Datenbeschreibung

RWI – Leibniz-Institut für Wirtschaftsforschung

FDZ Data Description: Regional Real Estate Price Indices for Germany (RWI-GEO-REDX) - Version 14.1: 2008-05/2024

Patrick Thiel

April 2025

Impressum

Herausgeber:

RWI – Leibniz-Institut für Wirtschaftsforschung

Hohenzollernstraße 1–3 | 45128 Essen, Germany

Postanschrift

Postfach 10 30 54 | 45030 Essen, Germany

Fon: +49 201–81 49-0 | E-Mail: rwi@rwi-essen.de

www.rwi-essen.de

Vorstand

Prof. Dr. Dr. h.c. Christoph M. Schmidt (Präsident)

Prof. Dr. Thomas K. Bauer (Vizepräsident)

Dr. Stefan Rumpf (Administrativer Vorstand)

Prof. Dr. Kerstin Schneider (Mitglied des erweiterten Vorstands)

RWI Datenbeschreibung

Schriftleitung: Prof. Dr. Dr. h. c. Christoph M. Schmidt

Gestaltung: Magdalena Franke, Claudia Lohkamp

FDZ Data Description: Regional Real Estate Price Indices for Germany (RWI-GEO-REDX) - Version 14.1: 2008-05/2024

April 2025

Patrick Thiel

Leibniz-Institut für Wirtschaftsforschung

Leibniz Gemeinschaft

Das RWI wird vom Bund und vom Land Nordrhein-Westfalen gefördert.

Contents

List of Figures

List of Tables

Abstract

The FDZ Ruhr at RWI provides price indices for apartments and houses (rentals and sales) in Germany starting in 2008. The price indices are based on the real estate dataset RWI-GEO-RED, which combines real estate listings from ImmoScout24. We perform three types of hedonic regressions to allow comparisons across time (annual or quarterly), across regions (labor market regions, counties, or municipalities), and in combination across time and regions. The data are publicly available and can be requested from the FDZ Ruhr website. A weakly anonymized version is also available for researchers on request.

This data report is an updated version of previous reports and refers to RWI-GEO-REDX V14.1, which covers data until May 2024.

1. Introduction

An important topic in the current public and political discourse in Germany is the development of housing prices. Rising rents, housing shortages in urban areas, and rural-urban migration are some of the problems being discussed. Despite its topicality, little data is available on the recent development of house prices in Germany. To fill this gap, the RWI-GEO-REDX dataset quantifies regional differences in house sales, apartment rents, and apartment sales 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).

Several methods are used to derive real estate indices. First, median sales prices are the easiest to construct and are used by the U.S. Census Bureau. However, they do not adjust for the quality of the properties on the market (Ghysels et al., 2013). Case and Shiller (1989) propose a repeated-sales price index, which is a constant-quality index that uses only information on houses that have been sold at least twice during the sample period. This index requires very little data. However, these properties may not be representative of the market (Clapp & Giaccotto, 1992), and the index relies on the crucial assumption that quality does not change due to renovation or modernization. This second method is used by the U.S. Federal Housing Finance Agency for the HPI index.

Third, hedonic price indices are based on hedonic price regressions, where the price is expressed as a function of the characteristics of the dwelling or house. The characteristics thus describe the quality of the property. For all three types of indices, it is important that the sale of a property can be correlated with local economic conditions, as more expensive homes tend to be put on the market during expansions. Gatzlaff and Haurin (1998) take into account the selectivity of using only sold houses. Englund et al. (1998) combine the hedonic index with the repeated sales index. Finally, there are stock market-based indices that are obtained by trading real estate investment trust shares.

While a simple study of regional average rents and prices is not sufficient to make reliable statements on housing prices, the RWI-GEO-REDX is based on hedonic price regressions. The hedonic price regressions take into account qualitative differences in housing supply, such as different amenities. Hedonic price regressions are a commonly used method for real estate pricing in Germany, e.g., in the hedonic EPX of Eurospace AG (2019), a commercial online distributor of real estate financing, and the real estate price index of the Verband Deutscher Pfandbriefbanken e V. (vdp) (2019), which processes transaction data from certain financial institutions. The real estate price index by bulwiengesa AG (2018) focuses on cities and larger towns, but does not include consistent information on data structure and methodology and is available for commercial use only. The AK-OGA, a cooperation of all German federal advisory committees on real estate issues, publishes housing price indices for different house types on an annual basis and at the federal and sub-district level. These price indices are partly derived by median sales observation and hedonic price regressions. In addition, the IMX Offer Index is a hedonic price index that is derived from the offers placed by the online real estate agency ImmobilienScout24 for the sale and rental of houses or apartments.

The first two indices report house price changes at the national level and do not provide further information for smaller regions. This study fills this gap by combining a comprehensive, up-to-date dataset and a hedonic price regression; it provides 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 analysis is based on a dataset of online real estate advertisements, the RWI-GEO-RED. This comprehensive dataset is obtained from ImmobilienScout24. The data is regularly updated, which makes it possible to analyze the latest developments. It is therefore possible to update the price indices on a regular basis.

