Understanding Spatial House Price Dynamics in a Housing Boom*

Leo Kaas† Georgi Kocharkov‡ Nicolas Syrichas§

March 10, 2024

Abstract

We examine the evolution of spatial house price dispersion during Germany's recent housing boom. Using a dataset of sales listings, we find that house price dispersion has significantly increased, which is driven entirely by rising price variation across postal codes. We show that both price divergence across labor market regions and widening spatial price variation within these regions are important factors for this trend. We propose and estimate a directed search model of the housing market to understand the driving forces of rising spatial price dispersion, highlighting the role of housing supply, housing demand and frictions in the matching process between buyers and sellers. While both shifts in housing supply and housing demand matter for overall price increases and for regional divergence, we find that variation in housing demand is the primary factor contributing to the widening spatial dispersion within labor market regions.

JEL classification: D83; R21; R31

Keywords: House price dispersion; Spatial housing markets; Search frictions in housing markets

---

*We thank SAFE (project 21529) for financial support. The views expressed in this paper represent the authors' personal opinions and do not necessarily reflect the views of the Deutsche Bundesbank or the Eurosystem.

†Goethe University Frankfurt, kaas@wiwi.uni-frankfurt.de

‡Deutsche Bundesbank, georgi.kocharkov@bundesbank.de

§Free University Berlin, nicolas.syrichas@fu-berlin.de

1 Introduction

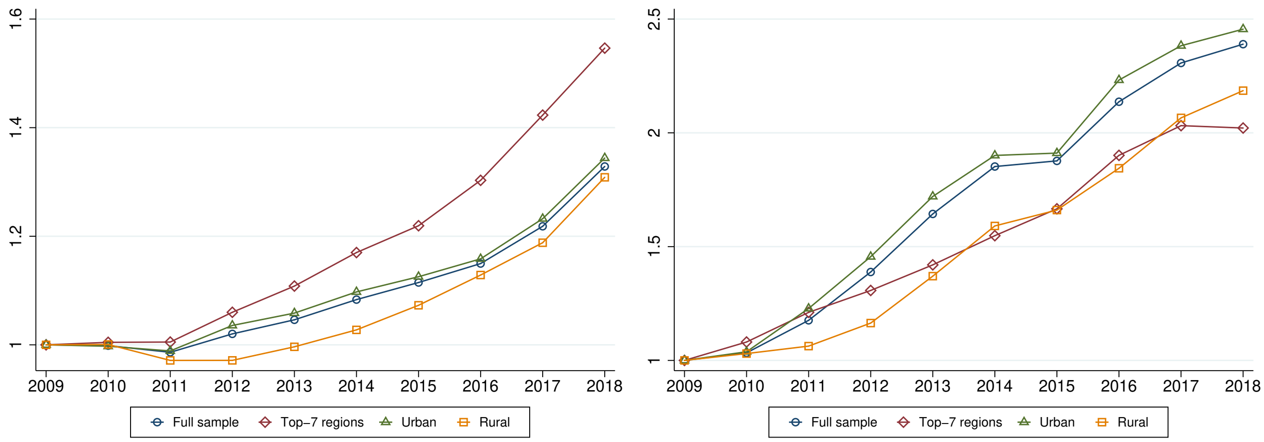
A striking feature of many housing markets is the large and often rising dispersion of house prices and rents across locations. Spatial dispersion of housing costs has several important social and economic consequences, such as widening wealth inequality between households, increasing residential segregation with spillovers on children's human capital (Fogli et al., 2023), or regional misallocation of capital and labor with detrimental effects on economic growth (Herkenhoff et al., 2018; Hsieh and Moretti, 2019). The existing literature on widening spatial price dispersion focuses on differences in house prices across metropolitan areas or municipalities (e.g. Van Nieuwerburgh and Weill, 2010; Gyourko et al., 2013), while house price dispersion at more granular levels remains largely unexplored.

This paper analyzes the trends and determinants of spatial house price dispersion during Germany's recent housing boom over the period 2009–2018. Different from other industrialized countries, real house prices in Germany did not exhibit any upward trends in the four decades prior to 2010.[1] Since then, however, real house prices increased overall and at varying speeds in different geographic subsamples, as is visible in panel (a) of Figure 1. At the same time, panel (b) illustrates that the spatial dispersion of house prices has widened sharply, even in rural regions where average house prices went up by much less than in urban regions. Also within the relatively homogeneous group of the largest seven metropolitan areas that saw the largest overall increase of house prices, a large increase of spatial price dispersion can be observed.

After documenting the main empirical patterns of Germany's housing boom and the simultaneous rise in spatial price dispersion, we build and estimate a simple spatial housing search model whose parameters can be identified from the price, contact-per-listing, and duration data at the postal code level. We use the estimated model to analyze the separate roles of housing supply, housing demand, and matching frictions for the observed price trends.

In Section 2 we describe a dataset of sales listings from Germany's largest housing online platform and document the contribution of location to the observed house price trends since the year 2009. We calculate inflation- and quality-adjusted house prices and find that the cross-sectional variance has increased substantially during 2009–2018. We first document that the entire increase of this variance is accounted for by an increase of dispersion between postal codes which we use as our granular location measure (cf. Figure 1.b), whereas within postal codes there is no change of house price dispersion. Second, we dissect spatial price dispersion into between and within labor market region components. For the full sample, the between-region components accounts for about two thirds of the between-location variance

[1] See Kindermann et al. (2021) for the historic house price development on the basis of different datasets. The doubling of nominal house prices during 1975–1995 shown in their paper is almost exactly offset by a doubling of the CPI during this period.



(a) Mean

(b) Spatial dispersion

Figure 1: Mean and spatial dispersion of house prices

NOTES: House prices are the residuals of hedonic regressions of inflation-adjusted prices in sales listings. Panel (a) shows the mean of these residuals, panel (b) shows the variance across postal codes, where all series are normalized to unity in the year 2009. See Section 2.1 for further details about the data, calculation of the variables and definition of the geographic subsamples.

and is responsible for about three quarters of the rise in spatial dispersion. We find similar results if we restrict the sample to only rural or urban regions. When focusing only on the more homogeneous Top-7 metropolitan regions, we find that rising dispersion within labor market regions accounts for about a half of the overall increase. We further document an overall tightening of the housing market, as evidenced by a decline of the duration of a listing and a substantial increase of the contact-per-listing-day ratio, which point to a surge of housing demand during the observation period.

In Sections 3 and 4, we propose and estimate a stylized housing search model that helps to understand the separate roles of demand, supply and matching frictions for rising spatial price dispersion, both between and within labor market regions. The model features homogeneous buyers and sellers whose house price valuations vary across space and over time. We further introduce time-invariant location premia that control the average market shares at the location level. While sellers choose the number of listings and the posted prices, buyers decide in which locations to search and which sellers to contact. In line with standard competitive search theory (cf. Moen, 1997), both sellers and buyers trade off prices and matching probabilities.

Importantly, our housing search model is a very stylized, not fully structural approach to describe price setting and price variation in spatial housing markets. While abstracting from the underlying reasons for demand or supply changes, all model parameters can be uniquely identified on the basis of our house listings data.

In the equilibrium of our model, prices, listing duration and tightness in a local housing market respond to the time and space variation of buyers' and sellers' house valuations and to a rent-sharing term that reflects housing market frictions. While the buyer valuation stands for the willingness to pay in certain locations, the seller valuation represents the outside option of a housing unit for sale which may reflect the construction cost of a new unit or the outside value of renting out an existing unit.[2] These two components capture the contributions of housing demand and supply, respectively. Additionally, differential trends in house prices could reflect differences in buyer-seller rent sharing between housing locations. Although ubiquitous in labor economics, this channel is mostly absent in the quantitative housing literature. Rent sharing here refers to the additional compensation that buyers pay in excess of the reservation price of sellers. In hot housing markets, sellers may exert more bargaining power over their buyers.

The model estimation uses a two-step procedure. First, we estimate matching functions on the basis of duration and contact-per-listing data. These parameters are estimated separately for each labor market region, where matching efficiency is additionally allowed to vary over time. The latter is required by our data which indicate an increase of matching efficiency in the second half of our observation period in most labor market regions. The second step is to jointly estimate the time- and location-specific valuations of buyers and sellers as well as the time-invariant location premia, using our data on prices, tightness, the estimates of matching functions, and the market shares. Within larger labor market regions, our model has several thousand parameters that include over 100 postal codes and 40 quarters. Nonetheless, this estimation step can be performed rather efficiently since our model is linear in nearly all parameters that are estimated at the second step.

In Section 5 we use the estimated model to quantify the driving forces behind the observed house price dynamics during the period 2009–2018. Through the lense of the model, three factors generate variation in house prices over time: housing supply via the valuation of sellers, housing demand via the valuations of buyers, or rent sharing between buyers and sellers which reflects trends in matching frictions and changes in market tightness. A simple counterfactual exercise is used to quantify the respective contribution of each of these factors for the increasing trend of house prices and their dispersion in the Top-7 metropolitan regions and for the between- and within-region variation.

We find that the majority of the rise of house prices in the Top-7 regions is accounted for by stronger housing demand, which accounts for around 80 percent of the price increase. Changes in supply have a secondary impact on the increase of prices, whose contribution

[2]Regulatory constraints and geographic barriers could impose hurdles in some premium locations driving sellers' valuations up (e.g. Saiz, 2010; Hsieh and Moretti, 2019).

to the overall increase varies between 5 and 30 percent. The rent-sharing factor has only a minor effect on the evolution of prices in any of the Top-7 labor market regions. House price dispersion increases in all but one region (Berlin) throughout the period 2009-2018. Rising within-region dispersion is mostly accounted for by differential changes in demand, while supply and rent sharing play a rather modest role.

We also use the estimated model to decompose the between-location variance into within- and between-region components, paralleling our data decomposition of Section 2. Similar to our findings for the Top-7 regions, the majority of within-region dispersion is attributed to demand-side changes. Nonetheless, a sizable share of between-region divergence is accounted for by housing supply, which possibly reflect the expansion of construction activity in relatively less demanded regions during this period. Changes in rent sharing have little impact on within-region dispersion and even a dampening effect on the rise of between-region price difference. The latter can be explained by regional convergence of housing market tightness over time.

1.1 Related Literature

Spatial dispersion. Our work is relates to Van Nieuwerburgh and Weill (2010) and Gyourko et al. (2013), who study reasons why house price dispersion across U.S. metropolitan areas increased over time. Van Nieuwerburgh and Weill (2010) use a dynamic spatial equilibrium model in the spirit of Rosen (1979) and Roback (1982) to show that high-ability households move into metropolitan areas with high wages and stringent regulatory housing supply. Likewise, Gyourko et al. (2013) argue that house price differentials in large metropolitan cities can be attributed to inelastic supply combined with an increasing sorting of high-income households. Our article differs from these two studies in two dimensions. First, we consider house price variation at a much more granular level. In particular, we show that house prices exhibit increasing dispersion over time, not only across labor market regions (i.e metropolitan areas) but also at the postal code level within labor market regions. Second, to use information on listing duration and contact-per-listing data, we employ a spatial directed search matching model that accounts for frictions in local housing markets instead of the frictionless island-type model of Van Nieuwerburgh and Weill (2010).

Our paper is related to recent empirical studies explaining differential house price trends during a housing boom. Kindermann et al. (2021) study regional disparities in house prices across German labor markets in the same ten-year period, focusing on the role of regional differences in expectation formation. Amaral et al. (2023) study the relationship between price and rent divergence across metropolitan areas in 15 advanced economies during a period of low-interest rates. They find that house prices have increased at a much faster pace compared to rents, both in major metropolitan areas but also on the national level.

Again the focus of this paper is on house price trends at a more granular spatial level. While we do not consider rents in our main analysis, we also document in the appendix that rent dispersion has also increased over time across postal codes, especially within labor market regions.

Housing market search. On the modelling side our paper relates to a literature employing directed search models to explain salient features of housing markets (Albrecht et al., 2016; Hedlund, 2016; Rekkas et al., 2022; Moen et al., 2021; Kotova and Zhang, 2021; Garriga and Hedlund, 2020). Closest to our work is Rekkas et al. (2022) who use a directed search model with heterogeneous buyers which they estimate using listings data from the Vancouver area. Similar to us, they find that heterogeneous tastes of buyers explain much of house price dispersion, whereas search frictions matter only little for dispersion (although contribute to the price stickiness observed in their data). Our paper mainly differs in two dimensions. First, we use our model to disentangle the respective contributions of buyers' and sellers' valuations, next to search frictions, for house price dynamics. Second, we seek to explain the factors that account for spatial dispersion between and within labor market regions.

Our paper also relates to a literature that uses online listings data to study the role of imperfect and costly information frictions to house price variation (Ben-Shahar and Golan, 2022; Kotova and Zhang, 2021; Guren, 2018). Our paper differs from this literature in its focus on the structural factors that explain residual variation across locations, rather than frictions that generate variation in the prices of similar houses within locations.

2 Empirical Patterns

2.1 Data

We use sales listings of residential housing units in Germany that were posted at the online platform *ImmobilienScout24* during January 2009 and December 2018.[3] The raw data are further prepared, geo-referenced and labeled by the RWI Essen within the RWI-GEO-RED dataset which can be accessed for research purposes. Next to the posted prices, this dataset contains a large number of housing characteristics, including geographical location at the square-kilometer level. It further contains information on the duration of a listing in days, the number of views that a listing received and the number of contact attempts of potential buyers.

A limitation of these data is that only listed prices are available, but not the actual transaction prices. However, we examine transaction prices aggregated at the district (Kreis)

[3] *ImmobilienScout24* is the largest real estate listing website in Germany with a self-reported market share of over 50 percent (Georgi and Barkow, 2010).

level from a private provider (Bulwiengesa) and do not observe significant discrepancies between the trends of posted prices and transaction prices between the two datasets. In a similar comparison between posted prices aggregated at a city level and the newly created German Real Estate Indices (GREIXX) across cities covered by this measure, we find striking similarities of the levels and the evolution of these two series over time.[4] Moreover, earlier studies using both transaction and listing price data show that on average a property sells within 1.6% of its listed price (Guren, 2018). Nonetheless, we do observe if the same property has been listed multiple times within a short horizon with marginal changes. In those cases, we keep only the last listing.[5]

2.2 Hedonic Regressions

Since we are interested in spatial variation of house prices over time, rather than changes in the composition of housing units for sale, we control for any observable differences in the characteristics of these housing units. To this end, we estimate a standard hedonic house price regression for our sample of sales listings. We pool all observations and estimate the OLS regression

$$\log p_{ht} = \text{const} + X'_{ht}\beta_X + \varepsilon_{ht} \ ,$$

where $\log p_{ht}$ is the (log) inflation-adjusted listed price per m^2 of housing unit h posted at time t, X_{ht} is a vector of housing characteristics of that property which includes a set of categorical variables for the number of rooms, dummy variables for guest toilet and cellar, age of the property in five-year categorical intervals, 22 categories indicating the type of the property, and quarterly dummies to take care of seasonal variation. Appendix C provides further details about the control variables.

We are interested in the residuals ε_{ht} of this hedonic regression which we can aggregate at the location level in the quarterly panel. Note that these residuals include not only location premia but also their variation over time.

