README and Guidance

Lars Brausewetter   Stephan L. Thomsen*   Johannes Trunzer†

November 26, 2023

This README contains information about the data and R codes needed to replicate the results in the paper *"Regional Supply and Demand Fundamentals in the German Housing Price Boom"* by Lars Brausewetter, Stephan L. Thomsen, and Johannes Trunzer, published in the German Economic Review. Three master files run all of the code to generate the data for the 9 figures and 3 tables in the main paper, the 2 tables in the appendix, and the 9 figures and 25 tables in the online appendix. The replicator should expect the code to run for about eight minutes.

Data Availability and Provenance Statements

Statement about Rights

☒ We certify that the authors of the manuscript have legitimate access to and permission to use the data used in this manuscript.

Summary of Availability

☒ Some data cannot be made publicly available.

Details on each Data Source

Data Set Name	Provided	Citation
RWI-GEO-REDX (2008-06/2022, v9)	FALSE	Schaffner et al. (2022)
Regional Database Germany	TRUE	Destatis (2023)
INKAR	TRUE	BBSR Bonn (2023)
Employment Statistics	TRUE	BA Statistics (2023a, 2023b)
SchuldnerAtlas	TRUE	Creditreform (2022)
Naturgefahrenreport	FALSE	GDV (2023)
Risiko- und Kredit-Kompass	FALSE	SCHUFA (2022)
Wohnbauland-Umfragen	FALSE	NBank (2022)
BBSR-Wohnungsmarktbeobachtung	FALSE	BBSR Bonn (2020)
GREIX	FALSE	Amaral et al. (2023)
Ahlfeldt et al. (2023) Index	FALSE	Ahlfeldt et al. (2023)
Time on the Market	TRUE	ImmoScout24 (2019a, 2019b, 2019c)
Geodata	TRUE	GeoBasis-DE/BKG (2018)

RWI-GEO-REDX (2008-06/2022, v9) Dataset

- The paper uses annual housing price data from the RWI-GEO-REDX (2008-06/2022, v9) data set. The annual data are open for public use. Interested users need to register for data access via fdz@rwi.essen.de, including information about the applying department or person as well as the desired data format.

---

*Leibniz University Hannover, IZA Bonn, and ZEW Mannheim
†Leibniz University Hannover

- To ensure that the indices are not driven by small sample sizes, the data set only covers those indices that rely on at least 50 observations per year and region. Therefore, the information is limited for smaller districts.

  → For a detailed description of the data source, see Klick and Schaffner (2019).

- After granted access to the data set, save the files in the directory `data/RWIGEOREDX`.

→ Datafiles: `RWIGEOREDX/RWIGEOREDX_ApRent_v9.xlsx`, `RWIGEOREDX/RWIGEOREDX_ApPurc_v9.xlsx`, and `RWIGEOREDX/RWIGEOREDX_HouPurc_v9.xlsx` (not provided)

Regional Data from Destatis, BBSR Bonn, the Federal Employment Agency, and Other Sources

- A large set of regional variables at the district level was downloaded from the Regional Database of the Federal Statistical Office of Germany (Destatis 2023) and the INKAR database of the Federal Institute for Research on Building, Urban Affairs and Spatial Development (BBSR Bonn 2023). Additional labor market data were provided by the Statistics of the Federal Employment Agency (BA Statistics) through special evaluations with costs (BA Statistics 2023a, 2023b). Data on private debtors is from Creditreform (2022). Mortgage lending data for German savings banks were self-collected from balance sheets published in the Federal Gazette (*Bundesanzeiger*) (as of October 2023). We then allocated the volume of mortgage loans of the individual savings banks to the district level, as the majority of the savings banks' business areas are district based (Doerrbecker 2020). A copy of the regional data is provided as part of this archive.