We use a hedonic price regression for the price indices in order to capture various characteristics of sales and rental prices beyond the median. The repeat sales approach may not be representative for the German real estate market, as Voigtländer (2012) mentions. Here, too, it is not possible to link two sales of the same property due to data limitations.

While other published price indices are constructed using financial data, the following analysis deals with seller-listed listings. This means that only apartments and houses that are for rent or sale and listed on the listing website are available in the dataset. Consequently, the analysis is based on advertised prices and not transaction prices.

In addition, indices for home sales are also included in the dataset, complementing the regional price information for home sales and rentals. In addition to the municipality and district level, all three supply categories are also reported at the level of labor market regions according to the delineation of RWI (2018).

2. Data

We use the RWI-GEO-RED data (RWI - Leibniz Institute for Economic Research, 2024a, 2024b, 2024c) of the FDZ Ruhr at the RWI to generate the price indices. They are based on real estate listings published on Germany's largest real estate listing website, ImmobilienScout24. Here, property owners and estate agents can advertise their houses and apartments for a fee. All information is provided by the owner or agent selling or renting the property. While some information is required to place an ad online, most information is provided voluntarily. More information about the property helps to present it appropriately and is likely to increase the chances of it being sold or rented. A detailed description of the data source can be found in Thiel (2024).

The dataset includes information on the prices of property listings as well as various characteristics that determine the value of a property. Monthly data are used. The present dataset covers the period from January 2007 to June 2024. Since there are few observations in 2007, we restrict the data in the following to the years 2008 to May 2024. The restriction to May 2024 is to avoid a look-ahead bias in the time trend.

The RWI-GEO-RED provides information on the building at the level of the housing unit, municipality, district and federal state. Furthermore, the data includes information on the size of the house or apartment (e.g., living space, plot area, number of rooms), on its facilities (e.g., balcony, garden, bathrooms, level of facilities), on financial aspects (e.g., price and additional costs), as well as information on energy consumption. Unfortunately, some variables are characterized by many missing values, which has to be taken into account for the following analysis.

The selection of variables for the analysis is based on two considerations. First, we aim for coherence in the data set to ensure comparability between properties. Second, the set of characteristics used in the analysis must be comprehensive enough to capture the different characteristics of apartments and houses. In order to get as close as possible to the real market price of the property, we only include advertisements in their most recent month of publication, i.e., when they leave the listing website. Previously updated versions of the listed apartment or house are not included. This strategy aims to best approximate the actual sale price with the published, self-reported listing price. Further information on excluded homes from the original dataset is described in the Section 9.

We calculate price indices for districts and municipalities based on the regional definitions of 2019 (Federal Agency for Cartography and Geodesy, 2019). In 2019, there were 401 counties in Germany, varying in area and population size. In addition, there were 4,625 municipalities. As a supplement, we included labor market regions according to the delineation in RWI (2018) as a third type of region, defining 182 areas. This delineation is advantageous for modeling real estate price indices, as it follows the idea of labor market accessibility for commuters. The labor market boundaries are drawn from existing commuting relationships. The advantage of this delineation is that these interrelationships are a stronger determinant of residence decisions and real estate market developments than pure administrative boundaries.

3. Methodology

Regional price indices should take into account the characteristics of the property as well as regional and time differences. Therefore, we first develop a hedonic price regression that takes into account the characteristics of property advertisements. The regression is comparable to common hedonic price regressions (e.g., Sirmans et al., 2005), such as those applied for Germany in Bauer et al. (2011).

The listing price per square meter is used as the dependent variable for the housing unit for sale, and the exclusive rent per square meter (Kaltmiete) is used for rentals. Thus, the endogenous variable is independent of the size of the apartment or house. Other characteristics that determine the rental or listing price are included as exogenous variables in the hedonic price regression.

The purpose of the indices is to show the time evolution of housing prices for Germany, their regional differences, and the regional time evolution. Different models are needed to estimate these three types of indices. First, we use the following model to estimate price indices for the whole of Germany over time:

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

where the dependent variable *y* is the sale or rental price per square meter of dwelling unit *i* in region *g* (county, municipality, or labor market area) in year *t*. The characteristics of the property are included in vector *X*. The characteristics in the hedonic regressions vary between rental, owner-occupied, and home ownership; a list of the characteristics used as exogenous variables is given in Table A1 in Section 9.

This first model includes a time-independent price index u_g for region g (fixed effect). In addition, year fixed effects δ_t are included. The time evolution in Germany is described by the evolution of δ_t for each year *t = 2008*, ..., *2024*, respectively each quarter *t = 2008-Q1*, *2008-Q2*, ..., *2024-Q2*. Since all regions and years are examined together, it is assumed that the characteristics are valued in the same way for all years and all regions. The error term ϵ_{igt} is assumed to have a standard normal distribution.