[4]The GREIX are constructed on a quarterly frequency using transaction data of nominal prices for different types of properties across multiple cities. The indices are available for flats in 16 cities, for single-family houses in 13 cities, and for multi-family houses in 13 cities. For more details, see https://greix.de/.

[5]Another concern is the presence of phishing or fraud listings which usually look like legitimate listings, often at below market prices to attract potential buyers. *ImmobilienScout24* has developed a sophisticated algorithm to detect and remove those listings. It is also a fee-based platform, so that the cost for listing a fake offer is high. To alleviate remaining concerns, we remove ultra-popular offers (i.e. listings with hits or contacts beyond the 99th percentile) in our data cleaning process.

2.3 Baseline Sample

Since we are interested in the spatial distribution of house prices and its changes over time, we construct a quarterly panel which builds on postal codes as our main geographical unit.[6] We restrict the sample to those postal codes that contain at least ten listing observations in all quarters of our ten-year period.

As a larger aggregate geographic unit, we use the labor market regions categorised by Kosfeld and Werner (2012). These regions, which usually combine several municipalities and districts, are characterised according to commuter links to local labor market centers. Since some rural labor market regions are not well represented, we drop all labor market regions with less than 14 postal codes.

Both restrictions mitigate the impact of regions or postal codes which are sparsely populated and contain only few listings. In the following, we refer to postal codes as *locations*, while *regions* denote the labor market regions in our classification. The final balanced panel contains 2,161 locations in 99 regions over 40 quarters. It is important to note that none of our empirical findings is sensitive to these sample restrictions. For further details about our data across time and space see Appendix B.

2.4 Descriptive Statistics

Table 1 shows descriptive statistics of our baseline sample, reporting the means of selected variables, separate for five two-year periods. The first two rows illustrate the sharp rise in house prices over the ten-year horizon. The average inflation-adjusted house listed for sale in the 2009-2010 period cost around €1451 per m^2. Ten years later the posted sales price increased around 36% to €1978 per m^2. Note that this increase cannot be attributed to changes in housing characteristics, as the hedonic house prices $\bar{\varepsilon}_t$ exhibit a similar increase as the raw prices (in log points). When restricted to the largest seven labor market regions, house prices grew by 58% (from €1863 to €2951 per m^2), indicating a widening of cross-regional house price dispersion which we elaborate on in the next section.[7]

The bottom four rows of Table 1 indicate a tightening of the German housing market over the same period. The average number of listings in a location per quarter decreased by 35 percent, while the average duration of a listing fell from 57 to 45 days, and the number of contacts (i.e. buyers clicking the contact button) increased by 72 percent. The last row

[6] Relative to the km^2 grid information provided in the RWI-GEO-RED data, postal codes are larger and more homogeneous in population size. Germany has about 40.9m households and 8,200 postal code locations, so that a postal code includes on average about 5,000 households.

[7] Tables 12 and 13 in Appendix D display similar patterns for the rental market. Listed rents per m^2 increased by 18 log points all over Germany and by 23 log points in the Top-7 labor market regions over the same period.

reports the number of contacts per listing day as a flow-based measure of housing market tightness. This number almost quadrupled which indicates a substantial tightening of the German housing market over this ten-year period.[8]

Table 1: Descriptive statistics

Variable	2009-10	2011-12	2013-14	2015-16	2017-18
Log price $\ln p$	7.28	7.29	7.35	7.48	7.59
Price residual ϵ	-0.13	-0.12	-0.07	0.03	0.17
Listings S	71	69	73	58	46
Duration in days d	56	52	44	48	45
Contacts C	169	209	280	305	292
Flow tightness $\frac{C}{dS}$	0.05	0.07	0.11	0.16	0.19
Observations	17,288	17,288	17,288	17,288	17,288

NOTES: Means of selected variables for the baseline sample of location-quarter observations. Prices are in euros and adjusted for inflation using the CPI of the federal states in Germany.

2.5 House Price Dispersion Across Space and Time

Not only has the average house price gone up during 2009-2018, there is also a substantial widening of house price dispersion over the same period. To document this phenomenon, we go back to the level of individual listings and consider the residual posted price per m^2, denoted ε_{ht} for listing h at time t, as obtained from the hedonic regression described above. Across listings, the variance of residual prices has increased by over 50 percent, see Table 2.

To understand the spatial dimension of rising dispersion, we first decompose the variance of residual prices into within- and between-location components. Suppressing the time index, the variance of residual prices is split into

$$\underbrace{\operatorname{var}\varepsilon_h}_{\text{total variance}} = \underbrace{\sum_{i\in L} s_i \operatorname{var}_i(\varepsilon_h)}_{\text{within locations}} + \underbrace{\sum_{i\in L} s_i(\bar{\varepsilon}_i - \bar{\varepsilon})^2}_{\text{between locations}} , \tag{1}$$

where L is the set of locations (postal codes) with index i, $\bar{\varepsilon}_i = \frac{1}{n_i}\sum_{h=1}^{n_i}\varepsilon_h$ is the aver-

[8] Trends in the absolute number of listings S and contacts C may principally reflect changes in the market share of *ImmobilienScout24* over this ten-year period that may also vary between locations. However, to the extent that this platform is representative of the German housing market, such changes in market shares should not matter for the other two measures, namely listing duration d and flow tightness $C/(dS)$. Our identification strategy in Section 4 uses only these latter two variables. Hence, it builds on the assumption of representativeness of the platform, regardless of potential changes in its market share (between locations or over time).

age residual price in location i with number of listings n_i, and $s_i = n_i/(\sum_{j\in L} n_j)$ is the listing share of location i. $\bar{\varepsilon}$ is the average residual price across all of Germany. The within-location term on the right-hand side is the listing-weighted average of the variances $\text{var}_i(\varepsilon_h) = \frac{1}{n_i}\sum_{h=1}^{n_i}(\varepsilon_h - \bar{\varepsilon}_i)^2$ over all locations i. The second term is the listing-weighted variance of location-level prices, i.e. the between-location variance. We calculate this additive decomposition separately for each year.

Table 2: Between- and within location variance decomposition

	Total variance			Within locations			Between locations		
	2009	2013	2018	2009	2013	2018	2009	2013	2018
Full Sample	0.190	0.237	0.290	0.115	0.113	0.111	0.075	0.123	0.179
West-Germany	0.187	0.234	0.283	0.114	0.112	0.107	0.073	0.122	0.176
East-Germany	0.188	0.239	0.295	0.132	0.136	0.161	0.055	0.103	0.134
Top-7 regions	0.184	0.199	0.230	0.115	0.101	0.091	0.069	0.098	0.139
Urban	0.193	0.246	0.298	0.117	0.114	0.109	0.077	0.132	0.189
Rural	0.180	0.208	0.265	0.111	0.113	0.114	0.069	0.095	0.151

NOTES: "Full sample" builds on the listings in all quarter-location observations in our baseline sample. "West Germany" and "East Germany" include all listings located in Kreis (NUTS-3) that belonged to the FRG or GDR respectively before German reunification. The "Top-7 regions" comprise the labor market regions of Berlin, Munich, Hamburg, Frankfurt am Main, Cologne, Stuttgart and Dusseldorf. "Urban" denotes all units belonging to a Kreis indicated either as "Kreis", "Kreisfreie Stadt" or "Stadtkreis" and "Rural" all housing units located in a "Landkreis".

Table 2 reports the three terms in equation (1) separately for the years 2009, 2013 and 2018. Starting from the full sample, we see that the entire increase in variance is accounted for by the between-location component which increased steeply during 2009–2018, whereas the average within-location variance has not changed over time. In fact, while the within-location variance accounts for about 60 percent of the total variance in 2009, it merely contributes 38 percent to overall house price variation in 2018. Focusing on different geographic subsamples, this result is largely robust with some minor differences. In East German locations, house price dispersion has also gone up within locations, possibly reflecting rising disparities between unrenovated and modernized housing units (a housing characteristic that we cannot control in the hedonic regressions). In contrast, within urban and Top-7 locations, within-location dispersion has fallen, so that more than the entire increase of the variance is due to the between-location component.

The rising spatial dispersion of house prices is also documented in Figure 2 which shows the distribution of residual posted prices in the four years 2009, 2012, 2015 and 2018. During

2009–2015, the mode of these distributions remains rather stable, while the rise of the average house price is driven by a widening of house prices in the upper half of the distribution. During 2015–2018, the bottom half of the distribution has also widened substantially.

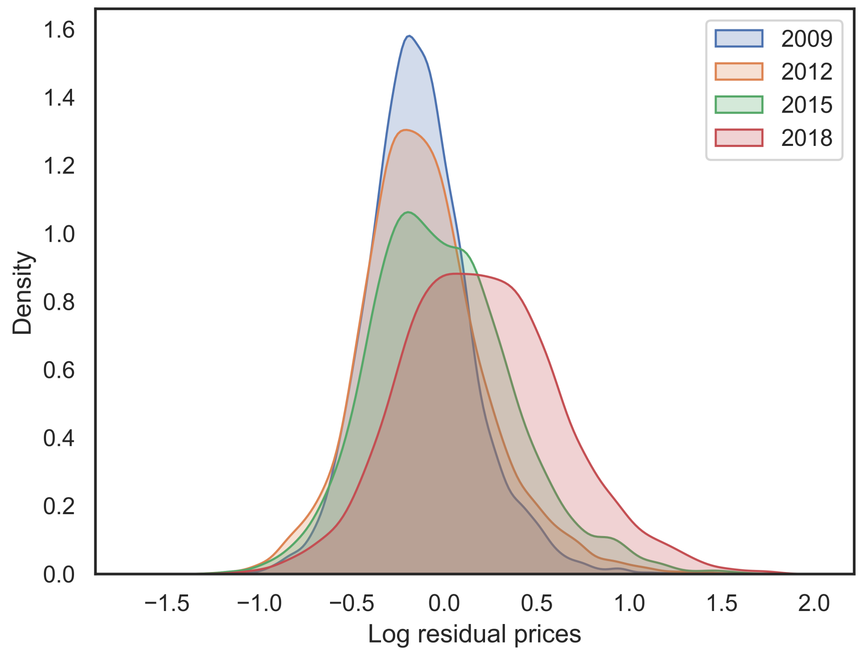


Figure 2: Distribution of residual prices across locations

NOTES: Between-location distributions of residual log prices in the years 2009 (blue), 2012 (orange), 2015 (green) and 2018 (red). The residuals are obtained from hedonic house price regressions as described in the main text and averaged in each location (postal code).

In light of the important role of location for rising house price dispersion, we are now asking to what extent these trends are driven by house price divergence between labor market regions or rising differences between locations within these regions. The variance decomposition of interest is

$$\underbrace{\sum_{i\in L} s_i(\bar{\varepsilon}_i-\bar{\varepsilon})^2}_{\text{between location variance}} = \underbrace{\sum_{r\in R}\sigma_r\,\mathrm{var}_r(\bar{\varepsilon}_i)}_{\text{within regions}} + \underbrace{\sum_{r\in R}\sigma_r(\bar{\varepsilon}_r-\bar{\varepsilon})^2}_{\text{between regions}}\ , \tag{2}$$

where R is the set of regions, $\sigma_r = \sum_{i\in r} s_i$ is the listing weight of region r, and $\bar{\varepsilon}_r \equiv \sum_{i\in r}\frac{s_i}{\sigma_r}\bar{\varepsilon}_i$ is the mean residual price of the region r. The first term is the listing-weighted average of the within-region variances $\mathrm{var}_r(\bar{\varepsilon}_i) \equiv \sum_{i\in r}\frac{s_i}{\sigma_r}(\bar{\varepsilon}_i-\bar{\varepsilon}_r)^2$, so that this term measures to what extent spatial house price differences are accounted for by differences between locations within labor market regions. The second term is the listing-weighted variance of average regional prices, i.e. the between-region variance. As before, this decomposition is calculated separately for each year. See Appendix F for derivations of the variance decompositions in equations (1) and (2).

Table 3: Within- and between-region variance decomposition

	Between-location			Within regions			Between regions		
	2009	2013	2018	2009	2013	2018	2009	2013	2018
Full Sample	0.075	0.123	0.179	0.032	0.048	0.054	0.043	0.076	0.125
West-Germany	0.073	0.122	0.176	0.032	0.047	0.055	0.041	0.075	0.121
East-Germany	0.055	0.103	0.134	0.031	0.053	0.048	0.024	0.049	0.086
Top 7 regions	0.069	0.098	0.139	0.044	0.060	0.073	0.025	0.037	0.066
Urban	0.077	0.132	0.189	0.034	0.049	0.053	0.043	0.083	0.136
Rural	0.069	0.095	0.151	0.018	0.027	0.033	0.051	0.068	0.118

NOTES: See the notes to Table 2 for definitions of the different samples.

Table 3 shows the results of this decomposition for different geographic units. Two interesting patterns emerge. First, in the full sample about 70 percent of the house price variance between locations in the year 2018 is accounted for by the between-region variance. Moreover, over three quarters of the rise in house price dispersion during 2009 and 2018 is driven by an increase in the variance of house prices between labor market regions, while less than a quarter of the increase is attributed to greater house price dispersion within labor market regions. Similar results are observed for the West and East German subsample, and also if we divide the sample into rural and urban regions. On the other hand, zooming into the Top-7 subsample, we find that around half of the increase in variance is driven by diverging house prices within the labor market regions. Furthermore, the within-location component accounts for the majority of overal spatial dispersion. Intuitively, labor market regions in this subsample are more comparable, so that a greater share of the variance (and its increase) is accounted for by the within-region variance (and its increase).[9]

These empirical patterns do not, of course, settle the question of what caused rising house prices and a widening of spatial house price dispersion in the first place. In Section 5

[9] In Appendix E we repeat the analysis of this section for the rental market where the increase in dispersion over this ten-year period is less pronounced than in the sales market. Similarly to house prices, we find that most of the increase in variance is attributed to rising disparities across locations (postal codes), although rental dispersion also increases within locations. Furthermore, as for house sales, the increase in cross-location variance is attributed to both within-region and between-region components where the latter plays a more important role.

we revisit this decomposition through the lens of our structural model that we estimate on our data and that sheds light on the relative role of demand, supply and rent-sharing factors for the observed house price developments.

3 Model

We propose a simple model that can be estimated on our data so as to analyze the driving forces behind the diverging house price trends documented in the last section. In particular, we aim to pin down the relative role of supply, demand and rent-sharing shifters in house prices at the location, region and aggregate level during the ten-year horizon covered in our data. The model describes a given labor market region that is divided into locations (postal codes). In each location, potential sellers decide about entry and the posted price of the housing unit for sale. Buyers decide in which location to search and which sellers to contact at their posted prices where trade is subject to search frictions. The housing market is characterized by directed search (Moen, 1997; Wright et al., 2021), while location decisions respond to taste shocks that are common in spatial dynamic choice models (e.g. Aguirregabiria and Mira, 2010; Caliendo et al., 2019). House prices and housing market tightness endogenously depend on the time and space variation of buyers' and sellers' housing valuations.

We deliberately keep the model parsimonious, abstracting from tenure choice, mortgage financing, differentiation of housing units by size or quality, and migration between labor market regions. While these simplifications leave out many important aspects of housing markets, they permit estimation of all key parameters on the basis of the listings data described in the previous section.

