

Der Zensus 2011 als Datenquelle für die Forschung: Potenziale und Restriktionen

Heike Wirth

Konferenz für Sozial- und Wirtschaftsdaten 5
13.-14. Januar 2011, Wiesbaden

- Zensus: von der empirischen Forschung in Deutschland bislang wenig genutzte Datenquelle
 - ‚Seltenes‘ Ereignis
 - einfaches Merkmalspektrum
 - Datenzugang
- Vorwiegende Nutzung Zensusdaten
 - kleinräumige Regionalforschung , Bevölkerungsforschung, Subpopulationen
 - Referenz für eigene Primärerhebungen (Hochrechnung, Auswahlgrundlage)

- Zensus 2011: Methodenwechsel
 - bisher: Primärstatistische Totalerhebung
 - neu: multiple source mixed mode design
 - Kombination: Register, Primärerhebung (Totalerhebung, Stichprobe) persönliche Interviews, postalische/online Befragung

- Zielsetzung Zensus unabhängig von Methode
 - administrativ
 - statistisch
 - inhaltlich

=> Hat Methodenwechsel Effekte (positiver oder negativer Art) auf Nutzen des Zensus als Datenquelle für die Forschung?

Analysepotenzial, Verfügbarkeit Zensus 2011 ?

Vorbehalte im Vorfeld des Zensus 2011

- Grundsätzliche Bedenken hinsichtlich der Verwendung von Registerdaten, z.B.
 - Qualität der Registerdaten
 - Abdeckung Population
 - Merkmalspektrum Register
 - Qualität der Datenverknüpfung

=> Groß angelegtes methodisches Projekt der amtlichen Statistik im Vorfeld Zensus 2011: Zensustest 2001

Zentrale Befunde Zensus 2001

- Einwohnerzahl um circa 1,3 Millionen erhöht
- Registerfehler regional & nach Bevölkerungsgruppen ungleich verteilt

⇒ Nutzung Register nur in Kombination mit Primärstatistischen Erhebungen

- Überprüfung; statistische Korrektur Registerbefunde
- Ermittlung zusätzlicher Merkmale, die nicht in Registern verfügbar

=> Verbesserung Verfahren & Register in Hinblick auf Zensus 2011

⇒ aber vgl. Beitrag Ulli Rendtel: Nach dem Zensus 2011 ist vor dem Zensus 2021: Bedarf an weiterführender methodischer & statistischer Forschung

Analysepotenzial - Merkmalspektrum

Merkmalskatalog Zensus 2011 setzt sich zusammen aus

- Registermerkmalen (z.B. Nationalität, Familienstand, Geburtsjahr)
 - primärstatistischen Merkmalen
 - Wohn- und Gebäudezählung (Vollerhebung)
 - Haushaltsstichprobe ($\approx 10\%$)
 - Primärstatistische Erhebungen
 - Empfehlungen Zensuskommission zur Aufnahme einzelner Merkmale (z.B. Migrationshintergrund; im Haushalt vorwiegend gesprochene Sprache; Pendlerverhalten; Energiequelle Heizung) siehe:
http://www.zensuskommission.de/docs/Stellungnahme_2009_01_22.pdf
- ⇒ letztendlich realisierte Prinzip: *„nur das unbedingt Nötige“*

Analysepotenzial: EU-Pflichtprogramm

| Demografische Merkmale | Erwerbs- & Bildungsmerkmale | Haushalts- /Familienmerkmale | Gebäude- & Wohnungsmerkmale |
|--|---|---|---|
| <ul style="list-style-type: none"> •Geschlecht •Alter •Familienstand •Staatsangehörigkeit •Geburtsland/-ort •Üblicher Aufenthaltsort •Vorheriger Aufenthaltsort /Datum der Ankunft oder Aufenthaltsort ein Jahr vor dem Zensus •Zuzug aus Ausland/ Jahr der Ankunft in Deutschland | <ul style="list-style-type: none"> •Erwerbsstatus •Beruf •Wirtschaftszweig •Stellung im Beruf •Arbeitsort •Bildung & Ausbildung | <ul style="list-style-type: none"> •Haushaltstyp •Haushaltsgröße •Stellung im Haushalt •Stellung in der Familie •Typ der Kernfamilie •Größe der Kernfamilie | <ul style="list-style-type: none"> •Art der Unterkunft •Gebäudetyp •Belegungsstatus •Baujahr Gebäudes •Nutzfläche und/oder Zahl der Räume •Zahl der Bewohner •Besitzstatus •Ausstattung der Wohnung •Heizung |