The second regression describes a yearly cross-sectional approach with a regional price index u_g for region *g* and a specific year t_0:

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

In this approach, it is assumed that the characteristics are valued in the same way throughout Germany at the respective time t_0 and in the respective region *i*. It is given annually from 2008 to 2024, as well as for all three region types. The price index *u* indicates the price differences between the regions at time *t*. The indices derived from this regression describe the regional price discrepancy to the German average for all properties offered in this specific time. In a nutshell: What is the regional premium for a square meter of living space in a property of the same quality and features as all listed properties in the respective year?

The first approach provides insights into the overall development over time, while the second provides insights into year-specific regional differences. However, in addition to absolute price differences, regional differences over time are also of interest for research and policy discussion. Therefore, we use a third approach to measure the respective time development of region *g* via the year-region fixed effects η:

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

The specific time evolution of region g between year t_0 and t_1 can be derived by

$$\eta_{gt_1} - \eta_{gt_0} \tag{4}$$

The main assumptions in this analysis are similar to those in Equation (2), namely that the characteristics are valued the same in each region and across years. In addition, the time evolution may differ between regions. This approach is used for all three types of regions on a yearly basis as the development compared to the base year 2008.

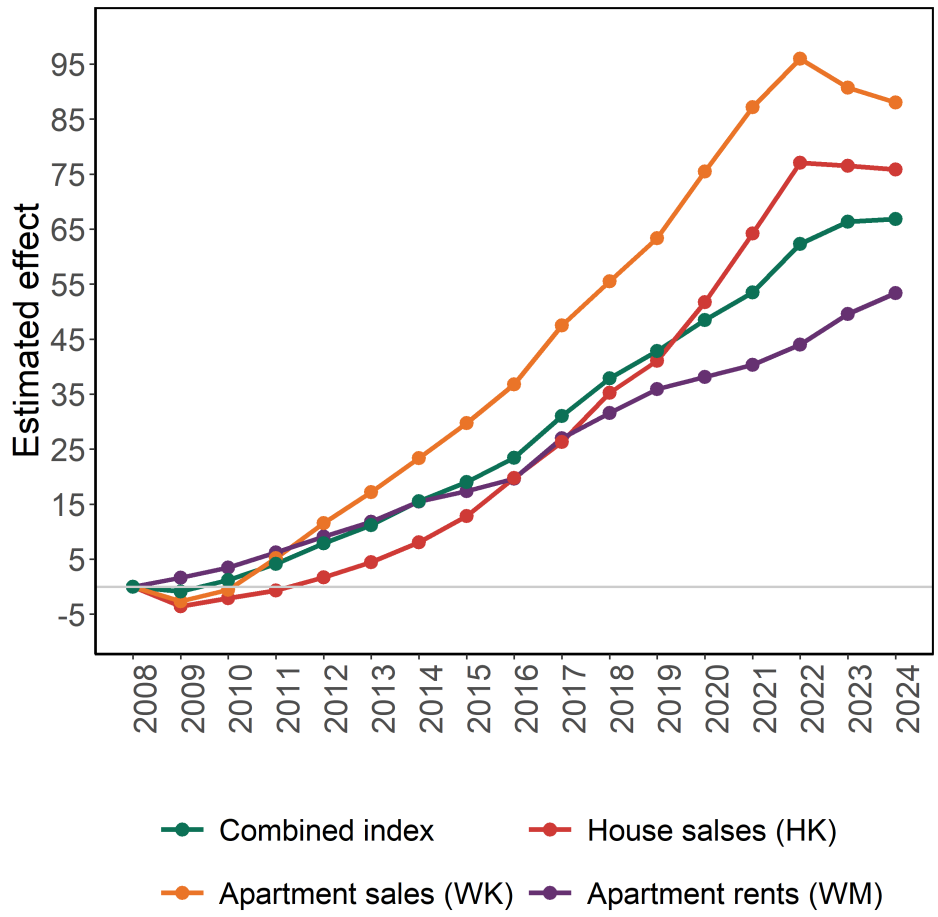
To ensure the coherence of the price indices, only price indices for regional units with more than 50 observations are shown in Equation (2) and Equation (3) in the Public Use File (PUF) of the RWI-GEO-REDX data. The Scientific Use File (SUF) version requires at least five observations. It is up to the researcher to decide whether this lower threshold is reliable in the considered research design.

4. Descriptive Evidence

Figure 1 shows the time evolution of the German price indices derived from Equation (1). House and apartment prices declined or stagnated at the beginning of the period. While apartment rents rose steadily, apartment sales began to rise in 2011 and house sales followed suit in 2014. All three types then experienced strong price increases. The ever-increasing trend for sales (both apartments and houses) was only interrupted in 2022. Rising inflation and interest rates, the war in Ukraine, and a lack of investment in the construction sector led to a shortage of housing, making it difficult for people to finance new homes. Apartment rents, on the other hand, remain on an upward trend.

The figure also shows the combined index, which is a combination of the individual indices as described in Section 3. It follows the evolution of the other indices and ranks between them in terms of the estimated effect, which is constructed.