3.1 Environment

We consider a labor market region with a finite number of locations i (postal codes) over discrete time periods $t \geq 1$ (quarters). The region is populated by house buyers and sellers whose trade is subject to search frictions. Buyers and sellers aim to maximize discounted utility values with common quarterly discount factor β. All prices, values and costs in the model are understood as quality- and inflation-adjusted prices, values and costs per square meter.

3.1.1 Sellers

There is a free entry of sellers whose housing unit has exogenous outside value K_{it} in location i in period t, which represents either the construction cost of a new unit or the value of an existing unit under alternative use, such as the discounted value of a lease or the monetized value of owner occupancy. Free entry requires that the endogenous value of a seller V_{it}^S equals K_{it} in all local markets (i, t). A housing unit for sale involves cost c per period, reflecting the utility costs of a vacant unit as well as sales costs which are assumed to be constant across time and space.

3.1.2 Buyers

There is an exogenous inflow of new buyers into the region at time t, denoted B_t^n, so that the total number of buyers in the region, denoted B_t, is composed of unmatched buyers from the last period and the new buyers, where the stock of buyers in the first period B_1 is predetermined. Every buyer chooses in which location i to search in period t.[10] Search in location i yields utility value $V_{it}^B + \varphi_{it} + \tau_i$ where φ_{it} is an idiosyncratic (buyer-specific) taste shock which is type-I extreme value distributed with zero mean, and τ_i is a time-invariant location premium for location i that is common for all buyers and constant over time. V_{it}^B is the discounted utility value of a buyer searching in market i at time t, net of the taste shock and the location premium. If a buyer remains unmatched in market i, she decides in which location to search next period after drawing new idiosyncratic taste shocks. If a buyer is matched in period t, she pays the posted price and leaves the market with discounted utility value A_{it}. These values are exogenous to the model and represent the values that buyers attach to a (quality and size adjusted) housing unit in location i when bought at time t. In any period of search, we assume that the buyer pays a cost r_t which represents the rental cost in the region.

3.1.3 Search and Matching

Sellers post prices and buyers direct search to the sales listings, so that the housing market in a given location potentially segments into submarkets that are differentiated by posted prices and buyer-seller ratios. Both sides of the housing market trade off matching probabilities and prices, as is standard in markets with competitive search (Moen, 1997). When θ is the buyer-seller ratio (tightness) in a submarket, a seller is matched with probability $q_t(\theta)$ and a buyer is matched with probability $f_t(\theta) = q_t(\theta)/\theta$. q_t is a strictly increasing and strictly concave function, so that f_t is decreasing in tightness. We allow matching efficiency to vary over time which is why both functions are indexed by the time index t. Since all buyers and sellers searching in a given market (i, t) share the same respective values, only one submarket is active in this market which has posted price p_{it} and market tightness θ_{it}, both of which

[10] We rule out simultaneous search in multiple locations as we do not have enough information to discipline such a model feature.

are equilibrium outcomes as described below.[11]

3.2 Value Functions and Equilibrium

The Bellman equations of sellers and buyers in market i and period t are

$$\begin{aligned} V_{it}^S &= -c + \beta V_{i,t+1}^S + q_t(\theta_{it})\left(p_{it} - \beta V_{i,t+1}^S\right) \ , \\ V_{it}^B &= -r_t + \beta \bar{V}_{i,t+1}^B + f_t(\theta_{it})\left(A_{it} - p_{it} - \beta \bar{V}_{i,t+1}^B\right) \ . \end{aligned}$$

A seller pays flow cost c in the current period and is matched with probability $q_t(\theta_{it})$ in which case she leaves the market with continuation value p_{it}. Otherwise, she either continues to search in the same market or stops searching, yielding in both cases continuation utility $V_{i,t+1}^S = K_{i,t+1}$. A buyer pays flow cost r_t and is matched with probability $f_t(\theta_{it})$, yielding continuation utility $A_{it} - p_{it}$. Otherwise, an unmatched buyer has continuation utility

$$\bar{V}_{t+1}^B = \mathbb{E} \max_j [V_{j,t+1}^B + \varphi_{j,t+1} + \tau_j] = \ln\left[\sum_j e^{V_{j,t+1}^B + \tau_j}\right] \ , \tag{3}$$

where the expectation is over the realization of next period's idiosyncratic taste shocks $\varphi_{j,t+1}$.

In every local market (i,t), sellers post prices and buyers direct their search to the posted prices. Let (p,θ) denote the price-tightness combination in a potential submarket. Let Ω_{it} denote the expected buyer surplus from searching in market (i,t) which is identical for all (homogeneous) buyers in that market. Buyers must be offered at least surplus Ω_{it} to be willing to search in submarket (p,θ). A seller chooses (p,θ) to maximize the expected gain from trade,

$$\max_{p,\theta} q_t(\theta)[p - \beta V_{i,t+1}^S] \qquad \text{s.t.} \qquad f_t(\theta)[A_{it} - p - \beta \bar{V}_{i,t+1}^B] \geq \Omega_{it} \ .$$

The constraint says that sellers must offer at least surplus Ω_{it} to attract buyers to the submarket. Substituting the price and the matching function $f_t(\theta) = q_t(\theta)/\theta$ yields the first-order condition

$$\Omega_{it} = q_t'(\theta)[A_{it} - \beta \bar{V}_{i,t+1}^B - \beta V_{i,t+1}^S] \ .$$

Because the matching function q_t is strictly concave, all sellers choose the same price p_{it}, so that only one submarket is active with tightness θ_{it}. Using $\Omega_{it} = f_t(\theta_{it})[A_{it} - p_{it} - \beta \bar{V}_{i,t+1}^B]$

[11] Although dispersion in residual prices exists *within* locations in our data, the within-location component exhibits no time trend, see Table 2 in Section 2. Given our interest in widening house price dispersion over time, our model abstracts from this feature in the data.

and $f_t(\theta)\theta = q_t(\theta)$ gives the equilibrium price

$$p_{it} = \zeta_t(\theta_{it})\beta V^S_{i,t+1} + (1 - \zeta_t(\theta_{it}))[A_{it} - \beta \bar{V}^B_{i,t+1}] \ , \tag{4}$$

with matching function elasticity $\zeta_t(\theta) = q'_t(\theta)\theta/q_t(\theta) \in (0, 1)$. This equation demonstrates how the posted price in market (i, t) depends on housing supply (the sellers' valuation $\beta V^S_{i,t+1}$), housing demand (the buyers' gain from trade $A_{it} - \beta \bar{V}^B_{i,t+1}$), and the rent sharing factor $\zeta_t(\theta_{it})$ which responds to features of the matching technology and housing market tightness in market (i, t). We build on this equation for our decomposition analysis in Section 5.

Substituting the equilibrium price into the Bellman equations gives

$$V^S_{it} = \ -c + \beta V^S_{i,t+1} + (q_t(\theta_{it}) - \theta_{it} q'_t(\theta_{it})) \left[A_{it} - \beta \bar{V}^B_{i,t+1} - \beta V^S_{i,t+1}\right] \ , \tag{5}$$

$$V^B_{it} = \ -r_t + \beta \bar{V}^B_{i,t+1} + q'_t(\theta_{it}) \left[A_{it} - \beta \bar{V}^B_{i,t+1} - \beta V^S_{i,t+1}\right] \ . \tag{6}$$

At the beginning of a period, all buyers B_t in a labor market region draw idiosyncratic taste shocks φ_{it} after which fraction

$$\pi_{it} = \frac{e^{V^B_{it} + \tau_i}}{\sum_j e^{V^B_{jt} + \tau_j}} \tag{7}$$

decide to search in location i. Over time, the number of buyers in the labor market region adjusts according to

$$B_{t+1} = \sum_i [1 - f_t(\theta_{it})]\pi_{it} B_t + B^n_{t+1} \ , \tag{8}$$

where B^n_{t+1} is the exogenous inflow of new buyers into the labor market region in period $t+1$ which adds to the number of unmatched buyers from the previous period.

Equilibrium Definition

Given an initial stock of buyers B_1 and buyer inflow B^n_t in periods $t \geq 2$, a *spatial competitive search equilibrium* describes, for all periods $t \geq 1$ and locations i, posted house prices p_{it}, market tightness θ_{it}, discounted values of sellers and buyers V^S_{it}, $\bar{V}^B_{it}$, V^B_{it}, location choices π_{it} and a total buyer stock B_t satisfying equations (3)–(8) and the free-entry conditions of sellers $V^S_{it} = K_{it}$.

4 Estimation

In this section, we explain how we estimate the parameters of this model for a given labor market region with $i = 1, \ldots, N$ locations (postal codes) and $t = 1, \ldots, T$ periods (quarters).[12] We restrict the analysis to those locations in which we observe at least ten listings in at least two thirds of the quarters in the period 2009Q1–2018Q4. We use the variables from the RWI-GEO-RED data which are averaged or aggregated at the location-quarter level as described in Section 2. These are the residualized average hedonic price p_{it},[13] the number of listings S_{it} which we identify with the number of sellers,[14] average duration of a listing in days d_{it} and the number of buyer contacts C_{it}. Note that the stock of buyers, and therefore market tightness, is not observed. We explain below our identifying assumptions that allow us to back out these values and to estimate a matching function from information on listing duration d_{it} and the numbers of contacts C_{it} and listings S_{it}.

Three model parameters are calibrated outside the model. β is a standard discount factor at quarterly frequency that equals 0.995 to match an annual interest rate of 2%. c is set to an estimate for service charges per square meter, e.g. 6.50 Euros per quarter.[15] For the quarterly costs of an unmatched buyer r_t, we use the average, inflation-adjusted rental rate per m^2 in the region which we take from the rental listings in the RWI-GEO-RED dataset.

The model estimation proceeds in two steps. First, we estimate a matching function, separate for each region, using the data on listing duration, number of listings and contacts. Second, we back out the buyer and seller valuations A_{it} and K_{it} that are consistent with the observed variation of prices, tightness and matching rates across time and space, and we estimate the location premia τ_i which control the distribution of buyers across locations.

4.1 Matching Function

In the data, we measure the stock of sellers by the number of listings S_{it}, but we do not observe the stock of buyers in a given market (i, t), denoted B_{it}. Hence market tightness

[12] We choose a quarterly period length to smooth out very short-term (monthly) volatility at the local level that might partly arise due to a low number of observations. A quarter also seems to reasonably reflect a typical planned transaction time for buyers and sellers in the housing market. We assign the day of first listing to a specific quarter so that some days of an active listing may fall into the next quarter.

[13] Specifically, we take the residual of the hedonic, inflation-adjusted log prices per m^2 at the listing level ε_{ht} as defined in Section 2.1, delog and multiply them with the average price in Euros, and average to the listing level to obtain p_{it}.

[14] Multiple listings of the same seller should not raise concerns since they show up as independent listings on the *ImmobilienScout24* platform. Note that we only consider listings of single housing units in our analysis.

[15] See https://www.mieterbund.de/service/betriebskostenspiegel.html which reports a monthly cost of 2.17 Euros per month and square meter in 2018. This includes utilities, insurance, property tax, among others, and may be a lower bound as it does not include any additional sales costs.

$\theta_{it} = B_{it}/S_{it}$ is unobserved. However, we build on the assumption that the search intensity of every active buyer is the same so that every buyer contacts a given number of listings per day.[16] Therefore, we estimate an *auxiliary matching function* using the contact-per-listing-day ratio $\vartheta_{it} = C_{it}/(d_{it}S_{it})$ as an auxiliary flow-based measure of market tightness, which is then transformed into a matching function that depends on the buyer-seller ratio, as will be further explained below.

We first regress log duration of a listing on the contact-per-listing-day ratio, pooling all locations and quarters in a labor market region. That is, we estimate

$$\ln d_{it} = a_0 + a_1 \ln \vartheta_{it} + g_t + \varepsilon_{it} \ , \tag{9}$$

where ε_{it} is an error term and g_t is a time fixed effect. The latter takes care of any time trends in the listing duration relationship, as well as seasonality in the housing markets as documented in previous literature (cf. Ngai and Tenreyro, 2014).

Table 4: Matching function estimation

y=ln(d_{it})	Berlin	Munich	Hamburg	Frankfurt	Stuttgart	Dusseldorf	Cologne
a_1	-0.32***	-0.41***	-0.31***	-0.25***	-0.35***	-0.28***	-0.24***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
a_0	2.70***	2.70***	2.97***	3.04***	3.06***	3.02***	3.21***
	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.216	0.419	0.330	0.306	0.501	0.361	0.433
N	5,440	3,440	3,760	3,960	2,800	3,680	2,720

NOTES: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4 shows the estimates of parameters a_0 and a_1 for each of the Top-7 labor market regions. As expected, in all cases parameter a_1 is negative, showing that more contacts per listing day relate negatively with the duration of the listing. The estimates show that a doubling of contacts per listing day goes together with a decrease of duration between 24 and 41 percent, depending on the labor market region. Figure 3 shows the distribution of the estimates of a_1 across all labor market regions in Germany.

The regression constant also varies between regions, showing that listing duration is about 58 percent longer in Cologne than in Berlin in the first quarter of 2009 (the reference category for the time fixed effect) for a given contacts-per-listing-day ratio. The time

[16]Here we follow the logic of matching function estimation in labor market models where typically the stock of unemployed workers is observed, but flow measures of search intensity (e.g. the number of applications sent per day) are not observed. In our data, we only observe the flow of contacts, but not the stock of buyers.

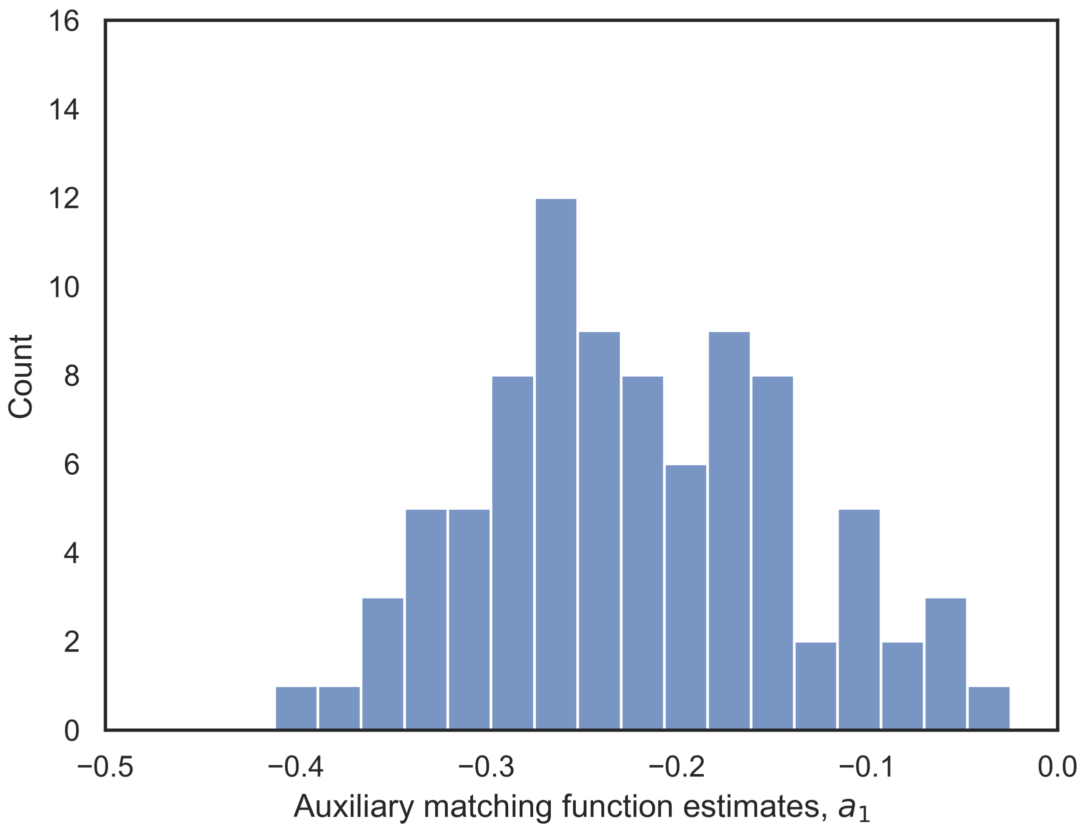


Figure 3: Distribution of estimates of a_1 across labor market regions

NOTES: The distribution includes all estimates of a_1 which are statistically different from zero.

trends, which are shown in Table 16 in the Appendix for the Top-7 regions, also show some heterogeneity between regions, but generally remain rather flat until 2014 after which listing duration has increased, conditional on the same contacts-per-listing-day ratio.

