

Housing Yields

Stefano Colonnello[†] Roberto Marfe[‡] Qizhou Xiong^{§*}

June 10, 2026

Abstract

We build a granular dataset of residential property yields using rental and sale listings from a major German real estate platform. With more than 1.5 million property-level rent-to-price ratios, we document a novel *heterogeneity puzzle*. About two-fifths of yield dispersion remains unexplained after controlling for extensive observable property characteristics and postal code-level time-varying factors using fixed effects. When we absorb finer neighborhood amenities through richer fixed effects, the unexplained share remains close to one-third. This evidence points to the importance of investors’ beliefs and preferences—rather than several alternative explanations we test—as sources of heterogeneity in risk premia.

JEL Classification: G12, G51, R31

Keywords: Housing, Rent-to-Price Ratio, Geographic Heterogeneity

*We thank the Editor, George Pennacchi, and two anonymous referees for constructive feedback and guidance. Moreover, we are indebted to Ufuk Akcigit, Pietro Dindo, Dmitry Kuvshinov (discussant), Francesco Lippi, Yiran Mao (discussant), Juan S. Morales, Tommaso Oliviero, Lorian Pelizzon, Alberto Plazzi, Geert Rouwenhorst, Marti G. Subrahmanyam, Martin Simmler, Viktor Slavtchev, Gregor von Schweinitz, Michael Ungeheuer (discussant), Mungo Wilson, and seminar participants at Collegio Carlo Alberto, the 2021 European Meeting of the Urban Economics Association, the 2021 Annual Meeting of the German Finance Association, the 2021 Paris December Finance Meeting, the workshop “The socioeconomics of housing and finance” at FU Berlin, and the “Finance workshop” at Ca’ Foscari University of Venice for insightful comments. We are grateful to staff at FDZ Ruhr of RWI Essen for granting us access to the RWI-GEO-RED database. The authors used ChatGPT-5.5 to assist with proofreading, language editing, and improving the clarity and flow of the manuscript. After utilizing this tool, the authors reviewed and edited the AI-assisted text and take full responsibility for the final content.

[†]Corresponding author. Ca’ Foscari University of Venice (Department of Economics) and Halle Institute For Economic Research (IWH), Cannaregio 873, Fondamenta San Giobbe, 30121 Venezia, Italy. Phone: +39 041 234 9198. E-mail: stefano.colonnello@unive.it.

[‡]University of Turin (ESOMAS Department) and Collegio Carlo Alberto, Italy. E-mail: roberto.marfe@carloalberto.org.

[§]University of Oxford (Said Business School), United Kingdom. E-mail: qizhou.xiong@said.oxford.edu.

I. Introduction

The main residence is the largest component of wealth for many households (e.g., [Flavin and Yamashita \(2002\)](#)). Also in Germany, the country we focus on, despite a relatively low homeownership rate of 44% as of 2017 (vs. 64% in the US), real estate assets dominate the portfolio of the average household by a wide margin ([Deutsche Bundesbank \(2019\)](#)).¹ Nonetheless, the traditional view based on the permanent income hypothesis places little emphasis on the consequences of house price fluctuations for aggregate consumption and, in turn, the business cycle. Once the assumption of complete markets—i.e., of households’ insurance against idiosyncratic shocks—is relaxed, changes in housing wealth become important ([Berger, Guerrieri, Lorenzoni, and Vavra \(2018\)](#)). In line with this conjecture, responses of consumption to local shocks to house prices are substantial in the US, with relevant consequences for the amplification of business cycles ([Mian, Rao, and Sufi \(2013\)](#), [Guren, McKay, Nakamura, and Steinsson \(2021\)](#)) and the effectiveness of monetary policy ([Beraja, Fuster, Hurst, and Vavra \(2019\)](#)).²

Understanding the drivers of house valuations is thus key to designing credible macroeconomic models. We focus on the rent-to-price ratios—the *housing yields*—, which are a metric widely used by investors and policy makers to gauge the conditions of the housing market because it incorporates market participants’ expectations about properties’ future discount and rent growth rates ([Plazzi, Torous, and Valkanov \(2010\)](#)).³

¹Home ownership in Germany is especially low in (Eastern) urban areas; yet, private landlords own around two-thirds of rental properties ([Savills \(2019\)](#)). [Kaas, Kocharkov, Preugschat, and Siassi \(2021\)](#) investigate the drivers of German low homeownership (which is coupled with a high rate of house ownership for investment purposes), pointing to a high property transfer tax rate, tax deductions of mortgage interest payments available to landlords but not to owner-occupiers, and the accessibility of social housing. [Kohl \(2016\)](#) and [Kohl and Sørvoll \(2021\)](#) provide an historical perspective on the roots of the homeownership gap between Germany and other countries like the US and the Nordic ones.

²[Guerrieri and Mendicino \(2018\)](#) show that the effect of housing wealth changes on consumption is more modest for European countries, but is persistent and stronger in the long-run than in the short-run.

³The housing yield is a slow-moving variable, constituting a key measure to characterize the state of local housing markets, over and above the dividend-to-price ratio for stocks. Indeed, as pointed out by [Plazzi et al. \(2010\)](#), both the property price and the rent are observed market prices. By contrast, dividends also reflect to a large extent short-term managerial decisions ([Vuolteenaho \(2002\)](#)). Note that throughout the remainder of the paper, we use the terms “housing yield”, “rent-to-price ratio”, “valuation ratio”, and

Our main contribution is to study the distribution and determinants of housing yields over a recent, highly granular, and comprehensive database on the market for residential properties of a large economy like Germany. We document a novel *heterogeneity puzzle*: a substantial degree of variation in housing yields can be explained neither by postal code-level time-varying factors nor by a rich set of property-level characteristics. Such heterogeneity is economically sizable, with the 90th–10th percentile range of unexplained yields in our dataset corresponding to a variation of EUR 71,837 in the value of the median flat, against a mean household net wealth estimated at EUR 202,541 according to [Deutsche Bundesbank \(2019\)](#). Whereas neighborhood amenities explain away a non-negligible part of this residual heterogeneity, a large fraction of it may stem from unobservable property-specific traits as well as from dispersion in investors’ beliefs and preferences. These forces, in turn, presumably generate heterogeneity in both risk premia and expected rent growth, with some evidence pointing to a more prominent role for the former. By contrast, proxies for local agglomeration effects, housing supply rigidities, informational frictions, regulation, climate risk, and ownership structure do not appear to command sizable variation in housing yields over and above postal code-level time-varying factors.

To examine rent-to-price ratios, we use sale and rental prices for flats listed on a major German online real estate platform between 2007 and 2017. One challenge is that we generally observe the market price of a property either as a sale price or as a rental price. To work around this problem, we build a synthetic measure of the property-level gross rent-to-price ratio, which relies on matching each rental property to a counterfactual property for sale based on a comprehensive set of observable property traits. The rent-to-price ratios so obtained vary greatly across regions and their dispersion is remarkable not only across states or districts, but even across postal codes within the same city. Time-series variation in rent-to-price ratios is largely property-specific or at the postal code level, with macroeconomic fluctuations accounting only for up to half of it.

“valuation” interchangeably.

At the same time, cross-sectional dispersion in housing yields is remarkably stable with respect to the level of regional aggregation considered (federal state-, district-, or postal code-level). Even after controlling for an extensive set of observable property-level characteristics and fine fixed effects at the postal code-calendar quarter level in a regression framework, unexplained heterogeneity in yields remains large at 41% based on the adjusted R^2 . Such a heterogeneity originates from a notably symmetric distribution of residuals, in which outliers appear to have a limited role.

Then, we verify that neither potential matching errors inherent to our synthetic yields nor the unobservability—in our setting—of costs borne by owners are likely to drive the 41% of variation that we cannot explain. We also implement an alternative procedure improving the matching precision with respect to the location of properties. At the cost of obtaining a significantly smaller sample, we are indeed able to include an even finer layer of fixed effects capturing neighborhood-level time-varying factors. As a result, we can explain an additional 12% of variation in yields, which speaks to the importance of local amenities. By contrast, we do not detect any quantitatively relevant role of proxies for agglomeration economies, informational frictions, rigidities in housing supply, regulation, climate risk, or ownership structure in explaining within-postal code-quarter dispersion in yields.

The 41% of variation in rent-to-price ratios that remains unexplained—or 29% when accounting for neighborhood amenities in an alternative matched sample—is noteworthy. This is even more striking given that observable property characteristics alone deliver an adjusted R^2 of only 30%, while the remaining explained variation is captured by fixed effects that subsume factors ultimately unobservable in our dataset.

Taking one step back, such heterogeneity is puzzling because residential properties offer a relatively homogeneous service to households, thus, once filtering out obvious differences in key dwelling traits (e.g., size and number of rooms, presence of a balcony, quality of facilities, etc.) and any time-varying postal code- or neighborhood-level traits (e.g., distance from schools, hospitals or restaurants, quality of local services, number

of nearby shops, etc.), one may expect that rent-to-price ratios exhibit little variation.⁴ Although a clear benchmark for a “satisfactory” adjusted R^2 in housing yield regressions is hard to define, we argue that the explanatory power of our richest specification is remarkably low, which calls into question common wisdom about the drivers of residential property valuations.

From an asset pricing perspective, indeed, a property’s rent-to-price ratio should respond to changes in expected returns and rent growth (e.g., [Plazzi et al. \(2010\)](#)). Although theory does not predict that yields should be equal across properties, our results pose a challenge about the drivers of house valuations. The substantial variation in yields at the very local level after accounting for a host of property-level traits may stem from heterogeneity in properties’ risk premia and/or expected rent growth. Dispersion in investors’ beliefs and preferences, in this respect, emerges as a possible crucial force.

Distinguishing between these channels—risk premia vs. expected rent growth—would ideally require observing the same properties over time, which is challenging because our listings data come as repeated cross-sections. We therefore construct a pseudo-panel by aggregating properties into cohorts defined by location, number of rooms, and size category, which allows us to compute quarterly rent-to-price ratios, excess returns, and rent growth rates. Consistent with the present-value logic of housing valuation ([Plazzi et al. \(2010\)](#)), we show that rent-to-price ratios predict both future excess returns and rent growth, corroborating the economic content of our synthetic yield measure. Moreover, predictability is stronger for returns in cohorts with high cross-sectional dispersion in yields, while rent-growth predictability is relatively more pronounced in low-dispersion cohorts. This pattern suggests that local yield dispersion is more closely related to heterogeneity in risk premia than to differences in expected rent growth.

Finally, it is worth discussing the generalizability of our findings. The German housing

⁴Moreover, we document that the total variation of housing yields is comparable to, if not smaller than, that of yields for a widely researched asset class like US equities, pointing to the importance of the granular structure of our data and the documented unexplained heterogeneity.

market has a number of peculiarities that set it apart from those of other advanced economies, such as the US and the UK. In particular, homeownership is particularly low and unequally distributed; private contracts dominate the rental market over our sample period, against the backdrop of the declining importance of—historically widespread—social housing (e.g., [Dustmann, Fitzenberger, and Zimmermann \(2022\)](#), [Voigtländer \(2009\)](#)); property taxation is comparatively low (e.g., [Bach and Eichfelder \(2021\)](#)). Moreover, real house prices exhibited a distinctively stable pattern over the last three decades, without a boom-bust cycle around the 2007–2008 financial crisis (e.g., [Dustmann et al. \(2022\)](#)). Possibly as a result of these market features, yield dispersion is arguably lower in Germany than in other markets—[Voigtländer \(2009\)](#) shows it is indeed the case for aggregate time-series variation. Moreover, flats—the type of properties we focus on—are probably characterized by smaller yield dispersion than detached houses or commercial properties. Then, the fraction of unexplained variation in yields we find may qualify as a lower bound. Put differently, the heterogeneity puzzle is probably more pronounced in other economies and/or other segments of the real estate market.

A growing asset-pricing literature studies the cross-section of real estate valuations. [Sinai and Souleles \(2005\)](#) analyze the determinants of residential rent-to-price ratios using MSA-level data from the US, focusing on the trade-off between rent risk for tenants and house price risk for owners. [Han \(2013\)](#) studies how the risk-return relation for residential properties varies across US local markets. Using US postal code-level data on housing returns, [Eiling, Giambona, Lopez Aliouchkin, and Tuijp \(2020\)](#) document substantial dispersion across MSAs in the sources of risk priced in housing returns, with idiosyncratic risk playing a sizable role. [Giacoletti \(2021\)](#) and [Sagi \(2021\)](#) use residential and commercial property-level data to study the idiosyncratic component of real estate prices, showing that it does not follow a random walk, contrary to a common assumption in asset-pricing models. [Kantak \(2025\)](#) examines the relation between local expected economic growth and price-to-rent ratios at the MSA level in the US, documenting a negative cross-sectional association and explaining

it through an asset-pricing model with housing consumption commitments. Relatedly, [Feng, Jaimovich, Rao, Terry, and Vincent \(2023\)](#) show that local industry structure, and the manufacturing share in particular, helps explain inequality in housing valuations. [Demers and Eisfeldt \(2022\)](#) decompose total returns to U.S. single-family rentals into net rental yields and house price appreciation, showing that yields and capital gains are negatively correlated across cities. Then, [Halket \(2025\)](#) documents large within-city variation in total returns to single-family houses across U.S. neighborhoods, driven primarily by differences in net yields. We contribute to this literature by studying the distribution and drivers of property-level rent-to-price ratios using a large and granular dataset for Germany. We show that a substantial and economically meaningful share of their variation cannot be explained by local factors or by observable property characteristics. In the Online Appendix, we further discuss how our contribution relates to other strands of the literature.

II. Data and Housing Yields Construction

Through the RWI-GEO-RED database maintained by the Research Data Center Ruhr (FDZ Ruhr) at RWI Essen ([Breidenbach and Schaffner \(2020\)](#)), we obtain information on prices and characteristics of residential properties for the period January 2007-October 2017 from ImmobilienScout24, a major German real estate listings website. The platform covers about 50% of all real estate properties listed for sale or rent in Germany ([an de Meulen, Micheli, and Schaffner \(2014\)](#)), which speaks to the representativeness of the data provided by the platform. We restrict the analysis to flats excluding detached houses, because of the more standardized nature of the former properties. The raw data contain 16,429,909 listings of flats for rent, and 7,122,908 for sale. The standardization of these properties translates into liquid rental and sale markets, which eases the matching exercise we conduct below to recover synthetic rent-to-price ratios. By contrast, the German rental market for detached houses is thin, which would adversely impact the reliability of the matching procedure below. By focusing on flats, we arguably over-represent urban relative to rural areas in our sample.

Whereas RWI-GEO-RED data come in monthly vintages of listings, instances of flats reappearing on the platform for multiple months are relatively infrequent. We thus narrow down the analysis to listings appearing only once or at their first appearance in the dataset. We then remove observations for which information on any of the following traits is missing: price (for rental or for sale), surface, rooms, bathrooms, bedrooms, floor number, or location (district, municipality, postal code, 1-km raster cell). We also remove observations with sale (monthly rental) price below EUR 10,000 (EUR 50) and surface below 10 square meters (sqm). We exclude observations in the top 0.5% of price, surface, number of rooms, bathrooms, bedrooms, and floor number. As a result of this screening, the sample goes down to 5,100,753 rental listings, and 1,982,461 sale listings.

It is worth noting that we observe *listed* and not *transaction* prices ([an de Meulen et al. \(2014\)](#)). As such, they reflect the supply-side assessment of property value rather than the actual market price, meaning that they are likely to be upward biased. This is an objective limitation of our empirical setting, but two reasons alleviate concerns on the soundness of the analysis. First, the analysis below mostly focuses on rent-to-price ratios, so that the biases in rental and sale prices should to some extent cancel out. Second, contributors to the platform are generally professional estate agents, which should ensure a degree of rationality in reported prices, being possibly based on the opinion of qualified real estate appraisers. Indeed, although transaction prices are a more credible gauge of property valuations, it is not uncommon to supplement them with listed prices to build longer time series (e.g. [Amaral, Dohmen, Kohl, and Schularick \(2025\)](#)). Throughout the paper, we consistently control for property-level demand pressure—as proxied by the number of visits, clicks, and time online of the listings—to account for those listed prices that are more likely to be upward biased.

