

This readme file provides basic information on the data sources, variables and do-files used in Dummert, S.; Leber, U. and Schwengler, B., Unfilled training positions in Germany – regional and establishment-specific determinants, forthcoming in Journal of Economics and Statistics.

Data sources:

The data of the Institute for Employment Research (IAB) are available via the Research Data Centre (FDZ) of the German Federal Employment Agency (BA) at the IAB. The FDZ provides researchers access to micro data for non-commercial empirical research in the fields of social security and employment.

We used the following datasets:

1.) The cross-sectional version of the Linked Employer-Employee dataset (LIAB QM) from the Institute for Employment Research (IAB)

- ➔ **Brief data description:** The Linked Employer-Employee Data from the IAB (LIAB) allow for simultaneous analysis of the supply and demand sides of the German labour market. For this purpose, the IAB Establishment Panel data are matched with individual data from the IAB, thus providing a linked dataset which comprises data from a representative annual establishment survey as well as personal data generated in labour administration and social security data processing.
- ➔ The LIAB QM is available via On-site Use (incl. Remote Data Access)
- ➔ Detailed information about the dataset and data access can be found here: https://fdz.iab.de/en/Integrated_Establishment_and_Individual_Data/LIAB.aspx

2.) The Establishment History Panel (BHP) (Full sample)

- ➔ **Brief data description:** The Establishment History Panel (BHP) is composed of cross sectional datasets since 1975 for West Germany and 1992 for East Germany. Every cross section contains all the establishments in Germany which are covered by the IAB Employment History (BeH) on June 30th. These are all establishments with at least one employee liable to social security on the reference date. Establishments with no employee liable to social security but with at least one marginal part-time employee are included since 1999. The cross sections can be combined to form a panel.
- ➔ The BHP is available as a 50% simple random sample at the FDZ via Remote Data Access and On-site Use (incl. Remote Data Access)
- ➔ Detailed information about the dataset and data access can be found here: https://fdz.iab.de/en/FDZ_Establishment_Data/Establishment_History_Panel.aspx

Regional variables:

- 1.) Unemployment rate of young people (15-24 years):
→ Statistics of the Federal Employment Agency (on request), date of creation: 20.06.2018

- 2.) Number of school leavers:
→ The Federal Statistical Office and the statistical offices of the Länder: The Regional Database Germany (www.regionalstatistik.de)

- 3.) Area in square kilometers:
→ The Federal Statistical Office and the statistical offices of the Länder: The Regional Database Germany (www.regionalstatistik.de)

- 4.) Number of basic / intermediate secondary level graduates:
→ The Federal Statistical Office and the statistical offices of the Länder: The Regional Database Germany (www.regionalstatistik.de)

Do-Files:

The do-files provided are:

- 1.) Preparation.do to merge and prepare the data set and

- 2.) Multilevel.do for the descriptive statistics and the multilevel mixed-effects regressions.