  → Datafile: `data/regdata/reg_data.xlsx`

- Further data were obtained upon request without the permission for distribution to third parties: data on subprime credit scores from the SCHUFA, data on hazard rates from the German Insurance Association (GDV), estimated data on housing vacancies (*Leerstandsabschätzung*) from the BBSR Bonn, and data on residential zoning provided by the *Investitions- und Förderbank Niedersachsen* (NBank). The data from the GDV and the BBSR Bonn were only provided in intervals, which we replaced with the interval mean and the upper and lower bounds, respectively. In order to get access to these data, please contact the individual institutions. Once the raw data are obtained, the replicator has to extract the respective variables as defined in the paper (see Appendix Table A-1), bring them into the appropriate shape (krs × year), and linearly interpolate a few missing values. Name the variables as specified in the script where the panel data set is built (`programs/01_dataprep/01_panel_krs.R`).

  → Datafile: `data/regdata/data_descr.xlsx` gives an overview about all used indicators, their definitions, and sources.

Alternative Price Indices and Additional Housing Market Information

- For robustness checks, we use two alternative price indices. The theory-consistent index by Ahlfeldt et al. (2023) is publicly available at the FDZ Ruhr and can be found here. After downloading the dataset, save the files for purchase prices and rents at the district level in the directory `data/ahlfeldt_index`.

  → Datafiles in `.../INDEX_AHL_IMMO_county_PURCH.xlsx` and `.../INDEX_AHL_IMMO_county_RENT.xlsx` (not provided)

- The German Real Estate Index (GREIX) by Amaral et al. (2023) can be downloaded here. Please use the following settings:

  - Property type: Single-family house, Multi-family house, Apartment
  - Data type: Indexes
  - Data frequency: Yearly
  - Inflation adjusted: No
  - Time period: 2008-2021
  - Data format: Excel

- Download separate files for each of the following cities: Berlin, Frankfurt, Dortmund, Cologne, Duesseldorf, Muenster, Duisburg, Wiesbaden, Stuttgart, Munich, Dresden, Hamburg, Luebeck, Chemnitz, Bonn, Potsdam, Leipzig, Erfurt. Name all files accordingly and save them in the directory `data/GREIX/download/`. We provide a key to match the cities to the district level.

  → Datafiles: `data/GREIX/download/` (not provided) and `data/keys/GREIX_IDs.xlsx`

- Furthermore, we provide a copy of data on time on the market, published on the webpage of ImmoScout24 for a selected set of districts (ImmoScout24 2019a, 2019b, 2019c).

  → Datafiles: `data/regdata/TOM.xlsx`

Geodata and Labor Market Region Keys

- Geodata on German administrative areas (VG250) is derived from GeoBasis-DE/BKG (2018). The data can be freely downloaded here. We provide a copy of the data in UTM32 projection (vg250_12-31.utm32s.shape.ebenen) as part of this archive.

  → Datafiles in `data/geodata/vg250_12-31.utm32s.shape.ebenen`

- Finally, we provide a key file that assigns counties to 182 labor market regions (AMR1) defined by RWI (2018).

  → Datafiles: `data/keys/krs_amr_rwi_key.xlsx`

Dataset List

Table 2: Dataset list

Data file	Source	Notes	Provided
`data/RWIGEOREDX/` `RWIGEOREDX_ApRent_v9.xlsx`	Schaffner et al. (2022)	Public use	No
`data/RWIGEOREDX/` `RWIGEOREDX_ApPurc_v9.xlsx`	Schaffner et al. (2022)	Public use	No
`data/RWIGEOREDX/` `RWIGEOREDX_HouPurc_v9.xlsx`	Schaffner et al. (2022)	Public use	No
`data/regdata/reg_data.xlsx`	see `data_descr.xlsx`	-	Yes
`data/regdata/TOM.xlsx`	ImmoScout24 (2019a, 2019b, 2019c)	Public use	Yes
`data/GREIX/download/...`	Amaral et al. (2023)	Public use	No
`data/ahlfeldt_index/` `INDEX_AHL_IMMO_county_PURCH.xlsx`	Ahlfeldt et al. (2023)	Public use	No
`data/ahlfeldt_index/` `INDEX_AHL_IMMO_county_RENT.xlsx`	Ahlfeldt et al. (2023)	Public use	No
`data/geodata/` `vg250_12-31.utm32s.shape.ebenen`	GeoBasis-DE/BKG (2018)	Public use	Yes
`data/keys/krs_amr_rwi_key.xlsx`	RWI (2018)	Public use	Yes
`data/keys/GREIX_IDs.xlsx`	Own assignment	Public use	Yes