Nationwide data on inflation and interest rates are from Federal Reserve Economic Data (FRED) of the St. Louis Federal Reserve Bank. Valuation ratios for stocks listed on the New York Stock Exchange (NYSE) are from Kenneth French’s website. Regional data, used for tests in the Online Appendix, are from the German Federal Statistical Office. Additional

datasets used for specific tests, along with a detailed discussion of the corresponding results, are described in the Online Appendix.

All monetary variables are expressed in 2007 euros (EUR). Similarly, all returns and growth rates are expressed in real terms. Moreover, to reduce the impact of outliers, we trim variables at the 0.5% and 99.5% level.

A. Matching Approach

The rent-to-price ratio of a given property is an unobservable quantity as we typically do not simultaneously observe the rental and sale price of the same property. We follow different matching schemes to obtain counterfactual sale prices for flats for rent.

Baseline Matching Scheme. For the baseline analysis, we adopt a parsimonious set of matching covariates balancing the number of attainable matches against their quality, namely: flat surface (distance minimization), number of rooms, number of bedrooms, number of bathrooms, floor category, five-digit postal code, and calendar quarter (exact matching). The floor category indicates whether the flat is in the basement, at ground floor, at floors 1 to 3, or at higher floors. Note that we do not observe whether a flat is furnished or not, potentially a key information for rent determination. However, the German rental market is largely dominated by unfurnished flats, which should mitigate this limitation. At the same time, we do not match on the year of construction of the property or for the quality of facilities because these pieces of information, while important for housing valuations, are often missing in our sample. Nonetheless, below we perform robustness tests in which we discard those matches characterized by sizable distances on those dimensions.

Alternative Matching Schemes. We then implement four alternative matching exercises.

- (a) First, we carry out a propensity-score matching (PSM) approach by estimating a logit model for the rental indicator on a comprehensive set of flat-level covariates across the full sample of listed properties. We then perform 1:1 nearest-neighbor matching with

replacement and a caliper of 0.1, imposing exact matching on postal code and calendar quarter.

- (b) Second, we use a matching scheme that follows the baseline algorithm but allows properties to be matched across different time periods. To ensure comparability, we adjust matched prices for local house price inflation and include, among the matching criteria, the distance in listing quarters between paired properties.
- (c) Third, we augment the baseline exact matching covariates with a categorical variable capturing the conditions of the property—going from flats needing renovation to new ones. Requiring that the flat for rent and the matched sale flat are in the same conditions allows us to better control for quality differences.
- (d) Finally, we impose a more precise matching in terms of flat location. Indeed, whereas in Germany five-digit postal codes are quite narrowly defined—especially in large cities—, in several instances they cut through different municipalities or even different labor market regions (as defined by [Kosfeld and Werner \(2012\)](#)). We thus exploit geo-referenced information on the location of properties at the level of 1-km raster cells, which we supplement with information on the five-digit postal code, municipality, and labor market region of properties. By considering all possible combinations of such location traits, we define highly granular geographic areas. Imposing exact matching on these areas greatly improves the precision in terms of property location, reducing the errors in rent-to-price ratios stemming from potential differences in neighborhood-level amenities within matched flats. The dataset resulting from this alternative matching scheme features matched properties from over 20,000 of these highly granular geographic areas as opposed to less than 5,000 different postal codes. Based on Germany’s population of 81.7M as of 2015, these geographic areas have on average around 4,000 inhabitants, a size that is pretty close to that of US census tracts, but, unlike the latter, are not specifically defined to be homogeneous in terms

of socioeconomic characteristics.⁵ Nonetheless, forcing matched properties to be in the same narrow area should ensure that they have access to similar amenities (e.g., restaurants, shops, entertainment, parks), which have been shown to be important determinants of households’ location choices and house prices (e.g., Couture, Gaubert, Handbury, and Hurst (2024)).

The schemes (a) and (b) substantially expand the set of matched observations and mitigate sample bias, likely at the cost of match quality. In contrast, the schemes (c) and (d) unambiguously improve matching precision, but pose a trade-off when it comes to sample bias by restricting the analysis to those flats for which a very high quality match is available. For these reasons, below we use these alternative schemes only for supplementary analyses.

Admittedly, any matching scheme relies on the assumption that properties for rent and for sale are ultimately comparable. However, the choice to rent or to sell a flat is not random but is instead likely to reflect also its unobservable characteristics. Hence, we conduct additional tests on a subsample of properties for which both the rental and the sale price are observable.

B. Computation of Housing Yields

In the baseline matching, we remove any match for which we do not obtain exact matching on each considered discrete variable or for which the absolute distance in terms of surface is larger than 10 sqm (in robustness tests below, we also consider a lower threshold for such a distance). We take one match, among flats for sale, for each flat for rent. However, because of “ties”, around one-fifth of rental properties have multiple matches: the average (resp., maximum) number of matches is 1.56 (resp., 20).⁶ Given each match between the flat for rent and the counterfactual flat for sale, we compute the natural logarithm of the

⁵See, e.g., <https://www2.census.gov/geo/pdfs/education/CensusTracts.pdf>.

⁶Allowing for multiple matches helps us have less volatile estimates of the sale price of properties for rent. At the same time, however, we remove extremes (and rare) cases, such as those properties with more than 20 matches.

annual rent-to-price ratio (in %)—the quantity whose variation we seek to explain in our main regression analysis—for the “synthetic flat” f as

$$(1) \quad \ln \left(\frac{H_{f,t}}{P_{f,t}} \right) = \ln \left(100 \times \frac{12 \times H_{r,t}}{P_{\bar{s},t}} \right),$$

where r , s , and t denote the flat for rent, the flat for sale, and the calendar quarter, respectively.⁷ Following the notation of [Plazzi et al. \(2010\)](#), $H_{r,t}$ is the monthly rent per sqm exclusive of expenses (*Kaltmiete* in German). Such a measure of the rental price is arguably a purer house pricing measure and is more comparable to sale prices per sqm, $P_{\bar{s},t}$, than the rent inclusive of expenses (e.g., utilities), which we nonetheless use for robustness purposes below. In other words, $H_{r,t}$ better approximates the period income the property owner gets from his/her investment. The notation $\bar{s}$ indicates that—in case more than one flat for sale is matched to the flat for rent—we average out their prices.

It is worth pointing out that the valuation ratio in equation (1) is a gauge of the owner’s *gross* yield on the property because it does not reflect maintenance costs (not covered by the RWI-GEO-RED database), which can be sizable and typically increase the volatility of the owner’s income stream ([Chambers, Spaenjers, and Steiner \(2021\)](#)). By the same token, we are not able to reliably adjust the gross yield for taxation on the property itself, inheritance, capital gains, or rental income (e.g., in a progressive income tax system such as the German one, the effective tax rate on rental income depends on the overall—and unobservable to us—income of the specific owner). Yet, below we also conduct robustness tests on *net* yield measures accounting for nontax costs, vacancy losses, and property taxes, which we approximate by means of regional variables and available estimates from the literature (e.g., [Chambers et al. \(2021\)](#), [Eichholtz, Korevaar, Lindenthal, and Tallec \(2021\)](#), [Demers and Eisfeldt \(2022\)](#)).

⁷We express the rent-to-price ratio in annual percentage terms to favor its readability in the cases in which we do not consider its logarithmic transformation. In baseline regressions, however, we use this transformation, under which the multiplication by 100×12 in equation (1) boils down to a constant, leaving the other coefficient estimates unchanged.

III. Heterogeneity in Housing Yields

This section documents the extent of heterogeneity in housing yields and progressively assesses how much of it can be attributed to observable property characteristics and local geographic factors.

A. Descriptive Statistics

We start with a naive examination of the degree of heterogeneity in synthetic rent-to-price ratios from equation (1). Table 1 reports summary statistics for flat characteristics under the baseline matching approach.⁸ The final sample contains 1,625,131 flats for rent matched to counterfactual flats for sale.⁹ By construction, the differences between these two groups of flats are statistically indistinguishable from zero for matching covariates—exactly zero in the case of covariates on which we impose exact matching, such as the number of rooms. Still, based on *t*-tests allowing for district-level clustering, flats for rent display statistically significant differences from flats for sale with respect to most of the other covariates, although such differences are typically modest in magnitude. Nonetheless, below we augment rent-to-price ratio regression specifications with these observable differences to absorb possible systematic patterns in ratios arising artificially from the matching exercise.

[Insert Table 1 about here]

With this caveat in mind, Figure 1 visualizes the empirical distribution of the synthetic rent-to-price ratio together with the national and local macroeconomic conditions over the sample period. In Graph A, we examine the distribution conditional on the size

⁸Variable definitions are presented in Table A.1. Variables used for additional tests in the Online Appendix are defined in Table OA.1.

⁹The number of observations for matched flats for sale is mechanically equal to that of flats for rent. As we allow for matching with replacement, the number of distinct sale listings that are actually used is lower at 753,386 units, reflecting their lower frequency in the raw data relative to flats for rent. In our baseline matching exercise, the average (median) sale listing is used 3.6 (2) times as a match.

category of the property, uncovering a positive relation between valuations and flat surface. Variation across categories is limited with a median roughly ranging between 6% for small flats and 4.5% for very large flats. In Graph B, in which we condition on the federal state where properties are located, both between and within-group variation is more pronounced. States such as Bavaria and Hamburg exhibit substantially lower ratios than Eastern states like Saxony-Anhalt or areas that underwent massive de-industrialization like Saarland. Not surprisingly, the economic success of states appears to correlate negatively with rent-to-price ratios. Graph C highlights that the median ratio moderately increases up to 2010 to then decline. This pattern is mirrored, although less markedly, by within-period heterogeneity, which exhibits an initial upward trend followed by a decrease toward the end of the sample period.¹⁰ This is broadly consistent with the economic conditions observed both nationally and locally in Germany over 2007-2017. Over the sample, the only recession was the 2008-2009 whereas around the European debt crisis a mere slowdown took place (Graph D). Periods of stronger economic conditions are associated with sales prices growing faster than rents—implying lower housing yields—and with a concurrent compression of cross-sectional dispersion. While, in general, this pattern could mechanically arise from nominal rigidities in rents, this explanation is less compelling in our setting since we focus on new rental contracts, which should adjust more flexibly.

[Insert Figure 1 about here]

This first aggregate evidence also suggests that variation in rent-to-price ratios is to a large extent cross-sectional, relating to both to property-specific and regional features. Federal states display substantial median differences in rent-to-price ratios, but patterns become more and more nuanced as we consider finer geographical subdivisions, as highlighted by Figure 2. Graph A documents within-state median disparities in ratios that are more remarkable than those between states (e.g., focusing on Bavaria, the Munich area vs. the

¹⁰Consistently, Figure OA.1 shows a hump-shaped behavior of both the level and dispersion of rent-to-price ratios.

districts on the Czech border). If we zoom in on the seven “global” German cities—i.e., those with an advanced service sector and that serve as hubs of international transportation networks—and consider variation at the postal code level in Graphs B-H, we observe that media ratios greatly vary even within some of the most thriving metropolitan areas like Hamburg or Frankfurt.¹¹

[Insert Figure 2 about here]

B. A Geographical Decomposition of Time-Series Variation in Housing Yields

In the spirit of Piazzesi, Schneider, and Tuzel (2007), we then decompose time-series variation of valuation ratios into postal code-, district-, federal state-, and national level variation. For each geographic area, we compute the average of rent-to-prices ratios quarter-by-quarter so to obtain time-series, for which we estimate standard deviation, coefficient of variation, and interquartile range. Table 2 reports the average of these statistics across geographic units. Similarly to Piazzesi and Schneider (2016), we find that national variation in rent-to-price ratios only explains up to half of postal code-level variation. These results do not change qualitatively whether, in computing average measures of variation, we weight geographic areas equally or by their number of listed properties. While we cannot evaluate property-specific time variation in ratios as our data are not longitudinal, we can still compare the average time-series standard deviation across postal code areas (around 1%) against property-level overall standard deviation of rent-to-price ratios (around 2%), which suggests that idiosyncratic patterns in pricing are pervasive. The decomposition of yield variation points to the relevance of looking at single properties within fine geographic subdivisions to investigate house pricing, which resonates with the above finding that cross-sectional heterogeneity is prevalent. The analysis below centers on

¹¹We identify global cities as those with a rating ranging between “Alpha” and “Gamma” according to the 2020 ranking by the Globalization and World Cities Research Network (see <https://www.lboro.ac.uk/gawc/world2020t.html>). Other major cities are those with a “Sufficiency” rating in the same ranking. Even starker within-city differences—though at generally lower levels of the rent-to-price ratio—emerge if we look at such cities in Figure OA.2.

providing bounds to the fraction of variation in yields that can be explained by local as opposed to idiosyncratic factors.

[Insert Table 2 about here]

Our estimates of variation in rent-to-price ratios for Germany can then be compared to those from other countries. [Halket \(2025\)](#), in an exercise akin to our Table 2, compute postal code-level time-series standard deviations of yields on single-family homes (net of property taxes) for 21 US cities, obtaining estimates that range between 0.62% and 2.13%. Our postal code-level estimate for Germany is around 1%, which hints at a similar degree of variability in yields, but on the low-end of the distribution across US cities. Also the levels of our yield estimates appear to be comparable to the US ones: [Demers and Eisfeldt \(2022\)](#) estimate an average yield (net of property taxes, maintenance costs, and vacancy losses) of 4.2% for major cities; the average yields obtained by [Halket \(2025\)](#) range between 4.52% and 8.01%. The average gross yield in our sample is 5.51%. Below, when adjusting for likely—yet unobservable to us—costs, we obtain an average net yield of around 3.5%, only slightly below US levels. Moving to other economies, [Voigtländer \(2009\)](#) provides a direct comparison, based on aggregate data, between Germany and selected European countries (UK, Netherlands, and Spain) for the period 1970-2009, illustrating that time-series variability of both returns and yields is lowest in Germany, which, contrary to other countries, experienced negative real growth of house prices over such a period. It is possible that in the subsequent period characterized by house price appreciation—the one we consider—the distribution of German yields became more aligned with other European countries. Overall, it seems safe to argue that variation in rent-to-price ratios is not particularly pronounced in Germany. Put differently, despite its features contributing to a peculiarly low homeownership rate, the German housing market appears to be a useful setting to learn about general patterns in the pricing of residential properties.

C. A Comparison Against Equity Yields

It is also instructive to compare the heterogeneity of rent-to-price ratios against that of valuation ratios for a well-known asset class like the US equities. To this end, we retrieve information on the quantiles of ratios over the cross-section of NYSE stocks. Table 3 reports selected percentiles for (actual) rent-to-price ratios for German residential properties vs. dividend-, earnings-, and cash flow-to-price ratios for US stocks over 2007-2017.¹² Focusing on the interquartile range and on the spread between the 95th and the 5th percentile, rent-to-price ratios exhibit a degree of cross-sectional dispersion in line with stock dividend-to-price ratios, but lower than earnings- and cash flow-to-price ratios.

Moreover, based on the coefficient of variation of the time-series of average nationwide rental and sale prices per sqm, the former appear stickier than the latter (0.065 vs. 0.166), even though our rents refer to newly listed prices and are thus less subject to the nominal rigidities that characterize contracts in place. In other words, though new ask rental prices are more akin to market prices than stock dividends—which are subject to managerial discretion—they exhibit similarly contained variation over time. We confirm this finding below in a pseudo-panel setting, in which we can more reliably track the time-series variation of rents and prices for similar flats. Importantly, because we observe only newly posted rents rather than the stock of ongoing rental contracts, our estimates likely represent an upper bound on the true time-series variability of rents.