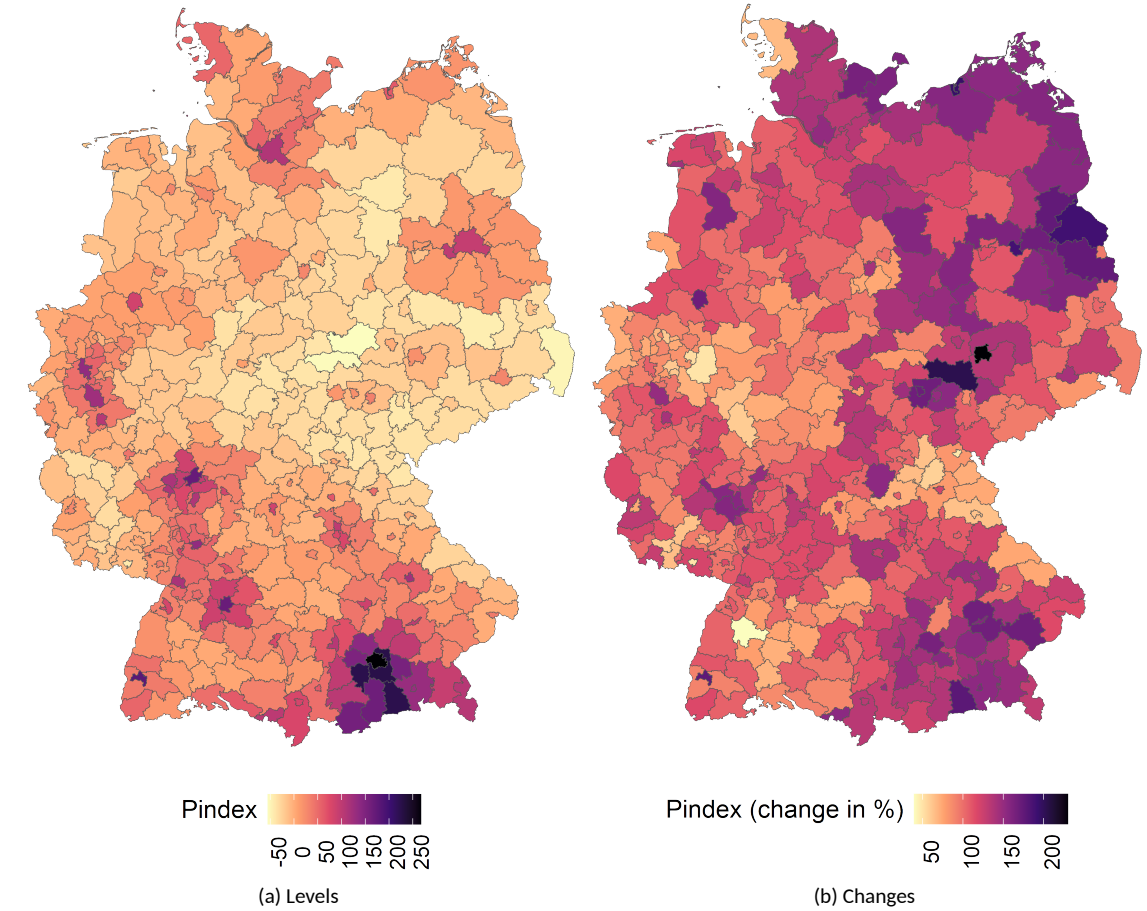
Figure 1
Time Development of the German Price Indices for House Sales, Apartment Rentals, Apartment Sales and Combined Index
Change in percentage points to base year 2008



Notes: The time effects are measured in changes in δ_t in percentage points.
Source: Authors' graph based on RWI-GEO-REDX (RWI - Leibniz Institute for Economic Research, 2024d).

To illustrate the possibilities offered by the RWI-GEO-REDX dataset, a two examples of descriptive analyses are given below. The same can be done for the indices for rental housing and apartment sales. We limit our results here to home sales. These have a wide coverage, even in more rural areas, while rental and apartment ownership tend to be concentrated in urban areas.

Figure 2
Regional Pattern of the Price Index for Houses in 2024
Level Values and Change Values Relative to 2008 at the County Level



(a) Levels (b) Changes

Notes: The figure shows the geographical distribution of the price index for house sales in 2024 represented as levels (Panel A) and changes relative to 2008 (Panel B) at the county level.
Source: Authors' graph based on RWI-GEO-REDX (RWI - Leibniz Institute for Economic Research, 2024d).

Figure 2a, which shows the regional price index for home sales (levels) at the county level in 2024. The areas with the highest prices are located in or around major cities, which makes sense as these locations are typically hot spots for working and living. It also reflects the tight housing market with limited housing supply in urban areas.

Figure 2b, which reflects the change in house prices between 2008 and 2024, shows that house prices have risen sharply in most regions over time. Areas with particularly strong increases are again cities such as Berlin and Munich, but Leipzig is also visible.

5. Changes Between Versions

This section provides a concise overview of the modifications incorporated between the current version (V14.1) and the previous version (V14).

