October 17, 2023

Readme: Data and Software Code for "Grain Prices in Pre-industrial Germany, Fifteenth to Nineteenth centuries"

This readme file explains the code that merges the city files of the dataset introduced in:

Hakon Albers and Ulrich Pfister: "Grain Prices in Pre-industrial Germany, Fifteenth to Nineteenth Centuries." *Vierteljahrschrift für Sozial- und Wirtschaftsgeschichte*.

When using data and/or code please cite the paper. In case of questions on data and code, which cannot be resolved with paper, Supplementary Appendix or this file, please contact the authors.

Dataset
The folder 'data.zip' includes all grain price data as Microsoft Excel files 'city_grain_r.xlsx'. Herein, 'city' is a placeholder for the corresponding names of the 70 cities. See main paper and Supplementary Appendix for further details.

Software code to aggregate data to national-level file (see Section 5.2 of main paper)
To make use of the supplementary software code 'merge_city_files.r' for information aggregation (see Section 5.2 in main paper), users need a minimum working knowledge on the software R. You need to install the software R (https://www.r-project.org/) and the package 'openxlsx' (https://ycphs.github.io/openxlsx/index.html).

We cannot guarantee any upward/downward compatibility of the code. The code files were executed without any errors on two different PCs (with Windows 10); installed R-versions were: 4.3.1, and 4.0.2, respectively. Older/newer R-versions could potentially create problems.

The following text explains how to create the national-level file 'Germany_rye.xlsx'. The code can also be adapted to create 'Germany_wheat.xlsx' etc.

Producing Germany_rye.xlsx

1. Download the 'data.zip' folder and unpack it into your working directory. Download the R-code file 'merge_city_files.r' into the same working directory.
2. Start R and make sure the working directory is set appropriately.
3. Run code: 'merge_city_files.r' with the command `source("merge_city_files.r")`. Depending on the speed of your computer this may take some time (typically below 1 minute).
4. The most important output is the file: 'Germany_rye.xlsx'. This is the national-level file containing all rye series from the city files contained in 'data'. Additional files with the same information in alternative formats include: 'Germany_rye.csv' and 'Germany_rye.Rdata'. The file 'Summary_stats_rye.csv' contains the summary statistics reported for rye in the main paper in Table 2 in Section 5.2. The file 'merge_city_files.Rdata' is the R-workspace.

Producing national-level files for other grains and further details

5. In the standard version, the code produces the dataset for rye. To produce the file for wheat, specify the grain type to 'wheat' via adjustment of the object 'grain_type' in code line 13, save, and run the code again. 'Germany_wheat.xlsx' is produced. The same procedure is possible for barley and oats.
6. Further details are available in the script files and the comments therein. One point is worth to be noted: If you like to adjust the cities that are included in the national-level dataset, this is possible by switching from procedure A to procedure B when loading the data from the city files (see line 17). Version A uses a 'whitelist' of cities (code line 21) to construct the national-level file. Version B is called 'Read files with pattern from folder data' (line 104). This first creates a list of files in 'data' corresponding to an expected pattern and then loads data from all files. With the 70 files of the dataset both approaches yield the same result. If you add city files (see details below), procedure B adds the corresponding data to the national-level file.

How to add city files

To add your own files with grain prices to the dataset and to the procedure to create aggregate-level files for individual grain types proceed as follows:

1. Place data for each city in an excel file following the naming convention 'city_grain_r.xlsx'. Add this file to the sub-folder 'data' in your working directory.
2. Make sure values in grams of silver per litre are placed in a sheet with the name '29 calendar year metric'.
3. Format of price data in sheet: '29 calendar year metric': (i) calendar year must be placed in column 1. (ii) Column names of grain prices must consist of 'barley', 'oats', 'rye' and 'wheat'. The order and the number of different types of grain price series (one to a maximum of four) are irrelevant. See files 'Aachen_grain_r.xlsx' and 'Duisburg_grain_r.xlsx' as examples.
4. Use procedure B to aggregate prices (see point 5 above). If procedure B is used, no changes to the R code are required. If you prefer procedure A, add city name(s) to the whitelist starting on line 21 of the code (correct alphabetical order is necessary).