Country report Germany

Workshop ‘Integration Global Census Microdata’
Durban, August 15th, 2008

Dr. Markus Zwick, Research Data Centre
Federal Statistical Office Germany
RDC of official statistics

- interface organisation between data producers and empirical science
- consulting and service for the use of official microdata
- possibility for access to microdata with low anonymisation level
RDC: Advantages for the data producer

- More research on our own data
- Higher data quality
- Greater network on researchers
- More competence on data and research knowledge
- International acceptance of RDC researcher
- Better reputation of German research
Level of Anonymisation

- Complete microdata
  - Delete direct identifier
    - Confidential microdata
      - Anonymisation method
        - De-facto anonymised microdata
          - Stronger anonymisation method
            - Fully anonymised microdata

Degree of confidentiality

Degree of analysis potential
Possibilities for microdata access

- de-facto anonymised microdata (Scientific Use Files)
- fully anonymised microdata (Public Use Files)
- Visiting Researcher‘s Desktop
- Controlled Remote Data Processing (Remote Execution)
- Special Data Processing
German Microdata as Public Use File for the IECM (IPUMS-Europe) project

nine anonymised microdata files
- census 1970 and 1987 for the Federal Republic of Germany
- census 1971 and 1981 for the former German Democratic Republic
German Microdata for the IECM project

Federal Statistical Office Germany, Dr. Markus Zwick

www.forschungsdatenzentrum.de
Census of the former GDR 1971
Characteristics and metadata

two data files
- Person file (demography, income, education, employment etc.)
- Dwelling and building file (state of repair, occupancy, etc.)
16,4 mio. persons, 6,2 mio. households, 6 mio. dwellings

metadata
- no codebooks at FSO and Federal Archive
- Archives of regional statistical offices in the former GDR states
  (Field of study, occupation codes)
Census of the Federal Republic Germany 1987

Sample microdata source: Statistical Offices of the Länder
Type of field work: standardized interview
Census day: 1987 May 25
Population: total population entitled to reside
Coverage: 100%
Enumeration unit: household
Respondent: all persons in households and communal establishments
Census of the Federal Republic Germany 1987

Size:
63,2 mio. persons
26,7 mio. households
25,9 mio. dwellings
177 variables

Special populations: foreigners
Census of the Federal Republic Germany 1987

Variables characteristics

two questionnaires: population and occupation census
- questions on person, sources of livelihood, economic activity, education/training, commuting

and census of buildings and housing Person file
- questions on demography, income, education, employment

dwelling and building file
- questions on occupancy, equipment
Public Use File Census 1987 (West)

Sampling:
- size: 1% household sample
- design: Systematic Random Sampling
- sorting of households and geographic variables
- deletion of vacant dwellings and dwellings used for other purposes
- adding household number
- first household selected randomly, then selecting every 100 households
Public Use File Census 1987 (West)

- deletion of geographic details (except for state and size of place)
- top and bottom coding
- principle: every value of a variable should have at least 10,000 observations in the original file
- size of place: every value should have at least 400,000 observations in the original file
- citizenship: every value should have at least 100,000 observations in the original file
Microcensus for the FRG

- annual 1% revolving household sample with obligation to give information by law
- 800,000 person, 380,000 households with nearly 750 variables
- Scientific Use File as 70% subsample for researcher in Germany
Microcensus as Public Use File

- 35% subsample
- ca. 300 variables
- anonymisation by local suppression and by top and bottom coding
## Timetable German contribution to IECM

<table>
<thead>
<tr>
<th>Development Public-Use-Files</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>Census 1987 FRG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MC 2001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MC 1991</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MC 1982</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MC 1973</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MC 1987</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Census 1970 FRG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Census 1981 GDR</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Project grant from the Federal Ministry of Education and Research**
- 7 8 9 10 11 12 1 2 3 4 5 6 7 8 9 10 11 12 1 2 3 4 5 6 7 8 9
Structure of the census 2011 in Germany

- Census of buildings and housing
- Special buildings
- Population registers
- Registers of employment statistics
- Sample surveys

Household generating procedure

Census data
Census Germany 2011

Characteristics of EU/ECE – recommendation

- Demographic and geographic characteristics
  (e.g. age, sex, marital status, citizenship, current and former domicile, place of birth)

- Economic and education characteristics
  (e.g. current activity status, occupation, industrial sector, status in employment, highest educational attainment)

- Household and family characteristics
  (e.g. type and size of household, type and size of family nucleus, family status)

- Building and dwelling characteristics
  (e.g. type of building, occupancy status, period of construction, number of occupants, number and space of rooms, Ownership, standard of equipment, type of heating)
Household generating procedure

- formation of core households
- identification of the owners or main tenants of dwellings in the population register
- formation of households on the basis of hard generation criteria (record linkage)
- formation of households on the basis of statistical generation criteria (statistical matching)
Model of census 2011

- Federal Statistical Office Germany, Dr. Markus Zwick

- Labour statistics register
  - 35 mil. sets

- Population register
  - 88 mil. sets

- Census of dwellings and buildings
  - 17.5 mil. owners of dwellings/buildings

- Data acquisition per special buildings
  - 2 mil. individuals

- Additional sample
  - 5.9 - 7.2 mil. individuals

- Clarification of implausible cases

- Household generating process
  - 38.5 mil. households

- Correction of registration errors

- Result: census-typical data set

- Combination and control for multiple cases

- Over-/under-coverage
- Take over of characteristics
Country report Germany

Workshop ‘Integration Global Census Microdata’
Durban, August 15th, 2008

Dr. Markus Zwick, Research Data Centre
Federal Statistical Office Germany