Computational Requirements

Software Requirements

- R (code was last run with version 4.2.2 (2022-10-31 ucrt)). The following packages have to be installed:
    - `binsreg` (0.9)
    - `ggh4x` (0.2.4)

- ggpubr (0.6.0)
- ggtext (0.1.2)
- glmnet (4.1-7)
- grid (4.2.2)
- lmtest (0.9-40)
- openxlsx (4.2.5.2)
- plm (2.6-3)
- plotmo (3.6.2)
- rgdal (1.6-6)
- sf (1.0-12)
- spdep (1.2-8)
- stargazer (5.2.3)
- tidyverse (2.0.0)
- tmap (3.3-3)
- zoo (1.8-12)

$\rightarrow$ the file `00_setup.R` will install all dependencies (latest version), and should be run once prior to running other programs.

Memory and Runtime Requirements

Summary Approximate time needed to reproduce the analyses on a standard (2023) desktop machine:

☒ <10 minutes
☐ 10-60 minutes
☐ 1-8 hours
☐ 8-24 hours
☐ 1-3 days
☐ 3-14 days
☐ > 14 days
☐ Not feasible to run on a desktop machine, as described below.

Details The code was last run on a **4-core Intel-based laptop with Windows10 Pro 64-Bit**.

Description of programs/code

- Programs in `programs/01_dataprep/` will clean the data and prepare them for analysis. It constructs the analysis panel data at the district level and a panel at the labor market region level. It also cleans the Ahlfeldt et al. index and the GREIX. The file `programs/01_dataprep/master.R` will run them all.
- Programs in `programs/02_analysis/` generate all tables and figures in the main body of the article. The file `programs/02_analysis/master.R` will run them all. Each program called from master.R identifies the table or figure it creates (e.g., table-1.R). Output files are called appropriate names (table-1.html, figure-1.png) and should be easy to correlate with the manuscript.
- Programs in `programs/03_appendix/` will generate all tables and figures in the appendix and online appendix. The file `programs/03_appendix/master.R` will run them all.

Instructions to Replicators

- Open the R-Project `GermanHousingMarket.Rproj` that will automatically adjust the default path.
- Run `programs/00_setup.R` to install all dependencies (latest version).
- Download and request the data files referenced above. Each file should be stored in the prepared subdirectories of `data/`, in the format that you download them in. However, the requested data must first be brought into the appropriate shape to merge with the data provided in this archive.
- Run `programs/01_master.R` to run all steps in sequence.

Details

- `programs/00_setup.R`: will create all output directories, install needed R packages.
- `programs/01_dataprep`:
    - If running programs individually, note that ORDER IS IMPORTANT.
    - The programs were last run top to bottom on November 26, 2023.
    - A `programs/01_dataprep/master.do` will run them all in sequence.
- `programs/02_analysis`:
    - If running programs individually, note that ORDER IS IMPORTANT.
    - The programs were last run top to bottom on November 26, 2023.
    - A `programs/02_analysis/master.do` will run them all in sequence.
- `programs/03_appendix`:
    - If running programs individually, note that ORDER IS IMPORTANT.
    - The programs were last run top to bottom on November 26, 2023.
    - A `programs/03_appendix/master.do` will run them all in sequence.

List of tables and programs

The provided code reproduces:

☒ All numbers provided in text in the paper
☒ All tables and figures in the paper
☐ Selected tables and figures in the paper, as explained and justified below.