[Insert Table 3 about here]

D. Observable Property Characteristics

Next, we assess the role of observable property-specific hedonic characteristics in explaining equilibrium valuation ratios. We consider all property characteristics observable

¹²As we highlight below, the actual rent-to-price ratio is available for a subsample of properties on sale for which a rental income is reported.

to us, including those not used to obtain the matched rent-to-price ratio from equation (1):

$$(2) \quad \ln \left(\frac{H_{f,t}}{P_{f,t}} \right) = \gamma \mathbf{m}_{r,\bar{s}} + \boldsymbol{\eta} \mathbf{x}_{r,t} + \boldsymbol{\theta} \mathbf{x}_{\bar{s},t} + \boldsymbol{\zeta} \mathbf{z}_{r,\bar{s},t} + \tau_t + \epsilon_{f,t}.$$

$\mathbf{m}_{r,\bar{s}}$ is the vector of covariates on which we match, either by minimizing distance (surface) or exactly (number of rooms, number of bedrooms, number of bathrooms, floor category), with the addition of squared surface.¹³ We control for matching covariates as they constitute some key dimensions along which our matched pairs of flats differ *across* each other; at the same time, because of the matching, there is typically no variation *within* each pair along such dimensions. $\mathbf{x}_r$ contains covariates available only for rental flats (flat expenses, heating expenses, an indicator for rent inclusive of heating expenses, deposit). $\mathbf{x}_{\bar{s}}$ includes covariates available only for sale flats (community charge, an indicator for holiday properties, an indicator for rented out properties). $\mathbf{z}_{r,\bar{s}}$ is a set of distances between flat r and synthetic flat $\bar{s}$ for characteristics on which we do not match (number of floors in the building, energy source, flat expenses, etc.): $\mathbf{z}_{r,\bar{s}} = \mathbf{w}_r - \mathbf{w}_{\bar{s}}$. If a flat trait is missing, we set it to its median value. To mitigate the bias potentially arising from this adjustment, for any incompletely reported variable, we include a corresponding missing value indicator. To absorb variation in nationwide macroeconomic conditions, we control for calendar quarter fixed effects τ_t .¹⁴ The standard errors are clustered at the district level.

Table 4 presents coefficient estimates from specification (2). Column 1 includes the covariates on which we match properties. The coefficients on flat size imply a hump-shaped relationship between surface and housing yields, indicating that the latter increase at low levels of surface but decline as surface becomes larger, after controlling for other characteristics. Similarly, the correlation of the rent-to-price ratio with the number of rooms is positive. Interestingly, bedrooms and bathrooms exhibit a negative, although not statistically significant for the latter, association with rent-to-price ratios, possibly

¹³We take the average between the surface of the rented flat and matched flats for sale. Where appropriate, we use log-transformed covariates (e.g., surface).

¹⁴Regressions models in the main analysis are estimated via the Stata package REGHDFE by [Correia \(2018\)](#).

pointing to a value-increasing role of other types of rooms. However, pairwise correlations among surface, number of rooms, and number of bedrooms are above 70% (this does not extend to the number of bathrooms), which suggests caution in interpreting their regression coefficients. The floor on which the flat is located does not load significantly.

Columns 2 and 3 introduce variables specific to rental and sale listings, respectively. Rental contracts requiring a higher deposit or higher heating expenses, and holiday properties for sale come with higher valuations. Properties for sale that are already rented are valued significantly less: these properties, for instance, may have not undergone modernization for a longer time, may be occupied by a defaulting tenant, or the owner may be forced to sell the property at fire-sale price.¹⁵

The correlation patterns described so far remain overall robust once we include all remaining observable covariates in column 4, with adjusted R^2 rising by 18% to 30% (from 12% in column 3).¹⁶ Heterogeneity in the rent-to-price ratio is unlikely to be uniquely a by-product of observable flat traits.

[Insert Table 4 about here]

E. Local vs. Idiosyncratic Factors

The analyses conducted so far suggest that cross-sectional variation in rent-to-price ratios is substantial. The importance of cross-sectional variation is corroborated by the fact that most of time-series variation in yields originates within highly granular geographic units, such as postal code areas. The idiosyncratic component of variation also appears to be crucial, with observable (to us) property-specific covariates together with time fixed effects explaining about one-third of the variation of property-level rent-to-price ratios.

¹⁵Matched and actual log rent-to-price ratios are strongly correlated (63.77%). Figure OA.3 shows that the two measures are similarly distributed, although matched yields are slightly lower than actual yields for the same properties (median 5.09% vs. 5.65%), consistent with some selection of lower-quality properties into the rental market. The PSM approach below helps tackle possible concerns about this issue.

¹⁶In Table OA.2, we augment specification (2) with interactions between hedonic matching covariates and national macroeconomic variables to allow for time-varying coefficients; the adjusted R^2 remains at 30%.

One of the most prominent features setting the housing market apart from other financial asset markets is its pronounced geographic segmentation. Therefore, a substantial fraction of heterogeneity in rent-to-price ratios may be explained by local differences in factors such as the age structure of the population, unemployment, income per capita, and the like. In Table 5, we thus seek to provide an upper bound to the fraction of rent-to-price ratio variability that can be explained by local geographic factors. In particular, we saturate the baseline specification in column 4 of Table 4—including all observable property-level covariates—with progressively finer geographical fixed effects. Relative to the baseline (adjusted R^2 of 30%), federal state and federal state-calendar quarter fixed effects can explain 5% and 6% more rent-price ratio variation (columns 1 and 2), respectively. In columns 3 and 4, we examine fixed effects at the district or district-quarter level. In this case, the fraction of explained variation rises to 40% and 44%, respectively. In columns 5 and 6, we introduce postal code and postal code-quarter fixed effects, reaching an adjusted R^2 of 47% and 59%, respectively.¹⁷

[Insert Table 5 about here]

Figure 3 sheds light on the distribution of residuals from these regressions, starting from a specification without any fixed effect up to one featuring postal code-quarter fixed effects. No matter the fixed effects scheme, residuals are symmetrically distributed if we consider the full regression sample. As we saturate the specification with finer fixed effects, however, the distribution becomes less dispersed and more peaked at the mean of 0. If we split the regression sample based on rent-to-price ratios quartiles, we observe that richer specifications produce more centered residuals. Especially for the top (bottom) quartile yield subsample, the residuals from the model without postal code-quarter fixed effects exhibit substantially negative (positive) mean and median. Hence, the 41% unexplained variation of

¹⁷Table OA.3 re-estimates the baseline specifications without log-transforming the dependent or explanatory variables. The adjusted R^2 upper bounds decline, indicating that the unexplained variation is not driven by the log specification.

yields from column 6 of Table 5 stems from a remarkably symmetric distribution, in which outliers seem to play a limited role.¹⁸

[Insert Figure 3 about here]

Time-varying factors at the postal code level, together with observable property-level traits, are *potentially* able to explain a sizable fraction of heterogeneity in rent-to-price ratios. But variation that cannot be even potentially explained by those variables is still at a staggering 41%. And the challenge is not only about identifying the possible sources of the 41% of unexplained variation, but also about pinning down the *actual* local time-varying factors behind the 29% ($= 59\% - 30\%$) of variation captured by postal code-quarter fixed effects.

To attack this problem, we use the historical record of social and economic indicators at the district level to study their impact on rent-to-price ratio regional heterogeneity. After selecting relevant demographic, economic, and local housing market factors (Tables OA.5, OA.6, and OA.7), we assess how much of the local fixed effects they explain (Table OA.8). These district-level variables account for only about two-fifths of the 14% ($= 44\% - 30\%$) of yield variation captured by district-quarter fixed effects.

To sum up, although specifications featuring postal code-quarter fixed effects explain around three-fifths of variation in yields, the part that we can re-conduct to specific property- or regional-level factors is substantially lower.

IV. Exploring Unexplained Variation

Observable property-level characteristics and fine geographic fixed effects cannot span around two-fifths of variation in housing yields. This unexplained heterogeneity in German rent-to-price ratios goes hand in hand with the prominent role of idiosyncratic shocks for the variance of property-level returns documented in the US market (Giacoletti

¹⁸We confirm the limited presence and impact of outliers in Figures OA.4 and OA.5, and in Table OA.4.

(2021)). But unexplained variation in yields may not be strictly idiosyncratic (e.g., related to random liquidity shocks to landlords selling or renting the property at fire-sale prices, to investors’ heterogeneity in terms of risk/consumption preferences, etc.). Apart from idiosyncratic factors, we thus evaluate several possible non-mutually exclusive stories behind unexplained cross-sectional variation in rent-to-price ratios, such as measurement issues, local amenities and agglomeration economies, and housing market frictions (illiquidity, supply rigidities, regulation).

A. Measurement Issues

Despite the richness of the observable property-level traits we match on or control for, several sources of measurement error may still contribute to the heterogeneity in housing yields that we document. First, listed gross yields may differ from the returns relevant to market participants because of expenses faced by tenants that are not fully reflected in listed rents, or because of costs borne by owners that drive a wedge between gross and net returns. Second, measurement error may arise from imperfect matching between rental and sale properties, either because relevant property-level dimensions such as finishes, noise, and views from the flat are omitted, or because matched properties still differ along observed characteristics.

In Table 6, we seek to assess some of these sources of “artificial” heterogeneity. In column 1, we start by considering rent-to-price ratios adjusted for expenses faced by tenants—namely, by using rent inclusive of utilities at the numerator—in a specification otherwise identical to the most saturated one of Table 5. Not surprisingly, the adjusted R^2 decreases from 59% to 52%, because the type and contractual features of included expenses are likely to introduce a further layer of heterogeneity in our valuation ratios, possibly exacerbating measurement problems. In other words, the choice to focus on rents exclusive of expenses for the computation of housing yields is warranted. Moreover, we do not observe the costs faced by owners. Although such costs are likely to exacerbate rather than mitigate

idiosyncratic dispersion in yields by inducing higher volatility in rental income (Chambers et al. (2021)), it is useful to tentatively quantify them. In Table OA.9, we compute measures of net yields by approximating maintenance costs, vacancy losses, and property taxes in a variety of ways. Net yields are on average roughly 30% lower than gross ones and exhibit a very similar fraction of unexplained variation in specifications including postal code-quarter fixed effects.

We then consider imperfect matching due to omitted property-level dimensions, such as layout, exposure, views, decoration, cultural heritage status, or proximity to amenities. These characteristics are unobservable to us but may be relevant for pricing, especially in the presence of search and matching frictions in the housing market (Genesove and Han (2012)). To assess their importance, column 2 of Table 6 estimates the specification with postal code-quarter fixed effects using the actual, rather than synthetic, rent-to-price ratio available for a subset of flats, probably non-owner-occupied flats for sale. Since no rental-sale matching is required in this case, this exercise avoids matching errors due to omitted property-level dimensions. The adjusted R^2 increases to 71%. However, this higher explanatory power appears to reflect sample selection rather than the absence of matching error: when we use the synthetic rent-to-price ratio for the same subsample in column 3, the adjusted R^2 remains very similar, at 70%.

Yet, sample selection may itself be associated with lower matching errors. We therefore further scrutinize the role of matching quality. We next examine whether residual differences between matched properties along observed characteristics can account for the unexplained dispersion in yields. In column 4, we augment the specification with fixed effects for each percentile of a matching quality score, defined as an equally weighted average of re-scaled distances between the rental flat and the matched sale flat(s) for covariates not used in the matching procedure; when a trait is missing for one or both properties, we assign the maximum distance. As an additional check, we also address the possibility that repeated listings inflate measured heterogeneity. In column 5, we exclude properties

that, despite having distinct identifiers in the RWI-GEO-RED database, display the same pricing and characteristics within a given postal code-quarter. These “likely duplicates” may reflect properties listed more than once on the platform, or by multiple realtors. In neither specification does the adjusted R^2 exceed the baseline value of 59%.

Next, we turn to imperfect matching due to residual differences between matched properties along observed characteristics. To address this concern, we impose additional restrictions in the matching algorithm that further improve comparability across matched properties, beyond controlling for covariate distances. In column 6, we retain only matches with an absolute difference in living surface of at most 5 sqm, compared with the 10 sqm threshold used in the baseline analysis. In column 7, we apply the same restriction to usable surface.¹⁹ In column 8, we require matched properties to have construction years no more than 10 years apart. In column 9, we require them to have the same quality of facilities, based on a four-category variable ranging from “simple” to “deluxe.” The stricter restriction on living surface excludes more than 450,000 matched properties and leaves the adjusted R^2 unchanged at 59%. The adjusted R^2 rises to 69% when focusing on matched properties with negligible differences in usable surface, but this comes at the cost of losing almost 1.5M matched properties. Filtering on year of construction and quality of facilities increases the adjusted R^2 to 64% and 65%, respectively; in both cases, however, the sample shrinks to around 300,000 observations. This large reduction in sample size is mainly due to missing information on these variables for either the rental flat or the matched sale flat, rather than necessarily reflecting genuine matching errors in our algorithm. Hence, the increase in explanatory power may simply reflect stronger sample selection.

[Insert Table 6 about here]

Table 7 further explores the trade-off between sample size and match quality by implementing four alternative matching schemes. Schemes (a) and (b) relax some features

¹⁹According to ImmobilienScout24, usable surface is typically larger than living surface, as it also includes ancillary usable areas, such as balconies, cellars, storage rooms, or attics, while excluding circulation areas such as hallways and staircases.

of the baseline algorithm and substantially expand the sample of matches, while schemes (c) and (d) impose additional restrictions to improve comparability along flat quality and location, respectively. Scheme (a) implements a 1:1 nearest-neighbor PSM algorithm with replacement and a caliper of 0.1, while imposing exact matching on postal code and calendar quarter. Propensity scores are estimated from a logit model for the probability that a listed property is for rent, using a comprehensive set of flat characteristics.²⁰ PSM increases the number of matched flats from around 1.6M in the baseline sample to more than 3.6M, thereby alleviating concerns that our results are driven by selection induced by the baseline matching procedure. In column 1 of Table 7, we estimate our preferred saturated specification for the log-transformed rent-to-price ratio—obtained via PSM—with postal code-quarter fixed effects. The adjusted R^2 is 57%, close to the baseline value of 59% in column 6 of Table 5. This similarity confirms again that the unexplained heterogeneity documented in the baseline analysis is not an artifact of the exact matching scheme. It also mitigates concerns that the modest bias in synthetic yields induced by selection of lower-quality properties into the rental market, discussed above with respect to Figure OA.3, drives our main finding.

Matching scheme (b) follows the baseline algorithm but allows rental and sale properties to be matched across different listing periods. To make prices comparable over time, we adjust sale prices using district-level house price inflation before computing synthetic rent-to-price ratios. Relaxing exact matching on the listing period expands the set of feasible matches and improves matching on other covariates, increasing the matched sample to more than 4M flats. However, the larger sample appears to come at the cost of lower explanatory power—the adjusted R^2 falls to 45%—possibly due to lower match quality.

In scheme (c), we extend the baseline algorithm by also imposing exact matching on flat condition, which reduces the sample to fewer than 0.7M matched properties. In column

²⁰Table OA.10 shows that rental and sale listings differ in observables before matching, but that PSM substantially improves balance across the two samples. Figure OA.6 shows distinct propensity-score distributions, yet with substantial common support across rental and sale flats.

3, we re-estimate our usual specification obtaining an adjusted R^2 of 56%, lower than in the baseline sample, which suggests that mismatches in flat quality are not a major driver of the dispersion in housing yields.

Scheme (d), instead, refines the baseline approach by imposing exact matching on highly granular geographic areas, defined by all possible combinations of 1-km raster cells, five-digit postal codes, municipalities, and labor market regions. The resulting dataset includes around 0.6M matched properties. This approach minimizes the spatial distance between each rental flat and its matched sale flat, thereby improving comparability in terms of exposure to local amenities (e.g., restaurants) and agglomeration economies, such as the pooling of highly skilled human capital. Despite this more accurate matching along spatial dimensions, the adjusted R^2 in column 4 remains at 61%.

[Insert Table 7 about here]

Overall, the alternative matching schemes point to a clear trade-off. Less restrictive algorithms, such as schemes (a) and (b), substantially expand the sample but appear to do so at the cost of lower match quality, as reflected in unreported weaker correlations with both the baseline and actual rent-to-price ratios. By contrast, the more restrictive schemes (c) and (d) yield rent-to-price ratios that remain closely aligned with the baseline. Taken together, these findings suggest that measurement issues are unlikely to be a major source of unexplained variation in rent-to-price ratios under our baseline matching scheme. Put differently, the observed dispersion in housing yields does not appear to be an artifact of our methodology.