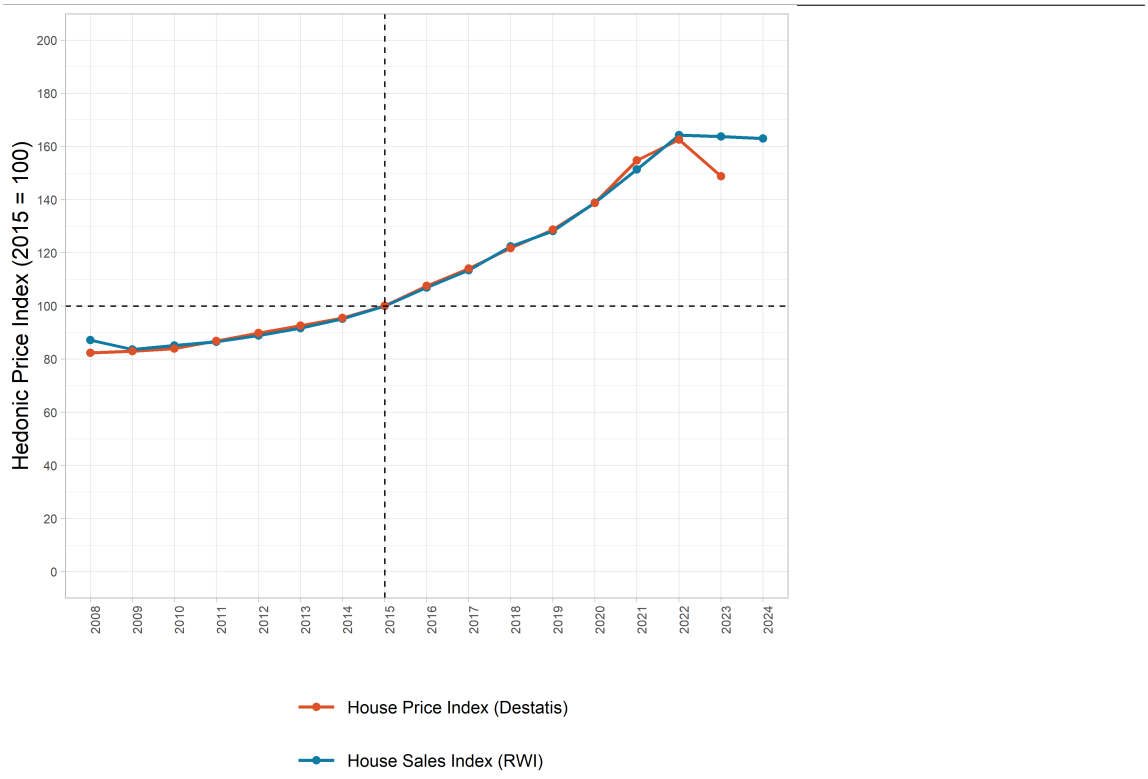
V14.1 fixes an error in the export of the final dataset. In V14, the values of the apartment rental index and the combined index are identical in the exported file *"RWIGEOREDX_GRIDS_v14_*"*. This is due to a typo in the exporting procedure and affects the sheets *"Grids_RegionEff_yearly"* and *"Grids_RegionEff_Change_yearly"*. Note that all the other sheets in the file *"RWIGEOREDX_GRIDS_v14_*"* are not affected. The same holds for the individual indices at higher aggregation level.

6. Data Validation

This section presents a brief comparison between the indices included in RWI-GEO-REDX and the official price/rent indices offered by the Federal Statistical Office (Destatis). In this comparison, we focus exclusively on temporal developments, given the limitations of data from official sources at the regional level. It is important to note that the comparison is limited by data differences. As the Federal Statistical Office employs alternative raw data and constructs slightly different indices, the final price/rent values are not entirely identical to our estimates. However, the trend pattern should be comparable, indicating that our indices capture significant economic developments.

House price index. In order to facilitate a comparison of our estimates for houses for sale, we have recourse to the house price index published by Federal Statistical Office (2024a) for the year 2024. It should be noted that the Federal Statistical Office employs 2015 as the reference period; thus, our estimates have been recalibrated to the 2015 baseline for the purposes of this comparison. In our distributed dataset, we utilize 2008 as the reference period. Figure 3 depicts the house price index as provided by the Federal Statistical Office (in red) and our own index for house sales (in blue). It is evident that both indices exhibit a high degree of correlation, with only a slight divergence in recent times. This divergence may be attributed to the significant economic, social, and political challenges currently facing Germany.

Figure 3
Comparison House Price Index - Federal Statistical Office and RWI
Estimates relative to 2015 (= 100)



Notes: The figure shows the house price index by the Federal Statistical Office (Destatis) (in red) and the RWI-GEO-REDX indices for houses for sale (in blue). *Source*: Authors' graph based on RWI-GEO-REDX (RWI - Leibniz Institute for Economic Research, 2024d) and Federal Statistical Office (2024a).

Apartment price index. In examining the dynamics of the apartment price index, we employ data provided by Federal Statistical Office (2024b), specifically the data set designated "Price index for owner-occupied housing". It should be noted that only the values indicated by the label "Price index for owner-occupied housing" are used, despite the raw data offering other distinctions. Moreover, it is not entirely evident from the raw data whether the "Price index for owner-occupied housing" aligns with our definition of apartment sales as outlined by ImmoScout. It is therefore necessary to exercise caution when interpreting the results of this comparison.

