KombiFiD - Combined Firm Data for Germany

Anja Spengler, Tanja Hethey
Institute for Employment Research, Nuremberg

Co-operation partner are:
Institute for Employment Research
Research Data Centre of the Federal Statistical Office of Germany
Institute of Economics of the University of Lüneburg
German Central Bank
Mainz University of Applied Sciences
Motivation

Separate Data Producers:

German Federal Employment Agency

Federal Statistical Office of Germany

Deutsche Bundesbank
Motivation

Sorts of datasets:
- Survey
- Administrative Data

Data access: possible at the Research Data Centres (RDC)
- On-Site Use
- Remote Data Access
- Scientific Use Files
New

Combining firm datasets of the single data producers for the first time in Germany; time horizon 1993 - 2006

New dataset for researchers with:
• a variety of variables,
• a variety of data generating processes
• a variety of economic sectors

Reduce respondent burden for the firms
Innovative

Four years ago:
No possibility to match datasets and also no data exchange between the single data producers
Prohibited by law

Today:
Change of the Federal Statistics Act
Matching is (technical) possible

Combined Firm Data are unprecedented until KombiFiD
KombiFiD - approach

Federal Data Protection Act:
match data with a written agreement only

„Survey“:
60.000 firms
Only one question

Matching ...
the data of firms, which gave their written agreement
Our Challenges

Unit non-response:
Number of answering firms?

Matching difficulties:
Datasets with different levels of aggregation!
(establishment, firm)
Data consistency

Legislative initiative
Selected datasets

1. Federal german business register (URS)

- Implemented in late 90s
- Regulated by european law (Council regulation No 2186/93)
- Master file

- Identifier:
  - Unique business identifier
  - All corresponding establishment identifiers
  - All corresponding tax numbers
2. Mix of official (sample) surveys

- Cost structure surveys
  Information about added value, turnover, subsidies,…

- Salary and wage-earning set up surveys
  Information about wages, salaries, working hours, vacation entitlement

3. Tax data

- Turnover tax
- Corporate income tax
1. **Process-generated data:**

Personal data from the employment notification
annual aggregation on establishment level (establishment number)

**Establishment-History-Panel (BHP)**

2. **Survey data:**

*IAB Establishment Panel*

- annual panel (10,000 establishments)
- Conducted by the IAB since 1993
- Information about various and changing topics (e.g. labour demand)
Process generated data:

- **Microdatabase Direct Investment (MiDi)**
  annual firm data on foreign direct investment stocks
due to German trade and payment regulation

- **Corporate Balance Sheet Statistics (Ustan)**
  annual balance sheet statistics of non-financial enterprises
  (main economic sectors: manufacturing, construction,
  wholesale and retail trade)
Establishment-history-panel
IAB Establishment Panel

Tax data
Official (sample) surveys
Corporate balance sheet statistics

German business register

KombiFiD dataset
KombiFiD Dataset

Dataset features:

- variety of variables
- variety of data generating processes
- variety of economic sectors (Manufacturing, Construction, Wholesale and retail trade, Service Sector)

enhanced firm tracking:

- identifying firm entries, exits
- identifying changes in firm structures (mergers, aquisitions)
Key dates

- **May 2008**: KombiFiD survey
- **Summer '08**: Kick-off workshop
- **Spring '09**: Call for papers
- **Summer '09**: First release, Notification of accepted papers
- **Summer '10**: User conference
- **End of 2010**: Final release

Data access:
- On-site use
- Remote data access
- Funds for foreign researchers
Further information

KombiFiD webpage:

www.kombifid.de (under construction)

Contact:

anja.spengler@iab.de
tanja.hethey@iab.de
Thank you for your attention
Back up
Three ways of data access

**On-Site Use:** At this researchers may analyze weakly anonymous data at separate workplaces for guest researchers during a research visit.

**Remote Data Access:** Therefore researchers develop programs on the basis of test data. These programs are run through the original data at the research data centre. After verification of compliance with data protection legislation the results are sent to the researcher.

**Scientific Use Files:** This are factually anonymous datasets offered for researchers of scientific institutions for analysis.