B. Local Amenities and Agglomeration Economies

Two complementary explanations for high residual heterogeneity in yields relate to amenities and agglomeration economies operating below the postal code level. A growing body of evidence highlights how better access to amenities induces households to sort into

certain neighborhoods based on their income or skills (e.g., [Handbury and Weinstein \(2015\)](#), [Ferreira and Wong \(2023\)](#), [Couture et al. \(2024\)](#), [Diamond and Moretti \(2021\)](#)), potentially affecting property valuations. [Rosenthal and Strange \(2020\)](#) discuss how agglomeration economies—due, for instance, to knowledge spillovers and access to concentrated skilled workforce—are spatial in nature and can be present at different levels, even within neighborhoods, blocks or buildings. [Leonardi and Moretti \(2023\)](#) uncover agglomeration effects in the within-city location choices of amenities. Therefore, a substantial part of the observed heterogeneity in rent-to-price ratios may be driven by such spatial effects and not to stand-alone traits of properties or of landlords/households.

To explore this possibility, we continue to work with the matching scheme (d). In column 5 of Table 7, we saturate the rent-to-price ratio regression with geographic area-quarter fixed effects. Relative to the specification with postal code-quarter in Table 5, the adjusted R^2 increases by 12% reaching 71%. Time-varying factors at the level of these narrow geographic areas thus meaningfully impact dispersion in housing valuations. This is consistent with a relevant role of local amenities as well as of local agglomeration economies. We seek to disentangle the two stories by splitting the sample between global cities (columns 6 and 7) and other regions (columns 8 and 9), conjecturing that the former are more conducive to agglomeration effects. Across the two subsamples, the adjusted R^2 exhibits analogous patterns when moving from postal code-quarter to geographic area-quarter fixed effects. More intuitively, the explanatory power of these neighborhood-level factors appears to be more consistent with within-city disparities in access to amenities—a phenomenon likely at work also outside big cities—than with local agglomeration effects. Nonetheless, around 30% of variation in housing yields remains unexplained even after accounting for any possible difference in amenities available across our highly granular geographic areas. Research is needed to verify whether this residual variation is at least partially spatial in nature—i.e., at even more local level, like streets or blocks—or rather mostly reflects the property- or investor-specific factors we discuss below, as the study of [Eichholtz et al. \(2021\)](#)

on Amsterdam and Paris suggests.

C. Housing Market Frictions

Several local housing-market frictions could contribute to unexplained heterogeneity in housing valuations. One such friction is limited information diffusion. Unlike stock markets, housing markets are decentralized, illiquid, and lack an infrastructure comparable to that of equity analysts. As a result, investors, especially private individuals, may search only locally and trade infrequently, limiting price discovery. Market segmentation with limited search ([Piazzesi, Schneider, and Stroebe \(2020\)](#)) and property illiquidity ([Head and Lloyd-Ellis \(2012\)](#)) can therefore prevent localized supply and demand shocks from being fully arbitrated away, generating cross-sectional dispersion in valuation ratios. A second set of frictions concerns rigidities in local housing supply, due for instance to zoning laws or geographical constraints. When supply is inelastic and demand is strong, even small differences in property quality may translate into large valuation wedges. Finally, local regulation may further amplify yield heterogeneity. Housing-market regulation can take several forms, including zoning, taxation, and rent controls. To the extent that these rules affect properties differentially even within the same postal code, they may contribute to the unexplained dispersion in rent-to-price ratios.

In [Table OA.11](#), we assess the role of housing market frictions by estimating our saturated specification separately across districts with high versus low liquidity, housing supply, construction activity, and property taxation. The explanatory power of the specification remains close to the baseline across all subsamples, suggesting that these frictions do not account for a quantitatively large share of the residual dispersion in rent-to-price ratios. However, these tests should be interpreted with caution, as our proxies for local frictions are relatively coarse and may fail to capture more granular variation within districts.

Rent regulation is a relevant form of housing-market regulation in the German

context. Traditionally, however, German rent controls have mainly restricted rent increases in ongoing contracts, rather than rents on newly listed units such as those in our data. This is also the case for the 2013 *capping limit*, which tightened limits on rent increases for existing contracts in designated municipalities. The 2015 *rental brake* represents instead a more relevant case for our setting, as it extended rent regulation to new contracts in tight markets, while exempting new and substantially modernized properties. In Table OA.12, we therefore assess whether the rental brake affected the dispersion of rent-to-price ratios using a difference-in-differences design based on its staggered implementation across German municipalities, while controlling for the capping limit. We find no statistically significant increase in yield dispersion, suggesting that rent controls are unlikely to be a major driver of the heterogeneity we document.²¹

To sum up, we fail to find compelling evidence of a major role of measurement issues, agglomeration economies, informational frictions, supply rigidities, or regulation for postal code-level unexplained dispersion of rent-to-price ratios. By contrast, within-postal code differences in access to amenities seem to account for a relevant share of dispersion. Nonetheless, even after controlling for those differences, about 30% of variation in housing yields cannot be captured.

Abstracting from rent growth expectations, rent-to-price ratios may exhibit substantial variation within narrow geographic areas because of heterogeneous risk premia.²² In turn, these risk premia may reflect properties' heterogeneous exposure to systematic risk factors but, due to widespread underdiversification of homeowners, also idiosyncratic risk, which in turn crucially depends on yield risk (Eichholtz et al. (2021),

²¹We also examine whether dispersion in rent-to-price ratios is related to local climate risk and energy efficiency. Climate-risk proxies in Table OA.13 display limited explanatory power, while dispersion in energy consumption in Table OA.14 is positively associated with yield dispersion but accounts for only a small share of the overall heterogeneity.

²²According to the static Gordon growth model, the expected rent-to-price ratio of flat f can be decomposed as $\mathbb{E}_t(H_{f,t+1})/P_{f,t} = r_t^f + \mathbb{E}_t(r_f^e) - \mathbb{E}_t(g_{H,f})$, where $\mathbb{E}_t(H_{f,t+1})$ is expected next-period rent, $P_{f,t}$ is the current price, r^f is the risk-free rate, $\mathbb{E}_t(r_f^e)$ is the risk premium, and $\mathbb{E}_t(g_{H,f})$ is expected rent growth.

Eiling et al. (2020)). The cross-sectional nature of our data limits our ability to distinguish between idiosyncratic and systematic risk as determinants of such risk premia (and housing yields dispersion). The pseudo-panel constructed in the next section helps overcome this problem by introducing a time-series dimension: by testing whether rent-to-price ratios predict future excess returns and/or rent growth in the presence of high or low yield dispersion, we can assess whether such dispersion is more closely related to risk premia or to expected rent growth.

Moreover, idiosyncratic variation in yields relates to the growing literature on heterogeneity across housing-market investors (e.g., Fischer, Füss, Ruf, and Stehle (2022), Gurun, Wu, Xiao, and Xiao (2023), Lambie-Hanson, Li, and Slonkosky (2022)). In the presence of market frictions, unexplained local dispersion in yields may reflect heterogeneity in investors’ beliefs about both discount rates and future rent growth. For instance, a particularly pessimistic seller may ask for a low price, inducing a high rent-to-price ratio; more generally, investor disagreement about future rent growth can generate idiosyncratic dispersion in yields.²³

Property-specific risk premia and expected rent growth rates—potentially originating from dispersion in investors’ beliefs and preferences—are thus likely to underlie unexplained variation in yields. This speaks to the work of Ludwig, Mankart, Quintana, and Wiederholt (2024), who show that heterogeneity in household expectations about future house price growth has sizable implications for the level and dynamics of equilibrium house prices, highlighting the importance of belief dispersion for housing valuations. At the same time, spatial effects related to amenities and agglomeration economies may be at work even below the neighborhood level.

²³As an indirect test of this channel, we examine whether institutional ownership is associated with yield dispersion, using municipality-level measures of the share of dwellings owned by property companies and other institutional investors. We find little evidence that ownership structure explains the dispersion in housing yields (Table OA.15). However, this test is limited by the granularity of the available ownership data.

V. Risk Premia vs. Expected Rent Growth

Because flat listings in our database are available as monthly vintages, akin to repeated cross-sections, we construct a pseudo-panel by aggregating cohorts of similar properties to introduce a time-series dimension for prices and rents of comparable properties. We use this pseudo-panel to study whether rent-to-price ratios predict future excess returns or rent growth, also shedding light on whether local yield dispersion is more closely related to risk premia or to expected rent growth.²⁴

The rent-to-price ratio, analogous to the dividend-to-price ratio for stocks (Campbell and Shiller (1988)), serves as a gauge of market participants' expectations about future housing returns and rent growth (Plazzi et al. (2010)). In particular, we estimate predictive regressions at the cohort-level

$$(3) \quad \begin{aligned} r_{c,t+1 \rightarrow t+k}^e &= \beta_k \ln \left(\frac{H_{c,t}^q}{P_{c,t}} \right) + \tau_c + \nu_{c,t+1 \rightarrow t+k}, \\ \Delta h_{c,t+1 \rightarrow t+k} &= \lambda_k \ln \left(\frac{H_{c,t}^q}{P_{c,t}} \right) + \tau_c + \varsigma_{c,t+1 \rightarrow t+k}, \end{aligned}$$

where $r_{c,t+1 \rightarrow t+k}^e$ and $\Delta h_{c,t+1 \rightarrow t+k}$ are the k -quarter-ahead excess return and rent growth rates, respectively. To focus on time-series variation, we include cohort-level fixed effects, τ_c , which capture any time-invariant difference across cohorts (and, therefore, districts). We account for heteroskedasticity and correlation in error terms $\nu_{c,t+1 \rightarrow t+k}$ and $\varsigma_{c,t+1 \rightarrow t+k}$ by adjusting standard errors following Driscoll and Kraay (1998).²⁵

In Table 8, we estimate the predictive specifications in (3). In columns 1 to 3, we look at excess housing returns at horizons ranging from one quarter to three years. The rent-to-price ratio is invariably significant and positively related to the current rent-to-price

²⁴Table OA.16 reports pseudo-panel summary statistics, confirming that variation in yields and returns is mostly local and that rents vary less over time than sale prices. Figure OA.7 shows that housing performed strongly over the sample period, supporting the use of the pseudo-panel to study return and rent-growth dynamics.

²⁵In unreported tests, we verify that our findings are not sensitive to removing cohort fixed effects or to using alternative standard errors.

ratio. The within R^2 suggests that the discount rate expectations impounded in the ratio explain up to 16% of the time-series variation in cohort-level housing premia. The rent-to-price ratio loads significantly and negatively in the case of rent growth specifications, explaining up to 9% of variation. Findings on both returns and rent growth are consistent with the traditional present-value relationship: a higher rent-to-price ratio reflects higher expected discount rates and/or lower expected rent growth.²⁶

[Insert Table 8 about here]

Table 9 shows that the predictive content of rent-to-price ratios depends on the degree of cross-sectional dispersion in housing valuations. In cohorts with higher dispersion, the rent-to-price ratio more strongly forecasts future returns, while in lower-dispersion cohorts its predictive power is relatively stronger for rent growth. This result is robust to measuring dispersion using either raw yields (Panel A) or filtered yields (Panel B). This pattern is consistent with the view that cross-sectional dispersion in rent-to-price ratios largely reflects heterogeneity in risk premia, rather than differences in expected rent growth, thereby shedding light on the heterogeneity puzzle illustrated in the main analysis.

[Insert Table 9 about here]

VI. Further Analyses

In this section, we study the behavior of cross-sectional dispersion in housing yields at different levels of geographic aggregation. Moreover, we quantify the relevance of unexplained heterogeneity in housing yields by benchmarking it against the distribution of wealth of German households.

²⁶Table OA.17 augments the predictive regressions with calendar-quarter fixed effects, absorbing aggregate fluctuations in the spirit of Fama and French (2025). While R^2 increases—especially for returns—the rent-to-price ratio retains predictive power, pointing to the role of local factors driving yield variation. Table OA.18 shows that the predictive power of the rent-to-price ratio is robust to controlling for local factors, although the within R^2 rises substantially, indicating that local conditions contain additional predictive information.

A. The Role of Geographic Aggregation

We investigate to which extent cross-sectional heterogeneity in rent-to-price ratios evolves as we move from finer to coarser geographic units. After removing likely duplicate property listings that may dampen variation as highlighted above, we compute the standard deviation of valuation ratios for each geographic unit-quarter (with at least 30 observations, to obtain meaningful estimates). Starting from non-log-transformed ratios to favor interpretation of economic magnitudes, we carry out this procedure for both raw ratios and ratios filtered for property-specific observables and postal code-quarter fixed effects, and at different levels of aggregation: postal code area, district, and federal state.

Table 10 contains information about the distribution of these cross-sectional standard deviations. Mean dispersion of raw ratios increases by 31.25% ($= 2.125/1.619 - 1$) going from postal code- to federal state-level data. After factoring out observables and postal code-quarter fixed effects, the degree of dispersion is de facto invariant to the coarseness of aggregation (as we would expect), with a mean standard deviation of roughly 1.4% (or around 25% of the mean raw ratio). This evidence corroborates the largely local nature of cross-sectional variation in rent-to-price ratios.

[Insert Table 10 about here]

We also provide additional evidence on the time-series and spatial structure of local yield dispersion. Table OA.19 documents that dispersion in filtered rent-to-price ratios is only weakly persistent over time at the postal code-quarter level, while Figure OA.8 shows that dispersion is positively correlated across nearby postal codes but that this correlation declines rapidly with distance. These results suggest that unexplained yield dispersion is not a fixed local characteristic, but that it nevertheless exhibits a localized spatial component.

B. Implications for Household Wealth

The high degree of heterogeneity—both explained and unexplained—in rent-to-price ratios maps into the distribution of household wealth around Germany. We attempt to put this relationship in perspective by means of some back-of-the-envelope calculations.

The median flat in our dataset, which we use as a benchmark, has a surface of 66 sqm and its annual rent per sqm (exclusive of expenses) is EUR 81.06. The 10th and 90th percentiles of the rent-to-price ratio (non log-transformed) are 3.16% and 8.29%, respectively.²⁷ These ratios correspond to housing valuations of EUR 169,303 ($= (81.06 \times 66)/0.0316$) and EUR 64,535 ($= (80.88 \times 66)/0.0829$) for the median flat, implying a variation in the housing wealth of its owners of EUR 104,768 ($= \text{EUR } 169,303 - \text{EUR } 64,535$).

If we repeat the same calculations on the unexplained component of rent-to-price ratios alone, a not-so-different picture emerges. To this end, based on the specifications in columns 4 and 6 of Table 5—where, besides controlling for property-level characteristics, we include district- and postal code-quarter fixed effects—, we obtain residuals of the log-transformed percentage rent-to-price ratio: $\hat{\epsilon}_{d,t}$ and $\hat{\epsilon}_{p,t}$ respectively. We then retrieve the filtered rent-to-price ratio as $\exp[\text{P50}_{\ln(H_{f,t}/P_{f,t})} + \hat{\epsilon}_{d,t}]$, where $\text{P50}_{\ln(H_{f,t}/P_{f,t})}$ is the unconditional median log-transformed percentage rent-to-price ratio. We proceed analogously for $\hat{\epsilon}_{p,t}$. This procedure allows us to abstract from any variation in the valuation ratio due to observable property characteristics and time-varying factors at the district or postal code level. Looking at the difference between the 10th and 90th percentiles of the filtered rent-to-price ratio, the implied variation in valuations of the median flat stands at EUR 88,531 and EUR 71,837 if we absorb time-varying district- or postal code-level factors, respectively. Magnitudes shrink—as we would expect—but remain sizable.

²⁷The 10th and 90th percentile of the rent-to-price ratio refer to its unconditional distribution. One may be concerned that in this way we are simply picking up variation across flats with very different characteristics. Nonetheless, the degree of heterogeneity in rent-to-price ratios decreases only by little if we focus on flats very similar to the median one. For instance, looking only at flats with a surface between 60 and 70 sqm and an annual rent per sqm between EUR 75 and EUR 85, the 10th and 90th percentiles of the ratio are 3.34% and 8.16%.

To put things in perspective, we retrieve information on German household balance sheets from [Deutsche Bundesbank \(2019\)](#). As of 2017, the mean (resp., median) net wealth across all households is EUR 202,541 (resp., EUR 61,597). For the 44% of home-owning households, the mean (resp., median) value of their main residence stands at EUR 225,161 (resp., EUR 173,308). Given these figures, explained and unexplained heterogeneity in rent-to-price ratios spans a substantial fraction of dispersion in household wealth. The forces at play at geographically granular level may thus translate into substantial disparities in household wealth via housing valuations. In turn, depending on households willingness and ability to consume out of housing wealth (e.g., [Berger et al. \(2018\)](#)), this may transmit to local business cycles and, ultimately, economic development.

