
Replication Manual for “Changes in Workplace Heterogeneity and How They Widen the Gender Wage Gap” by Benjamin Bruns

The main data used in the paper is the LIAB Mover Modell 9308 (Version 1, 1993-2008). It is a matched employer-employee data set provided by the Research Data Centre (FDZ) of the German Federal Employment Agency (BA) at the Institute for Employment Research (IAB). The data set can be accessed via on-site use and remote data access. I am not able to make this data set available. However, researchers can get access to the data by submitting a written application to the FDZ. Further details on the application procedure can be found online at http://fdz.iab.de/de/FDZ_Data_Access/FDZ_On-Site_Use.aspx.

Subject to being granted access to the data, the results in the manuscript can be replicated following the steps below. All programmes are in Stata-language (*.do) and were executed in Stata 13.

1. The data archive contains two folders:

- dofile
- ado

The “dofile”-folder includes all programme needed to replicate the results. The “ado”-folder includes several ado-files that are not available on the SSC archive, but used in the analyses.

2. The data paths are referenced to via global macros. Define the following globals, e.g., in a profile-dofile or in the master-dofile.
 - “orig” – path to folder with original data sets of the IAB
 - “prog” – path to dofile-folder
 - “data” – path to folder with using data
3. The dofiles can be run via the master.do, which also contains a short description of the main content of the dofile. Here is a brief overview what the programmes do. The dofiles contain more detailed comments on the procedures.

- 01_extract_iab_ep.do

Merge raw data files of the IAB-Establishment Panel and create consistent variable names across survey waves. The basic dofile is provided by the IAB. I extracted further variables as indicated in the programme.

- 02_merge_and_clean_iab_ep.do

Constructs firm level unbalanced panel data set. Some imputation steps are taken to fill missing observations.

- 03_prepare_ieb.do

Prepares the worker data set and merges the firm level variables.

- 04_impute_education.do

Runs the IP1 imputation procedure of Fitzenberger, Osikominu, and Voelter (2005)

- 05_addvars.do

Creates additional variables for the wage imputation.

- 06_impute_wages.do

Runs the wage imputation procedure by year, education, age group, and gender.

- 07_akm_build.do

Builds the AKM-estimation samples

- 08_estimate_akms.do

Runs the AKM and associated match-effects models. Saves the parameter estimates in separate output files (e.g., Table 2)

-
- 09_descriptives.do
Prepares summary statistics for various subsamples (e.g., Table 1)
 - 10_stylised_facts.do
Prepares stylised facts of gender wage gaps, wage inequality, and productivity (e.g., Figure 1 & 2)
 - 11_eventstudy.do
Performs the event study analyses of CHK and CCK. Output is contained in the Online Appendix.
 - 12_binning_fe.do
Estimates average AKM firm effects for men and women across the distribution of productivity e.g., Figure 3, Figure 6)
 - 13_gridsearch.do
Estimates the kinkpoint above which rent-sharing is assumed to start by searching for the minimum value of the sum of squared errors from the nonlinear system of equations (e.g., Figure 3).
 - 14_decompositions.do
Runs the baseline decompositions for various subsamples (e.g., Table 3, Figure 4, Figure 7, and Online Appendix material)
 - 15_dfl_wagegaps.do
Performs the DFL-decomposition of sorting and bargaining over time (e.g., Table 4)
 - 16_stayermodels.do
Analyses the relationship between wages and productivity for men and women using within-firm variation (e.g., Table 5)
 - 17_rentsharing_models.do
Estimates rent-sharing models à la CCHK using the AKM-wage decomposition (Online Appendix material)
 - 18_unions_rentsharing.do
Summarise changes in union coverage and the relationship between union coverage and firm fixed effects (e.g., Table 6)
 - 19_dfl_unions.do
Estimation of the DFL decomposition of the AKM firm effect gap conditional on the distribution of union coverage over time and between genders (e.g., Table 7)
 - 20_akms_lowhigh.do
Re-estimate AKM models for subsamples with high/low union coverage (Online Appendix material)
 - 21_identify_childbirth.do
Algorithm to identify childbirth in IEB data (based on Dana Mueller and Katharina Strauch)
 - 22_analyse_childbirth.do
Analyse the impact of childbirth on the wages and firm premiums of women using an event study design à la Kleven et al. (2018) (e.g., Figure 8, Table 8).
 - 23_gelbach.do
-

Performs Gelbach's exact decomposition based on gender pooled AKM models (Online Appendix material).

Note that figures and tables were prepared in Excel using Stata's output.