DATA AND DATA ACCESS
AT THE RESEARCH DATA CENTER

Universidad Carlos III de Madrid
April 11, 2019

Dana Müller
AGENDA

• BA, IAB and FDZ
• Administrative Data Sources
• Selected Data Products
• Data Access
• Additional projects
FEDERAL EMPLOYMENT AGENCY AND IAB

- **BA: Federal Employment Agency**
  - Job counselling and placement services
  - Unemployment insurance benefits
  - Welfare benefits
  - Active labor market policies
  - Labor market reporting and statistics
  - Collects data on social security contributions, unemployment benefits, …

- **IAB: Institute for Employment Research**
  - Research department of the BA
  - Scientifically independent
  - Labor market and occupational research
  - Political consulting
  - Research data for academic purposes
RESEARCH DATA CENTER (FDZ)

- Established in 2004 with a funding period of 3 years
  - Permanent department since 2007
- What do we do?
  - access to confidential micro data to external researchers for academic purposes
  - standardized data products
  - link administrative labour market data to (our own) survey data
  - write documentation for our data products
  - advice our data users during data selection, application and analysis
  - support remote access centers in Germany and abroad
  - do our own research
CORE FUNCTIONS OF THE FDZ

Data Production

Service

Research
NUMBER OF USERS AND USERS’ RESEARCH PROJECTS

<table>
<thead>
<tr>
<th>Month</th>
<th>Users</th>
<th>Projects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jun 2016</td>
<td>1003</td>
<td>470</td>
</tr>
<tr>
<td>Dec 2016</td>
<td>1036</td>
<td>483</td>
</tr>
<tr>
<td>Jun 2017</td>
<td>1150</td>
<td>539</td>
</tr>
<tr>
<td>Dec 2017</td>
<td>1242</td>
<td>575</td>
</tr>
<tr>
<td>Jun 2018</td>
<td>1241</td>
<td>591</td>
</tr>
<tr>
<td>Dec 2018</td>
<td>1280</td>
<td>613</td>
</tr>
</tbody>
</table>

Source: FDZ, own calculations
DATA AVAILABLE AT THE FDZ

- Micro labor market data on
  - individuals
  - households
  - Establishments
- Overview of available data on our website

Social Security Notifications

Process-generated data of the BA

Data products of the FDZ

Surveys
ADMINISTRATIVE DATA SOURCES
NOTIFICATION PROCEDURE TO THE SOCIAL SECURITY INSTITUTIONS

- Social Security Number
- Establishment Number
- Last Name
- First Name
- Address
- Reason for Notification
- Period of Employment
- Nationality
- School Education/ Vocational Training
- Type of Employment
- Wages
- Occupation (since 2011: new classification)
- Industry code
EMPLOYMENT AGENCIES

• Data from operative processes of BA:
  – Registered unemployment
  – Registered job search
  – Participation in labor market programs
  – Information on earnings replacement/unemployment benefit receipt
  – Information on welfare

• Daily information on status at the individual level
MAIN DATABASE –
THE INTEGRATED EMPLOYMENT BIOGRAPHIES (IEB)

Data and Data Access at the Research Data Centre // Page 11
DATA AVAILABLE AT FDZ - OVERVIEW

External Data

Surveys

Data and Data Access at the Research Data Centre // Page 12
DATA AVAILABLE AT FDZ (SELECTION)
DATA AVAILABLE AT FDZ - OVERVIEW

The diagram illustrates the data and data access at the Research Data Centre. It categorizes data into two main types: BA/IAB data and FDZ data.

**BA/IAB data** includes:
- Employment History
- Benefit Recipient History
- Participation in labor market programmes
- Jobseeker History
- Integrated Employment Biographies

**FDZ data** includes:
- Establishment History Panel (BHP)
- Sample of Integrated Labor Market Biographies (SIAB)
- Linked Employer-Employee Data (LIAB)
- IAB Establishment Panel (IABB)
- National Education Panel Study (NEPS-SC6-ADIAB)
- Panel Study 'Labor Market and Social Security' (PASS)

Surveys and external data sources are also indicated in the diagram.
ESTABLISHMENT HISTORY PANEL (BHP)

- Repeated cross sectional dataset on all establishments with
  - at least one employee liable to social security (until 1998)
  - and/or at least one marginal worker (since 1999)
- 50% random sample
- Time reference: June 30
- 640,000-1.5 million establishments p.a.
BHP TOPICS

- Establishment characteristics (location, industry)
- Employment structure (Total, male/female, FT/PT/marginal, Skill structure, task requirement structure, Age structure, Nationality, Occupational structure)
- Wage structure
- Worker inflows and outflows
- Classification of establishment entry and exit
- Bankruptcies (new!)
Big sister of the BHP (coming soon)

Cross sectional dataset on labor market flows and stocks for the universe of German establishments with at least one (full-time) employee liable to social security

50% random sample

Time reference:
- Quarterly frequency: Last day of quarter (March 31, June 30, Oct. 31, Dec. 31)
- Annual frequency: June 30

