

A Service of

ZBW

Leibniz-Informationszentrum Wirtschaft Leibniz Information Centre for Economics

Klick, Larissa; Schaffner, Sandra

Article — Published Version Regional Real Estate Price Indices for Germany, 2008 – 2019: RWI–GEO–REDX

Jahrbücher für Nationalökonomie und Statistik

Provided in Cooperation with: RWI – Leibniz-Institut für Wirtschaftsforschung, Essen

Suggested Citation: Klick, Larissa; Schaffner, Sandra (2020) : Regional Real Estate Price Indices for Germany, 2008 – 2019: RWI–GEO–REDX, Jahrbücher für Nationalökonomie und Statistik, ISSN 2366-049X, De Gruyter Oldenbourg, Berlin, Boston, Iss. Ahead of Print, pp. 1-11, https://doi.org/10.1515/jbnst-2020-0001

This Version is available at: https://hdl.handle.net/10419/225247

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.



https://creativecommons.org/licenses/by/4.0/

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



WWW.ECONSTOR.EU

Data Observer

Larissa Klick* and Sandra Schaffner Regional Real Estate Price Indices for Germany, 2008 – 2019: RWI-GEO-REDX

https://doi.org/10.1515/jbnst-2020-0001

1 Introduction

An important issue in the current German public and political discourse is the development of housing prices. Sharply increasing rent and purchasing prices, shortage of living space in urban areas and rural exodus are some of the discussed problems. Despite its topicality, few indices are available on the recent development of housing prices in Germany. To fill this gap, the RWI-GEO-REDX dataset quantifies regional differences in house purchase, apartment rent and purchase prices on the level of districts (*Kreise*, NUTS 3-level) and municipalities (*Gemeindeverband*, LAU 1-level) as well as labor market areas defined by RWI (2018).

The RWI-GEO-REDX combines a comprehensive up-to-date dataset and a hedonic price regression setup controlling for the property's quality to capture various features of the sales and rent prices beyond the observation of average prices. We end up with regional price indices relative to the German mean capturing regional differences, the region-specific time trend as well as the national development over time.

The RWI-GEO-REDX covers indices for apartment purchase, house purchase and rental apartments. All three categories are available on labor market region, district and municipality level. All indices are published on a yearly basis. Additionally, the overall German trends are also available on a quarterly basis. The current version of the RWI-GEO-REDX is not only an update but also an extension to the first version as described in Klick and Schaffner (2019a). Apart from these additions, the applied original dataset RWI-GEO-RED was improved by imputations, so we gained more observations in all three property categories and regional units. 9

^{*}Corresponding author: Larissa Klick, RWI – Leibniz Institute for Economic Research, Essen, Germany, E-mail: larissa.klick@rwi-essen.de

Sandra Schaffner, RWI – Leibniz Institute for Economic Research, Essen, Germany, E-mail: Sandra.schaffner@rwi-essen.de

³ Open Access. © 2020 Klick and Schaffner, published by De Gruyter. © BY This work is licensed under the Creative Commons Attribution 4.0 Public License.

Section 2 gives a detailed decription of the data, in Section 3 the methodology of the real estate price indices is depicted. Thereafter, the analytic potential and some descriptive references are made in Section 4 and the data access is referred to in Section 5.

2 Data

We use the RWI-GEO-RED data (Boelmann et al. 2019a, 2019b, 2019c) of the Research Data Center Ruhr at RWI (FDZ Ruhr) to generate the price indices. It is based on real estate offers published on the largest German listing website ImmobilienScout24. It gives real estate owners and estate agents the opportunity to advertise their houses and apartments for a fee. All information is provided by the owner or the agent who sells or rents out the property, most also on a voluntarily basis. The original data ranges from January 2007 to March 2019 and entitles regional information below to the 1 km² grid level. Further, the data covers information on the size of the house or the apartment, on its facilities and features, additional costs as well as on energy consumption. For a more detailed description see Boelmann and Schaffner (2019).

