

Haushaltstatistiken im Niederländischen Zensus

Eric Schulte Nordholt

Statistics Netherlands

Division Social and Spatial Statistics

Department Support and Development

Section Research and Development

ESLE@CBS.NL

Vorlesung in Düsseldorf (Januar 2010)



Contents

- **Geschichte des Niederländischen Zensus**
- **Introduktion des virtuellen Zensus 2001**
- **Dataquellen**
- **Quellen kombinieren: Microlinkage**
- **Quellen kombinieren: Microintegration**
- **Ableitung von Haushaltsvariablen**
- **Verbesserungen in der Ableitung**
- **Vergleichbarkeit in der Zeit**
- **Data-integrationsaktivitäten zwischen den 2001 Zensus und den 2011 Zensus**
- **Schlußbemerkungen**



Geschichte des Niederländischen Zensus

EIN TRADITIONELLER ZENSUS

Innenministerium:

1829, 1839, 1849, 1859, 1869, 1879 und 1889

Statistics Netherlands:

1899, 1909, 1920, 1930, 1947, 1960 und 1971

Nicht länger bereit teil zu nehmen (Nonresponse) und Kostenreduktion ⇒ keine traditionelle Zensussen mehr

ALTERNATIV: EIN VIRTUELLER ZENSUS

1981 und 1991: Bevölkerungsregister und Erhebungen

Entwicklung im 90'er Jahren: mehr Registers →

2001: ein integrierter Set von Registers und Erhebungen, SSD

Introduktion des virtuellen Zensus 2001

Warum ein Zensus?

Statistische Information für Forschungszwecken und politische Zwecke

Wie sieht die Information aus?

- Größe der Subpopulationen
- Demographische und Sozial-ökonomische Merkmale, auf nationaler und regionaler Ebene

Gentlemen's agreement

- Eurostat: Koordinator von EU-, Akzession- und EFTA-Länder in der 2001 Zensus Runde
- Zensus Tabellenprogramm, jede 10 Jahren

Dataquellen

Registerdaten:

- **Bevölkerungsregister**, 16 Million Records
Demographische Variablen: **Geschlecht, Alter, Haushaltsstatus u.s.w.**
- **Jobregister, Stellen von Arbeitnehmern**, 6.5 Million Records, **und Freiberuflichenregister**, 790 Tausend Records
Anfang (und Ende) der Stellen, Ökonomische Aktivitätskode (NACE)
- **Steuerregister (FIBASE), Stellen**, 7.2 Million Records, **und Pensionen/Renten und Lebensversicherungen**, 2.7 Million Records
- **Soziale Sicherheitsadministrationen**, 2 Million Records, **Hilfsinformation für den Integrationsprozess**

Erhebungen:

- **Arbeitsplätzen- und Lohnerhebung**, 3 Million Records, **(vertragliche) Arbeitszeit, Arbeitsstätte**
- **Arbeitskraftenerhebung**, 2 Jahren: 230.000 Records
Ausbildung, Beruf, (Ökonomische) Hauptaktivität



Quellen kombinieren: Microlinkage

- **Linkagekey:**
Registers
Soziale Sicherheit und Fiskal Nummer (SoFi), einmalig
Erhebungen
Geschlecht, Geburtsdatum,
Adresse (Postleitzahl und Hausnummer)
- **Linkagekey ersetzt durch RIN-Person**
- **Linkagestrategie**
Optimalisierung von der Zahl der Verknüpfungen
Minimalisierung von der Zahl der falsche
Verknüpfungen und fehlte Verknüpfungen

Quellen kombinieren: Microintegration

- Daten von verschiedenen Quellen werden **zusammengebracht** ⇒ **begreiflichere und kohärente Information über Aspekte im Lebenslauf**
- Vergleich Quellen
 - **Deckung der Bevölkerung**
 - **Quellen die mit einander in Konflikt geraten**
- Integrationsregeln
 - **Kontrollen**
 - **Anpassungen**
 - **Imputationen**
- **Optimaler Benutzung der Information** ⇒ **Qualitätsverbesserung**