VII. Conclusion

Relying on sale and rental prices for flats from a major German online real estate platform, we study the distribution and drivers of rent-to-price ratios. To this end, we compute property-level rent-to-price ratios by matching rental flats to comparable flats for sale. After accounting for a wide set of property-specific characteristics, much of the variation in rent-to-price ratios remains unexplained, even when we absorb time-varying postal code-level factors with fixed effects or if we account for measurement error in a variety of ways. This residual heterogeneity appears to be related to disparities in amenities across neighborhoods and to property-specific differences in risk premia and rent growth expectations, potentially driven by heterogeneity in investors’ beliefs and preferences. A predictability analysis suggests that risk premia play a more important role than expected rent growth in accounting for the dispersion of housing yields.

Overall, this paper points to the existence of a surprisingly large degree of heterogeneity in rent-to-price ratios, with the interdecile range of their unexplained component accounting for more than one-third of the mean German household’s wealth. Given the importance of real estate valuations both for the distribution of wealth across

households and for the amplification of business cycles through the consumption channel, further work is needed to rigorously pin down the origin(s) of such yet unexplained variation, be it related to local amenities and/or agglomeration effects, heterogeneous exposures to systematic risk factors, idiosyncratic risk, investors' disagreement about future rent growth, limited arbitrage due to transaction costs (e.g., informational and regulatory frictions), or to other economic mechanisms.

References

- Amaral, F.; M. Dohmen; S. Kohl; and M. Schularick. “Superstar Returns? Spatial Heterogeneity in Returns to Housing.” *Journal of Finance*, 80 (2025), 3057–3094.
- an de Meulen, P.; M. Micheli; and S. Schaffner. “Documentation of German Real Estate Market Data.” Working Paper (2014).
- Bach, S., and S. Eichfelder. “Real Estate Taxation Reform: Tax Land Values, Abolish Privileges.” *DIW Weekly Report*, 11 (2021), 199–205.
- Beraja, M.; A. Fuster; E. Hurst; and J. Vavra. “Regional Heterogeneity and the Refinancing Channel of Monetary Policy.” *Quarterly Journal of Economics*, 134 (2019), 109–183.
- Berger, D.; V. Guerrieri; G. Lorenzoni; and J. Vavra. “House Prices and Consumer Spending.” *Review of Economic Studies*, 85 (2018), 1502–1542.
- Breidenbach, P., and S. Schaffner. “Real Estate Data for Germany (RWI-GEO-RED).” *German Economic Review*, 21 (2020), 401–416.
- Campbell, J. Y., and R. J. Shiller. “The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors.” *Review of Financial Studies*, 1 (1988), 195–228.
- Chambers, D.; C. Spaenjers; and E. Steiner. “The Rate of Return on Real Estate: Long-Run Micro-Level Evidence.” *Review of Financial Studies*, 34 (2021), 3572–3607.
- Correia, S. “REGHDFE: Stata Module to Perform Linear or Instrumental-Variable Regression Absorbing Any Number of High-Dimensional Fixed Effects.” Boston College (2018).
- Couture, V.; C. Gaubert; J. Handbury; and E. Hurst. “Income Growth and the Distributional Effects of Urban Spatial Sorting.” *Review of Economic Studies*, 91 (2024), 858–898.
- Demers, A., and A. L. Eisfeldt. “Total Returns to Single-Family Rentals.” *Real Estate Economics*, 50 (2022), 7–32.

- Deutsche Bundesbank. “Household Wealth and Finances in Germany: Results of the 2017 Survey.” *Monthly Report March*, 57 (2019).
- Diamond, R., and E. Moretti. “Where Is Standard of Living the Highest? Local Prices and the Geography of Consumption.” Working Paper (2021).
- Driscoll, J. C., and A. C. Kraay. “Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data.” *Review of Economics and Statistics*, 80 (1998), 549–560.
- Dustmann, C.; B. Fitzenberger; and M. Zimmermann. “Housing Expenditure and Income Inequality.” *Economic Journal*, 132 (2022), 1709–1736.
- Eichholtz, P.; M. Korevaar; T. Lindenthal; and R. Tallec. “The Total Return and Risk to Residential Real Estate.” *Review of Financial Studies*, 34 (2021), 3608–3646.
- Eiling, E.; E. Giambona; R. Lopez Aliouchkin; and P. Tuijp. “Homeowners’ Risk Premia: Evidence from Zip Code Housing Returns.” Working Paper (2020).
- Fama, E. F., and K. R. French. “House Prices and Rents.” *Review of Financial Studies*, 38 (2025), 547–563.
- Feng, X.; N. Jaimovich; K. Rao; S. J. Terry; and N. Vincent. “Location, Location, Location: Manufacturing and House Price Growth.” *Economic Journal*, 133 (2023), 2055–2067.
- Ferreira, F. V., and M. Wong. “Estimating Preferences for Neighborhood Amenities Under Imperfect Information.” Working Paper (2023).
- Fischer, M.; R. Füss; D. Ruf; and S. Stehle. “Betting on Homes.” Working Paper (2022).
- Flavin, M., and T. Yamashita. “Owner-Occupied Housing and the Composition of the Household Portfolio.” *American Economic Review*, 92 (2002), 345–362.
- Genesove, D., and L. Han. “Search and Matching in the Housing Market.” *Journal of Urban Economics*, 72 (2012), 31–45.
- Giacoletti, M. “Idiosyncratic Risk in Housing Markets.” *Review of Financial Studies*, 34 (2021), 3695–3741.

- Guerrieri, C., and C. Mendicino. “Wealth Effects in the Euro Area.” Working Paper (2018).
- Guren, A. M.; A. McKay; E. Nakamura; and J. Steinsson. “Housing Wealth Effects: the Long View.” *Review of Economic Studies*, 88 (2021), 669–707.
- Gurun, U. G.; J. Wu; S. C. Xiao; and S. W. Xiao. “Do Wall Street Landlords Undermine Renters’ Welfare?” *Review of Financial Studies*, 36 (2023), 70–121.
- Halket, J. “The Cross-Section of Housing Returns.” Working Paper (2025).
- Han, L. “Understanding the Puzzling Risk-Return Relationship for Housing.” *Review of Financial Studies*, 26 (2013), 877–928.
- Handbury, J., and D. E. Weinstein. “Goods Prices and Availability in Cities.” *Review of Economic Studies*, 82 (2015), 258–296.
- Head, A., and H. Lloyd-Ellis. “housing Liquidity, Mobility, and the Labour Market.” *Review of Economic Studies*, 79 (2012), 1559–1589.
- Kaas, L.; G. Kocharkov; E. Preugschat; and N. Siassi. “Low Homeownership in Germany—A Quantitative Exploration.” *Journal of the European Economic Association*, 19 (2021), 128–164.
- Kantak, P. “Consumption Commitments and Housing Dynamics.” *Management Science*, 71 (2025), 8670–8697.
- Kohl, S. “Urban History Matters: Explaining the German–American Homeownership Gap.” *Housing Studies*, 31 (2016), 694–713.
- Kohl, S., and J. Sørvoll. “Varieties of Social Democracy and Cooperativism: Explaining the Historical Divergence Between Housing Regimes in Nordic and German-Speaking Countries.” *Social Science History*, 45 (2021), 561–587.
- Kosfeld, R., and A. Werner. “Deutsche Arbeitsmarktreionen–Neuabgrenzung Nach Den Kreisgebietsreformen 2007–2011.” *Raumforschung und Raumordnung—Spatial Research and Planning*, 70 (2012), 49–64.

- Lambie-Hanson, L.; W. Li; and M. Slonkosky. "Real Estate Investors and the US Housing Recovery." *Real Estate Economics*, 50 (2022), 1425–1461.
- Leonardi, M., and E. Moretti. "The Agglomeration of Urban Amenities: Evidence from Milan Restaurants." *American Economic Review: Insights*, 5 (2023), 141–157.
- Ludwig, A.; J. Mankart; J. Quintana; and M. Wiederholt. "Heterogeneity in Expectations and House Price Dynamics." Working Paper (2024).
- Mian, A.; K. Rao; and A. Sufi. "Household Balance Sheets, Consumption, and the Economic Slump." *Quarterly Journal of Economics*, 128 (2013), 1687–1726.
- Piazzesi, M., and M. Schneider. "Housing and Macroeconomics." *Handbook of Macroeconomics*, 2 (2016), 1547–1640.
- Piazzesi, M.; M. Schneider; and J. Stroebel. "Segmented Housing Search." *American Economic Review*, 110 (2020), 720–59.
- Piazzesi, M.; M. Schneider; and S. Tuzel. "Housing, Consumption and Asset Pricing." *Journal of Financial Economics*, 83 (2007), 531–569.
- Plazzi, A.; W. Torous; and R. Valkanov. "Expected Returns and Expected Growth in Rents of Commercial Real Estate." *Review of Financial Studies*, 23 (2010), 3469–3519.
- Rosenthal, S. S., and W. C. Strange. "How Close Is Close? the Spatial Reach of Agglomeration Economies." *Journal of Economic Perspectives*, 34 (2020), 27–49.
- Sagi, J. S. "Asset-Level Risk and Return in Real Estate Investments." *Review of Financial Studies*, 34 (2021), 3647–3694.
- Savills. "Eigentümerstruktur Am Wohnungsmarkt." SPOTLIGHT Savills Research (2019).
- Sinai, T., and N. S. Souleles. "Owner-Occupied Housing as a Hedge Against Rent Risk." *Quarterly Journal of Economics*, 120 (2005), 763–789.
- Voigtländer, M. "Why Is the German Homeownership Rate So Low?" *Housing Studies*, 24 (2009), 355–372.
- Vuolteenaho, T. "What Drives Firm-Level Stock Returns?" *Journal of Finance*, 57 (2002),

233–264.

Figure 1. Empirical Distribution of the Rent-to-Price Ratio vs. Macroeconomic Conditions

This figure shows the conditional empirical distribution of the rent-to-price ratio (obtained via the baseline matching) through box plots (Graph A to C) against the evolution of macroeconomic conditions in Germany (Graph D) between 2007 and 2017. In Graph A, the conditioning variable is the flat's surface category, going from small (S) to medium (M), large (L), and very large (VL). In Graph B, the conditioning variable is the federal state where the flat is located (ordered by the median rent-to-price ratio): Saxony-Anhalt (ST), Bremen (HB), Saxony (SN), Saarland (SL), Lower Saxony (NI), Thuringia (TH), Schleswig-Holstein (SH), North Rhine-Westphalia (NW), Rhineland-Palatinate (RP), Mecklenburg-Vorpommern (MV), Brandenburg (BB), Hesse (HE), Baden-Württemberg (BW), Berlin (BE), Hamburg (HH), and Bavaria (BY). In Graph C, the conditioning variable is the year. Graph D shows the annual GDP growth rate at national level, as well as the median together with 1st and 99th percentile of the district-level GDP growth rate.

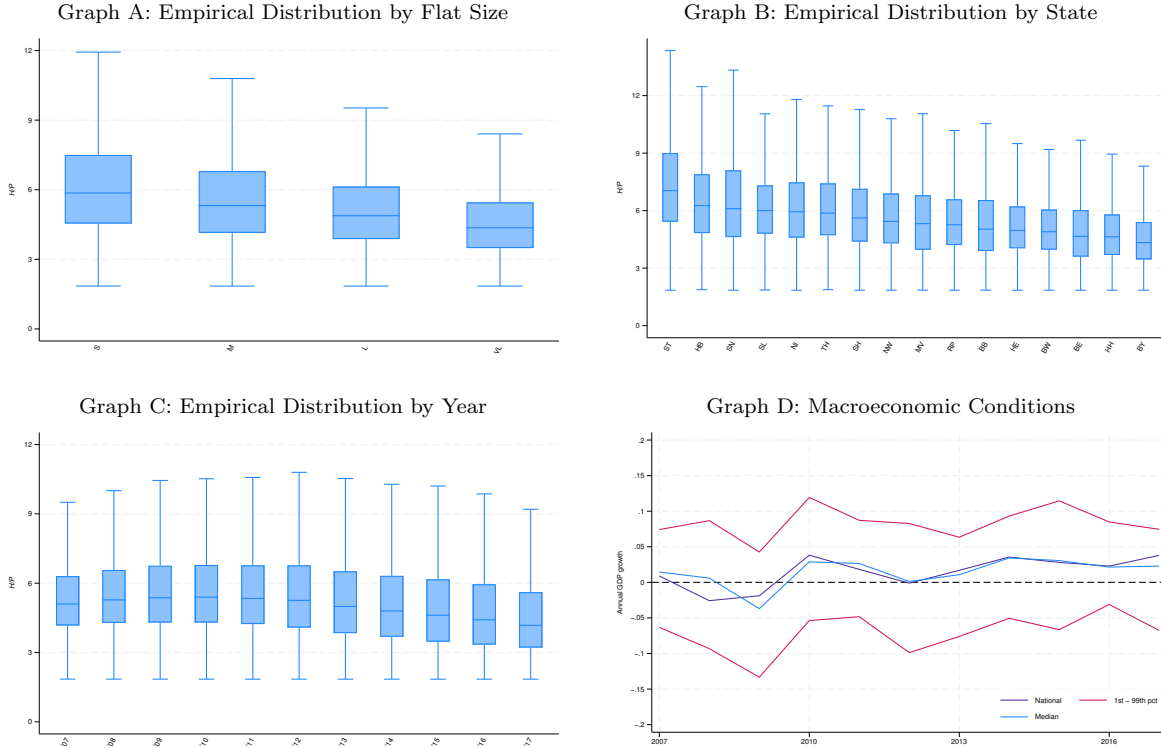


Figure 2. Geographical Dispersion in Median Rent-to-Price Ratios

This figure visualizes the median rent-to-price ratio (obtained via the baseline matching) across Germany over the period 2007–2017. Graph A reports district-level estimates nationwide, while Graphs B–H report five-digit postal code-level estimates for German global cities. Districts and postal codes with fewer than 10 observations are excluded (white colored). Global cities are those assigned a rating between “Alpha” and “Gamma” in the 2020 ranking by the Globalization and World Cities Research Network.

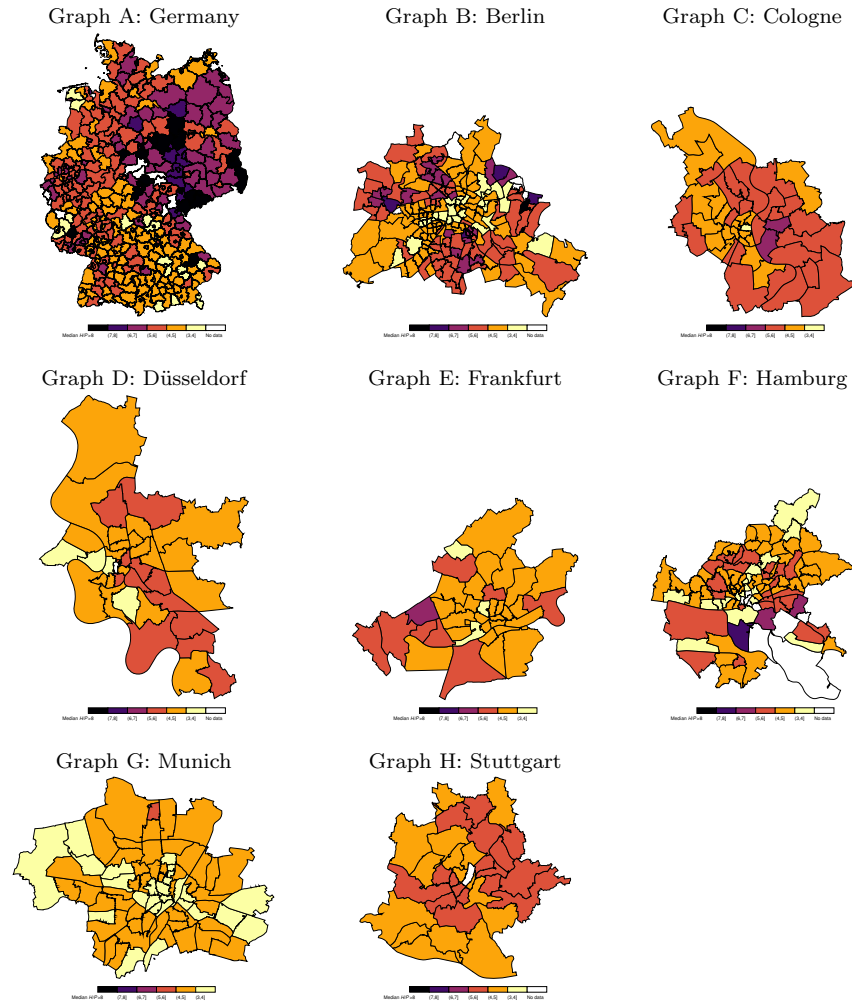


Figure 3. Distribution of Residuals from Rent-to-Price Ratio Regressions

This figure visualizes the distribution of residuals from rent-to-price ratio regressions under different fixed effects schemes by means of violin plots. The violin plot of each variable features a superimposed box plot. The residuals are first obtained from estimating equation (2) without calendar quarter fixed effects. The other specifications augment equation (2) with (i) state-quarter fixed effects as in column 2 of Table 5, (ii) district-quarter fixed effects as in column 4 of Table 5, and (iii) postal code-quarter fixed effects as in column 6 of Table 5. The violin plots are reported for the entire distribution of the rent-to-price ratio as well as conditioning on the quartile of the latter.