Zusätzlich zu EU-Standardprogramm

- Erweiterter Migrationshintergrund
- Zugehörigkeit zu öffentlich-rechtlicher Religionsgemeinschaft; Glaubensbekenntnis

Im Vergleich zur VZ87, z.B. nicht enthalten

- Praktische Berufsausbildung und ihre Dauer; Quelle überwiegenden Lebensunterhalt, Arbeitszeit (zumindest Teilzeit; Vollzeit); Pendlerverhalten; Wohnungsmiete, Energiequelle Heizung;

Analysepotenzial - Merkmalspektrum

- Zensus 2011: Beschränkung auf das ‚notwendige Minimum‘ (auch in Hinblick auf Registermerkmale)
 - steht in Tradition der klassischen Volkszählung und deren inhaltlich begrenzten Analysepotenzial
 - **erhebliche Einschränkung:** keine Daten zur Arbeitszeit & Quelle überwiegenden Lebensunterhalt
- **Positiv:**
 - Erhebung des erweiterten Migrationshintergrund
 - Differenzierte Erfassung ausgewählter Merkmale
 - Bildung, Ausbildung, Schulbesuch
 - Erwerbsverhalten, Arbeitssuche
 - Gemeinschaftsunterkünfte (Vollerhebung Basismerkmale + Stichprobe mit Fragenkatalog Haushaltsstichprobe)

Analysepotenzial – Regionale Differenzierung

- Klassische Stärke VZ: Bereitstellung sachlich und räumlich tief gegliederte Daten über die Grundgesamtheit
- Methodenwechsel Zensus => Einschränkung kleinräumiger Analysen
 - **Grundsätzlich:** Stichprobenmerkmale nicht unterhalb Gemeindeebene
 - **Spezifisch:** Stichprobenmerkmale für Gemeinden < 10.000 Einwohner nicht unterhalb Kreisebene (bzw. Gemeindeverbund)

=> Erhebliche Einbuße an Analysepotenzial für Stadt- & Regionalforschung

Analysepotenzial – Generierung von Haushalten

- Klassische VZ
 - Haushalt = Gemeinsames Wohnen & Wirtschaften
 - Haushaltsmitglieder = Definiert durch Befragte (Haushaltsmantelbogen)
- Zensus 2011
 - Haushalt = Gemeinsames Wohnen
 - Haushaltsmitglieder = Ermittelt über Angaben aus dem Melderegister (Verzeigerungen) & Wohn- und Gebäudezählung (Wohnungsinhaber)
 - Haushaltsstichprobe erfasst keine Informationen zu Beziehungen zwischen den Haushaltsmitgliedern

=> statistische Generierung von Haushalten stößt in Forschung auf Bedenken (Fokus: Nichteheliche Lebensgemeinschaften)

| VZ 1970 | | VZ 1987 | Zensus 2011 | | |
|--|---|---|--|--|--|
| Primärerhebungen (Total) | | Primärerhebung (Stichprobe) | Melderegister | Gebäude- & Wohnungszählung | |
| Haushalt = Gemeinsames Wohnen & Wirtschaften | | Haushalt = Gemeinsames Wohnen | | | |
| Haushaltsmantelbogen = Auflistung aller Personen in Haushalt | | - | | | |
| Personenbogen | | Personenbogen | | | |
| <ul style="list-style-type: none"> Familienstand | <ul style="list-style-type: none"> Familienstand | <ul style="list-style-type: none"> Familienstand (inkl. eingetragener Partnerschaft) | <ul style="list-style-type: none"> Familienstand (inkl. eingetragener Partnerschaft) | | |
| <ul style="list-style-type: none"> Stellung i. Haushalt (z.B. Hhltbezugsperson, Ehepartner, Kind,) | - | - | <ul style="list-style-type: none"> Verzeigerungen Ehegatten, Kinder, eingetrag. Partnerschaften | | |
| - | - | <ul style="list-style-type: none"> Gemeinsames Wohnen mit Partner (Nicht-EP; Nicht eingetr. Partnerschaft) | <ul style="list-style-type: none"> sonstige Merkmale (z.B. Einzugsdatum, frühere Anschrift) | <ul style="list-style-type: none"> Namen von maximal 2 Wohnungsinhabern Zahl der Bewohner je Wohnung | |