Ableitung von Haushaltsvariablen

- Vorher: Information aus Arbeitskraftenerhebung
- Seit 2001: abgeleitet aus Bevölkerungsregister
 - Information für alle Gemeinden im Land
 - Für 92,6% der Bevölkerung kann die Variable Haushaltsposition abgeleitet werden (auf Basis von Information über Heiraten und Kinder)
 - Für 7,4% der Bevölkerung konnte die Variable Haushaltsposition nicht abgeleitet werden (oft geht es um zwei unverheirate Personen auf der gleichen Adresse)
 - Stochastische Imputation (Methode Fellegi)
- Konsistente Ergebnisse in allen Projekten ⇒
Imputationen aufgenommen im Bevölkerungsregister

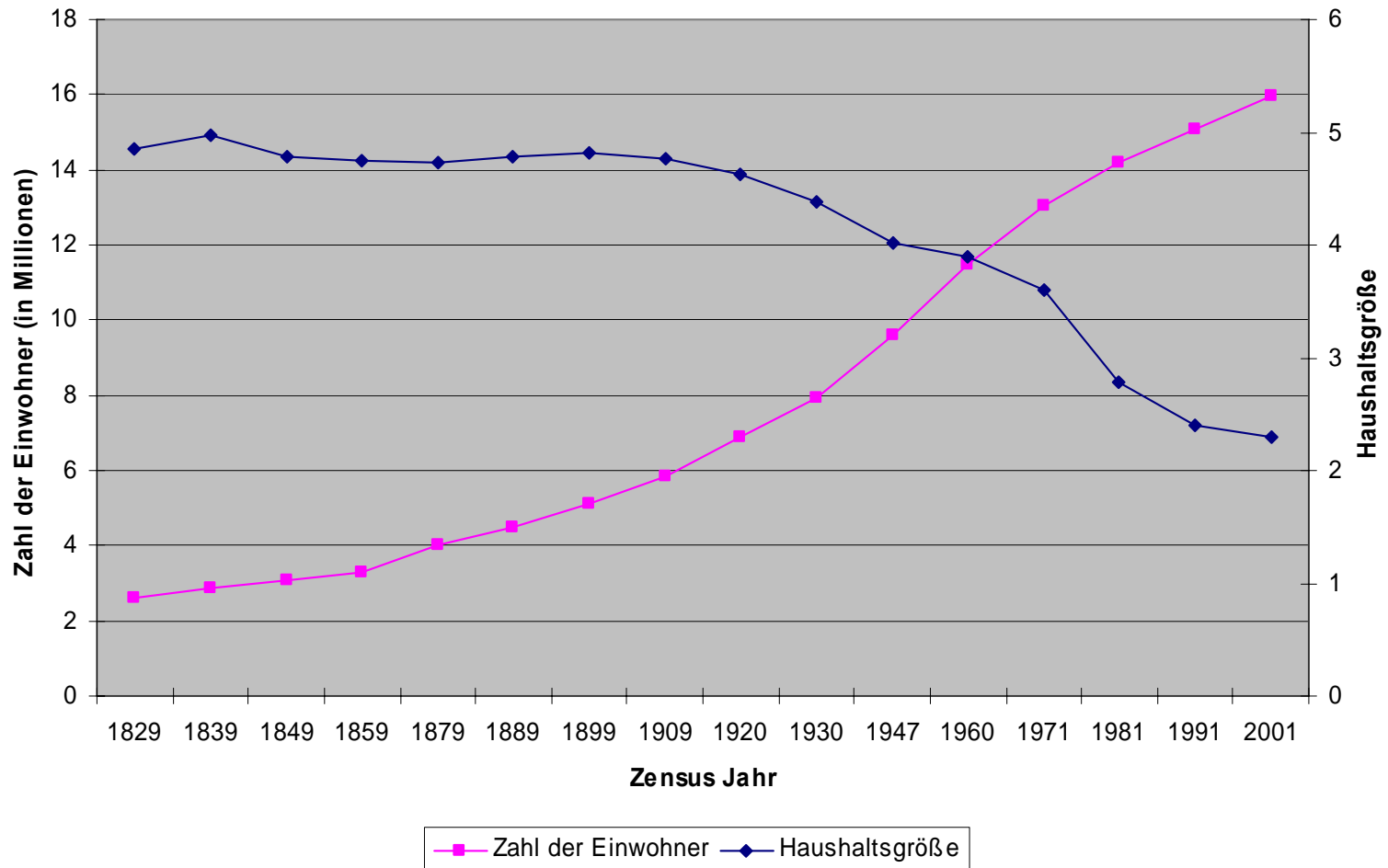


Verbesserungen in der Ableitung

- **Neue Information**
 - Information über Gesundheitszuschläge
 - Information über Mietzuschläge
 - Zurückgabe von Steuergeldern
- **Seit 2009: abgeleitet aus Bevölkerungsregister**
 - Für 94% der Bevölkerung kann die Variable Haushaltsposition abgeleitet werden (auf Basis von Information über Heiraten, Kinder und die neue Information)
 - Für 6% der Bevölkerung konnte die Variable Haushaltsposition nicht abgeleitet werden (oft geht es um zwei unverheiratete Personen auf der gleichen Adresse)
 - Stochastische Imputation (Methode Fellegi)

Vergleichbarkeit in der Zeit

Zahl der Einwohner und Haushaltsgröße



Data-integrationsaktivitäten zwischen den 2001 Zensus und den 2011 Zensus (1)

- Tabellen
(<http://www.cbs.nl/en-GB/menu/themas/dossiers/volkstellingen/publicaties/2005-virtual-dutch-census-art.htm>)
- Buch und extra Kapitel
(<http://www.cbs.nl/en-GB/menu/themas/dossiers/volkstellingen/publicaties/2001-b57-e-pub.htm>)



Data-integrationsaktivitäten zwischen den 2001 Zensus und den 2011 Zensus (2)

- Integrated Public Use Microdata Series (<https://international.ipums.org/international>)