1975-2014 West Germany, 1992-2014 East Germany

669-931 (577-912) thousand establishments per quarter (per year)
AWFP TOPICS AND DIFFERENCE BETWEEN BHP AND AWFP

• AWFP Topics
  – Establishment characteristics (location, industry)
  – Employment structure (full-time, skill structure, mean age, mean tenure)
  – Wage structure
  – Worker inflows and outflows

• Differences between BHP and AWFP
  – Availability at the quarterly frequency
  – Contains information for: partitions of the labor force according to employee characteristics and sub-groups of employees
  – Disaggregated flow information, e.g. inflows from / outflows to employment
  – Disaggregated wage information (mean daily wage for inflows/stayers/outflows)
DATA AVAILABLE AT FDZ - OVERVIEW
SAMPLE OF INTEGRATED LABOR MARKET BIOGRAPHIES (SIAB)

• Individual labor market biographies based on social security records and BA process data
• 2% random sample of the full population (based on individuals)
• Time reference: daily
• 1975-2014 West, 1992-2014 East
• 1,757,926 individuals
• Information on employment (since 1975), benefit receipt (SGB III since 1975, SGB II since 2005), registered job search (since 1997) and participation in labor market programs (since 2000)
SIAB TOPICS

- Socio-demographic characteristics (Gender, year of birth, nationality, marital status, number of children, schooling and vocational training, ….)

- Information on employment, benefit receipt and job search (Begin and end of employment, benefit receipt and job search, daily wage, occupation, full-time/part-time, employment status, begin/duration of unemployment, …)

- Information on location (Place of residence, place of work [federal state, district])

- Information on establishment (Branch of industry, total number of employees, employment structure, average wage at the establishment, …)
DATA AVAILABLE AT FDZ - OVERVIEW

Social security notifications -> Process-generated data

BA/IAB data:
- Employment History
- Benefit Recipient History
- Participation in labor market programmes
- Jobseeker History

Integrated Employment Biographies

FDZ data:
- Establishment History Panel (BHP)
- Sample of Integrated Labor Market Biographies (SIAB)
- Linked Employer-Employee Data (LIAB)
- IAB Establishment Panel (IABB)
- National Education Panel Study
  (NEPS-SC6-ADIAB)
- Panel Study 'Labor Market and Social Security' (PASS)

External Data

Surveys
IAB ESTABLISHMENT PANEL (BP)

- Annual representative survey (1993 - 2017; since 1996 in East Germany)
- Approx. 16,000 establishments per year
- Establishments with at least one employee liable to social security on June 30th the year before
- Stratified according to establishment size, industry and federal state
- General company information, development of employment, business policy and development, investments, in-house innovations, government subsidies, personnel structure, vocational training and apprenticeship places, recruitments and dismissals, personnel search, wages, working hours, training programmes
- Alternating special focus themes (financial crisis, job security, minimum wage)
DATA AVAILABLE AT FDZ - OVERVIEW
LINKED-EMPLOYER-EMPLOYEE DATA (LIAB)

- Worker information from the IEB database are linked to the IAB Establishment Panel survey data.

- Two versions of the LIAB data are available:
  - **Cross-sectional model**: all establishment surveys combined with administrative worker data as of June 30th (linked establishments: 4,188 – 14,981 per year; workers: 1,629,542 – 2,584,520 per year)
  - **Longitudinal model**: selection of repeatedly interviewed establishments combined with longitudinal worker biographies (linked establishments: 2,702 – 11,117 per year; workers: 1,090,728 – 1,536,665 per year)
DATA AVAILABLE AT FDZ - OVERVIEW
PANEL ‘LABOR MARKET AND SOCIAL SECURITY’ (PASS)

• Annual household survey since 2006/2007
• Labor market and poverty situation in Germany
• Situation of recipients of benefits in accordance with SGB II
• Recipients of Unemployment Benefit II (“Hartz IV”))
• Two random samples from two partial populations:
  – Persons and households in receipt of Unemployment II Benefits
  – All households with persons registered as residents of Germany
• 11 waves are currently available
• Combined data: PASS-ADIAB
DATA AVAILABLE AT FDZ - OVERVIEW
NATIONAL EDUCATIONAL PANEL STUDY (NEPS)

- Longitudinal survey by Leibniz Institute for Educational Trajectories e.V.
- Examine educational trajectories in different contexts and the development of competencies throughout the life course
- Multi-cohort-sequence study (six cohorts)
- Starting Cohort 6 (SC6)
  - “Adult Education and Lifelong Learning”
  - Based on the previous ALWA study
  - Pillars: Competence development, learning environments, educational decisions, migration background, returns to education
NATIONAL EDUCATIONAL PANEL STUDY (NEPS)