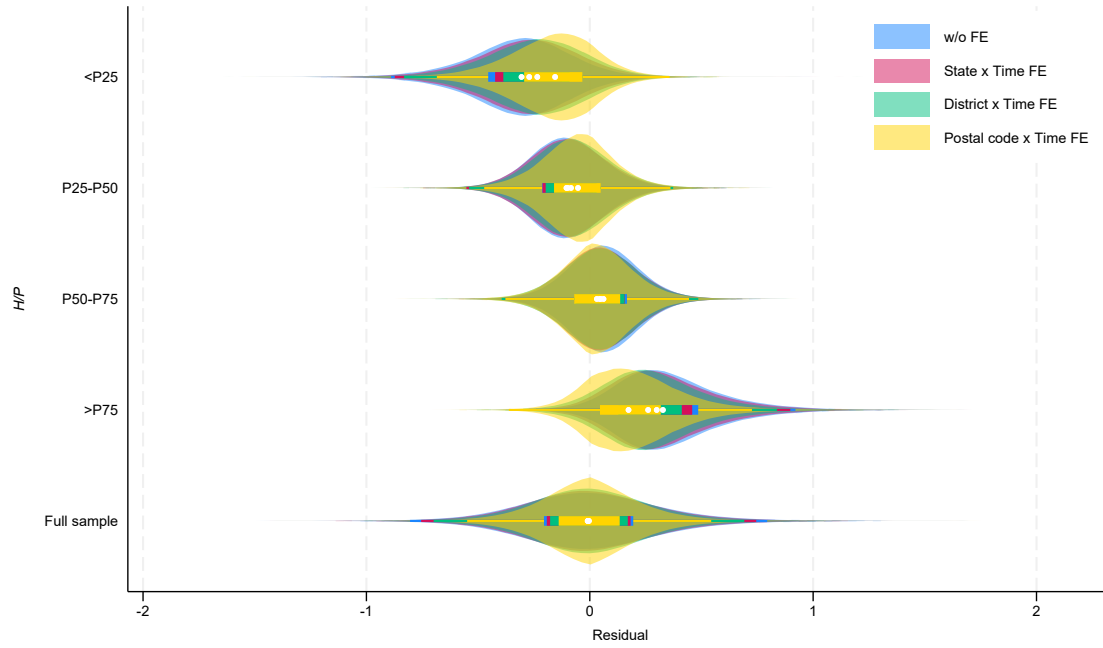


Table 1. Summary Statistics on Flats

This table shows summary statistics on property characteristics distinguishing between flats for rent and matched flats for sale between 2007 and 2017. The t -statistics for mean-comparison tests are based on standard errors clustered by district. Refer to Table A.1 for variable definitions.

	For rent			For sale (matched)			Mean-comparison test	
	Obs.	Mean	SD	Obs.	Mean	SD	Diff.	t -stat
<i>Rental listing</i>								
Annual rent (EUR)	1,625,131	5,916.965	2,678.813					
Annual inclusive rent (EUR)	1,538,530	7,376.999	2,787.831					
Expenses (EUR)	1,539,302	124.956	48.370					
Heating expenses (EUR)	338,398	66.399	24.522					
Heating included	1,431,683	0.694	0.461					
Deposit (EUR)	1,435,151	1,260.245	632.434					
<i>Sale listing</i>								
Sale price (EUR)				1,625,131	127,907.526	86,433.132		
Community charge (EUR)				1,039,854	185.498	61.848		
Rented				1,020,789	0.350	0.473		
Annual rent income (EUR)				343,019	5,391.668	3,471.617		
Holiday property				1,130,045	0.065	0.245		
<i>Matching covariates</i>								
Surface (sqm)	1,625,131	68.498	18.857	1,625,131	68.505	18.684	0.007	0.262
No. rooms	1,625,131	2.451	0.677	1,625,131	2.451	0.677	0.000	.
No. bedrooms	1,625,131	1.449	0.582	1,625,131	1.449	0.582	0.000	.
No. bathrooms	1,625,131	1.024	0.154	1,625,131	1.024	0.154	0.000	.
Floor no.	1,625,131	1.944	1.165	1,625,131	1.968	1.216	0.023	3.730

	For rent			For sale (matched)			Mean-comparison test	
	Obs.	Mean	SD	Obs.	Mean	SD	Diff.	t-stat
<i>Other characteristics</i>								
No. building floors	1,265,269	3.472	1.303	1,291,148	3.717	1.558	0.245	10.553
Usable surface (sqm)	592,967	73.090	20.650	680,945	75.848	21.014	2.759	11.018
Construction year	1,008,532	1,967.650	39.146	1,434,997	1,971.641	36.198	3.991	7.217
Year last modernization	428,778	2,009.473	5.419	464,245	2,006.766	6.578	-2.707	-9.158
Property conditions	1,369,343	2.675	0.688	1,409,938	2.617	0.790	-0.059	-5.375
Quality of facilities	884,447	2.471	0.597	903,791	2.479	0.607	0.007	1.355
Energy consumption (kWh/(sqm× y))	355,945	123.069	48.551	490,224	120.510	46.635	-2.559	-4.063
Energy rating	26,915	4.380	1.986	33,916	3.923	1.908	-0.457	-6.006
Warm water in energy exp.	876,453	0.151	0.358	888,588	0.227	0.415	0.076	15.270
Protected building	87,558	0.008	0.090	1,125,766	0.051	0.219	0.043	2.573
Balcony	1,419,668	0.794	0.404	1,501,780	0.873	0.330	0.079	22.408
Available parking	160,655	0.935	0.247	199,917	0.957	0.200	0.022	3.055
Accessible with wheelchair	16,320	0.792	0.406	24,623	0.915	0.278	0.123	4.961
Guest WC	1,116,741	0.183	0.387	1,179,567	0.214	0.407	0.031	9.223
Garden	1,017,953	0.295	0.456	1,014,797	0.336	0.469	0.041	2.988
Cellar	1,283,369	0.805	0.396	1,329,583	0.831	0.370	0.025	2.867
Kitchen	1,210,702	0.563	0.496	1,216,816	0.565	0.492	0.002	0.170
Elevator	1,096,689	0.313	0.464	1,153,651	0.448	0.495	0.135	19.212
Assisted living	518,881	0.056	0.229	492,693	0.125	0.328	0.069	19.263
Immediate availability	1,529,034	0.317	0.465	1,141,619	0.345	0.469	0.029	1.807
No. of days online	1,533,636	28.148	33.793	1,588,191	60.596	71.220	32.448	33.234
No. of hits	1,600,338	912.609	1,021.451	1,597,762	828.985	979.988	-83.623	-2.232
No. of clicks (contact button)	1,006,522	17.549	26.062	996,836	8.972	11.633	-8.577	-11.013
No. of clicks (customer profile)	295,469	7.005	7.747	299,107	3.989	4.576	-3.016	-14.432
No. of clicks (share button)	632,091	2.195	1.769	601,810	1.777	1.542	-0.418	-9.157
No. of clicks (customer URL)	797,971	6.575	7.190	925,134	6.127	7.337	-0.448	-1.881

Table 2. Time-Series Variation of Rent-to-Price Ratios at Different Levels of Geographic Aggregation

This table shows summary statistics on time-series variation of rent-to-price ratios—obtained via the baseline matching procedure of flats for rent to flats for sale—at postal code-, district-, and state- and national level. To this end, for each geographic area a time series of the rent-to-price ratio is obtained by computing its average quarter-by-quarter. Then, the standard deviation, the coefficient of variation, and the interquartile range of each of these time series is computed. Finally, based on these estimates, the average of these dispersion measures across geographic areas at a given level of aggregation are obtained and reported below. These averages are computed either giving equal weight to the geographic areas or by weighting them by the number of listed flats. Geographic area-quarters with fewer than 10 observations are removed from the sample. Refer to Table A.1 for variable definitions.

Level of aggregation:	Postal code	District	State	National
<i>Equally weighted</i>				
SD (H/P)	1.031	0.870	0.575	0.370
CV (H/P)	0.184	0.150	0.101	0.068
P75-P25 (H/P)	1.197	1.088	0.872	0.560
<i>Weighted by no. properties</i>				
SD (H/P)	0.971	0.632	0.450	0.370
CV (H/P)	0.177	0.119	0.086	0.068
P75-P25 (H/P)	1.264	0.965	0.749	0.560

Table 3. Valuation Ratios in the Housing vs. Stock Market

This table shows summary statistics regarding valuation ratios in the German housing market and in the US stock market between 2007 and 2017. Housing market valuation ratios comprise the synthetic rent-to-price ratio (obtained via the baseline matching procedure of flats for rent to flats for sale) as well as the actual one (available for a subsample of properties on sale for which a rental income is reported). Stock market valuation ratios comprise the dividend-to-price ratio (D/P), the earnings-to-price ratio (E/P), and the cash flow-to-price ratio (CF/P). The percentiles of stock market ratios are averages of annual Fama-French portfolio breakpoints over the sample period. Refer to Table A.1 for variable definitions.

	P5	P25	P50	P75	P95	P75–P25	P95–P5
<i>German housing market</i>							
H/P	2.777	3.979	5.094	6.528	9.723	2.549	6.946
Actual H/P	3.400	4.600	5.657	7.210	11.294	2.610	7.894
<i>US stock market</i>							
D/P	0.349	1.149	2.036	3.297	6.647	2.148	6.299
E/P	1.518	4.207	5.885	7.982	14.584	3.775	13.065
CF/P	2.304	5.892	8.169	11.394	19.827	5.502	17.524

Table 4. Rent-to-Price Ratio and Property Characteristics

This table reports estimates from regressions for property-level rent-to-price ratios on characteristics of the flat for a German sample between 2007 and 2017. The log-transformed rent-to-price ratio is obtained via the baseline matching procedure of flats for rent to flats for sale. Column 1 includes the covariates on which the matching exercise is performed. Column 2 adds flat traits observed only for flats for rent. Column 3 adds flat traits observed only for flats for sale. The covariates added in columns 2 and 3 are set to 0 if missing. To control for that, for each of those variable a missing value indicator (equal to 1 if such a variable is missing and 0 otherwise) is added. Column 4 augments the specification with a set of covariates capturing the differences between the matched flats for rent and for sale with respect to a host of traits (number of floors in the building, log usable surface, year of construction, year of the last modernization, property conditions, quality of facilities, protected building status, market segment, log energy consumption, energy rating, inclusion of hot water in energy consumption, balcony, availability of parking place, accessibility with wheelchair, guest WC, garden, cellar, installed kitchen, elevator, living assistance, immediate availability, log number of days online, and log number of clicks on different items of the listing), together with the corresponding missing value indicators. All specifications include calendar quarter fixed effects. Estimated coefficients are multiplied by 100 to favor readability. The t -statistics (in parentheses) are based on standard errors clustered by district. Significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively. Refer to Table A.1 for variable definitions.

	$\ln(H/P)$			
	1	2	3	4
$\ln(\text{Surface})$	38.992** (2.02)	40.652** (2.57)	32.283** (2.16)	8.080 (0.61)
$\ln(\text{Surface})^2$	-11.224*** (-4.64)	-10.227*** (-5.64)	-9.052*** (-5.40)	-6.185*** (-4.54)
No. rooms	12.610*** (9.57)	10.997*** (9.41)	10.910*** (9.98)	10.869*** (14.16)
No. bedrooms	-1.407* (-1.74)	-0.332 (-0.48)	-0.129 (-0.21)	-0.114 (-0.24)
No. bathrooms	-0.369 (-0.25)	-0.176 (-0.12)	-0.252 (-0.19)	0.065 (0.06)
Floor no.	0.536 (0.94)	0.725 (1.41)	0.666 (1.35)	0.698* (1.84)
$\ln(\text{Expenses})$		-1.287 (-1.15)	-1.368 (-1.26)	-3.589*** (-4.02)
$\ln(\text{Heating expenses})$		-4.312*** (-7.25)	-4.568*** (-7.61)	-3.047*** (-5.99)
Heating included		3.572** (2.59)	3.231** (2.43)	1.677* (1.82)
$\ln(\text{Deposit})$		-10.133*** (-3.18)	-9.684*** (-3.04)	-11.156*** (-4.25)
$\ln(\text{Community charge})$			2.733** (2.30)	0.549 (0.58)
Holiday property			-6.350*** (-6.35)	-4.821*** (-5.52)
Rented			11.561*** (8.09)	7.894*** (6.81)
Missing indicators	No	Yes	Yes	Yes
Covariate distances	No	No	No	Yes
Time FE	Yes	Yes	Yes	Yes
Unit of observation	Match. flat	Match. flat	Match. flat	Match. flat
Mean(y)	1.63	1.63	1.63	1.63
SD(y)	0.38	0.38	0.38	0.38
Observations	1,625,131	1,625,131	1,625,131	1,625,131
Adjusted R^2	0.09	0.11	0.12	0.30

Table 5. Local Variation of Rent-to-Price Ratios

This table reports coefficients of determination from regressions for property-level rent-to-price ratios on different fixed effect structures for a German sample between 2007 and 2017. The log-transformed rent-to-price ratio is obtained via the baseline matching procedure of flats for rent to flats for sale. Each column augments the specification of column 4 of Table 4 with progressively finer fixed effects (indicated below). Standard errors are clustered by district. Refer to Table A.1 for variable definitions.