Table 3: List of tables, figures, and programs

Figure/Table	Program	Output file
Table 1	02_analysis/table1.R	table-1.html
Table 2	02_analysis/table2.R	table-2.html
Table 3	02_analysis/table3.R	table-3.xlsx
Figure 1	02_analysis/figure-1.R	figure-1.png
Figure 2	02_analysis/figure-2.R	figure-2.png
Figure 3	02_analysis/figure-3.R	figure-3.png
Figure 4	02_analysis/figure-4.R	figure-4.png
Figure 5	02_analysis/figure-5.R	figure-5.png
Figure 6	02_analysis/figure-6.R	figure-6.png
Figure 7	02_analysis/figure-7.R	figure-7.png
Figure 8	02_analysis/figure-8.R	figure-8.png
Figure 9	02_analysis/figure-9.R	figure-9.png
Table A-1	n.a. (no data)	
Table A-2	03_appendix/table-a2.R	table-a2.html
Table B-1	03_appendix/table-b1.R	table-b1_1.xlsx; table-b1_2.xlsx; table-b1_3.xlsx
Table B-2	02_analysis/figure-3.R	table-b2_a.html; table-b2_b.html; table-b2_c.html; table-b2_d.html; table-b2_e.html; table-b2_f.html; table-b2_g.html
Table B-3	03_appendix/table-b3.R	table-b3.html
Table B-4	03_appendix/table-b4.R	table-b4.html
Figure B-1	03_appendix/figure-b1.R	figure-b1.png
Figure B-2	03_appendix/figure-b2.R	figure-b2.png
Figure B-3	03_appendix/figure-b3.R	figure-b3.png
Figure B-4	03_appendix/figure-b4.R	figure-b4_a.png; figure-b4_b.png; figure-b4_c.png
Figure B-5	03_appendix/figure-b5.R	figure-b5.png

Figure/Table	Program	Output file
Figure B-6	03_appendix/figure-b6.R	figure-b6.png
Figure B-7	03_appendix/figure-b7.R	figure-b7.png
Figure B-8	03_appendix/figure-b8.R	figure-b8.png
Table C-1	03_appendix/table-c1.R	table-c1.html
Table C-2	03_appendix/table-c2.R	table-c2.html
Table C-3	03_appendix/table-c3.R	table-c3.html
Table C-4	03_appendix/table-c4.R	table-c4.html
Table C-5	03_appendix/table-c5.R	table-c5.html
Table C-6	03_appendix/table-c6.R	table-c6.html
Table C-7	03_appendix/table-c7.R	table-c7.html
Table C-8	03_appendix/table-c8.R	table-c8.html
Table C-9	03_appendix/table-c9.R	table-c9.html
Table C-10	03_appendix/table-c10.R	table-c10.html
Table C-11	03_appendix/table-c11.R	table-c11.html
Table C-12	02_analysis/figure-4.R	table-c12.html
Table C-13	02_analysis/figure-4.R	table-c13.html
Table C-14	02_analysis/figure-4.R	table-c14.html
Table C-15	03_appendix/table-c15.R	table-c15.html
Table C-16	02_analysis/figure-4.R	table-c16.html
Table C-17	02_analysis/figure-4.R	table-c17.html
Table C-18	02_analysis/figure-4.R	table-c18.html
Table C-19	03_appendix/table-c19.R	table-c19.html
Table C-20	03_appendix/table-c20.R	table-c20.html
Table C-21	03_appendix/table-c21.R	table-c21.html
Figure C-1	03_appendix/figure-c1.R	figure-c1.png

References

Ahlfeldt, G.M., Heblich, S. and Seidel, T. (2023). Micro-geographic property price and rent indices. Regional Science and Urban Economics, 98, 103836.

Amaral, F., Dohmen, M., Schularick, M. and Zdrzalek, J. (2023). German Real Estate Index (GREIX). ECONtribute Discussion Paper, No. 231. University of Bonn and University of Cologne, Germany.

Doerrbecker, M. (2020). Geschäftsgebiete der Sparkassen und Sparkassenverbände in Deutschland (Karte). https://de.wikipedia.org/wiki/Datei:Karte_Gesch%C3%A4ftsgebiete_der_Sparkassen_und_Sparkassenverb%C3%A4nde_in_Deutschland.png (last accessed: 30 September 2023).

Federal Institute for Research on Building, Urban Affairs and Spatial Development [BBSR Bonn] (2020). BBSR-Wohnungsmarktbeobachtung, BBSR-Abschätzung des Wohnungsleerstands (Berechnungsverfahren zum Stand 31.1.2020), basierend auf Leerstandsdaten des Zensus 2011, Gebäude- und Wohnungszählung, Stand: 28.5.2014. BBSR, Bonn.