Auszug aus: Tabelle 7: Haushalte von nichtehelichen Lebensgemeinschaften laut Haushaltegenerierung und Haushaltebefragung
Hochgerechnetes Ergebnis des Zensusstests vom 5. Dezember 2001

| Haushaltstyp | Haushalte laut | | Abweichung | |
|--|----------------|--------------------|------------|-------|
| | Generierung | Haushaltebefragung | | |
| | 1 000 | | % | |
| Haushalte ohne Register- und GWZ-Fehler | | | | |
| Nichteheliche Lebensgemeinschaften ohne Kind | 951,2 | 903,6 | +47,6 | +5,3 |
| Nichteheliche Lebensgemeinschaften mit Kind(ern) | | | | |
| unter 27 Jahren | 516,4 | 448,3 | +68,1 | +15,2 |
| darunter mit ... unter 18 Jahren | | | | |
| 1 Kind | 294,0 | 262,0 | +32,0 | +12,2 |
| 2 Kindern | 138,5 | 111,2 | +27,3 | +24,5 |
| 3 Kindern | 34,5 | 26,1 | +8,3 | +31,9 |

GWZ = Gebäude- und Wohnungszählung.

Quelle: Vorndran (2004): Zensusstest – Verfahren und Ergebnisse der Haushaltegenerierung. In: *Wirtschaft und Statistik* 11/2004:1253

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

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

Analysepotenzial – Generierung von Haushalten

- Befunde basieren auf ersten Ergebnissen Zensus 2001 => Verfeinerung Verfahren
 - Bewertung Unschärfen abhängig von Untersuchungsinteresse
 - univariate Verteilungen, hochaggregiert
 - Bezugsgröße (z.B. Wohnfläche pro Kopf)
 - mehrdimensionale Verteilungen
 - Mikrodatenanalyse
 - Geringer Wissensstand bzgl. Ausmaß & Effekten von Unschärfen der Haushaltegenerierung
 - Möglichkeiten zur methodischen Begleitforschung der Haushaltegenerierung im Zensus 2011 begrenzt
 - Generierung erfolgt z.T. über Hilfsmerkmale MR & GWZ
 - Löschfrist Hilfsmerkmale (nach Abschluss Datenaufbereitung; spätestens 4 Jahre nach Berichtszeitpunkt)
- => Notwendigkeit einer datenschutzgerechten Lösung zur Nachvollziehbarkeit der Haushaltegenerierung

- Bedeutungsgewinn Zensus aufgrund des Methodenwechsels (vgl. Beitrag Ulli Rendtel zur Zensus-Begleitforschung)
 - Evaluation der neuen Methode mit Zielsetzung quantitativer & qualitativer Verbesserung Register- und Primärerhebungen und damit auch zukünftiger Zensen
 - z.B. Messungen aus unterschiedlichen Datenquellen: Analysen zu Ausmaß, Ursachen und Effekten von Abweichungen zwischen Messungen aus Register- und Befragungsdaten
 - Bsp. Haushalte nach Staatsangehörigkeit

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

| Haushaltstyp | Haushalte laut | | Abweichung | |
|---|----------------|-------------------------|------------|-------|
| | Generierung | Haushalte- befragung | | |
| | 1 000 | | | % |
| Haushalte ohne Register- und GWZ-Fehler | 25 398,8 | 25 499,4 | -100,6 | -0,4 |
| deutsche Haushalte | 23 717,2 | 23 996,8 | -279,6 | -1,2 |
| ausländische Haushalte | 751,1 | 765,0 | -13,9 | -1,8 |
| Mischhaushalte | 930,5 | 737,6 | +192,9 | +26,2 |

GWZ = Gebäude- und Wohnungszählung.