Figure 5 depicts the comparison between the two data sources. The data indicate a positive trend for both the index provided by Federal Statistical Office (2024b) and the RWI - Leibniz Institute for Economic Research (2024d) index. However, there are notable discrepancies, particularly in the period following 2015. The Federal Statistical Office index demonstrates a deceleration in growth, whereas the RWI-GEO-REDX index exhibits a consistent and uniform growth trajectory.

Figure 4
Comparison Apartment Price Index - Federal Statistical Office and RWI
Estimates relative to 2015 (= 100)

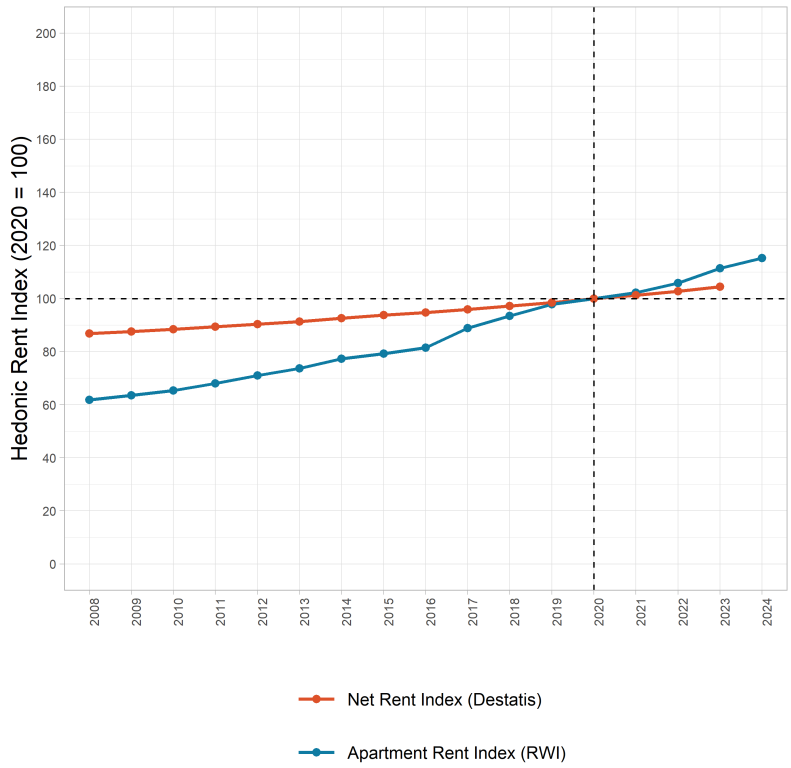
Notes: The figure shows the apartment price index by the Federal Statistical Office (Destatis) (in red) and the RWI-GEO-REDX indices for apartments for sale (in blue). *Source*: Authors' graph based on RWI-GEO-REDX (RWI - Leibniz Institute for Economic Research, 2024d) and Federal Statistical Office (2024b).

Apartment rent index. Finally, for the last comparison, we employ data given by Federal Statistical Office (2024c) which lists the rent index for the single federal states over time. It is based on the net rent without heating expenses similar to our own apartment rent index. Since the data set distinguishes between the federal states and we are interest in the national temporal pattern, we calculate the average index across all states. Note Federal Statistical Office (2024c) uses 2020

as reference period for its calculations, so we recalibrate our apartment rent index to this period as well. Note further that the Federal Statistical Office has a time lag of one year and reports only values up to 2023.

The data from both rent indices demonstrate a consistent upward trajectory over the entire time span under consideration. However, the Federal Statistical Office index demonstrates a notable uniformity and flatness of development, while our index exhibits a more pronounced acceleration over time.

Figure 5
Comparison Apartment Rent Index - Federal Statistical Office and RWI
Estimates relative to 2020 (= 100)



Notes: The figure shows the apartment rent index by the Federal Statistical Office (Destatis) (in red) and the RWI-GEO-REDX indices for apartments for sale (in blue). *Source*: Authors' graph based on RWI-GEO-REDX (RWI - Leibniz Institute for Economic Research, 2024d) and Federal Statistical Office (2024c).

7. Data Access and Additional Information

The data can be obtained as Public Use File (PUF) or Scientific Use File (SUF) from the FDZ Ruhr at RWI. The FDZ Ruhr is the research data center of the RWI - Leibniz Institute for Economic Research. In order to ensure that the indices are not driven by small sample sizes, the PUF dataset includes only those indices that are based on at least 50 observations per year and region. The indices based on less than 50 observations per year and region are also available on request as an SUF for scientific research purposes only. Since the RWI-GEO-REDX subsumes aggregated information, it does not contain any information that is restricted for data security reasons. The indices presented here are available as Excel (.xlsx) files.

Data access does not require a data use agreement, but users must register to access the data. Interested users should visit our website https://www.rwi-essen.de/en/research-advice/further/research-data-center-ruhr-fdz/data-access.