- Vorlesungen (Konferenzen, Universitäten, Forschungsinstitutionen, Statistische Ämter)
- ESTP-course Registers in Statistics (Oslo)
- International Statistical Seminar Eustat in Bilbao (http://www.eustat.es/prodserv/seminario_i.html)
- Digitalisierung (<http://www.volkstellingen.nl>)
- Recommendations und register-based statistics
- CENEX on ISAD (<http://cenex-isad.istat.it>)
- European census regulations



Data-integrationsaktivitäten zwischen den 2001 Zensus und den 2011 Zensus (3)

- Quellen (das Bevölkerungsregister als Rückgrat für den Zensus, Änderungen in Inhalt und Qualität der Registers, fehlende Information aus Erhebungen)
- Schätzungsmethode (repeated weighting, neue Version der Software, Rückfalloption: hochheben nach dem Bevölkerungsregister, weniger Imputationen)
- Statistische Sicherung der Hypercubes
- Web service (Tabellen in SDMX-Format)



Schlußbemerkungen (1)

- Ein Niederländischen virtuellen Zensus: yes, we can!
- Microintegration bleibt wichtig
- Repeated weighting war ein Erfolg

Vorteilen:

- Relativ billig (geringe Kosten pro Einwohner)
- Schnell (kurze Produktionszeit)

Nachteilen:

- Abhängigkeit von Registerinhaber (Statistik ist nicht ihr Priorität), rechtzeitig von Registern, Konzepte und Population von Registern kann abweichen von was notwendig ist (behalt gute Beziehungen mit den Registereinhabern!)
- Publikation von kleinen Subpopulationen ist manchmal schwer oder selbst unmöglich

Schlußbemerkungen (2)

Andere Aspekte:

- **Weniger Aufmerksamkeit für die Ergebnisse von einem virtuellen Zensus als für ein traditioneller Zensus**
- **Schwierig um die Kenntnis und Software up-to-date zu behalten (Zensus ist nur jede zehnte Jahr)**
- **Riesiges internationalen Interesse im virtuellen Zensus**
- **Viele interessante Arbeit in den nächsten Jahren!**

Vielen Dank!



Zeit für Fragen und Diskussion

