Linked Employer-Employee Data (LIAB)

IAB-Workshop
Data treasures at the Institute for Employment Research

2014-01-23
Overview

- Data overview
- LIAB in general
- Linkage
- Models of LIAB
- Research with LIAB
- Data access and further informations
1. Data Overview
2. LIAB in general

Combination of

- establishment data (BP)
  - yearly survey of approx. 16,000 establishments
  - information about interviewed establishments with changing topics

and

- individual data (IEB)
  - employment biographies of working individuals from social security and administrative data

→ rich set of socio-demographic variables
2. LIAB in general

Linked Employer Employee Data

→ simultaneous analyses of supply and demand side of labour market

How can we connect these two data sources?
3. Linkage

- **What is the establishment number (EID)?**
  - numerical identificator of notifying establishment, in which the respective individual is working

- **Inclusion in individual data (IEB):**
  - is required in notification process to the social security institutions in Germany

- **Inclusion in Establishment Panel (BP):**
  - is sampled from establishment file of Federal Employment Agency
  - contains universe of establishments including EID with at least one employee liable to social security
3.1 Possible Linkage problem

EID is only asked in first interview in the BP

- no adjustment of the EID for establishments asked a second time
- Example: Change of EID because of new owner in 2000

→ starting with wave 2000, false or no individual data is possibly linked with establishment data
3.2 Linkage quality

- Linked establishments: 4,188 bis 14,981 per year
4. Models of LIAB

Two different data models

- LIAB Cross-sectional Model (QM)
  - current model: Cross-sectional Model 2 1993 to 2010 (QM 2 9310)
  - updated every two years

- LIAB Longitudinal Model (LM)
  - new model: Longitudinal Model 1993 to 2010 (LM 9310)
  - Mover-Modell (MM 9308)
    - no more updates
5. Cross-sectional Model (QM2 9310)

- Cross-sectional Model 2 1993 to 2010 (QM 2 9310)
  - **annual** linkage of individual and establishment data on reference date (30 June) of respective year
  - Individual data base: BeH, LeH, (X)ASU, (X)LHG
5.1 Sampling Procedure QM2 9310

- Step 1: Selection of Establishments (from BP)
  - all establishments with valid interview in respective year

- Step 2: Selection of individual data (from IEB)
  - all individuals, employed at one of these selected establishments at least one day
  - for these individuals only episodes which contain 30 June of respective year

→ identical procedure for every year (1993-2010)

Linked Establishments: 4,188 to 14,981 per year
Individuals: 1,629,542 to 2,584,520 per year
5.2 Structure and size of QM2 9310

→ Linkage via EID

- **Storage:**
  - IAB Establishment Panel: 2 to 10 MB per year
  - Individual data: 180 to 270 MB per year
### 5.3 Example of linkage QM2 9310

#### Establishment Panel

<table>
<thead>
<tr>
<th>EID</th>
<th>wave 2005</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>interview</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>interview</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>no interview</td>
<td></td>
</tr>
</tbody>
</table>

#### Data on individuals

<table>
<thead>
<tr>
<th>EID</th>
<th>persnr</th>
<th>start_date</th>
<th>end_date</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20</td>
<td>01.01.2005</td>
<td>31.12.2005</td>
</tr>
<tr>
<td>2</td>
<td>21</td>
<td>15.02.2005</td>
<td>31.10.2005</td>
</tr>
<tr>
<td>.z</td>
<td>21</td>
<td>01.05.2005</td>
<td>30.08.2005</td>
</tr>
<tr>
<td>1</td>
<td>22</td>
<td>01.01.2005</td>
<td>31.12.2005</td>
</tr>
<tr>
<td>2</td>
<td>42</td>
<td>31.03.2005</td>
<td>17.09.2005</td>
</tr>
<tr>
<td>2</td>
<td>43</td>
<td>01.01.2005</td>
<td>31.12.2005</td>
</tr>
<tr>
<td>3</td>
<td>43</td>
<td>01.06.2005</td>
<td>31.07.2005</td>
</tr>
</tbody>
</table>