Analysepotenzial – in a nutshell

- Inhaltliche Forschung
 - Potenzial vergleichbar mit klassischer VZ
 - aber problembehaftet für Haushalts- und Familienforschung
 - erweitertes Analysepotenzial Migranten
 - Stadt- & Regionalforschung erhebliche Einschränkung Analysepotenzial im Vergleich zur klassischen VZ
 - => Wie kann Bedürfnissen nach kleinräumigen Analysen in zukünftigen Zensen besser entsprochen werden?
- Methodische Forschung (erweitertes Potenzial im Vergleich zur klassischen VZ)
 - Exkurs: Registergestützter Zensus Hochrechnungsrahmen?
 - Stichprobe = Grundlage für statistische Bereinigung Register = mittels Stichprobe bereinigte Register Grundlage für Hochrechnung Stichprobe?

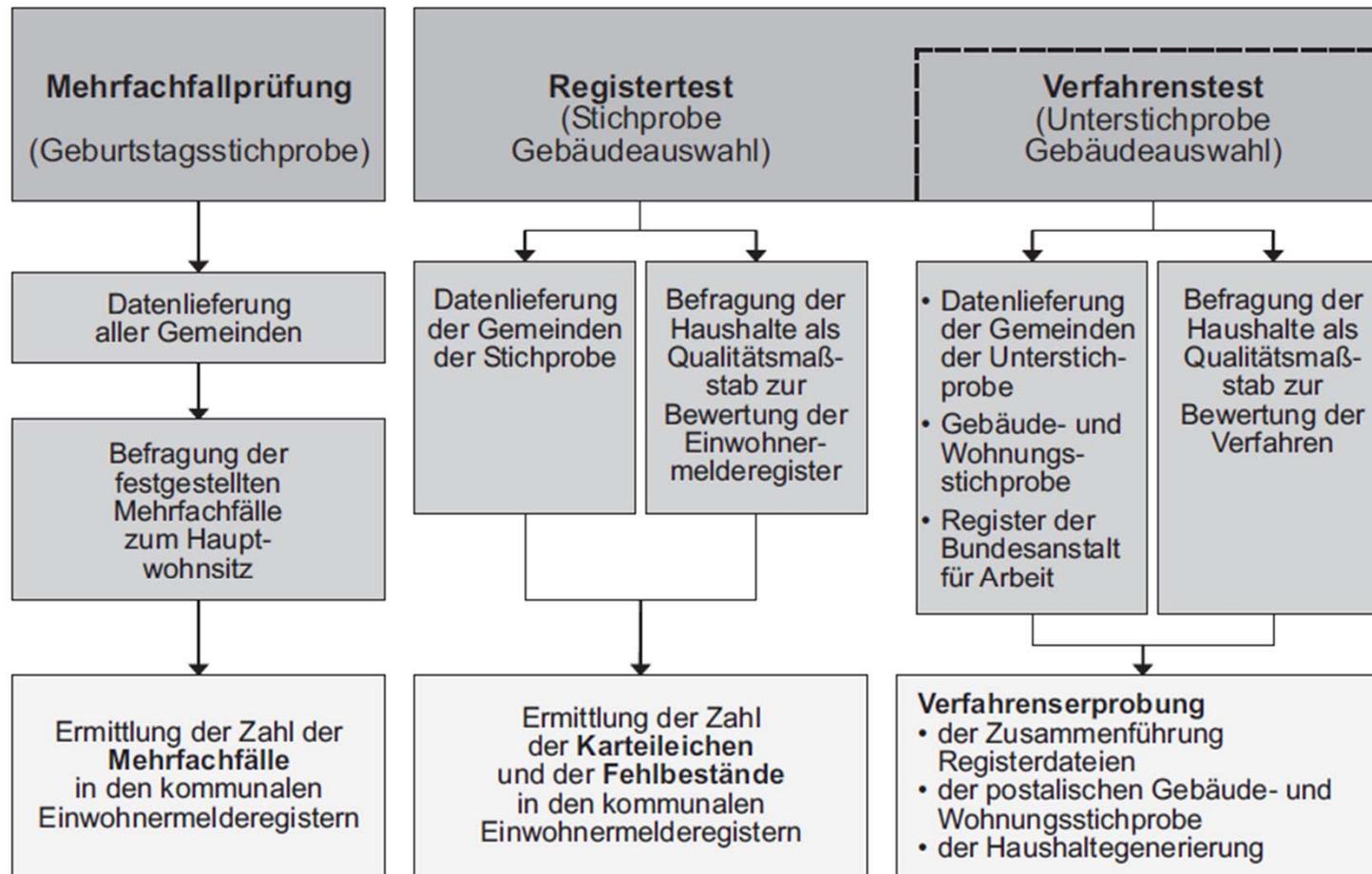
Datenzugang

- Nutzung Daten abhängig von Zugangsmöglichkeiten der Forschung
 - Aktualität
 - Merkmalsumfang (inhaltliche/methodische Merkmale)
 - Restriktionen Analysepotenzial durch Anonymisierung
 - Flexibilität Zugang
- Zugang kann sich im Vergleich zu VZ87 nur verbessern
 - Bereits im Vorfeld Zensus 2011 enge Kooperation amtliche Statistik & Forschung (Zensuskommission)
 - im Unterschied zu VZ87 Weitergabe Mikrodaten Zensus2011 durch BStatG §16(6) geregelt
 - Einrichtung einer beim RatSWD angesiedelten Arbeitsgruppe ‚Datenzugang der Forschung zu Mikrodaten des Zensus 2011‘

Datenzugang

- Datenzugang muss sich an den unterschiedlichen Forschungsinteressen und den entsprechenden Datenbedarfen (inhaltlich, methodisch, statistisch) orientieren
- Unterschiedliche Datenbestände