Processed with the aim of coherence and comparability, the RWI-GEO-REDX dataset reflects differences between the properties that determine the monetary outcome in the selling and renting progression. To come as close as possible to the real market price of the property we include only the advertisements in their last month published. We restrict the price index to the years 2008 to February 2019 in the following, as there are only little observations in 2007 and to prevent a look ahead bias. Further information on excluded offers form the original dataset are described in the Appendix.

We compute the price indices for districts (*Kreise*; 402) and municipalities (*Gemeindeverbände*; 4,542) based on the regional definitions of 2015 (BKG 2016). As supplement we include labor market areas (*Arbeitsmarktregionen*) according to the delineation of version 1 in RWI (2018) as a third region type defining 182 areas. This delineation is beneficial to model real estate price indices as it follows the idea of accessibility of labor markets for commuters. The labor market borders are drawn from existing commuting interrelations and are, thus, a strong determinant for the residence decision. For a detailed description see RWI (2018).

3 Methodology

The regional price indices account for characteristics of the facility as well as for regional and time differences and is comparable to common hedonic price regressions (e.g. Sirmans et al. 2005) as applied for Germany in Bauer et al. (2013), for example.

To provide reasonable housing price developments three different models are estimated.

3.1 Price indices for the overall time development in Germany

$$\ln(y_{igt}) = \beta X_{igt} + u_g + \delta_t + \varepsilon_{igt}, \qquad (1)$$

Equation 1 presents the basic model to identify overall time developments δ_t with the dependent variable y_{igt} representing the sale or rent price (*Kaltmiete*) per sqm of the single real estate advertisement *i* in region *g* (district, municipality or labor market area) at time *t*. The characteristics of the property are included in X_{igt} . The exogenous and endogenous variable remain unchanged throughout all equations. A detailed description of the features used as exogenous variable is given in Table 1 in the Appendix.

Moreover, δ_t defines year fixed effects describing the development in Germany for each year $t = \{2008, ..., 2019\}$, respectively each quarter $t = \{2008q1, 2008q2, ..., 2019q1\}$. Since all regions and time units are studied jointly, it is assumed that the characteristics are valued in the same way over time and region. The time-invariant regional fixed effect is given by u_g and the error term ε_{igt} is assumed to be standard normally distributed.

3.2 Regional differences

The second regression describes a yearly cross-sectional approach with a regional price index u_{gt_0} for a specific time t_0 as a fixed effect and is given in eq. (2):

$$\ln(y_{igt_0}) = \gamma X_{igt_0} + u_{gt_0} + \varepsilon_{igt_0}$$
⁽²⁾

This approach assumes that characteristics are valued the same way all over Germany during the respective time t_0 in region *i*. It is given yearly from 2008 to 2018 and for the last quarter of 2018, as well as for all three region types. The price index u_{gt_0} indicates the price differences between the regions at time t_0 . The indices from this regression describe the regional price discrepancy to the German mean of all properties offered in this specific time period.

3.3 Regional time developments

The estimation strategy to provide the within-region development η_{gt} of region *g* is given in eq. (3):

$$\ln(y_{igt}) = \theta X_{igt} + \eta_{gt} + \varepsilon_{igt}$$
(3)

The specific development of region *g* between year t_0 and t_1 can be derived by $\eta_{gt_1} - \eta_{gt_0}$. This model has similar assumptions as eq. (1), stating that characteristics are valued the same in every region and across years. Additionally, development can differ between regions. This approach is deployed for all three region types on a yearly basis as the development in comparison to the base year 2008.

4 Descriptive evidence

The data contains a variety of indices that can be used in different types of analyses concerning regional price discrepancies. Figure 1 presents the development of the German real estate price indices δ_t , based on eq. (1). This figure displays the German wide quarterly price trend. Although the real estate price is the largest asset in the scheme of the consumer price index (19.6%; Destatis 2019b), the prices for newly rented or sold apartments and houses outpaces the moderately increased consumer prices.