Users are requested to cite the source correctly and to inform the FDZ Ruhr about publications using the data. When using the dataset RWI-GEO-REDX, please cite the data as RWI (2025): RWI-GEO-REDX: Regional Real Estate Price Index for Germany, 2008-05/2024. Version: 1. RWI – Leibniz Institute for Economic Research. Dataset. http://doi.org/10.7807/immo:redx:v14.1. Further, we recommend citing this data description.

For more details on the index coding, please refer to our GitHub repository (https://github.com/PThie/RWI-GEO-REDX), which should be cited as: Patrick Thiel (2025). PThie/RWI-GEO-REDX: RWI-GEO-REDX V14.1 (V14.1). Zenodo. https://doi.org/10.5281/zenodo.15308279.

8. References

Bauer, T. K., Micheli, M., Schmidt, T., Kiefer, M., Wilke, L.-H., Feuerschütte, S., et al. (2011). *Ein hedonischer Immobilienpreisindex auf Basis von Internetdaten 2007-2011* (tech. rep.). RWI Projektberichte.

bulwiengesa AG. (2018). Bulwiengesa-Immobilienindex 1975-2017. https://www.bulwiengesa.de/sites/default/files/immobilienindex_2018.pdf. [Accessed: November 2019].

Case, K. E., & Shiller, R. J. (1989). The Efficiency Of The Market For Single-Family Homes. *The American Economic Review*, *79*(1), 125.

Clapp, J. M., & Giaccotto, C. (1992). Estimating Price Trends for Residential Property: A Comparison of Repeat Sales and Assessed Value Methods. *The Journal of Real Estate Finance and Economics*, *5*, 357–374.

Der Obere Gutachterausschuss für Grundstückswerte im Land Nordrhein-Westfalen. (2017). Grundstücksmarktbericht 2017 Nordrhein-Westfalen. https://www.boris.nrw.de/borisfachdaten/gmb/2017/GMB_000_2017_pflichtig.pdf. [Accessed: November 2019].

Englund, P., Quigley, J. M., & Redfearn, C. L. (1998). Improved Price Indexes for Real Estate: Measuring the Course of Swedish Housing Prices. *Journal of Urban Economics*, *44*(2), 171–196.

Eurospace AG. (2019). Index EPX Hedonic. https://www.europace.de/index-epx-hedonic/. [Accessed: November 2019].

Federal Agency for Cartography and Geodesy. (2019). Verwaltungsgebiete (2019) 1:25000 - VG250 und VG250-EW. https://gdz.bkg.bund.de/index.php/default/digitale-geodaten/verwaltungsgebiete.html.

Federal Statistical Office. (2024a). House Price Index, Price Index for Building Land: Germany, Years - Table: 61262-0001. [Data set].

Federal Statistical Office. (2024b). Price Indices for Owner-Occupied Housing: Germany, Years - Table: 61262-0003. [Data set].

Federal Statistical Office. (2024c). Index of Net Rents Exclusive of Heating Expenses: Länder, Years - Table: 61111-0020. [Data set].

Gatzlaff, D. H., & Haurin, D. R. (1998). Sample Selection and Biases in Local House Value Indices. *Journal of Urban Economics*, *43*(2), 199–222.

Ghysels, E., Plazzi, A., Valkanov, R., & Torous, W. (2013). Forecasting Real Estate Prices. *Handbook of Economic Forecasting*, *2*, 509–580.

Gutachterausschuss für Grundstückswerte in Sachsen-Anhalt. (2017). Grundstücksmarktbericht Sachsen-Anhalt 2017.

RWI. (2018). Überprüfung des Zuschnitts von Arbeitsmarktregionen für die Neuabgrenzung des GRW-Fördergebiets ab 2021. *RWI Projektberichte*. http://www.rwi-essen.de/media/content/pages/publikationen/rwi-projektberichte/rwi-pb_zuschnitt_arbeitsmarktregionen.pdf [Accessed: July 2019].

RWI - Leibniz Institute for Economic Research. (2024a). RWI-GEO-RED: RWI Real Estate Data - Apartments for Sale. https://doi.org/10.7807/immo:red:wk:v11 [Data set].

RWI - Leibniz Institute for Economic Research. (2024b). RWI-GEO-RED: RWI Real Estate Data - Apartments for Rent. https://doi.org/10.7807/immo:red:wm:v11 [Data set].

RWI - Leibniz Institute for Economic Research. (2024c). RWI-GEO-RED: RWI Real Estate Data - Houses for Sale. https://doi.org/10.7807/immo:red:hk:v11 [Data set].

RWI - Leibniz Institute for Economic Research. (2024d). RWI-GEO-REDX: Regional Real Estate Price Index for Germany, 2008-05/2024. https://doi.org/10.7807/immo:redx:v14 [Data set].

Sirmans, S., Macpherson, D., & Zietz, E. (2005). The Composition of Hedonic Pricing Models. *Journal of Real Estate Literature*, *13*(1), 1–44.