 - Zensustypischer Datensatz (vgl. Vortrag Sabine Bechtold)
 - Haushaltsstichprobe
 - [Zensustest 2001 (z.B. anonymisierte Daten aus Verfahrenstest)?]
- Prioritäten?
 - Welche Daten & welche Zugänge?
- Ressourcen für die Datenbereitstellung?

Ü Zensustest – Konzeption und Testmodule



Statistisches Landesamt Baden-Württemberg

140 04

Quelle: Braun, Ralf (2004): Zensustest 2001. Ergebnisse der Testerhebungen und Ausblick auf ein zukünftiges Zensusmodell. In: Statistisches Monatsheft Baden-Württemberg 5/2004; S.4.

Auszug aus: Tabelle 5: Haushalte von Ehepaaren ohne und mit Kind(ern) laut Haushaltegenerierung und Haushaltebefragung

Hochgerechnetes Ergebnis des Zensustests vom 5. Dezember 2001

| Haushaltstyp | Haushalte laut | | Abweichung | |
|--|----------------|--------------------|------------|------|
| | Generierung | Haushaltebefragung | | |
| | 1 000 | | % | |
| Haushalte ohne Register- und GWZ-Fehler | | | | |
| Ehepaare ohne Kind | 6458,0 | 6546,6 | -88,6 | -1,4 |
| Ehepaare mit Kind(ern) unter 27 Jahren | 6192,2 | 6371,0 | -178,8 | -2,8 |
| darunter mit ... unter 18 Jahren | | | | |
| 1 Kind | 2187,3 | 2256,1 | -68,8 | -3,0 |
| 2 Kindern | 2093,2 | 2156,6 | -63,4 | -2,9 |
| 3 Kindern | 543,9 | 541,4 | +2,4 | +0,5 |
| 4 Kindern | 86,0 | 93,0 | -7,0 | -7,5 |
| 5 oder mehr Kindern | 43,7 | 45,4 | -1,7 | -3,8 |

GWZ = Gebäude- und Wohnungszählung.

Quelle: Vorndran (2004): Zensustest – Verfahren und Ergebnisse der Haushaltegenerierung. In: *Wirtschaft und Statistik* 11/2004:1252