Figure 2 displays a depiction of regional house prices. Figure 2(a) shows especially high nominal increases in districts that are economically strong and densely populated. House prices among metropolitan regions seem to spread into adjacent districts and municipalities (Figure 2(b)). It becomes obvious which municipalities are part of the respective commuter belt of cities like Hamburg or Berlin since prices are almost as high as in the inner city.



Figure 1: Time development of the German price indices for house purchases, apartment rentals and apartment purchases change in percentage points to base quarter 2008q1, based on regression 1 on district level.

Source: RWI-GEO-REDX, Time effects gives change in δ_t in percentage points. Destatis (2019a) for yearly consumer price index.

However, the small regional level of municipalities comes at the cost of regional coverage. Particularly in Bavaria, municipalities (also unions of boroughs, *Gemeindeverbände*) are relatively small. It is only possible to derive price indices for these small municipalities at the cost of imprecision. Therefore, we display only price indices that rely on at least 50 advertisements per year. Although a large share of municipalities miss the threshold of 50 advertisements per year, these municipalities only cover about ten percent of the German population for house purchase, 35% for apartment purchase (which is quite rare in rural areas) and 25% for rental apartments between 2008 and 2018.

Figure 2(d) presents the development of housing prices between 2008 and 2018 on district level. It becomes apparent that most districts experienced no or only small nominal price increases that are below the consumer price development (light yellow and yellow) or increased only slightly above the consumer price index. The regional variation is quite in contrast to Figure 1 that implies a sharp price increase. However, this increase is mainly driven by some densely populated, high-priced regions e.g. Munich and its surroundings. These agglomerations influence immensely the German price trend due to the high share of population living there.



Figure 2: Regional price indices in 2018 on district level, municipality and labor market area level for house purchase as deviation in % from the German mean (a-c), Change in percentage points (d). Only regions with at least 50 observations in 2018, respectively in 2008 and 2018, are displayed. Indices in a, b, c base on cross-sectional regressions in eq. (2). Indices in d base on eq. (3).

The RWI-GEO-REDX has its main potential for analyses in its regional components. Thus, it can be applied as control or identifying variable in a vast of potential analyses. Most other consumer goods do not vary regionally as much as housing costs. Furthermore, the latter is the main asset in the individual consumption, which makes them a good proxy for regional cost disparities. For example Bauer et al. (2019) use the RWI-GEO-REDX data for analyzing drivers of internal migration in Germany.

5 Data access

The data can be obtained as a Public Use File by the FDZ Ruhr. The FDZ Ruhr is the research data center at the RWI – Leibniz-Institute for Economic Research. The data is open for public use. To ensure that the indices are not driven by small sample size, the dataset only covers those indices that rely on at least 50 observations per year and region. Hence, the information is restricted for smaller municipalities. We also provide the indices that base on less than 50 observations per year and region as Scientific Use File upon request for scientific research only. Since the RWI-GEO-REDX subsumes aggregated information it does not contain information with restricted use due to data security. The presented indices can be obtained as an Excel (.xlsx) file.

Data access does not require a data use agreement, but users need to register for data access. Interested users should register via email to fdz@rwi.essen.de. The email needs to include information on the applying department or person as well as the desired data format. The users are requested to cite the source correctly and to inform the FDZ Ruhr about publications with the data.

When using the dataset RWI-GEO-REDX, please cite the data as Klick, Schaffner, RWI, and ImmobilienScout24 (2019): RWI-GEO-REDX: Regional Real Estate Price Index for Germany, 2008-02/2019. Version: 2. RWI – Leibniz Institute for Economic Research. Dataset. http://doi.org/10.7807/immo:redx:v3. Further, we recommend citing this data description.

References

- Bauer, T.K., S. Feuerschütte, M. Kiefer, P. an de Meulen, M. Micheli, T. Schmidt, L.-H. Wilke (2013), Ein hedonistischer Immobilienpreisindex auf Basis von Internetdaten: 2007-2011.
 AStA Wirtschafts- und Sozialstatistisches Archiv 7: 5–30.
- Bauer, T.K., C. Rulff, M.M. Tamminga (2019), Berlin Calling Internal Migration in Germany. Ruhr Economic Papers 823, RWI.