* parallel employment in non-BP-establishments
** parallel employment in an BP-Establishment which was not selected for LIAB
6. Longitudinal Models

- **Longitudinal Model** LM 9310
  - Individual data of IEB *continuously* available across observation period
  - illustration of complete employment biographies
  - Individual data base: BeH, LeH

- **Mover Model** 9308 MM 9308
  - focus on persons who changed the employer
  - data base: BeH LeH (X)LHG (X)ASU
  - no more updates
6.1 Sampling procedure LM 9310

- **Step 1: Selection of Establishments (from BP)**
  - all panel cases from 2000 to 2008
  - What is a panel case?
    - establishments contained in the starting year
    - “new” establishments in the subsequent waves
    - establishments going out of operation

- **Step 2: Selection of individual data (from IEB)**
  - all individuals, employed at one of these selected establishments within 1999 and 2009 at least one day
  - for these individuals all episodes in the BeH and LeH from 1993 to 2010

Linked Establishments: 2,702 to 11,117 per year
Individuals: 1,090,728 to 1,536,665 per year
6.2 Structure of the sampling process LM 9310

Time period of establishments selected from the IAB Establishment Panel; 2000 - 2008

Period of employment of individuals in the selected establishments; 1999 - 2009

Time period of employment biographies; 1993 - 2010
6.3 Structure and size of LM 9310

→ Linkage via EID

- Storage:
  - Establishment Panel: 2 to 10 MB per year
  - Individual data: 4GB
7.1 Variables of LIAB

- **Individual data (IEB)**
  - Personal information
    - e.g. gender, family status, education and qualification
  - Information on employment and benefit receipt
    - e.g. wage, employment and occupational status
  - Establishment variables
    - e.g. economic activities, number of employees
  - Location data
    - e.g. place of residence and work
  - Generated biographical variables
    - first day and number of days in employment and establishment
    - Information on employment status and wage

- **Establishment data (BP)**
  - Information on establishments with changing topics
    - Development of employment
    - Business policy and development
    - Investments
    - In-house innovations
    - Government subsidies
    - Personnel structure
    - Vocational training and apprenticeship places
    - Recruitments and dismissals
    - Training programmes
    - Wages
    - …
7.2 Sensitive variables and extension modules

- **Sensitive variables:**
  - **Individual data**
    - nationality
    - health problems
    - sever disability status
    - place of residence and work
  - **Establishment data**
    - sub-classes of economic activity
    - type of local authority
    - district key number
    - place of work

→ due to data protection explicit justification in the application process

- **Extension files BHP:**
  - **Stock data**
    - establishment characteristics, employement structure
  - **Worker flows**
    - information about entries and exits of employees
  - **Entry and exit**
    - information about establishment entries and exits
8. Research with LIAB datasets

- **wage-specific research**
  - Peichl et al. (2012): “Accounting for labor demand effects in structural labor supply models.” in Labour Economics
  - Blien et. al (2013): “The institutional context of an 'empirical law': The wage curve under different regimes of collective bargaining.” in British Journal of Industrial Relations

- **age and gender specific research**
  - Göbel et al. (2013): “Are personnel measures effective in increasing productivity of old workers?” in Labour Economics
9. Data access

- On-site use and remote data access afterwards
- in Nuremberg or in further locations
9. Further information

- **Working tools**
  - Datenreport and Methodenreport of LIAB, BP and BHP
  - Frequencies and Test data

- **http://fdz.iab.de**

- **contact in the FDZ:**
  - Dr. Jörg Heining
  - Wolfram Klosterhuber
  - Contact via iab.fdz@iab.de
Thank you for your attention!

Wolfram Klosterhuber

wolfram.klosterhuber@iab.de

www.iab.de
http://fdz.iab.de