We can now explain how we measure buyer-seller ratios and the original matching function which maps buyer-seller ratios into matching probabilities for both market sides. Since the daily matching probability of a seller is the inverse of average duration, $q^d_{it} = 1/d_{it}$, the estimates of the auxiliary matching function relationship imply that $q^d_{it} = q_t \vartheta^{\mu}_{it}$ where $q_t = e^{-a_0 - g_t}$ and $\mu = -a_1$.

While the number of buyers B_{it} and their daily matching probability f^d_{it} are unobserved, we assume that a buyer contacts a given number of listing per day equal to k. Then the total number of contacts in market (i, t) is $C_{it} = kB_{it}\frac{1}{f^d_{it}}$ since a buyer searches on average $1/f^d_{it}$ days. It follows that the contacts-per-listing-day ratio is

$$\vartheta_{it} = \frac{C_{it}}{d_{it}S_{it}} = k\frac{B_{it}}{S_{it}}\frac{q^d_{it}}{f^d_{it}} = k\left(\frac{B_{it}}{S_{it}}\right)^2 ,$$

where the last equality uses that the number of matched sellers per day is identical to the number of matched buyers per day, $q^d_{it} S_{it} = f^d_{it} B_{it}$. Therefore, we can infer the unobserved buyer-seller ratio θ_{it} and the number of buyers from the auxiliary tightness measure as follows:

$$\theta_{it} = (\vartheta_{it}/k)^{1/2} \ , \tag{10}$$

$$B_{it} = S_{it}(\vartheta_{it}/k)^{1/2} \ . \tag{11}$$

Together with the estimated daily matching probability of a seller, we obtain quarterly matching probabilities for buyers and sellers, i.e. the matching function relationships used in Section 3:

$$q_t(\theta) = 1 - \left(1 - q_t k^{\mu} \theta^{2\mu}\right)^{90} \ ,$$
$$f_t(\theta) = q_t(\theta)/\theta \ .$$

We set parameter k such that $f_t(\theta)$ is a probability for all plausible data observations. Specifically we winsorize extreme observations of ϑ_{it} outside a large enough interval $[\vartheta_{min}, \vartheta_{max}]$ and set k such that $f((\vartheta_{min}/k)^{1/2}) = 0.99$.[17]

4.2 Location Premia and Valuations of Buyers and Sellers

The second step of our estimation procedure is to simultaneously estimate the time-invariant location premia τ_i and the time-varying buyer and seller valuations of a housing unit, A_{it} and K_{it}. The latter two objects can be uniquely pinned down to exactly match the observed prices p_{it} and tightness levels θ_{it} which are both measured from our data as described above. Location premia are set to match the average buyer market shares in all locations. We demonstrate that this joint estimation can be implemented as the solution of a high-dimensional, yet tractable equation system.

The buyer market shares differ in the data and in the model according to

$$\hat{\pi}_{it} = \pi_{it} e^{\eta_{it}} \ ,$$

where $\hat{\pi}_{it} = \frac{\hat{B}_{it}}{\hat{B}_t}$ is the share of buyers in market i at time t in the data and η_{it} is an error term. Buyer stocks in the data are inferred as explained in the previous subsection. We choose location premia τ_i to minimize $\sum_{i,t} \eta^2_{it}$ subject to the requirement that the average location premium is zero, $\sum_i \tau_i = 0$.

[17] We set $\vartheta_{min} = e^{M-3\sigma}$ and $\vartheta_{max} = e^{M+3\sigma}$ where M and σ are the mean and standard deviation of $\ln \vartheta_{it}$. Note that our estimates imply that $q((\vartheta_{max}/k)^{1/2}) < 1$, so that q is also a probability for all valid observations.

From (3) and (7) follows

$$V_{it}^B + \tau_i = \ln \pi_{it} + \bar{V}_t^B \ , \tag{12}$$

so that we can write

$$\eta_{it} = \ln \hat{\pi}_{it} - \ln \pi_{it} = \ln \hat{\pi}_{it} + \bar{V}_t^B - V_{it}^B - \tau_i \ .$$

Minimization of $\sum_{i,t} \eta_{it}^2$ subject to the constraint

$$\sum_i \tau_i = 0 \tag{13}$$

with respect to τ_i has the first-order conditions[18]

$$\tau_i = \frac{1}{T} \sum_{t=1}^{T} \left[\ln \hat{\pi}_{it} + \bar{V}_t^B - V_{it}^B \right] - \frac{\lambda}{2T} \ . \tag{14}$$

where λ is the multiplier on the constraint.

Since our model is set up as an infinite-horizon model, we need to make an assumption about the forecasts of continuation values of buyers and sellers in the last observation period T. We do this by linearly extrapolating the values during the observation period, namely V_{it}^S and V_{it}^B for $t = 1, \ldots, T$, to the first quarter thereafter. The analytic expressions of this extrapolation procedure are:

$$V_{i,T+1}^S = \frac{2}{T(T-1)} \left\{ \sum_{t=1}^{T} V_{it}^S \left[3t - (T+2) \right] \right\} \ , \tag{15}$$

$$V_{i,T+1}^B = \frac{2}{T(T-1)} \left\{ \sum_{t=1}^{T} V_{it}^B \left[3t - (T+2) \right] \right\} . \tag{16}$$

We are now ready to describe how the unknown valuation parameters (A_{it}, K_{it}) and location premia τ_i can be calculated. Let the number of locations in the labor market region be denoted N. From our data, we use the (quality- and inflation-adjusted, per square meter) prices p_{it}, tightness θ_{it} as calculated above, the estimated (time-varying) matching function relationships, and the buyer market shares $\hat{\pi}_{it}$ as defined above. Then the pricing equations (4), the two Bellman equations (5) and (6), the extrapolation equations (15) and (16), the optimality conditions (13) and (14), and the continuation utilities of unmatched buyers (3) constitute a system of $(3N+1)(T+1)+1$ equations in $(3N+1)(T+1)+1$ unknowns:

[18] This minimization takes the values $\bar{V}_t^B$ as given, and hence ignores the impact of τ_i on the values of unmatched buyers, see equation (3). This approximation is innocuous when the number of locations is large so that the impact of each τ_i on $\bar{V}_t^B$ is negligible.

$(A_{it})_{t=1}^{T}$, $(V_{it}^{B}, V_{it}^{S})_{t=1}^{T+1}$, τ_i for $i = 1, \ldots, N$, λ, and $(\bar{V}_t^B)_{t=1}^{T+1}$. Except the $T+1$ equations (3), these are all linear equations. Their joint solution is straightforward to implement, with further details described in Appendix H. The solution gives the buyer and seller valuations A_{it}, $K_{it} = V_{it}^{S}$, and location premia τ_i.

5 Results

In this section, we use the estimated model to study what factors contribute to house price dynamics in terms of the observed rise in house prices and their spatial dispersion. We distinguish between three contributing forces: *(i)* housing supply factors related to the location- and time-specific seller valuations K_{it}, *(ii)* housing demand factors represented by the buyer valuations A_{it}, and, *(iii)* rent-sharing factors associated with the region- and time-specific matching frictions through the matching function $q_{rt}(\cdot)$. We first focus in Section 5.1 on the Top-7 labor market regions and document the contribution of the three factors separately for each of these metropolitan regions. In Section 5.2 we turn to the between- and within-region variance decomposition considered in Section 2 and investigate the contribution of supply, demand, and rent sharing for the observed changes.

5.1 Model-Based Decomposition

We use the observed hedonic prices p_{it}, tightness levels θ_{it}, the estimated valuation parameters K_{it}, A_{it} and $\bar{V}_{rt}^B$, and the estimated matching functions $q_{rt}(\cdot)$ to isolate the role of factors *(i)*-*(iii)* for generating the observed house price dynamics. Building on the equilibrium pricing equation (4), we express the price in location i at time t into the following terms:

$$p_{it} = \underbrace{\zeta_{rt}(\theta_{it})}_{\text{Rent Sharing}} \underbrace{\beta K_{i,t+1}}_{\text{Supply}} + \underbrace{(1 - \zeta_{rt}(\theta_{it}))}_{\text{Rent Sharing}} \underbrace{[A_{it} - \beta \bar{V}_{r,t+1}^B]}_{\text{Demand}}.$$

To measure the respective contributions of the three factors, we proceed as follows. Regarding the contribution of housing supply, we fix housing demand via the buyers' gain and the rent-sharing factor to their initial 2009 values, $A_{i,1} - \beta \bar{V}_{r,2}^B$ and $\zeta_{r,1}(\theta_{i,1})$, while allowing housing supply via the sellers' valuations $\beta K_{i,t+1}$ to evolve. By doing so, we derive counterfactual prices p_{it}^{supply} through the pricing equation which reflect only the shifts in housing supply. Similarly, fixing housing supply and rent-sharing to their initial values, while letting housing demand evolve, we derive another set of counterfactual prices p_{it}^{demand} which reflect only shifts in housing demand. Finally, the constructed prices p_{it}^{rent} summarize only changes in the rent-sharing factor $\zeta_{rt}(\theta_{it})$.

Note that the rent-sharing factor varies across space and across time for different reasons:

First, the estimated matching function elasticity parameters are allowed to differ across regions, while they are identical across locations (postal codes) and over time. Hence, these parameters only matter for between-region components considered in Section 5.2. Second, the matching function elasticity $\zeta_{rt}(.)$ is not constant but decreases in market tightness which itself varies across locations and over time. Intuitively, a tighter housing market (from the buyers' point of view) intensifies the congestion externality on the buyers' side and relaxes the congestion externality on the sellers' side which contributes to a price increase without changes of buyers' or sellers' valuations.

Table 5: Decomposition of average price changes, 2009-2018

	$\bar{p}_T - \bar{p}_1$	$\bar{p}_T^{\text{supply}} - \bar{p}_1$	$\bar{p}_T^{\text{demand}} - \bar{p}_1$	$\bar{p}_T^{\text{rent}} - \bar{p}_1$
Munich	0.643	0.185	0.520	-0.014
	(100)	(29)	(81)	(-2)
Frankfurt	0.312	0.012	0.268	-0.044
	(100)	(4)	(86)	(-14)
Berlin	0.574	0.049	0.505	-0.073
	(100)	(9)	(88)	(-13)
Stuttgart	0.491	0.158	0.361	-0.015
	(100)	(32)	(74)	(-3)
Cologne	0.285	0.022	0.241	-0.030
	(100)	(8)	(85)	(-11)
Hamburg	0.446	0.115	0.369	0.009
	(100)	(26)	(83)	(2)
Dusseldorf	0.262	0.048	0.221	0.001
	(100)	(18)	(84)	(1)

NOTES: The supply, demand and rent-sharing contributions to the change of average log prices between 2008 and 2019 in Top-7 labor market regions are derived as described in the text. Percentages of the total log price change for each region are shown in parentheses.

Table 5 summarizes the contribution of factors *(i)*-*(iii)* to the average price change in each of the Top-7 labor market regions between 2009 and 2018, $\bar{p}_T - \bar{p}_1$. Housing supply in isolation produces a price change $\bar{p}_T^{\text{supply}} - \bar{p}_1$, while housing demand contributes $\bar{p}_T^{\text{demand}} - \bar{p}_1$. Finally, changes in rent-sharing account for $\bar{p}_T^{\text{rent}} - \bar{p}_1$. The numbers in parentheses show the percent of the overall price change generated by the different factors, separately for each region. Note that these percentages do not add to 100 since the three counterfactual scenarios build on a non-linear equation.

The highest average price increase over the period is 90 percent (0.643 log points) with respect to the initial price and occurs in Munich, while the lowest is in Dusseldorf, 30 percent. In all regions, the demand-driven price change contributes the most to the average price increase. For instance, in Berlin, shifts in housing demand have the highest contribution and account for around 86 percent of the overall price increase, while in Stuttgart this number is the lowest, 74 percent. Housing supply factors produce much smaller contributions to the average price changes. In Stuttgart, their contribution is 32 percent, while it is only 4 percent in Frankfurt. A stronger expansion of housing construction in Frankfurt relative to Stuttgart could be a likely explanation for these differences. Finally, changes in rent-sharing factors have a relatively small effect on price changes between 2009 and 2018. In most regions, their contribution is negative, i.e. market frictions alone produce a price decrease. Despite housing markets becoming tighter overall (see Section 2), the increase is relatively weaker in locations with initially high buyer valuations, so that ultimately rent sharing component contributes negatively. Only in Hamburg and Dusseldorf, changes in rent-sharing factors contribute positively to the price increase, although by a negligible amount.

Table 6: Decomposition of changes in house price dispersion, 2009-2018

	$\text{var}(p_T) - \text{var}(p_1)$	$\text{var}(p_T^{\text{supply}}) - \text{var}(p_1)$	$\text{var}(p_T^{\text{demand}}) - \text{var}(p_1)$	$\text{var}(p_T^{\text{rent}}) - \text{var}(p_1)$
Munich	0.003	-0.014	0.011	-0.007
	(100)	(-465)	(365)	(-227)
Frankfurt	0.035	0.002	0.031	-0.002
	(100)	(5)	(88)	-7
Berlin	-0.011	-0.018	-0.007	-0.007
	(100)	(167)	(66)	(65)
Stuttgart	0.076	0.002	0.021	0.001
	(100)	(8)	(95)	(3)
Cologne	0.086	0.016	0.071	0.007
	(100)	(19)	(83)	(8)
Hamburg	0.020	-0.003	0.022	-0.001
	(100)	(-14)	(110)	(-5)
Dusseldorf	0.062	0.011	0.053	0.003
	(100)	(18)	(86)	(5)

NOTES: The supply, demand and rent-sharing contributions to the change of price dispersion (variance of log prices) between 2008 and 2019 in Top-7 labor market regions are derived as described in the text. Percentage of the total variance change for each region are shown in parentheses.