	$\ln(H/P)$					
	1	2	3	4	5	6
Flat covariates	Yes	Yes	Yes	Yes	Yes	Yes
Missing indicators	Yes	Yes	Yes	Yes	Yes	Yes
Covariate distances	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	No	Yes	No	Yes	No
State FE	Yes	No	No	No	No	No
State×Time FE	No	Yes	No	No	No	No
District FE	No	No	Yes	No	No	No
District×Time FE	No	No	No	Yes	No	No
Postal code FE	No	No	No	No	Yes	No
Postal code×Time FE	No	No	No	No	No	Yes
Unit of observation	Match. flat	Match. flat	Match. flat	Match. flat	Match. flat	Match. flat
Mean(y)	1.63	1.63	1.63	1.63	1.63	1.63
SD(y)	0.38	0.38	0.38	0.38	0.38	0.38
Observations	1,625,131	1,625,131	1,625,129	1,624,137	1,624,481	1,606,359
Adjusted R^2	0.35	0.36	0.40	0.44	0.47	0.59
Adjusted within R^2	0.27	0.27	0.28	0.28	0.29	0.29

Table 6. Unexplained Variation of Rent-to-Price Ratios and Measurement Error

This table reports coefficients of determination from regressions for property-level rent-to-price ratios for a German sample between 2007 and 2017, exploring the role of potential measurement error in rent-to-price ratios. Each column builds on the specification of column 4 of Table 4 augmented with fixed effect structures (indicated below). In column 1, the dependent variable is the log-transformed inclusive rent-to-price ratio (i.e., adjusted for utilities), as obtained via the baseline matching procedure of flats for rent to flats for sale. In column 2, the dependent variable is the log-transformed actual rent-to-price ratio, as observed for a subsample of properties for sale. This specification, by construction of the dependent variable, does not control for traits observed only for flats for rent, for the differences between the matched flats, or for the corresponding missing value indicators. In columns 3 to 9, the dependent variable is the log-transformed rent-to-price ratio, as obtained via the baseline matching procedure of flats for rent to flats for sale. Column 3 restricts the sample to properties for which also the actual rent-to-price ratio is available. Column 4 augments the specification with fixed effects for each percentile of a matching quality measure. Column 5 removes from the sample potential duplicate listings. In columns 6 to 9, additional restrictions in the matching procedure are introduced. The sample in column 6 (column 7) discards any match for which the absolute difference in terms of (usable) surface between the flat for sale and the matched flat for sale is larger than 5 sqm. The sample in column 8 discards any match for which the absolute difference in terms of construction year between the flat for sale and the matched flat for sale is larger 10 years or for which the construction year is missing for any of the flats. The sample in column 9 discards any match for which the flat for sale and the matched flat for sale have different reported quality of facilities or for which this variable is missing for any of the flats. Standard errors are clustered by district. Refer to Table A.1 for variable definitions.

	$\ln(\text{Inclusive } H/P)$	$\ln(\text{Actual } H/P)$	$\ln(H/P)$						
	1	2	3	4	5	6	7	8	9
Flat covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Missing indicators	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Covariate distances	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Postal code×Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Matching quality FE	No	No	No	Yes	No	No	No	No	No
Likely duplicates	Included	Included	Included	Included	Excluded	Included	Included	Included	Included
Additional matching	–	–	–	–	–	Surface	Usable surface	Constr. year	Quality of fac.
Unit of observation	Match. flat	Match. flat	Match. flat	Match. flat	Match. flat	Match. flat	Match. flat	Match. flat	Match. flat
Mean(y)	1.88	1.77	1.74	1.63	1.63	1.63	1.63	1.64	1.61
SD(y)	0.39	0.39	0.36	0.38	0.37	0.37	0.34	0.33	0.36
Observations	1,530,719	330,696	330,696	1,606,359	1,356,087	1,138,189	123,077	348,225	283,989
Adjusted R^2	0.52	0.71	0.70	0.59	0.58	0.59	0.69	0.64	0.65
Adjusted within R^2	0.10	0.08	0.24	0.29	0.30	0.29	0.31	0.23	0.24

Table 7. Unexplained Variation of Rent-to-Price Ratios Under Alternative Matching Schemes

This table reports coefficients of determination from regressions for property-level rent-to-price ratios for a German sample between 2007 and 2017, under alternative matching schemes. The dependent variable is the log-transformed rent-to-price ratio, as obtained via four different matching procedures of flats for rent to flats for sale. Matching scheme (a) is based on a 1:1 nearest-neighbor PSM algorithm with replacement and a caliper of 0.1. To obtain the propensity scores, a logit model for a rental indicator is estimated. The covariates include traits available both for flats for rent and for sale (flat surface, number of rooms, number of bedrooms, number of bathrooms, floor, usable surface, year of construction, year of the last modernization, property conditions, quality of facilities, market segment, energy consumption, energy rating, inclusion of hot water in energy consumption, balcony, availability of parking place, accessibility with wheelchair, guest WC, garden, cellar, installed kitchen, elevator, living assistance, immediate availability, number of days online, and number of clicks on different items of the listing), together with the corresponding missing value indicators. Exact matching on postal code and calendar quarter is imposed. Matching scheme (b) follows the baseline matching but allowing for matching of properties for rent and for sale across periods. All prices are adjusted for district-level house price inflation to ensure comparability across periods. Matching scheme (c) augments the baseline matching procedure by imposing exact matching also on flat conditions. Matching scheme (d) in columns 4 to 9 imposes exact matching at a finer geographic level than the five-digit postal code, namely on the areas resulting from combinations of labor market regions, municipalities, five-digit postal codes, and 1-km raster cells. Columns 6 and 7 restrict the sample to global cities, whereas columns 8 and 9 focus on the remaining regions. Each column builds on the specification of column 4 of Table 4 augmented with fixed effect structures (indicated below). Standard errors are clustered by district. Refer to Table A.1 for variable definitions.

	$\ln(H/P)$								
	1	2	3	4	5	6 Global cities	7 Global cities	8 Other	9 Other
Flat covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Missing indicators	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Covariate distances	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Postal code $\times$ Time FE	Yes	Yes	Yes	Yes	No	Yes	No	Yes	No
Geo $\times$ Time FE	No	No	No	No	Yes	No	Yes	No	Yes
Unit of observation	Match. flat	Match. flat	Match. flat	Match. flat	Match. flat	Match. flat	Match. flat	Match. flat	Match. flat
Matching scheme	(a)	(b)	(c)	(d)	(d)	(d)	(d)	(d)	(d)
Mean(y)	1.66	1.65	1.64	1.64	1.64	1.55	1.55	1.70	1.70
SD(y)	0.38	0.36	0.33	0.36	0.37	0.35	0.35	0.36	0.37
Observations	3,631,269	4,006,429	685,663	591,881	550,090	230,341	221,831	361,526	328,259
Adjusted R^2	0.57	0.45	0.56	0.61	0.71	0.60	0.68	0.60	0.72
Adjusted within R^2	0.23	0.30	0.21	0.29	0.29	0.33	0.33	0.27	0.26

Table 8. Predictive Regressions

This table reports estimates from predictive regressions for the housing premium and rent growth on the log-transformed quarterly rent-to-price ratio. The regressions are estimated on a pseudo-panel constructed from a sample of German flats listed between 2007 and 2017. The unit of observation is at the cohort-calendar quarter level, where cohorts are defined by the district in which the flat is located, its number of rooms category, and its size category. The dependent variable in columns 1 to 3 (4 to 6) is the k -quarter ahead housing premium $r_{t+1 \rightarrow t+k}^e$ (rent growth $\Delta h_{t+1 \rightarrow t+k}$), with $k = 1, 4, 12$. All specifications include cohort fixed effects. Estimated coefficients are multiplied by 100 to favor readability. The t -statistics (in parentheses) are based on Driscoll-Kraay standard errors (number of lags equal to k). Significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively. Refer to Table A.1 for variable definitions.

	$r_{t+1 \rightarrow t+k}^e$			$\Delta h_{t+1 \rightarrow t+k}$		
	1 $k = 1$	2 $k = 4$	3 $k = 12$	4 $k = 1$	5 $k = 4$	6 $k = 12$
$\ln(H^q/P)$	32.784*** (12.06)	43.865*** (5.45)	59.476*** (4.40)	-4.820*** (-16.64)	-10.372*** (-7.91)	-16.784*** (-4.46)
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Unit of observation	Cohort-time	Cohort-time	Cohort-time	Cohort-time	Cohort-time	Cohort-time
Mean(y)	0.02	0.09	0.28	0.00	0.02	0.05
SD(y)	0.12	0.15	0.19	0.04	0.05	0.07
Observations	23,996	20,897	14,866	23,970	20,857	14,860
Within R^2	0.14	0.16	0.15	0.03	0.07	0.09

Table 9. Predictive Regressions and Local Dispersion in Rent-to-Price Ratios

This table reports estimates from predictive regressions for the housing premium and rent growth on the log-transformed quarterly rent-to-price ratio, accounting for local dispersion of rent-to-price ratios. The regressions are estimated on a pseudo-panel constructed from a sample of German flats listed between 2007 and 2017. The unit of observation is at the cohort-calendar quarter level, where cohorts are defined by the district in which the flat is located, its number of rooms category, and its size category. The dependent variable in columns 1-2 (3-4) is the 12-quarter ahead housing premium $r_{t+1 \rightarrow t+12}^e$ (rent growth $\Delta h_{t+1 \rightarrow t+12}$). All specifications include cohort fixed effects. Odd (even) columns restrict the estimation sample to cohorts with above-median (below-median) average cross-sectional standard deviation of rent-to-price ratios. In Panel A, such a standard deviation is based on the baseline synthetic annual rent-to-price ratios, in Panel B on filtered ones, namely the residuals from the specification in column 6 of Table 5. Estimated coefficients are multiplied by 100 to favor readability. The t -statistics (in parentheses) are based on Driscoll-Kraay standard errors (number of lags equal to k). Significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively. Refer to Table A.1 for variable definitions.

Panel A: Based on Local Dispersion in $\ln(H/P)$

	$r_{t+1 \rightarrow t+12}^e$		$\Delta h_{t+1 \rightarrow t+12}$	
	1	2	3	4
$\ln(H^q/P)$	73.426*** (5.97)	44.470*** (2.90)	-10.776*** (-3.05)	-24.052*** (-5.61)
Cohort FE	Yes	Yes	Yes	Yes
Local SD($\ln(H/P)$)	High	Low	High	Low
Unit of observation	Cohort-time	Cohort-time	Cohort-time	Cohort-time
Mean(y)	0.27	0.29	0.04	0.06
SD(y)	0.20	0.17	0.07	0.08
Observations	7,585	7,281	7,771	7,089
Within R^2	0.21	0.09	0.04	0.15

Panel B: Based on Local Dispersion in Filtered $\ln(H/P)$

	$r_{t+1 \rightarrow t+12}^e$		$\Delta h_{t+1 \rightarrow t+12}$	
	1	2	3	4
$\ln(H^q/P)$	67.634*** (5.23)	48.748*** (3.46)	-13.126*** (-4.00)	-22.010*** (-4.98)
Cohort FE	Yes	Yes	Yes	Yes
Local SD(Filterd $\ln(H/P)$)	High	Low	High	Low
Unit of observation	Cohort-time	Cohort-time	Cohort-time	Cohort-time
Mean(y)	0.27	0.29	0.04	0.05
SD(y)	0.19	0.18	0.07	0.07
Observations	8,237	6,629	8,385	6,475
Within R^2	0.19	0.10	0.06	0.13

Table 10. Cross-Sectional Variation of Rent-to-Price Ratios at Different Levels of Geographic Aggregation

This table shows summary statistics on cross-sectional standard deviations of rent-to-price ratios computed at different levels of geographic aggregation. The standard deviations are estimated at postal code, district, and state-quarter level. The rent-to-price ratio is obtained via the baseline matching procedure of flats for rent to flats for sale. The filtered rent-to-price ratio is the residual from a regression of non-log transformed rent-to-price ratios specified as in column 6 of Table 5. Suspect duplicate property listings as well as geographic area-quarters with fewer than 30 observations are removed from the sample. Refer to Table A.1 for variable definitions.

	Obs.	Mean	SD	Min	P25	Median	P75	Max
<i>Postal code-level aggregation</i>								
SD(H/P)	12,622	1.619	0.661	0.223	1.149	1.491	1.942	5.496
SD(Filterd H/P)	12,622	1.383	0.597	0.385	0.963	1.236	1.643	5.199
<i>District-level aggregation</i>								
SD(H/P)	6,063	1.878	0.606	0.481	1.456	1.805	2.213	4.982
SD(Filterd H/P)	6,063	1.360	0.470	0.400	1.017	1.296	1.611	5.262
<i>State-level aggregation</i>								
SD(H/P)	664	2.125	0.485	0.816	1.779	2.078	2.405	4.274
SD(Filterd H/P)	664	1.378	0.336	0.651	1.156	1.336	1.580	3.507

Appendix for “Housing Yields”

Table A.1. Definition of Variables

Variable	Databases	Definition
<i>Flat characteristics</i>		
Rent-to-price ratio (H/P)	RWI-GEO-RED	Ratio of the annual rent of a listed rental flat exclusive of expenses for utilities (H) to the sale price of a matched counterfactual flat for sale (P). Details on the matching procedure are provided in Section II.A.
Annual rent (H)	RWI-GEO-RED	Listed annual rent exclusive of expenses for utilities in EUR. Available only for flats for rent.
Annual inclusive rent	RWI-GEO-RED	Annual rent inclusive of expenses for utilities in EUR. Available only for flats for rent.
Expenses	RWI-GEO-RED	Expenses for utilities in EUR. Available only for flats for rent.
Heating expenses	RWI-GEO-RED	Heating expenses in EUR. Available only for flats for rent.
Heating included	RWI-GEO-RED	Indicator equal to 1 if heating expenses are comprised in the inclusive rent, and 0 otherwise. Available only for flats for sale.
Deposit	RWI-GEO-RED	Deposit in EUR. Available only for flats for rent.
Sale price (P)	RWI-GEO-RED	Listed sale price in EUR. Available only for flats for sale.
Rented	RWI-GEO-RED	Indicator equal to 1 if the flat is rented, and 0 otherwise. Available only for flats for sale.
Housing benefits	RWI-GEO-RED	Housing benefits in EUR. Available only for flats for sale.
Holiday property	RWI-GEO-RED	Indicator equal to 1 if the flat can be used as a holiday property, and 0 otherwise. Available only for flats for sale.
Surface	RWI-GEO-RED	Surface of the flat in sqm.
No. rooms	RWI-GEO-RED	Number of rooms in the flat.
No. bedrooms	RWI-GEO-RED	Number of bedrooms in the flat.
No. bathrooms	RWI-GEO-RED	Number of bathrooms in the flat.
Floor no.	RWI-GEO-RED	Floor on which the flat is located.
No. building floors	RWI-GEO-RED	Number of floors of the building where the flat is located
Usable surface	RWI-GEO-RED	Usable surface of the flat in sqm.
Construction year	RWI-GEO-RED	Year of construction of the property.
Year last modernization	RWI-GEO-RED	Year in which the last modernization of the property took place.
Property conditions	RWI-GEO-RED	Categorical variable (four categories) indicating the conditions of the flat.
Quality of facilities	RWI-GEO-RED	Categorical variable (four categories) indicating the quality of the facilities in the flat.
Protected building	RWI-GEO-RED	Indicator equal to 1 if the flat is located in listed building, and 0 otherwise.
Energy consumption	RWI-GEO-RED	Annual energy consumption in kWh per sqm.
Energy rating	RWI-GEO-RED	Categorical variable capturing the rating of the flat based on the Energy Performance Certificate.
Hot water in energy consumption	RWI-GEO-RED	Indicator equal to 1 if hot water is included in energy consumption, and 0 otherwise.
Balcony	RWI-GEO-RED	Indicator equal to 1 if the flat has a balcony, and 0 otherwise.
Available parking	RWI-GEO-RED	Indicator equal to 1 if the flat comes with a parking place, and 0 otherwise.
Access with wheelchair	RWI-GEO-RED	Indicator equal to 1 if the flat is accessible with a wheelchair, and 0 otherwise.
Guest WC	RWI-GEO-RED	Indicator equal to 1 if the flat has a guest WC, and 0 otherwise.
Garden	RWI-GEO-RED	Indicator equal to 1 if the flat gives access to a garden, and 0 otherwise.
Cellar	RWI-GEO-RED	Indicator equal to 1 if the flat gives access to a cellar, and 0 otherwise.
Kitchen	RWI-GEO-RED	Indicator equal to 1 if a kitchen is already installed in the flat, and 0 otherwise.

(Continued)

Table A.1. – *Continued*

Elevator	RWI-GEO-RED	Indicator equal to 1 if there is an elevator in the building in which the flat is located, and 0 otherwise.
Assisted living	RWI-GEO-RED	Indicator equal to 1 if the flat provides assisted living services, and 0 otherwise.
Immediate availability	RWI-GEO-RED	Indicator equal to 1 if the flat is immediately available, and 0 otherwise.
No. of days online	RWI-GEO-RED	Number of days the flat listing stays online on the platform.
No. of hits	RWI-GEO-RED	Number of hits the flat listing on the online platform.
No. of clicks (contact button)	RWI-GEO-RED	Number of clicks on the “Contact” button of the flat listing.
No. of clicks (customer profile)	RWI-GEO-RED	Number of clicks on the customer profile linked to the flat listing.
No. of clicks (share button)	RWI-GEO-RED	Number of clicks on the “Share” button of the flat listing.
No. of clicks (customer URL)	RWI-GEO-RED	Number of clicks on the customer URL linked to the flat listing.
<i>Housing return and rent growth</i> ²⁸		
r^e	RWI-GEO-RED, FRED	Quarterly logarithmic total return (i.e., reflecting also rental income) in excess of the nationwide 3-month interbank rate for the cohort of properties of the pseudo-panel dataset.
Δh	RWI-GEO-RED	Quarterly logarithmic rent growth rate for the cohort of properties of the pseudo-panel dataset.
H^q/P	RWI-GEO-RED	Ratio of the average quarterly rent per sqm to the average sale price per sqm within the cohort of properties of the pseudo-panel dataset.
<i>District-level characteristics</i>		
Global city	Globalization and World Cities Research Network	Indicator equal to one if a district has a rating between “Alpha” and “Gamma”.

