Replication Files „Freedom of Enterprise and Economic Development in the German Industrial Take-Off“ – Read me

Martin Dohmen[1]

1. Overview of files

This read me file contains an overview of all folders and files in the replication files. I ran all codes using Stata 14. To run the codes, just alter the pathes in the paths.do file and the master.do file and run the master.do. All Tables and Figures used in the paper will be generated automatically.[1]

1.1. Folder Structure

The replication files contain the following folders:

File	Purpose
data_input	Contains all data sets needed to run the code files and generate the Tables and Figures.
Figures_output	All Figures generated by the code files and used in the paper (Figure 3) will be saved here.
intermediate_stata_data	All intermediate data sets (in ".dta" format) generated by the code files will be saved here.
stata_code	This folder contains all necessary Stata code files.
Tables_output	All Tables generated by the code files and used in the paper will be saved here.

[1] As long as all necessary packages like *geoneor* or *geodist* for Stata are installed.

1.2. Code Files for Stata

All code files are located in the folder "stata_code".

File	Purpose
paths.do	Define working path for entire program
master.do	Runs all program codes in right order to generate all results - only this code needs to be run
merge_iPEHD_data.do	Load all data from iPEHD data base for Prussia and the additional variables for the Electorate on county level
add_Prussia_additional _controls_and_data _Hessen_Kassel.do	Load data for counties in the Electorate and additional control variables for Prussian counties generated by hand
summary_statistics_main _county_data.do	Generate summary statistics for main data set used in later regressions (Table 2)
DiD_regressions _with_panel_data.do	Run main difference-in-difference panel regression to generate Table 3 and Table C.8-C.10
quantile_regressions.do	Run quantile regressions as robustness test and generate Table C.11
alternative_outcome _regressions _kingdom_westphalia.do	Run regressions for former Kingdom of Westphalia using alternative outcome variables and generate Table 4
regressions_external_validity _incl_Nassau_DAR.do	Run regressions for inclusion of additional states and generate Table 6
regressions_RDD_munici pality_level_border_nw.do	Calculate border distance, run regressions for RDD specification and generate Figure 3, Table 5 and Table C.12

1.3. Input Data Sets

All input data sets are located in the folder "data_input". The following table lists the input data sets containing the raw data in the order used in the code files.

File	Purpose
data_ipehd_*.csv	Data from the ifo Prussian Economic History Database (iPEHD; Becker et al. (2014)), downloaded at 21.10.2021, for details please refer to the data appendix and the documentation for iPEHD (https://www.ifo.de/en/iPEHD).
data_ipehd_merge _county.csv	Data set that helps merging the Prussian county-level data from iPEHD over time, for details please refer to the data appendix and the documentation for iPEHD (https://www.ifo.de/en/iPEHD).
data_Prussia_1837 _Sachsen_Westphalen _Rheinprovinz.xlsx	Data set containing population data for Prussian counties for 1837.
data_Hesse_plus_extra _variables_Prussia.xlsx	Data set containing the population and area data for counties in the Electorate with additional location controls (sheet "data_Hessen_Kassel") and the same additional location controls for counties in Prussia (sheet "additional_controls_Prussia").
data_county_level _NAS_DAR.xlsx	Data set containing the population and area data for counties in the Duchy of Hesse and Nassau.
data_municipalities_border _Electorate_Minden _Pru.xlsx	Data set containing the population data for municipalities in Prussia and the Electorate that are close to the common border in the north-west with Dummies indicating to which former Kanton the municipalities belonged.

Additionally, the within the folder "data_input", the folder "data_geocoding_results" contains Stata data sets containing the latitude and longitude of counties and municipalities. These were generated using the Stata package *opencagegeo* that relies on *OpenCage Geocoder API*. I rely on the county name for geocoding, which in most cases coincides with the name of the administrative centre of the county, and add information about the current federal state and country (some counties in the Prussian data are located today in Belgium). Only in unclear cases, when either the name of the county changed over time or does not point to a specific city, I instead use the name of the administrative centre of the former county. To geocode municipality data, I use the name of the municipality and additionally add the county name. I correct obvious mistakes using Google Maps. Details about the geocoded data sets are listed in the table below:

Moreover, the folder "data_input" contains a subfolder "data_Landesgrenzen_shape",

File	Purpose
dataset_Prussia_Hessen _Kassel_lat_long.dta	County-level data of latitude and longitude of Prussian counties used in main specification and counties in the Electorate.
dataset_Prussia_KAS_full _lat_long.dta	County-level data of latitude and longitude of all Prussian counties and counties in the Electorate.
dataset_NAS_DAR _lat_long.dta	County-level data of latitude and longitude of counties in the Duchy of Hesse and Nassau.
dataset_municipalities_RDD _geocodes_lat_long	Municipality-level data of latitude and longitude of municipalities in Prussia and the Electorate of Hesse located close to the common border in the north-west.

which contains a shape-file with current federal state borders in Germany. This is used to calculate the border distance of municipalities close to the border between the Electorate of Hesse and Prussia, which nowadays is the border between Hessen and North Rhine-Westphalia. Source: "Bundesamt für Kartographie und Geodäsie", date 2011; © GeoBasis-DE / BKG (2011)

References

Becker, S.O., Cinnirella, F., Hornung, E., Woessmann, L., 2014. ipehd - the ifo prussian economic history database. Historical Methods: A Journal of Quantitative and Interdisciplinary History 47, 57–66. doi:10.1080/01615440.2013.852370.