Table 6 displays the contribution of factors *(i)*-*(iii)* to the 2008-2019 change in price dispersion as measured by the variance of prices within a region. There is a large heterogeneity in the change of dispersion across labor market regions. In Berlin, the price dispersion has gradually declined over time (see also Figure 8 in the Appendix), while dispersion widened significantly in Cologne, Stuttgart, Dusseldorf, Frankfurt and Hamburg (in descending order of the increase). In Munich, dispersion did not change by much in between 2008 and 2019. This fact masks an inverted U-shaped pattern in the evolution of the price dispersion in this region, which also can be seen in Figure 8.

There is a significant heterogeneity in the contribution of each factor across the Top-7 regions too. Changes in housing demand accocunt for a sizable fraction of the overall dispersion change in all regions. In Munich and Hamburg, housing demand can account for more than the observed increase in price dispersion. In Stuttgart, Frankfurt, Dusseldorf and Cologne changes in demand can generate between 95 and 83 percent of the observed increase in dispersion. In Berlin, demand factors contribute to the decline in dispersion but they are not the leading factor.

Both in Berlin and in Munich, changes in housing supply contribute to declining dispersion. A likely explanation is that supply expanded relatively more in high-price locations which had a dampening impact on spatial dispersion. In the other regions, supply factors play only a secondary role. In Frankfurt, Stuttgart, Hamburg, Dusseldorf and Cologne, shifts in supply can generate less than a fifth of the observed rise in dispersion. Finally, in most cases changes in rent-sharing factors relating to changes in spatial dispersion of market tightness have a small effect on the evolution of price dispersion over time.

5.2 Within- and Between-Region Price Dispersion in the Model

We further use the counterfactual model-generated house price series due to changes in demand, supply and rent-sharing factors to extend the decomposition of the variance of house prices into within- and between-region components as in equation (2). First, we perform the original variance decomposition exercise of Section 2.5 for the same subsets of different geographic units used in Table 3.[19] Second, using the counterfactual prices which capture the contribution of factors *(i)*-*(iii)* to the 2008-2019 change in price dispersion, we can identify the share of within- and between-region price dispersion stemming from each of these factors alone.

Table 7 summarizes these results. The data decomposition is presented in a bold text for each georgraphical aggregation. The rows "Percent" depict the percent of the variance in each position in the prior bold rows relative to the initial level of the between-location

[19]These numbers are not exactly identical to those of Table 3 which is due to three reasons. First, the sample for this decomposition is slightly larger than in the one used in Section 2.5 because here we also include locations with a few quarters with missing observations which are imputed. Second, the decomposition in Section 2.5 is implemented at an annual frequency whereas here we use quarterly observations. Third, the data in the model implied decomposition are smoothed to reduce short-run volatility.

Table 7: Model-based within- and between-region decomposition

	Between-location variance			Within regions			Between regions		
	2009	2013	2018	2009	2013	2018	2009	2013	2018
Full Sample	**0.0758**	**0.1205**	**0.1701**	**0.0312**	**0.0455**	**0.0529**	**0.0446**	**0.0749**	**0.1172**
Percent	100.0	159.0	224.5	41.1	60.1	69.8	58.9	98.9	154.7
Demand	100.0	144.0	195.9	41.1	55.6	66.0	58.9	88.5	130.0
Supply	100.0	102.2	118.6	41.1	42.5	41.2	58.9	59.7	77.6
Rent Sharing	100.0	92.8	93.8	41.1	41.3	40.7	58.9	51.4	53.0
West-Germany	**0.0737**	**0.1203**	**0.1711**	**0.0311**	**0.0453**	**0.0532**	**0.0426**	**0.0750**	**0.1179**
Percent	100.0	163.2	232.0	42.2	61.5	72.2	57.8	101.7	159.9
Demand	100.0	147.4	202.1	42.2	56.8	68.2	57.8	90.6	133.9
Supply	100.0	102.8	118.2	42.2	43.4	42.0	57.8	59.4	76.3
Rent Sharing	100.0	92.6	92.3	42.2	42.2	41.5	57.8	50.3	50.6
East-Germany	**0.0565**	**0.0883**	**0.0998**	**0.0320**	**0.0499**	**0.0478**	**0.0245**	**0.0384**	**0.0521**
Percent	100.0	156.3	176.7	56.6	88.3	84.6	43.4	68.0	92.2
Demand	100.0	142.4	157.9	56.6	81.5	79.4	43.4	60.9	78.6
Supply	100.0	98.6	109.8	56.6	62.4	60.1	43.4	35.6	49.8
Rent Sharing	100.0	93.7	97.1	56.6	61.2	59.9	43.4	31.7	36.7
Top 7 Regions	**0.0693**	**0.0956**	**0.1381**	**0.0435**	**0.0575**	**0.0693**	**0.0258**	**0.0380**	**0.0688**
Percent	100.0	138.0	199.4	62.8	83.1	100.0	37.2	54.9	99.3
Demand	100.0	127.0	182.2	62.8	79.1	98.6	37.2	47.9	83.7
Supply	100.0	96.5	116.2	62.8	62.1	62.4	37.2	34.3	54.1
Rent Sharing	100.0	90.0	101.3	62.8	62.0	65.5	37.2	27.8	35.7
Urban	**0.0747**	**0.1195**	**0.1700**	**0.0327**	**0.0471**	**0.0539**	**0.0420**	**0.0724**	**0.1162**
Percent	100.0	160.1	227.7	43.7	63.1	72.1	56.3	97.0	155.6
Demand	100.0	144.9	198.7	43.7	58.5	68.5	56.3	86.4	130.3
Supply	100.0	102.6	119.8	43.7	45.0	43.8	56.3	57.6	76.3
Rent Sharing	100.0	93.2	94.8	43.7	43.9	43.6	56.3	49.2	51.1
Rural	**0.0456**	**0.0722**	**0.1174**	**0.0156**	**0.0279**	**0.0443**	**0.0301**	**0.0443**	**0.0730**
Percent	100.0	158.3	257.3	34.1	61.2	97.2	65.9	97.0	160.1
Demand	100.0	141.5	219.6	34.1	53.4	87.3	65.9	88.1	132.3
Supply	100.0	98.2	124.2	34.1	35.1	39.5	65.9	63.1	84.8
Rent Sharing	100.0	88.8	94.6	34.1	32.1	34.3	65.9	56.7	60.2

NOTES: See the notes to Table 2 for definitions of the different geographic units subsamples.

variance in 2009. For instance, the between-location variance in 2018 for the full sample is 224.5 percent of the initial 2009 variance (compare 0.0758 and 0.1701).The with-regions variance for the full sample in 2009 is 41 percent of the between-location variance (compare 0.0758 and 0.0312).

The subsequent rows “Demand”, “Supply” and “Rent” for each geographical aggregation show the percent of the variances based on each of the corresponding counterfactual model-generated house price series due to changes in demand, supply and rent-sharing factors realtive to the initial level of the between-location variance in 2009. For instance, for the full sample changes in demand factors between 2009 and 2018 alone can generate 196 percent of the initial between-location variance. The results point that changes in demand factors are the most important generator of price dispersion over time at any level of geographical aggregation, between regions, or within regions.

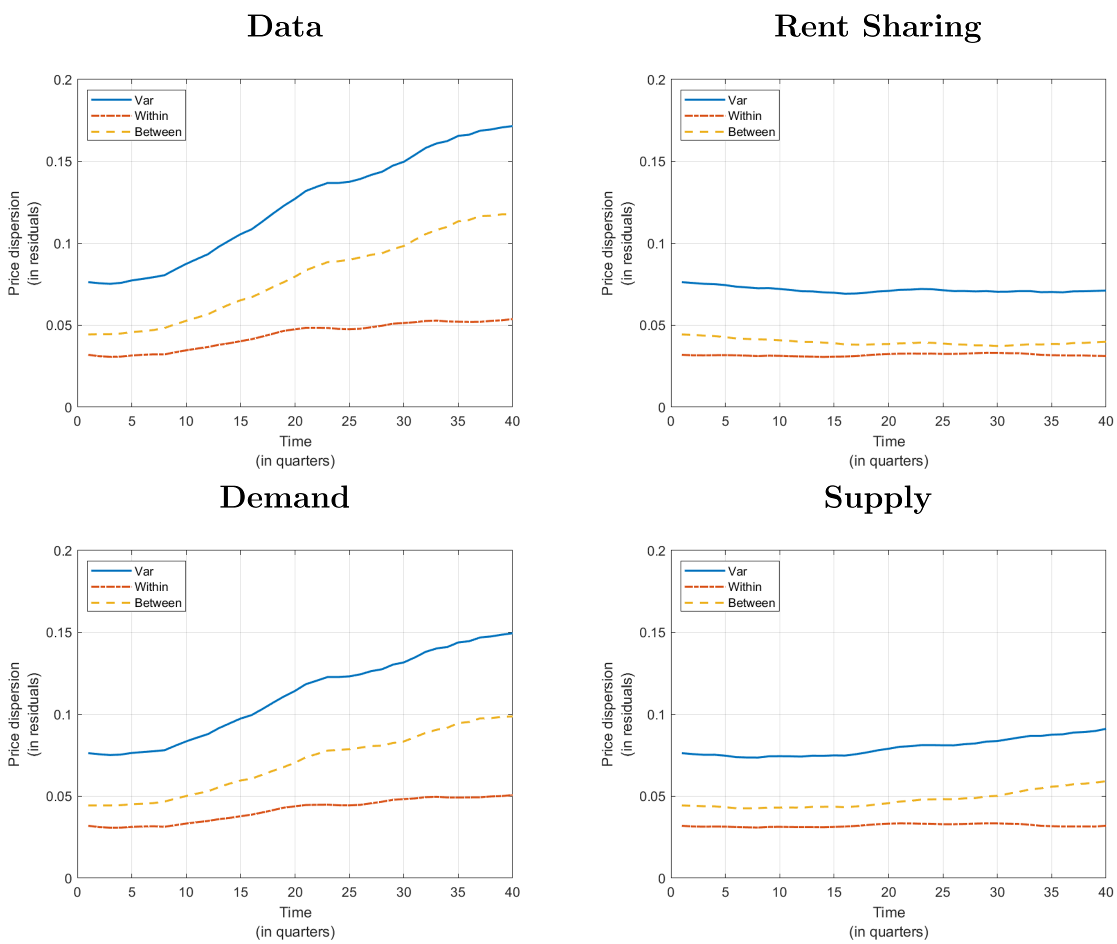


Figure 4: Variance decomposition of within- and between-region price changes

Notes: Model-based variance decomposition of equation (2) for all the regions in 2009-2018. Within (red line) depicts the within-region dispersion, whereas Between (yellow line) refers to dispersion coming from across labor market regions. The sum of within- and between-regions dispersion equals the total variance (blue line).

Figure 4 reports graphically the results for each quarter for the Top-7 regions of the model-based variance decomposition. The top-right plot shows the variance decomposition using the actual location- and time-specific prices. It reiterates the results from Table 3 and Table 7. Both within- and between-regions dispersion increase over time. However, within-regions dispersion contributes more to the overall variance increase than the between-regions dispersion.

The top-left plot of Figure 4 displays the time evolution of the overall variance as well its within- and between-regions components coming from changes in the rent-sharing factor, while the two bottom panels depict the same thing but in the cases in which only housing demand or housing supply changes are at work. The results clearly show that for the Top-7 regions only changes in housing demand contribute to the rise of price dispersion, while the other two factors have a negligible effect.

6 Conclusions

Using a large online dataset of real estate lisitings this study documents the trends and determinanants of spatial house price disparities across 2161 postcodes in 99 labor market regions during the recent German housing boom.

In the first part of our analysis, we compile a new dataset comprising -inflation and quality- adjusted real estate listings posted on Germany's largest online housing platform. Using this dataset, we uncover a pronounced heterogeneity in price movements over time across narrowly defined geographic areas such as postocodes but also across broader geographical units such as labor market regions. A simple variance decomposition reveals that the lion's share -approximately three-quarters- of the observed increase in house price heterogeneity across all postal codes in Germany is accounted for by an increase in price dispersion of houses between labor market regions. The results are robust to several sample restrictions, that is considering only rural or urban regions, or excluding East Germany. Focusing only on the more homogenous top 7 metropolitan areas in Germany, we find that between labor market dispersion accounts for half of the overall increase.

In the second part of this paper, we propose and estimate a simple frictional spatial housing search model. Through the lens of the model, we aim to disentangle the relative contributions of demand, supply, and rent-sharing factors to the observed increase in average price and dispersion both across and within labor market regions in Germany. The model features homogeneous buyers and sellers who trade-off prices with matching probabilities. Sellers rationally choose the number of listings and prices, whereas buyers select which sellers to contact based on preferences directed by a randomly drawn Gumbel distributed taste shock. The equilibrium prices and durations across time and space are then determined through the buyers' and sellers' valuations and a rent-sharing term reflecting housing market frictions. A key feature of our model approach is that we can uniquely pin down all model parameters using our novel housing listing data. We find that demand side factors contributed to over 70% of the increase in average house price and dispersion across locations in Germany in the 2009-2018 period. In contrast, supply-side factors played a less significant role, contributing to between 5 and 30% of the price increase, while changes in rent sharing had a minimal impact. Zooming into the top 7 labor market regions we find that house price dispersion increase in all regions but in Berlin and once again most of the increase is driven by differential changes in demand.

In our analysis here, we identify aggregate demand as the primary factor driving the increasing dispersion within Germany's housing markets. However, limitations inherent to our model and dataset restrict a deeper exploration of the fundamental drivers behind aggregate demand. Notably, the diverse impact of monetary policy on house prices, as evidenced in the US by Gorea et al. (2022), offers one potential explanation. Additionally, the influx of refugees into Germany during 2015 and 2016 has been associated with a notable decrease in nearby neighborhood listing prices in Berlin by 3 to 4% (Hennig, 2021), suggesting another contributing factor to demand dynamics. Moreover, a recent wave of papers highlight the trend of assortative matching - high-ability workers are increasingly sorting into highly productive firms-, predominantly in large urban areas (Dauth et al., 2022). This phenomenon has contributed to the geographical disparities in earnings in Germany, potentially influencing housing market behavior . In order to quantify the relative contribution of these factors, future research could leverage on alternative detailed micro-level datasets and develop richer structural quantitative models.

References

Aguirregabiria, Victor and Pedro Mira (2010), "Dynamic discrete choice structural models: A survey." *Journal of Econometrics*, 156, 38–67.

Albrecht, James, Pieter A Gautier, and Susan Vroman (2016), "Directed search in the housing market." *Review of Economic Dynamics*, 19, 218–231.

Amaral, Francisco, Martin Dohmen, Sebastian Kohl, and Moritz Schularick (2023), "Interest rates and the spatial polarization of housing markets." CEPR Discussion Paper No. 17780.

Ben-Shahar, Danny and Roni Golan (2022), "Price dispersion and time-on-market in the housing market." *Journal of Housing Economics*, 58, 101875.

Caliendo, Lorenzo, Maximiliano Dvorkin, and Fernando Parro (2019), "Trade and labor market dynamics: General equilibrium analysis of the China trade shock." *Econometrica*, 87, 741–835.

Dauth, Wolfgang, Sebastian Findeisen, Enrico Moretti, and Jens Suedekum (2022), "Matching in cities." *Journal of the European Economic Association*, 20, 1478–1521.

Fogli, Alessandra, Veronica Guerrieri, Mark Ponder, and Marta Prato (2023), "The end of the American dream? Inequality and segregation in US cities." Mimeo.

