungsbelegung,
- Klärung der Namensabgleiche zur Feststellung von Namensübereinstimmungen bei Ausländern.

Wir werden zu gegebener Zeit über weiter gehende Erkenntnisse berichten.

Manfred Fest / Petra Herzog