Thiel, P. (2024). FDZ Data Description: Real-Estate Data for Germany (RWI-GEO-RED v11) - Advertisements on the Internet Platform ImmobilienScout24 2007 – 06/2024. RWI Projektberichte, Essen.

Verband Deutscher Pfandbriefbanken e V. (vdp). (2019). INDEX: Der Immobilienpreisindex des vdp. https://www.pfandbrief.de/site/dam/jcr:120e19a8-cfd8-499d-8611-82141b5b0043/vdp_Index-Broschuere_Q3_2019_DE.pdf. [Accessed: November 2019].

Voigtländer, M. (2012). Real Estate Data Sources in Germany. In T. Just & W. Maennig (Eds.), *Understanding German Real Estate Markets* (pp. 3–17). Springer Berlin Heidelberg.

9. Appendix

9.1. Information on data restriction in the RWI-GEO-REDX Price Indices

From the original RWI-GEO-RED dataset, we exclude luxury apartments for the estimation of the RWI-GEO-REDX price indices. So that, rental apartments with rents exclusive utilities above 5,000 Euro per month are omitted. The living area is restricted from 15 to 400 sqm and up to 7 rooms per rental apartment. These restrictions still cover a very large range of rent prices and living space, but it discards the extremely small or very exclusive apartments from the portfolio. It is likely that the luxury apartment market is not fully integrated into the aver-age rental market, which makes comparisons over regions across these markets difficult. We further set a bottom limit of 15 sqm living space to include only self-efficient apartments and to exclude typing errors.

House sales are restricted in a similar fashion; 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. The aim, here, is to ensure self-efficiency but rule out possibly faulted on the one hand and extremely luxurious houses on the other hand, too. The focus of the analysis of house sales lies on single-family homes, thus, apartment buildings are excluded. Furthermore, holiday homes are excluded if declared explicitly by the seller as well as houses with more than five floors. The constraints are imposed on the offers for apartment sales likewise. Offered apartments with a price higher than 2 Million Euro, more than eight rooms and an adverted living area below the 1st percentile and above the 99th percentile are not accounted for in the following estimation.

When placing the advertisement online, the user decides which information to publish on the advertised real estate. Working with these self-declared information leads to many 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”.

Table A1
Explanatory Variables in Analysis of House, Apartment and Rent Prices

Variable	Description	House sales (HS)	Apartment rents (AR)	Apartment sales (AS)	Restrictions
		Numerical variable			
Number of rooms	Number of rooms in the unit	x	x	x	Restricted to: 15 rooms (HS), 7 rooms (AR), and 8 rooms (AS)
		Categorical variable			
Number of total floors	1 := Missing 2 := 1-3 Floors 3 := 4-5 Floors 4 := 6-10 Floors 5 := More than 10 floors			x	
Floor number of housing unit	0 := Missing 1 := Ground floor (UG) 2 := First floor (EG) 3 := 2nd to 3rd floor 4 := 4th to 5th floor 5 := 6th to 10th floor 6 := Above 10th floor			x	
Endowment	0 := Missing 1 := Simple 2 := Normal 3 := Sophisticated 4 := Deluxe	x	x	x	
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	x	x	x	
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]	x			Restricted to 2,500 sqm. due to possible agricultural use
		Binary variable			
First occupancy	1 if new owner or renter move in as first occupancy	x	x	x	
Detached house	1 if house is detached	x			
Semi-detached house	1 if house is semi-detached	x			
Terraced house	1 if house is a terrace house	x			

Continued on next page

Table A1 – *Continued from previous page*

Variable	Description	House sales (HS)	Apartment rents (AR)	Apartment sales (AS)	Restrictions
Exclusive house	1 if the property is declared as a mansion or castle	x			
Other house type	1 if house is categorized differently	x			
Balcony	1 if property has a balcony		x	x	
Garden	1 if apartment has access to a private garden		x	x	
Guest toilet	1 if housing unit includes a guest toilet	x	x	x	
Built-in kitchen	1 if housing unit comes with a built-in kitchen		x	x	
Granny flat	1 if the property contains a separate “granny flat” or secondary suite	x			
Cellar	1 if cellar room is available		x	x	
Common charge	1 if common charge is declared in offer			x	
Lift	1 if property contains a passenger lift			x	

Notes: In the report on the property market of the federal state of North Rhine-Westphalia (Der Obere Gutachterausschuss für Grundstückswerte im Land Nordrhein-Westfalen, 2017) the referees do not count sales of undeveloped rural plot area under 2,500 sqm. in their statistics of farmland sales. This is also the case in the report on the property market for the scarcely populated state of Saxony-Anhalt (Gutachterausschuss für Grundstückswerte in Sachsen-Anhalt, 2017). This gives a notion that plot areas above the margin of 2,500 sqm. can be of interest for agricultural and not only residential purposes. To focus on house sales for living purposes without further commercial use, only plot areas smaller than 2,500 sqm. are included in the following.
Source: Authors’ table based on RWI-GEO-REDX (RWI - Leibniz Institute for Economic Research, 2024d).