- Boelmann, B., R. Budde, L. Klick, S. Schaffner, RWI, et al. (2019a), RWI-GEO-RED. RWI Real Estate Data- Apartments for Sale. Version: 1. RWI – Leibniz Institute for Economic Research. Dataset. http://doi.org/10.7807/immo:red:wk:v1.
- Boelmann, B., R. Budde, L. Klick, S. Schaffner, RWI, et al. (2019b), RWI-GEO-RED. RWI Real Estate Data- Apartments for Rent. Version: 1. RWI – Leibniz Institute for Economic Research. Dataset. http://doi.org/10.7807/immo:red:wm:v1.
- Boelmann, B., R. Budde, L. Klick, S. Schaffner, RWI, et al. (2019c), RWI-GEO-RED. RWI Real Estate Data- Houses for Sale. Version: 1. RWI – Leibniz Institute for Economic Research. Dataset. http://doi.org/10.7807/immo:red:hk:v1.
- Boelmann, B., S. Schaffner (2019), FDZ Data Description: Real-Estate Data for Germany (RWI-GEO-RED V1). Advertisements on the Internet Platform ImmobilienScout24 2007-03/2019. RWI Projektberichte, Essen.
- Bundesamt für Kartographie und Geodäsie (BKG) (2016), Verwaltungsgebiete 1: 250 000. VG250 und VG250-EW. Frankfurt/Main.
- Der Obere Gutachterausschuss für Grundstückswerte im Land Nordrhein-Westfalen (2017), Grundstücksmarktbericht 2017. Nordrhein-Westfalen. Available at: https://www.boris.nrw. de/borisfachdaten/gmb/2017/GMB_000_2017_pflichtig.pdf. (Effective September 2017, Online accessed 26/11/2019).
- Destatis (2019a), Verbraucherpreisindex (inkl. Veränderungsraten): Deutschland. Available at: https://www-genesis.destatis.de/genesis/online?operation=result&code=61111-0001&deep=true (Effective 26/11/2019, Online accessed 26/11/2019).
- Destatis (2019b), Verbraucherpreisindex für Deutschland. Wägungsschema für das Basisjahr 2015. Available at: https://www.destatis.de/DE/Themen/Wirtschaft/Preise/ Verbraucherpreisindex/Methoden/Downloads/waegungsschema-2015.pdf?__blob= publicationFile (Effective 26/11/2019, Online assessed 26/11/2019).
- Gutachterausschuss für Grundstückswerte in Sachsen-Anhalt (2017), Grundstücksmarktbericht Sachsen-Anhalt 2017.
- Klick, L., S. Schaffner, RWI, ImmobilienScout24 (2019), RWI-GEO-REDX. Regional Real Estate Price Index for Germany, 2008-2/2019. Version: 2. RWI – Leibniz Institute for Economic Research. Dataset. http://doi.org/10.7807/immo:redx:v3.
- Klick, L., S. Schaffner (2019a), FDZ Data Description. Regional Real Estate Price Indices for Germany (RWI-GEO-REDX). RWI Projektberichte. Essen.
- RWI (2018), Überprüfung des Zuschnitts von Arbeitsmarktregionen für die Neuabgrenzung des GRW-Fördergebiets ab 2021. RWI Projektberichte. Essen.
- Sirmans, G., D. Macpherson, E. Zietz (2005), The Composition of Hedonic Pricing Models. Journal of Real Estate Literature 13: 3–43.