Garriga, Carlos and Aaron Hedlund (2020), "Mortgage debt, consumption, and illiquid housing markets in the Great Recession." *American Economic Review*, 110, 1603–1634.

Georgi, Sabine and Peter Barkow (2010), "Wohnimmobilien-Indizes: Vergleich Deutschland – Grossbritannien." ZIA Projektbericht, Zentraler Immobilienausschuss, Berlin.

Gorea, Denis, Oleksiy Kryvtsov, and Marianna Kudlyak (2022), "House price responses to monetary policy surprises: Evidence from the US listings data."

Guren, Adam M (2018), "House price momentum and strategic complementarity." *Journal of Political Economy*, 126, 1172–1218.

Gyourko, Joseph, Christopher Mayer, and Todd Sinai (2013), "Superstar cities." *American Economic Journal: Economic Policy*, 5, 167–199.

Hedlund, Aaron (2016), "Illiquidity and its discontents: Trading delays and foreclosures in the housing market." *Journal of Monetary Economics*, 83, 1–13.

Hennig, Jakob (2021), "Neighborhood quality and opposition to immigration: Evidence from german refugee shelters." *Journal of Development Economics*, 150, 102604.

Herkenhoff, Kyle F, Lee E Ohanian, and Edward C Prescott (2018), "Tarnishing the golden and empire states: Land-use restrictions and the US economic slowdown." *Journal of Monetary Economics*, 93, 89–109.

Hsieh, Chang-Tai and Enrico Moretti (2019), "Housing constraints and spatial misallocation." *American Economic Journal: Macroeconomics*, 11, 1–39.

Kindermann, Fabian, Julia Le Blanc, Monika Piazzesi, and Martin Schneider (2021), "Learning about housing cost: Survey evidence from the german house price boom." NBER Working Paper No. 28895.

Kosfeld, Reinhold and Alexander Werner (2012), "Deutsche Arbeitsmarktregionen – Neuabgrenzung nach den Kreisgebietsreformen 2007–2011." *Raumforschung und Raumordnung*, 70, 49–64.

Kotova, Nadia and Anthony L Zhang (2021), "Liquidity in residential real estate markets." Mimeo.

Moen, Espen R (1997), "Competitive search equilibrium." *Journal of Political Economy*, 105, 385–411.

Moen, Espen R, Plamen T Nenov, and Florian Sniekers (2021), "Buying first or selling first in housing markets." *Journal of the European Economic Association*, 19, 38–81.

Ngai, L Rachel and Silvana Tenreyro (2014), "Hot and cold seasons in the housing market." *American Economic Review*, 104, 3991–4026.

Rekkas, Marie, Randall Wright, and Yu Zhu (2022), "How well does search theory explain housing prices?" Mimeo.

Roback, Jennifer (1982), "Wages, rents, and the quality of life." *Journal of Political Economy*, 90, 1257–1278.

Rosen, Sherwin (1979), "Wage-based indexes of urban quality of life." In *Current issues in urban economics* (P Mieszkowski and M Straszheim, eds.), 74–104, Johns Hopkins Univ. Press.

Saiz, Albert (2010), "The geographic determinants of housing supply." *Quarterly Journal of Economics*, 125, 1253–1296.

Van Nieuwerburgh, Stijn and Pierre-Olivier Weill (2010), "Why has house price dispersion gone up?" *Review of Economic Studies*, 77, 1567–1606.

Wright, Randall, Philipp Kircher, Benoît Julien, and Veronica Guerrieri (2021), "Directed search and competitive search equilibrium: A guided tour." *Journal of Economic Literature*, 59, 90–148.

Appendix

A Data

In this Appendix, we provide additional information and descriptive statistics of the dataset for analysis which we derive from the original online listings data. In Appendix A.1, we briefly describe the Immobilienscout24 website and the RWI-GEO-RED dataset. We explain how we construct the final dataset for analysis in Appendix A.2. We provide information about the dataset's geographic breadth in Appendix B and about the hedonic regressions in Appendix C.

A.1 ImmobilienScout24

Background: Immobilienscout24 stands as the largest online real estate listing platform in Germany, catering to real estate providers, owners, tenants, and buyers. Operating in three countries - Germany, Austria, and Spain - the platform and its mobile app collectively attract approximately 20 million visitors per month. As of the close of 2019, Immobilienscout24 boasted around 450 million active listings, underscoring its prominent position in the real estate market.

The online portal can be accessed at https://www.immobilienscout24.de. Upon entering the German-language website, users are presented with the interface illustrated in Figure 5. The platform prompts users to select their country, specify the location for their search (city, address, or postal code), indicate the transaction type (purchase (*Kaufen*) or rental (*Mieten*)), and, lastly, define the property type -house (*Haus*), flat (*Wohnung*) or other types.

Additionally, the platform offers a range of filtering options, allowing users to refine their search by specifying property characteristics beyond geographical constraints. Users have the flexibility to set price ranges by providing a lower bound, an upper bound, or both. Furthermore, there is an option to specify the desired number of rooms.

Dataset: Our analysis relies on version 5.1 of the RWI-GEO-RED dataset, curated by the Research Data Centre (*Forschungsdatenzentrum* or FDZ Ruhr) at the Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI Essen), covering the period from January 2007 to July 2021. This comprehensive dataset comprises asked prices and rents for residential properties across Germany, categorized into four classes: house sales, flat sales, house rents and flat rents. These values are sourced from listings on the ImmobilienScout24 website.[20]

[20] ImmobilienScout24 claims a market share of approximately 50% of all advertised real estate objects in Germany (Georgi and Barkow (2010)).

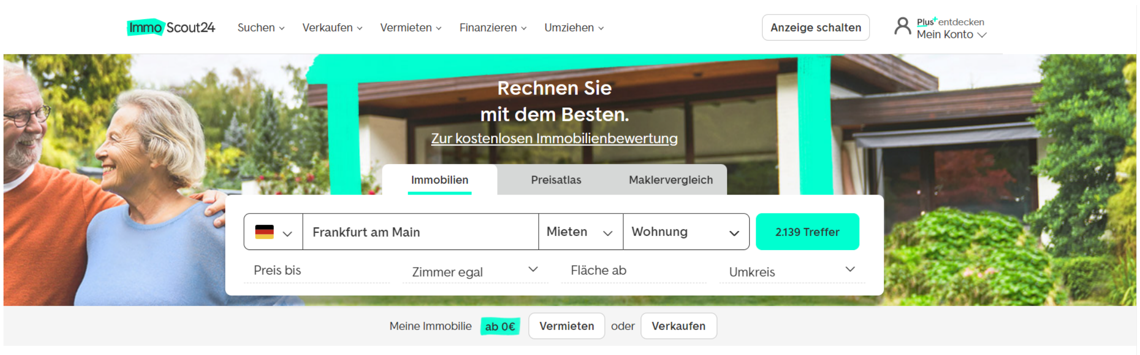


Figure 5: The Immobilienscout24 portal

In addition to asked prices and rents, the dataset incorporates user-contributed information that influences the valuation and location of each listing. Users actively furnish details about the real estate listings through a guided online questionnaire, subsequently transforming their input into an advertisement on the ImmobilienScout24 website. While essential information such as location, price (rent), and space of the listed property is mandatory, the remaining questionnaire fields are optional. There are a total of 76 distinct entries available for users to provide information, categorized into eight groups by RWI Essen.[21] Table 8 illustrates all the variable entries, along with their classification by RWI Essen. For an in-depth understanding of optional variables, refer to the RWI Essen technical manual.

Table 8: ImmobilienScount24 list of variables

Data report: Immoscout | RWI

1.3 List of variables

The dataset contains the following variables where some variables are only relevant for certain types of real estate. Note that we print the availability of a given variable for the most recent data delivery. Please refer to section 4 to see changes over deliveries.

Table 1
List of variables

Category	Variable name	Description	House rent	House sale	Flat rent	Flat sale
Identifier	obid	Object identifier	1	1	1	1
	uniqueID_gen	Unique object identifier (generated)	1	1	1	1
Time period	ajahr	Beginning of ad, year	1	1	1	1
	amonat	Beginning of ad, month	1	1	1	1
	ejahr	Ending of ad, year	1	1	1	1
	emonat	Ending of ad, month	1	1	1	1
Object features	aufzug	Elevator in object	1	0	1	1
	ausstattung	Facilities of object	1	1	1	1
	badezimmer	Number of bathrooms	1	1	1	1
	balkon	Balcony at object	0	0	1	1
	denkmalobjekt	Protected historic building	0	1	0	1
	einbaukueche	Kitchenette in object	1	0	1	1
	etage	Floor on which object is located	0	0	1	1
	ferienhaus	Usable as holiday home	0	1	0	1
	freiab	Available from	1	1	1	1
	gaestewc	Guest toilet in object	1	1	1	1
	garten	(Shared) garden available	0	0	1	1
	haustier_erlaubt	Pets allowed	1	0	1	0
	kategorie_Haus	House type	1	1	0	0
	kategorie_Wohnung	Flat type	0	0	1	1
	keller	Cellar in object	1	1	1	1
	parkplatz	Garage/parking space available	1	1	1	1
	zimmeranzahl	Number of rooms	1	1	1	1
	anzahletagen	Number of floors	1	1	1	1
	bauphase	Construction phase	0	1	0	0
	betreut	Assisted living for the elderly	0	0	0	0
	einliegerwohnung	Granny flat in object	0	1	0	0
	foerderung	Public housing – certificate of eligibility is needed	0	0	1	0
	immobilientyp	Type of real estate	1	1	1	1
	kaufvermietet	Rented when sold	0	1	0	1
	mieteinnahmenpromonat	Rental income per month in EUR	0	1	0	1
	nebenraeume	Number of ancillary rooms	0	0	0	0
	rollstuhlgerecht	Accessible, no steps	1	1	1	1
	schlafzimmer	Number of bedrooms	1	1	1	1
	wohngeld	Common charge for community association in EUR/month	0	0	0	1
	grundstuecksflaeche	Plot area	1	1	0	0
	nutzflaeche	Usable floor space	1	1	1	1
	wohnflaeche	Living area	1	1	1	1
Energy and structure information	baujahr	Year that object was built	1	1	1	1
	energieausweistyp	Type of Energy Performance Certificates (EPCs)	1	1	1	1
	energieeffizienzklasse	Energy Efficiency Rating	1	1	1	1
	ev_kennwert	Energy consumption per year and square meter	1	1	1	1
	ev_wwenthalten	Warm water consumption included in energy consumption	1	1	1	1
	heizkosten	Heating costs	1	0	1	0
	heizungsart	Type of heating	1	1	1	1
	letzte_modernisierung	Year of last modernisation of object	1	1	1	1
	objektzustand	Condition of object	1	1	1	1
Price information	courtage	Brokerage at contract conclusion	1	1	1	1
	heizkosten_in_wm_enthalten	Heating costs covered by inclusive rent	1	0	1	0
	kaufpreis	Purchasing price in EUR	0	1	0	1
	mietekalt	Exclusive rent in EUR	1	0	1	0
	mietekaution	Security deposit	1	1	1	1
	mietewarm	Inclusive rent in EUR	1	0	1	0
	nebenkosten	Utilities in EUR	1	0	1	0
	parkplatzpreis	Price of parking space in EUR	1	1	1	1
Regional information	blid	German state	1	1	1	1

7 | 8

[21] Most entries are only partially filled.

Geo-referenced locations: ImmobilienScout24 does not provide the address of the offered real estate. Instead, they geo-code addresses when available according to their own Mercator.[22] In turn, the RWI Essen converts the projected locations into the European standard ETRS89-LAEA based on Inspire(2014). This is a grid of 1-km^2 raster cells covering the whole territory of Germany. Subsequently, the grids are then assigned to broader administrative regions, in particular postal codes (*Postleitzahlen* or PLZ), municipalities (*Gemeinden*), districts (*Kreise*) or local labor market regions. This is done based on the 2015 geographical shapefiles provided by the Federal Agency for Cartography and Geodesy (BKG).

To compare the geographic house price/rent dispersion across time, we pool the housing units together in terms of PLZ codes. We choose PLZ codes rather than 1-km^2 cells because the former are sufficiently large to contain enough housing units but also small enough to exhibit the spatial heterogeneity within city boundaries. Nonetheless, since some remote PLZ codes do not contain sufficient housing units, we drop all the PLZ codes that do not contain at least 10 new listings within a quarter. The highest level of geographical aggregation we use is the labor market regions categorized by Kosfeld and Werner (2012). Labor market regions combine one or more districts and are characterised according to the commuter links to local labor centres. Some rural labor market regions might not be well-represented in the dataset. Therefore, we drop all labor market regions with less than 14 PLZ codes.

A.2 Construction of the Dataset for the Hedonic Regressions

Basic cleaning: We allow for a 2-year burn-in period at the beginning and end of the sample. This allows us to properly identify new listings and to also exclude the possibility of active listings at the end of the sample. To this end, we include in our dataset all listings that appear on the Immobilienscout24 platform between January 1, 2009 and December 31, 2018. Then we erase multiple entries that correspond to the same property within a short window.[23] In particular, we only keep the last price and we drop all previous listings for the

[22] In the initial years covered by the dataset, it was not mandatory for users to provide the address of the real estate. They could show only urban districts or municipalities for public use. Only for the most recent years, it is obligatory to provide the property address in the offer.

[23] According to the RWI-GEO-RED data manual duplicate entries occur for several reasons: *"First, since we obtain spells that have not been concluded at the time of data delivery, these will also occur in the next delivery which continues from the time of the previous delivery. Moreover, users can make small changes to the advertisement in order to attract more people. In the data, we only observe the status of the advertisement at the time of data delivery. Hence, the same advertisement might appear twice but with slightly different features in the data when a change was made after the delivery date. Fourth, users can temporarily set an object as inactive. This may be reasonable when a prospective buyer has committed to buying an object, but the deal has not yet been finalized. While inactive, objects will not be included in queries of potential buyers and will thus not be included in the dataset. However, if the potential buyer withdraws their offer to buy, the user might decide to activate the advertisement again. Lastly, users might decide to use an old advertisement as a template for a new ad, e.g. when renting two similar flats in the same house with only a short period in between."*

same item if it was posted more than once within a six-month period. We treat spells with starting dates at least six months from each other as different postings. [24] Second, we drop properties with missing mandatory information such as the geo-coded location, number of rooms, size or the age of the property. We also drop properties classified as "castles" or properties built before 1900. Finally, we remove all postings listed for less than a day.

Censoring: We exclude all postings with unreasonable price/rent entries. These entries include ultra-luxurious properties that form a market of their own and are likely to contaminate our analysis. We drop all units with a sale price of more than €6,000,000 or a rental price that exceeds €6000 per month. On the other hand, under-market value properties might be indicative of fraudulent listings or an attempt of the sellers to manipulate in their favor the Immobilienscout24 listing algorithm. This can happen only in the case the potential buyers list the property by price/rent in ascending order.[25] We remove all listings with a sale price of less than €10000 and a rental price of less than €130.

Moreover, we censor the price of a property per m^2. House and flats for sales are censored between €150 and €20000 per m^2 and rental units between €2.5 and €25 per m^2. The living area is restricted between 25 and 400 m^2 for flats and between 45 and 800 m^2 for houses. On top, we omit flats with more than 8 rooms and houses with more than 15 rooms. Finally, we drop all properties where the number of contacts or the number of clicks is beyond the 99-th percentile. Lastly, we drop listings with a duration longer than the 99-th percentile separately for sale and rental houses and flats.