- Currently 8 waves between 2007 and 2016
  - Ca. 7,000 to 14,000 respondents per wave
- Target population: Working age (born between 1944 and 1986)
- Linkage between NEPS SC6 survey and IAB admin data
  - Record linkage based on name, address, sex and birthdate
  - 92% linkage consent rate for waves 1-5
  - 90% successful linkage (12,660 cases)
- Admin. biographies: 1975-2014
- Survey-information on biography, education, further education and competency tests can be combined with data from Establishment History Panel and long run labor market biographies
IAB-SOEP MIGRATION SAMPLE (IAB-SOEP MIG)

• Household survey (joint collaboration with DIW Berlin)
• It is a stratified random sample of selected regional units, based on the administrative records in the IEB
• Personal and household questionnaire (all family members aged 16 years or older)
  – questionnaires are similar to the survey “Living in Germany” from the GSOEP at DIW Berlin
• Five waves (2013/2014) with 3,400/5,000 persons in 2,700/1,900 households
• Information on migration history, education history, employment history, labour market background, return migration, Remittances to home countries; life satisfaction; risk preferences; social integration and acceptance
• Linked with administrative data (1975-2014)
OTHER RESEARCH DATASETS

ACCESS MODES

- Raw microdata with direct identifier
- De-identified microdata
- de-facto anonymized microdata
- Anonymized microdata

Availability

Degree of analysis potential

On-site

Off-site
ACCESS MODES

**Campus File, Scientific Use File**

On-site

Remote data processing

FDZ

Download
DATA ACCESS STEPS

1. Non-technical project proposal (http://fdz.iab.de/en.aspx)
   - Advice: One research question, One research goal, Choose one dataset
   - For SUF applications: data security concept
2. Approval by FDZ
3. Use agreement with the institution of the researcher
   - No fees
   - Duration depends on demand and clarity of research proposal
4. Waiting for passwords – you will get it via mail (client certificate - SUF, user name via email)
5. On-site use (nearest data access point Palaiseau in France)
   - Make appointment well in advance
   - Availability depends on demand
6. Off-site use via JoSuA
REMOTE ACCESS CENTRES
REMOTE DATA PROCESSING

- Researchers prepare scripts (on-site or with artificial data sets) → Guidelines
- Scripts are submitted to FDZ via web interface
- JoSuA: Application to automate remote data processing
- Developed by the Institute for the Study of Labor (IZA)
- Processing of jobs in a secure environment (using Stata)
- Provision of output files after disclosure review
THE WORKFLOW OF REMOTE DATA ACCESS WITH JOSUA

1. Step:
   Test the operability of the programs at your own working place

2. Step:
   View, check and select the output in JoSuA (Mode Internal Use)

3. Step:
   Generate selected results in mode Presentation/Publication to obtain manually reviewed output

- Download artificial test data sets from FDZ homepage
- Take a look at our template Do-Files
RESEARCH PROCESS

DATA: BY THE NUMBERS

NUMBER OF YEARS TO GET DATA: 3

YES! FINALLY!

NUMBER OF YEARS TO INTERPRET DATA: 2

what does it all mean??

NUMBER OF YEARS TO WRITE ABOUT DATA: 1.5

blah blah blah blah...

NUMBER OF SLIDES TO PRESENT DATA: 1

RESULTS

that's it?

www.phdcomics.com

weeks
JOSUA’S TWO JOB SUBMISSION MODES:
INTERNAL USE VS. PRESENTATION/ PUBLICATION

Result Window for Internal Use Mode
Cannot be downloaded

Results file in Presentation/Publication mode
Can be downloaded after manual disclosure review
ADDITIONAL PROJECTS
CURRENT PROJECTS

Data production

Service (incl. Infrastructure)

Research (sometimes combined with data production)
THANK YOU!

Dana Müller
CONTACT US

+49-(0)911-179-1752
iab.fdz@iab.de
http://fdz.iab.de/en.aspx
SOON TO COME…

- Update of SIAB (1975 - 2017) scheduled for 2nd quarter of 2019
- Update of LIAB scheduled for 3rd quarter of 2019
- Release of „Administrative Wage and Labor Market Flow Panel (AWFP)“
- New datasets currently developed
  - Sample of Integrated Basic Security Biographies (SIG)
  - Linked Employer Employee dataset without survey part
SAMPLE OF INTEGRATED WELFARE BENEFIT BIOGRAPHIES (SIG)

- Individual labor market biographies based on social security records and BA process data
- Focus lies on welfare benefit recipients (Social Code Book II) and their household members
- Random sample from the full population of welfare recipients between 2007 and 2017
- …enriched by information on all household members during that time
- Time reference: daily
- Information on employment, benefit receipt (SGB III & SGB II), registered job search, participation in labor market programs, and financial information on entitlements, sanctions, etc.
- Release: 3th quarter 2019
SIG STRUCTURE AND TOPICS

Unemployment Benefit II Recipient History (LHG) 2005 – 2017

Information on employment

Sample of recipients

Information on unemployment, job search and measure participation

Information on employment

Information on members of benefit commun. (employment, unemployment, etc.)

Financial information (requirements, entitlements, income, and sanctions)

Information about establishments

SIG core

Extension modules