Appendix

Appendix A: Information on Data Restriction in the RWI-GEO-REDX Price Indices

To further enhance the coherence, very luxurious and unrealistically modest observations are dropped from the original RWI-GEO-RED dataset. Rental apartment with rents exclusive utilities above 5 000 Euro per month, more than 7 rooms and a living area outside the range of 15 to 400 sqm are omitted. House purchases are restricted to single-family houses; the living area ranges from 50 to 600 sqm and the house price varies up to 5 Million Euro. The number of rooms is restricted to 15 and holiday homes, apartment buildings or houses with more than five floors are excluded. Apartments for purchase with prices higher than 2 Million Euro, more than eight rooms and an adverted living area below the 1st percentile (27 sqm) and above the 99th percentile (230 sqm) are not accounted for in the following estimation. Differently to the first published version of the indices, weakly georeferenced data was imputed, which increases the regional coherence of the data.

Working with these self-declared and in many cases voluntary information leads to missing values in many variables that need to be handled with care. For the binary variables a missing is accounted for as a zero, so the offer does not meet the feature in question. This seems reasonable to the extent that the owner or agent tends to publish benefits of the real estate to attract searchers with certain preferences. Furthermore, in some years many characteristics are collected using checkboxes which means that there is no difference between "no" and "no answer". Examples are especially positive characteristics of the property, such as a balcony or guest toilet. In the analysis, we deal with missing values as a separate category for categorical variables. In the considered metric variable, number of rooms, missing values are given as "zero rooms".

ä
Ľ.
Ρ
Ð
Q
<u>a</u>
◄

prices.
l rent
and
tment
, apar
iouse,
<u>م</u>
/sis o
anal
s in
P
variab
latory
lar
Ä
÷
e
ē
ĥ

Variable	Description	House purchase	Apartment rent	Apartment purchase
Metric variable				
number of rooms	number of rooms in property	×	x	×
Categorical variables				
number of total floors	1:= missing, 2:=1-3 floors, 3:=4-5 floors, 4:=6-10 floors,			×
floor number	5:= more than 10 floors 0:= missing, 1:= ground floor (UG), 2:= first floor (EG), 3:= 2nd to 3rd floor, 4:= 4th to 5th floor,			×
equipment	5:= 6th to 10th floor, 6:= above 10th floor 0:= missing, 1:= Simple, 2:= Normal, 3:= Sophisticated, 4:= Exclusive	×	×	×
Binary variables				
year of construction	1:= missing, 2:= before 1900, 3:= 1900–1945, 4:= 1946–1959, 5:= 1960–1969, 6:= 1970–1979, 7:= 1980–1989, 8:= 1990–1999, 9:= 2000–2009, 10:= after 2009	×	×	×

plot area	[in sqm]: 0:= missing, 1:= (0-200], 2:= (200-400], 3:= (400-600], 4:= (600-800], 5:= (800-1 200], 6:= (1 200-2 500]. ^a	×		
first occupancy	1 if new owner or renter move in as first occupancy	×	×	×
detached house	1 if house is detached	×		
semi-detached house	1 if house is semi-detached	×		
terraced house	1 if house is a terrace house	×		
exclusive house	1 if property is declared as a mansion or castle	×		
other house type	1 if house is categorized differently	×		
balcony	1 if property has a balcony		×	×
garden	1 if apartment has access to a private garden		×	×
guest toilet	1 if object includes a guest toilet	×	×	×
fitted kitchen	1 if object comes with a fitted kitchen		×	×
granny flat	1 if property contains a separate "granny flat" or secondary	×		
	suite			
cellar	1 if cellar room is available		×	×
assisted living	1 if object is declared as assisted living			×
common charge	1 if common charge is declared in offer			×
lift	1 if property contains a passenger lift			×
^a The variable plot area is r the following. As undevelo further commercial and ag Gutachterausschuss für Gr	estricted to 2 500 sqm. To focus on house sales for living purposes only plot a ped rural plot area above 2 500 sqm are accounted for as farmland sales in r ricultural use (Der Obere Gutachterausschuss für Grundstückswerte im Land 1 undstückswerte in Sachsen-Anhalt 2017).	reas smaller th eports of the Vordrhein-Wes	nan 2.500 sqm are property market s :tfalen 2017;	included in uggesting