Finally, we restrict the dataset to PLZ codes that contain at least 10 postings within a quarter and labor market regions that contain at least 14 postal codes. We run this procedure separately for the rental and sales market.

Inflation adjustments: The house prices and rents in our dataset are in nominal terms. We compute the inflation-adjusted prices and rents by deflating the nominal values with the respective county-specific consumer price index at the monthly level obtained from the Federal Statistical Office.

Geo-location adjustments: The vast majority of geo-code coordinates and their respective administrative match are consistent but some challenges remain. First, some administrative districts have been merged or changed over time. To address this problem, we obtain from https://www.geodaten-deutschland.de an updated file that contains up-to-date geo-referenced administrative information.[26]

Several districts (Kreise) have changed names or were merged into a different district in

[24] RWI Essen has developed an automatised procedure to identify multiple entries at the same time.
[25] An example for this are properties listed with very low rent but then much higher than normal utilities.
[26] We use the 2019 version

2011. Table 9 shows the mapping from these changed 2011 districts to their 2015 versions.

Table 9: Changes of districts, 2011-2015

2011 District	2011 District Number	2015 District	2015 District Number
SK Aachen and LK Aachen	5313, 5354	Städteregion Aachen	5334
Nordvorpommern	13107	Vorpommern-Rügen	13073
Südvorpommern	13108	Vorpommern-Greifswald	13075
Bremerhaven	4021	Bremerhaven, Stadt	4012
Rostock	13101	Rostock	13003
Mittleres Mecklenburg	13104	Landkreis Rostock	13072
Mecklenburgische Seenplatte	13103	Mecklenburgische Seenplatte	13071
Nordwestmecklenburg	13106	Nordwestmecklenburg	13074
Schwerin	13102	Schwerin	13004
Südwestmecklenburg	13105	Ludwigslust-Parchim	13076

Finally, we drop listings without information regarding the PLZ code (0.2% of all listings). For the remaining listings, we matched the PLZ code and the municipality of the RWI Essen dataset with the https://www.geodaten-deutschland.de updated dataset. [27] We notice that around 98% of the listings do match perfectly in both dimensions. All the unmatched entries are dropped.

[27] One might expect that the PLZ code areas are coherent and disjoint. However, this is not the case. There are PLZ code areas where one area lies entirely inside another area (e.g. 53879 in Euskirchen is enclosed by 53881). There are even cases where an area contains more than one other area.

B Breadth of the Immobilienscout24 Dataset

The following section describes the coverage of the Immobilienscout24 dataset across time and space. First, we show that we have sufficient new listings every year across the different property classes. Second, we show that we have sufficient number of listings across the labor market regions.

B.1 Listings over Time

Table 10 shows the listings of our final dataset across the years for each of the four property classes. Overall, there are more than 400,000 new listings for rent and sale each year. Flat rents are the largest category, reflecting the fact that more than 50% of German households are tenants. There is a declining trend in the number of listings over time, especially for units for sale. We argue that this is attributed mostly to the high demand for house prices during this period as documented throughout our paper.

B.2 Listings across Space

Next, we present the number of observations across space for the 10-year horizon in our dataset. Figure 6 presents the number of listings across districts in Germany.

The left-hand side map shows the number of listings for properties offered for sale and the right-hand side map shows them for the rental market. Urban districts surrounding the big cities of Berlin, Frankfurt, Stuttgart, Hamburg, Cologne, Dusseldorf, and Munich are colored dark blue whereas the less populous rural areas in the center of Germany are yellow to white. The districts of Wittenberg, Würzburg, and Holzminden contain the lowest number of listings over the 10-year horizon, 878, 1189, and 1240, respectively.

The right-hand side map shows the listings for the rental market. As expected many sparsely populated regions do not contain sufficient rental units and most listings are concentrated around the major urban hubs. Nonetheless, we still have in disposal 193 districts in 58 labor market regions with a sufficient number of listings (>1000) over the 10-year horizon.

Table 10: Immobilienscout24 Listings-Balanced Sample

	House Sales	Flat Sales	House Rents	Flat Rents
2009	337,837	291,180	27,830	457,210
2010	324,249	281,170	27,651	487,036
2011	306,922	285,790	26,081	469,685
2012	298,577	306,469	28,720	456,756
2013	290,145	328,521	30,958	481,183
2014	286,395	361,004	31,350	626,024
2015	271,651	294,087	22,256	531,284
2016	211,567	224,835	16,397	421,648
2017	207,644	202,533	15,717	370,768
2018	189,633	187,725	15,087	356,733
N	2,724,620	2,763,314	242,047	4,658,327

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

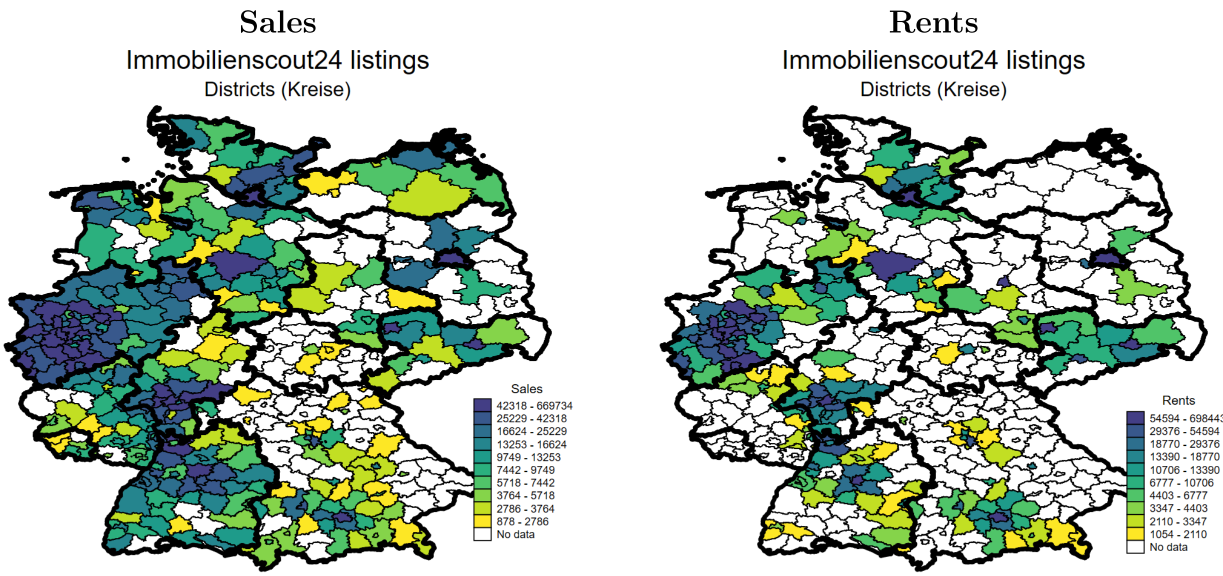


Figure 6: Immobilienscout24 listings across geographical units

C Hedonic Regressions

This section describes the construction of the hedonic regressions. We start with the variable selection based on our listings dataset and then display the regression results.

We would like to control for observed variable characteristics that determine the quality of the listed property in the hedonic regressions. Although we have at our disposal a set of 76 variables, most of them contain missing entries or are sparsely filled. We would like to include as many observable characteristics as possible without losing too many observations. As already mentioned in Appendix A, the sellers are guided by an online questionnaire when they file their ad. We use the information provided in this questionnaire to construct the variables for the hedonic regressions.

First, we divide the nominal listed price or rent by the monthly CPI and then by the property area to create inflation-adjusted property price/rent per m^2, which are the dependent variables. Second, we use the following explanatory variables:

- **Number of rooms:** We keep all listed houses with less than 15 rooms and flats with less than 8 rooms. In Germany, the number of rooms excludes kitchens, baths, or corridors. In several cases, the number of rooms is not a natural number, which is not necessarily due to a faulty entry. In Germany, there is the concept of half rooms. Following the DIN 283 norm, a half room is defined as a room with a size between 6 and 10 m^2. While this definition is outdated, it is still frequently in use. In this case, we rounded up to the nearest natural integer. Then we created 14 separate dummies (excluding properties with 1 room).

- **Age of the property:** We deduct the year the listing was posted from the year it was built. Then we create 5-year age dummies.[28]

- **Type of property:** We control for 22 detailed types of property: Not specified house, Single-family house (detached), Two-family house, Semi-detached house, Terraced house, Terraced house (middle unit), Terraced house (end unit), Bungalow, Farm-house, Mansion, Block of flats, Other property for living, Special property, Attic flat, Flat, Raised ground floor flat, Maisonette, Penthouse, Souterrain, Flat with terrace, Other Flat, and Not specified flat.

- **Cellar:** A dummy variable that takes the value of 1 indicating that the seller has answered yes in the questionnaire, otherwise is 0.

- **Guest toilet:** A dummy variable that takes the value of 1 indicating that the seller has answered yes in the questionnaire, otherwise is 0.

- **Quarterly dummies:** A set of dummies indicating the quarter the ad was listed in the Immobilienscout24 portal.

[28] On several occasions the seller lists the price before the property is constructed. We include these entries in the first age category.

D Additional Descriptive Statistics

Sales Market

Table 11: Descriptive statistics for dwelling sales (top-7 regions)

Variable	2009-10	2011-12	2013-14	2015-16	2017-18
Log price $\ln \bar{p}$	7.53	7.59	7.70	7.85	7.99
Price residual $\bar{\epsilon}$	0.08	0.12	0.22	0.34	0.52
Listings S	101	96	102	72	50
Duration d	48	43	38	40	40
Contacts C	281	335	429	437	404
Flow tightness $\frac{C}{dS}$	0.08	0.11	0.16	0.24	0.27
Observations	5,160	5,160	5,160	5,160	5,160

NOTES: Means of selected variables for the baseline sample of location-quarter observations. Prices are in euros and adjusted for inflation using the CPI of the federal states in Germany.

Rental Market

Table 12: Descriptive statistics for dwelling rents

	2009-10	2011-12	2013-14	2015-16	2017-18
Log rent $\ln \bar{r}$	1.89	1.91	1.94	1.99	2.07
Rent residual $\bar{\epsilon}$	-0.04	-0.03	-0.01	0.03	0.10
Listings S	74	73	87	73	56
Duration d	32	27	25	23	22
Contacts C	484	628	945	1,254	1,490
Flow tightness $\frac{C}{dS}$	0.30	0.47	0.66	1.30	2.02
Observations	13,520	13,520	13,520	13,520	13,520

NOTES: Means of selected variables for the baseline sample of location-quarter observations. Rents are in euros and adjusted for inflation using the CPI of the federal states in Germany.

Table 13: Descriptive statistics for dwelling rents (top-7 regions)

	2009-10	2011-12	2013-14	2015-16	2017-18
Log rent ln $\bar{r}$	2.05	2.08	2.13	2.18	2.28
Rent residual $\bar{\epsilon}$	0.09	0.12	0.15	0.20	0.28
Listings S	92	86	96	71	51
Duration d	28	24	22	20	18.73
Contacts C	744	969	1,389	1,718	1,898
Flow tightness $\frac{C}{dS}$	0.43	0.68	0.95	1.83	2.75
Observations	5,888	5,888	5,888	5,888	5,888

NOTES: Means of selected variables for the baseline sample of location-quarter observations. Rents are in euros and adjusted for inflation using the CPI of the federal states in Germany.

E Variance Decomposition: Rents

Table 14: Within and between location variance decomposition (rents)

	Total variance			Within locations			Between locations		
	2009	2013	2018	2009	2013	2018	2009	2013	2018
Full Sample	0.088	0.093	0.106	0.030	0.034	0.035	0.058	0.059	0.071
West-Germany	0.085	0.090	0.104	0.031	0.034	0.037	0.054	0.056	0.068
East-Germany	0.033	0.042	0.043	0.025	0.027	0.025	0.008	0.016	0.018
Top 7 regions	0.093	0.084	0.098	0.032	0.037	0.040	0.061	0.047	0.058
Urban	0.091	0.098	0.107	0.030	0.034	0.034	0.062	0.064	0.072
Rural	0.063	0.066	0.098	0.032	0.032	0.038	0.031	0.034	0.060

NOTES: “Full sample” builds on the listings in all quarter-location observations in our baseline sample. “West Germany” and “East Germany” include all listings located in Kreis (NUTS-3) that belonged to the FRG or GDR respectively before German reunification. The “Top-7 regions” comprise the labor market regions of Berlin, Munich, Hamburg, Frankfurt am Main, Cologne, Stuttgart and Dusseldorf. “Urban” denotes all units belonging to a Kreis indicated either as “Kreis”, “Kreisfreie Stadt” or “Stadtkreis” and “Rural” all housing units located in a “Landkreis”.

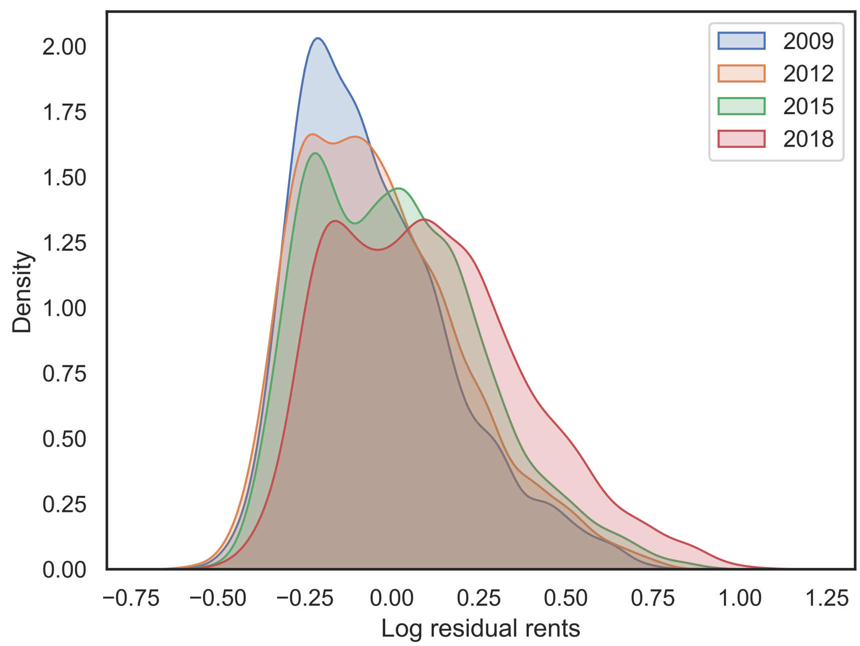


Figure 7: Distribution of residual rents across locations

NOTES: Between-location distributions of residual log rents in the years 2009 (blue), 2012(orange), 2015 (green) and 2018 (red). The residuals are obtained from hedonic regressions of posted rents per m^2 as described in the main text and averaged in each location (postal code).