Federal Institute for Research on Building, Urban Affairs and Spatial Development [BBSR Bonn] (2023). INKAR - Indikatoren und Karten zur Raum- und Stadtentwicklung. Federal Institute for Research on Building, Urban Affairs and Spatial Development (BBSR Bonn). https://www.inkar.de/ (last accessed: 30 September 2023).

Federal Statistical Office [Destatis] (2023). Regional database Germany. https://www.regionalstatistik.de/ (last accessed: 30 September 2023).

GeoBasis-DE/BKG (2018). Verwaltungsgebiete 1:250 000 (Ebenen), Stand 01.01. (VG250 01.01.). https://gdz.bkg.bund.de/index.php/default/digitale-geodaten/verwaltungsgebiete/verwaltungsgebiete-1-250-000-ebenen-stand-31-12-vg250-ebenen-31-12.html (last accessed: 19 November 2022).

German Insurance Association [GDV] (2023). Serviceteil zum Naturgefahrenreport 2022. GDV, Berlin.

ImmoScout24 (2019a). Datenauswertung zur Laufzeit von Wohnungsinseraten belegt hohen Nachfragedruck. https://www.immobilienscout24.de/unternehmen/news-medien/news/default-title/datenauswertung-zur-laufzeit-von-wohnungsinseraten-belegt-hohen-nachfragedruck/ (last accessed: 23 May 2023).

ImmoScout24 (2019b). Immobilienkauf: Hier sind Eigentumswohnungen am schnellsten weg! https://www.immobilienscout24.de/baufinanzierung/news/so-schnell-sind-eigentumswohnungen-weg.html (last accessed: 23 May 2023).

ImmoScout24 (2019c). Immobilienkauf: Hier sind Häuser am schnellsten weg! https://www.immobilienscout24.de/baufinanzierung/news/so-schnell-sind-haeuser-weg.html#par_titleblock (last accessed: 23 May 2023).

Investitions- und Förderbank Niedersachsen [NBank] (2022). Wohnbauland-Umfragen 2008-2022. NBank, Hannover.

Klick, L. and Schaffner, S. (2020). FDZ Data description: Regional Real Estate Price Indices for Germany (RWI-GEO-REDX) - Version 2: 2008-02/2019. RWI Projektberichte. RWI – Leibniz Institute for Economic Research.

RWI – Leibniz Institute for Economic Research (2018). Überprüfung des Zuschnitts von Arbeitsmarktregionen für die Neuabgrenzung des GRW-Fördergebiets ab 2021. Endbericht: Studie im Auftrag des Bundesministeriums für Wirtschaft und Energie (BMWi) (RWI Projektberichte). RWI – Leibniz Institute for Economic Research, Essen.

Schaffner, S., Thiel, P., RWI and ImmobilienScout24 (2022). RWI-GEO-REDX: Regional Real Estate Price Index for Germany, 2008-06/2022. Version: 1. RWI – Leibniz Institute for Economic Research. Dataset. http://doi.org/10.7807/immo:redx:v9.

SCHUFA Holding AG [SCHUFA] (2022). Risiko- und Kredit-Kompass 2009-2022. SCHUFA, Wiesbaden

Statistics of the Federal Employment Agency [BA Statistics] (2023a). Begonnene sozialversicherungspflichtige Beschäftigungsverhältnisse nach Merkmalen. Statistics of the Federal Employment Agency.

Statistics of the Federal Employment Agency [BA Statistics] (2023b). Bestand an gemeldeten Arbeitsstellen. Statistics of the Federal Employment Agency.

Verband der Vereine Creditreform e.V. [Creditreform] (2022). SchuldnerAtlas Deutschland. Überschuldung von Verbrauchern. 2012-2022. Creditreform.

Vilhuber, L., Kóren, M., Llull, J., Connolly, M. and Morrow P. (2021). A template README for social science replication packages. https://social-science-data-editors.github.io/template_README/ (last accessed: 19 November 2022).

Acknowledgements