Table 15: Within and between region variance decomposition (rents)

	Between locations			Within regions			Between regions		
	2009	2013	2018	2009	2013	2018	2009	2013	2018
Full Sample	0.058	0.059	0.071	0.018	0.021	0.021	0.040	0.039	0.050
West-Germany	0.054	0.056	0.068	0.019	0.022	0.024	0.035	0.034	0.044
East-Germany	0.008	0.016	0.018	0.006	0.010	0.006	0.003	0.006	0.012
Top 7 regions	0.061	0.047	0.058	0.025	0.031	0.037	0.036	0.017	0.022
Urban	0.062	0.064	0.072	0.016	0.019	0.019	0.046	0.045	0.054
Rural	0.031	0.034	0.060	0.009	0.010	0.011	0.022	0.024	0.049

NOTES: "Full sample" builds on the listings in all quarter-location observations in our baseline sample. "West Germany" and "East Germany" include all listings located in Kreis (NUTS-3) that belonged to the FRG or GDR respectively before German reunification. The "Top-7 regions" comprise the labor market regions of Berlin, Munich, Hamburg, Frankfurt am Main, Cologne, Stuttgart and Dusseldorf. "Urban" denotes all units belonging to a Kreis indicated either as "Kreis", "Kreisfreie Stadt" or "Stadtkreis" and "Rural" all housing units located in a "Landkreis".

F Variance Decomposition Derivations

Proof of Decomposition (1)

Write H for the set of listings and H_i for the set of listings in location i. Write n for the cardinality of H and n_i for the cardinality of H_i

$$
\begin{aligned}
\operatorname{var} \varepsilon_h &= \frac{1}{n} \sum_{h \in H} (\varepsilon_h - \bar{\varepsilon})^2 \\
&= \frac{1}{n} \sum_{i \in L} \sum_{h \in H_i} \left[(\varepsilon_h - \bar{\varepsilon}_i)^2 + 2(\varepsilon_h - \bar{\varepsilon}_i)(\bar{\varepsilon}_i - \bar{\varepsilon}) + (\bar{\varepsilon}_i - \bar{\varepsilon})^2 \right] \\
&= \sum_{i \in L} \frac{n_i}{n} \underbrace{\frac{1}{n_i} \sum_{h \in H_i} (\varepsilon_h - \bar{\varepsilon}_i)^2}_{=\operatorname{var}_i(\varepsilon_h)} + \frac{2}{n} \sum_{i \in L} (\bar{\varepsilon}_i - \bar{\varepsilon}) \underbrace{\sum_{h \in H_i} (\varepsilon_h - \bar{\varepsilon}_i)}_{=0} + \sum_{i \in L} \frac{n_i}{n} (\bar{\varepsilon}_i - \bar{\varepsilon})^2 \\
&= \sum_{i \in L} s_i \operatorname{var}_i(\varepsilon_h) + \sum_{i \in L} s_i (\bar{\varepsilon}_i - \bar{\varepsilon})^2 \ .
\end{aligned}
$$

Proof of Decomposition (2)

$$
\begin{aligned}
\sum_{i \in L} s_i (\bar{\varepsilon}_i - \bar{\varepsilon})^2 &= \sum_{r \in R} \sum_{i \in r} s_i \left[(\bar{\varepsilon}_i - \bar{\varepsilon}_r)^2 + 2(\bar{\varepsilon}_i - \bar{\varepsilon}_r)(\bar{\varepsilon}_r - \bar{\varepsilon}) + (\bar{\varepsilon}_r - \bar{\varepsilon})^2 \right] \\
&= \sum_{r \in R} \sigma_r \underbrace{\sum_{i \in r} \frac{s_i}{\sigma_r} (\bar{\varepsilon}_i - \bar{\varepsilon}_r)^2}_{=\operatorname{var}_r(\bar{\varepsilon}_i)} + 2 \sum_{r \in R} (\bar{\varepsilon}_r - \bar{\varepsilon}) \underbrace{\sum_{i \in r} s_i (\bar{\varepsilon}_i - \bar{\varepsilon}_r)}_{=0} + \sum_{r \in R} \sigma_r (\bar{\varepsilon}_r - \bar{\varepsilon})^2 \\
&= \sum_{r \in R} \sigma_r \operatorname{var}_r(\bar{\varepsilon}_i) + \sum_{r \in R} \sigma_r (\bar{\varepsilon}_r - \bar{\varepsilon})^2 \ .
\end{aligned}
$$

G Matching Functions

G.1 Matching Function Derivations

We are using the matching function

$$
q_{rt}(\theta) = 1 - \left(1 - q_{rt} k^{\mu} \theta^{2\mu}\right)^{90} \ ,
$$

with region-specific and time-varying q_{rt}, the estimated μ and a given value for k.

Then

$$
q'_{rt}(\theta) = \left(1 - q_{rt} k^{\mu} \theta^{2\mu}\right)^{89} 180 \mu q_{rt} k^{\mu} \theta^{2\mu - 1} \ ,
$$

and

$$
\hat{q}_{rt}(\theta) \equiv q_{rt}(\theta) - \theta q'_{rt}(\theta) = 1 - \left(1 - q_{rt} k^{\mu} \theta^{2\mu}\right)^{90} - \left(1 - q_{rt} k^{\mu} \theta^{2\mu}\right)^{89} 180 \mu q_{rt} k^{\mu} \theta^{2\mu} \ .
$$

G.2 Matching function estimates across cities

H Numerical Solution of the Model

We repeat the main equations of the model and the estimation of location premia as explained in Sections 3 and 4:

$$\bar{V}_t^B = \ln\left[\sum_j e^{\gamma V_{j,t}^B + \tau_j}\right] , \tag{3}$$

$$p_{it} = \zeta(\theta_{it})\beta V_{i,t+1}^S + (1-\zeta(\theta_{it}))[A_{it} - \beta\bar{V}_{i,t+1}^B] , \tag{4}$$

$$V_{it}^S = -c + \beta V_{i,t+1}^S + (q(\theta_{it}) - \theta_{it}q'(\theta_{it}))\left[A_{it} - \beta\bar{V}_{i,t+1}^B - \beta V_{i,t+1}^S\right] , \tag{5}$$

$$V_{it}^B = -r_t + \beta\bar{V}_{i,t+1}^B + q'(\theta_{it})\left[A_{it} - \beta\bar{V}_{i,t+1}^B - \beta V_{i,t+1}^S\right] , \tag{6}$$

$$V_{i,T+1}^S = \frac{2}{T(T-1)}\left\{\sum_{t=1}^T V_{it}^S\,[3t-(T+2)]\right\} , \tag{15}$$

$$V_{i,T+1}^B = \frac{2}{T(T-1)}\left\{\sum_{t=1}^T V_{it}^B\,[3t-(T+2)]\right\}, \tag{16}$$

$$\sum_i \tau_i = 0, \tag{13}$$

$$\tau_i = \frac{1}{T}\sum_{t=1}^T\left[\ln\hat{\pi}_{it} + \bar{V}_t^B - V_{it}^B\right] - \frac{\lambda}{2T} . \tag{14}$$

This is a high-dimensional system of $(3N+1)(T+1)+1$ equations that contain $(3N+1)(T+1)+1$ unknowns. Nonetheless, all equations except (3) are linear. So for a given guess of $\bar{V}_{t+1}^B$ for $t = 1, ..., T+1$, we back out all the remaining unknowns by elementary linear algebra.

Stepwise the procedure unfolds as follows:

1. Start with an arbitrary guess of $\bar{V}_{t+1}^B$ for $t = 1, ..., T+1$.
2. Solve equations (4)–(6) , (15)–(16) and (13)–(14) with matrix inversion.
3. Use the values of τ_i together with $V_{i,t}^B$ for all i and $t = 1, \ldots, T+1$ to obtain new values of $\bar{V}_{t+1}^B$ for $t = 1, ..., T+1$
4. Using the new values of $\bar{V}_{t+1}^B$, repeat steps 2-3 until the routine converges.

I Additional Figures

Table 16: Estimates of time fixed effects in equation (9)

t	Top-7	Berlin	Munich	Hamburg	Frankfurt	Stuttgart	Dusseldorf	Cologne
1	0	0	0	0	0	0	0	0
2	0.04	0.02	0.07	0.06	0.05	-0.06	0.06	0.06
3	-0.01	-0.05	0.06	0.04	-0.01	0	-0.03	-0.04
4	-0.1	-0.1	-0.01	-0.01	-0.12	-0.15	-0.16	-0.13
5	-0.02	-0.06	0.02	0.05	0	-0.1	0.02	0.02
6	0.03	-0.04	0.11	0.17	0.09	-0.06	0.03	0.01
7	0.07	-0.02	0.08	0.21	0.17	-0.02	0.13	0.07
8	0.07	-0.02	0.1	0.25	0.14	-0.07	0.14	0.08
9	0.03	-0.02	0.11	0.09	0.03	-0.02	0.11	0.04
10	0.01	-0.06	0.06	0.09	0.03	-0.08	0.11	0.04
11	-0.01	0	0.02	0.06	-0.05	-0.15	0.08	0.1
12	-0.01	-0.07	-0.04	0.08	0.06	-0.14	0.08	0.07
13	0.02	-0.04	0.08	0.06	0.13	-0.09	0.03	0.1
14	0.01	-0.06	0.08	0.07	0.07	-0.15	0.11	0.08
15	0.02	-0.09	0.12	0.11	0	-0.08	0.15	0.12
16	0.01	-0.09	0.12	0.15	-0.03	-0.1	0.1	0.15
17	-0.04	-0.15	0.07	0.13	0	-0.19	0.03	0
18	-0.03	-0.14	0.06	0.16	0.01	-0.13	0.03	-0.04
19	-0.05	-0.17	0.09	0.08	0.01	-0.16	0	-0.01
20	-0.05	-0.16	0.1	0.12	0.02	-0.09	-0.1	0
21	-0.04	-0.15	0.16	0.11	-0.03	-0.01	-0.04	-0.01
22	0.01	-0.05	0.19	0.1	0.01	0.01	0.05	-0.06
23	0.04	-0.03	0.23	0.16	0.11	0.03	0.03	-0.05
24	0.06	-0.03	0.23	0.15	0.21	0.07	0.05	-0.04
25	0.11	0.07	0.21	0.24	0.11	0.1	0.21	0.09
26	0.09	0.05	0.13	0.2	0.17	0.04	0.2	0.13
27	0.07	-0.02	0.18	0.24	0.14	0.1	0.14	-0.04
28	0.08	0.01	0.3	0.25	0.04	0.06	0.18	0
29	0.13	0.07	0.35	0.21	0.19	0.07	0.15	0.12
30	0.19	0.2	0.36	0.26	0.21	0.1	0.27	0.18
31	0.24	0.26	0.46	0.3	0.15	0.16	0.33	0.24
32	0.22	0.37	0.36	0.24	0.17	0.14	0.24	0.19
33	0.2	0.26	0.39	0.22	0.2	0.09	0.29	0.12
34	0.21	0.26	0.39	0.2	0.28	0.18	0.21	0.12
35	0.18	0.24	0.45	0.22	0.21	0.09	0.12	0.09
36	0.18	0.22	0.44	0.16	0.23	0.13	0.18	0.07
37	0.2	0.28	0.42	0.24	0.28	0.11	0.12	0.1
38	0.25	0.37	0.41	0.17	0.37	0.17	0.23	0.11
39	0.22	0.36	0.45	0.19	0.24	0.17	0.17	0.11
40	0.21	0.39	0.4	0.24	0.23	0.13	0.14	0.04

NOTES: This table shows the g_t (time-fixed effects) estimates of equation (9) separately for each city.

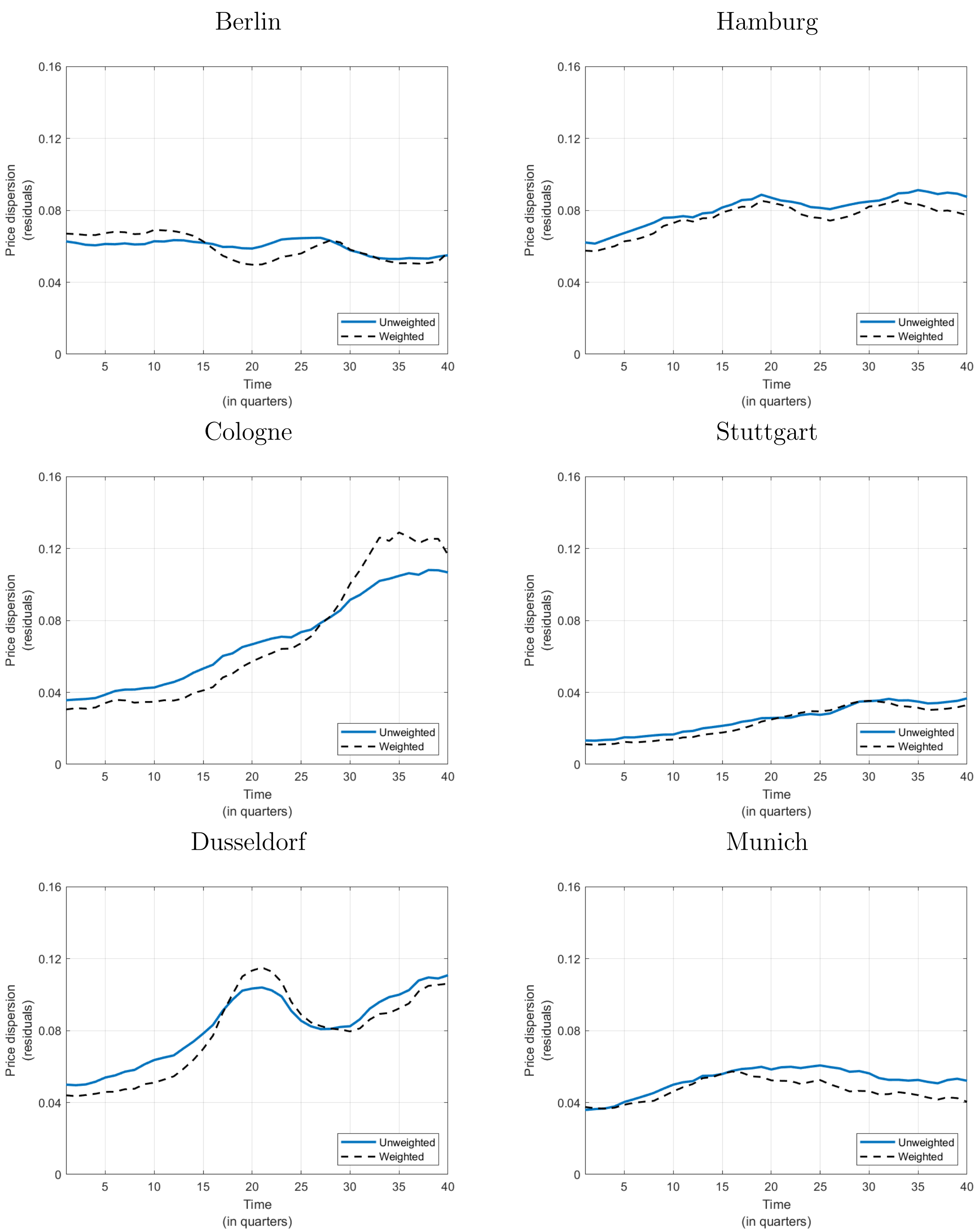


Figure 8: Price dispersion in selected regions

NOTES: The Figure shows the dispersion of the residualised log prices from 2009Q1 up to 2018Q4. The blue line shows the unweighted dispersion and the black dashed line the weighted dispersion based on the number of listings in each postal code.