²⁸Details on the construction of the pseudo-panel are provided in the Online Appendix.

Online Appendix

“Housing Yields”

OA.1. Further Literature

This section complements the literature review in Section I. While the main text focuses on our contribution to the asset-pricing literature studying the cross-section of housing valuations, here we discuss how our analysis also relates to other approaches to housing valuation and to additional strands of research.

More broadly, our paper contributes to the literature on the pricing of housing assets (for a recent survey, see [Duca, Muellbauer, and Murphy \(2021\)](#)). Real estate, especially residential real estate, has a dual nature as both a durable consumption good and an investment asset. As a result, the literature has traditionally relied on three main approaches to housing valuation: hedonic pricing models (e.g., [Hill \(2013\)](#)), user-cost models of homeownership (e.g., [Himmelberg, Mayer, and Sinai \(2005\)](#)), and the asset-pricing analogy (e.g., [Case and Shiller \(1989\)](#)). Our analysis mainly follows the latter approach, which is standard in the finance literature and views a property as an asset paying periodic cash flows in the form of rents.

A strand of work studies housing returns, rents, and rent-to-price ratios over time. [Jordà, Knoll, Kuvshinov, Schularick, and Taylor \(2019\)](#), in a study of the aggregate rate of return on assets available in the economy, compare housing as an asset class against other

[†]Corresponding author. Ca' Foscari University of Venice (Department of Economics) and Halle Institute For Economic Research (IWH), Cannaregio 873, Fondamenta San Giobbe, 30121 Venezia, Italy. Phone: +39 041 234 9198. E-mail: stefano.colonnello@unive.it.

[‡]University of Turin (ESOMAS Department) and Collegio Carlo Alberto, Italy. E-mail: roberto.marfe@carloalberto.org.

[§]University of Oxford (Said Business School), United Kingdom. E-mail: qizhou.xiong@said.oxford.edu.

forms of investment over a long time span across countries, and find that its country-level returns are akin to those on equities but exhibit lower volatility. By contrast, using data from UK portfolios of real estate investments between 1901 and 1983, [Chambers, Spaenjers, and Steiner \(2021\)](#) provide evidence of much lower long-run returns after adjusting for costs linked to owning properties. [Eichholtz, Korevaar, Lindenthal, and Tallec \(2021\)](#), relying on historical data from Paris and Amsterdam, highlight the primary role of property-level yields in explaining total housing returns.^{1,2} Using a pseudo-panel constructed from property-level German listings, we characterize residential real estate returns and rent growth for a recent sample at a relatively granular level. We find that as an asset class, residential real estate performed well compared to the stock market.

A number of papers then applies the present-value relationship approach of [Campbell and Shiller \(1988\)](#) to dissect the role of discount and rent growth rate expectations as captured by the rent-to-price ratio in predicting housing returns. Among others, [Campbell, Davis, Gallin, and Martin \(2009\)](#) and [Plazzi, Torous, and Valkanov \(2010\)](#) focus on MSA-level data on residential and commercial properties from the US, respectively. [Engsted and Pedersen \(2015\)](#) extends the analysis to a cross-country setting. Evidence is overall supportive of some degree of predictability. We confirm this finding over our pseudo-panel dataset.

Fluctuations and cross-sectional dispersion in house prices attracted significant attention also outside of the asset pricing literature. A line of research points to credit booms as a key driver of price and rent fluctuations in the US (e.g., [Mian and Sufi \(2009\)](#), [Duca, Muellbauer, and Murphy \(2011\)](#), [Saadi \(2020\)](#), [Reher \(2021\)](#)). This mechanism,

¹[Eichholtz et al. \(2021\)](#) also study housing yields dispersion. Differently from our analysis of the entire housing market in Germany, they focus on time trends in two large cities and find that time-invariant neighborhood fixed effects capture most of the spatial heterogeneity in yields, whereas more detailed demarcation of submarkets or more granular neighborhood-level indexes do not improve the empirical fit.

²Other studies consider long-time series about the housing market. [Eichholtz, Korevaar, and Lindenthal \(2025\)](#) use more than 500 years of data to construct rental price indexes for seven European cities, documenting that they exhibited modest growth. [Amaral, Dohmen, Kohl, and Schularick \(2025\)](#) analyze returns for 27 large cities around the world over 150 years, showing that properties located in those cities are safer than those from more peripheral areas.

however, is to some extent muted for Germany, where the real estate boom over our sample period was not coupled with a credit boom ([Bednarek, Kaat, Ma, and Rebucci \(2021\)](#)). More relevant for the German case are potentially agglomeration economies (e.g., [Combes and Gobillon \(2015\)](#)), which indeed have been shown to be at work also in Germany, both across cities ([Ahlfeldt and Feddersen \(2018\)](#)) and within cities ([Ahlfeldt, Redding, Sturm, and Wolf \(2015\)](#)). Such agglomeration effects impact the cross-sectional dispersion of house prices. [Gyourko, Mayer, and Sinai \(2013\)](#) document how superstar US cities attract high income individuals because of location preferences, crowding out poorer households and triggering housing booms. [Van Nieuwerburgh and Weill \(2010\)](#) theorize and show empirically that price dispersion comes together with increased wage and productivity dispersion. Unequal access to amenities in different locations and the ensuing spatial sorting of households are another potentially important mechanism behind dispersion in housing valuations (e.g., [Couture and Handbury \(2020\)](#), [Couture, Gaubert, Handbury, and Hurst \(2024\)](#)). Leveraging the size and granularity of our dataset, we complement this body of work by examining the impact of within-city agglomeration economies and neighborhood amenities on the heterogeneity of housing yields.

From a methodological perspective, this paper adds to a strand of research aimed at recovering rent-to-price ratios and returns for single properties or for granular geographic areas. A number of studies use a matching approach similar to ours over more limited datasets. [Smith and Smith \(2006\)](#) focus on ten US metropolitan statistical areas (MSAs) in 2005. [Bracke \(2015\)](#) focuses on the London area between 2006 and 2013 and implements an algorithm that matches rental and sale information for the very same property. [Clark and Lomax \(2020\)](#) apply the same methodology to a sample of listings from England in 2014-2015. The approach used in these two studies is highly precise but restricts the analysis to properties that transact both on the sale and the rental market within a short period of time, thus potentially increasing sample bias. [Hill and Syed \(2016\)](#) adopt instead hedonic imputation to obtain property-level rent-to-price ratios for the Sydney area in Australia.

Again for the Sydney area, [Waltl \(2018\)](#) combines the matching and the hedonic imputation approach to develop quality-adjusted rent-to-price ratios over the cross-section. [Melser and Hill \(2019\)](#) show how the same large-scale dataset on residential sales and rents in Sydney can be used to estimate property-level risk and return profiles. [Ahlfeldt, Heblich, and Seidel \(2022\)](#) use the same listing data for Germany as in this paper to develop rental and sale price indexes for arbitrarily defined geographic areas, highlighting some stylized facts about ongoing trends in rent-to-price ratios across such areas and calling for more research on their drivers. Our paper responds to this call by looking into the determinants of dispersion of property-level rent-to-price ratios for a wide sample of matched properties.

OA.2. Further Tests on the Heterogeneity in Housing Yields

To better explore our key stylized fact on heterogeneity in housing yields, we provide additional descriptive evidence, validate our baseline measure of rent-to-price ratios, consider alternative econometric specifications, and assess the role of outliers. Detailed definitions for variables used in the Online Appendix are provided in Table [OA.1](#).

OA.2.1. Visual Evidence

Figure [OA.1](#) shows that rent-to-price ratios increased slightly in the early part of the sample but declined after the mid-2010s. Dispersion follows a similar pattern, though with a lag, suggesting growing heterogeneity in rent-to-price ratios for much of our sample period, followed by a mild decline after 2015. Figure [OA.2](#) confirms the presence of substantial heterogeneity in rent-to-price ratios even in smaller and less global cities than those shown in Panels B–H of Figure [2](#).

OA.2.2. Validation of Housing Yields

We then validate our baseline rent-to-price ratios using properties listed for sale that are already rented. For these properties—which account for roughly one fourth of the raw

sample of sale listings—we observe not only the sale price but also the rental income. This is admittedly a selected subsample, and we therefore do not use it in the main analysis. Nevertheless, it provides a useful benchmark for assessing the quality of our matching procedure.

The correlation between the matched and actual log rent-to-price ratios is 63.77%. Figure OA.3 provides further evidence of this close correspondence. Graph A plots the two measures against each other and supports the validity of the matching procedure. Graphs B and C compare the densities and cumulative distribution functions of the baseline and actual yields, showing that the two measures are similarly distributed. The matched yields are, however, slightly lower than the actual yields for the same properties. One possible interpretation is selection into the rental market: rented properties may be of lower quality along dimensions on which we do not impose exact matching, or along unobserved dimensions [Halket, Nesheim, and Oswald \(2020\)](#), [Chen, Chen, Hill, and Hu \(2022\)](#). If, holding observable traits fixed, matched properties listed for sale are of higher quality and therefore have higher prices, the resulting synthetic yields would be mechanically lower. The magnitude of this bias appears limited: median matched yields are 5.09%, compared with 5.65% for actual yields. Overall, the close alignment between the two distributions is the main takeaway. Moreover, the alternative PSM approach in Table 7 helps mitigate concerns that our key result on unexplained variation in yields may be driven by selection of lower quality properties into the rental market by matching them (imperfectly) to properties on sale on a wide set of covariates.

OA.2.3. Alternative Specifications

In Table OA.2, we augment specification (2) with interaction terms between hedonic matching covariates and key national macroeconomic variables—such as the 10-year Bund yield-to-maturity, the growth of GDP per capita, and the growth of housing stock per capita—to allow for time-varying coefficients and thus better capture changes in expectations

about the housing market. Even after including these terms, the adjusted R^2 remains at 30% as in column 4 of Table 4.

In Table OA.3, to address possible concerns about the use of the log-transformation (e.g., Cohn, Liu, and Wardlaw (2022)), we re-estimate the specifications in column 4 of Table 4 and in column 6 of Table 5.³ The upper bounds for adjusted R^2 even decline, suggesting that the substantial fraction of unexplained variation in housing yields we uncover in the main analysis is not an artifact of the chosen specification.

OA.2.4. *The Role of Outliers*

Next, we expand on the analysis of residuals conducted in Figure 3. First, in Figure OA.4 we look at the role of covariates in generating outliers in residuals. In Graph A, we compare the density of actual rent-to-price ratios against that of fitted values (excluding estimated postal code-quarter fixed effects) from the model in column 6 of Table 5. The latter density is more peaked but, at the same time, exhibits only a modest presence of outliers on the left tail. To understand the origin of these outliers, we re-estimate the model excluding property-specific covariates, namely $\mathbf{m}_{r,\bar{s}}$, $\mathbf{x}_r$, $\mathbf{x}_{\bar{s}}$, and $\mathbf{z}_{r,\bar{s}}$ in equation (2). Graphs B, C, and D show the densities of estimated postal code-quarter fixed effects, fitted values (including estimated fixed effects), and residuals, respectively, for both the baseline specification and the one without flat covariates. It emerges that outliers in all these quantities appear to be linked to the inclusion of covariates.

Next, we identify outliers as those observations for which either the estimated fixed effect or the residual are more than five standard deviations away from their mean. Figure OA.5 confirms that the exclusion of such observations allows us to obtain fixed effects (Graph A) and residuals (Graph B) that are less plagued by outliers.

Finally, Table OA.4 repeats the analysis of Table 5 but excluding observations generating such extreme values. Reassuringly for the robustness of our baseline analysis,

³We also use non-log-transformed explanatory variables here (e.g., surface).

their exclusion has no appreciable effect on the adjusted R^2 .

Overall, outliers do not appear to be a major issue for our baseline specification, both in terms of presence in the empirical distribution of residuals and in terms of impact on our results if we exclude them from the sample.

OA.3. Regional Factors

We obtain national and regional economic and social statistics from the German Federal Statistical Office (*Statistisches Bundesamt* and *Statistische Ämter des Bundes und der Länder*) for the period 2007-2017.⁴

The housing market is often regarded as highly segmented across regions. We therefore reach the lowest administrative level for which a comprehensive set of economic and social indicators are publicly available in Germany, namely the district level. German districts are aggregations of municipalities (*Gemeinde*). These districts are akin to US counties and are categorized as rural districts (*Landkreis*) or urban districts (*Stadtkreis* or *Kreisfreie Stadt*). Table OA.5 reports summary statistics for district-level variables.

To ensure consistency of regional variables, we account for those instances in which districts changed codes over our sample period (e.g., because of statewide reforms such as those of Sachsen-Anhalt in 2007, Sachsen in 2008, and Mecklenburg-Vorpommern in 2011). To merge regional data with housing data, we use the 2015 vintage of district codes provided by the RWI-GEO-RED database for listed properties.

Next, we use these regional data to better pin down the specific economic forces driving variation in rent-to-price ratios.

⁴Regional data are available at annual frequency, at year end. As our main dataset is at quarterly frequency, we assume that regional variables stay constant between the fourth quarter of year y and the fourth quarter of year $y + 1$. More generally, if a regional variable is missing in between two dates, we assume it stays constant until a new non-missing observation is available. Note then that data on local property tax rates are available only up to 2015.

OA.3.1. Explaining Local Fixed Effects

The analysis in Section III.E, by means of the inclusion of geographic area-quarter fixed effects, provides possible upper bounds to the fraction of variation in housing yields that can be explained by local time-varying factors. Here, we make a step forward in trying to pin down actual local factors that matter for housing valuations. To this end, we use the historical record of social and economic indicators at the district level to study their impact on rent-to-price ratio regional heterogeneity by estimating the following regression of rent-to-price ratios on regional factors:

$$(OA.1) \quad \ln \left(\frac{H_{f,t}}{P_{f,t}} \right) = \nu \mathbf{p}_{d,t} + \gamma \mathbf{m}_{r,\bar{s},t} + \eta \mathbf{x}_{r,t} + \theta \mathbf{x}_{\bar{s},t} + \zeta \mathbf{z}_{r,\bar{s}} + \tau_t + \epsilon_{f,t},$$

where d denotes the district of the flat. The variables of interest in this step of the analysis are contained in $\mathbf{p}_d$, which is a vector of district-level covariates. This specification nests the most saturated model of Table 4. In this way, we focus on the component of rent-to-price ratio variation that relates neither to the observable traits of the flat for rent r nor to those of the synthetic counterfactual flat for sale $\bar{s}$, and is therefore (at least partially) attributable to different conditions across regions.⁵

To select the district-level covariates to be included in $\mathbf{p}_d$, we look at the relation between yields and single regional factors, discussing their possible role in shaping expectations about discount rates and rent growth. In particular, we group such factors in two categories: 1) demographic and economic fundamentals, and 2) local housing market characteristics.

⁵The remaining rent-to-price variation may still reflect unobservable differences between the flats in each match. But estimates of ν from equation (OA.1) are less likely to suffer from this problem than those from a simple regression of the matched rent-to-price ratio on $\mathbf{p}_d$ alone.

Demographic and Economic Fundamentals.

In Table [OA.6](#), we verify if housing yields associate with district-level demographic and economic fundamentals via regressions on single regional characteristics specified as in equation ([OA.1](#)). Both standard overlapping generations models and existing evidence on the stock dividend-to-price ratio indicate that demographics, especially the age profile of the population, correlate with the valuation of assets (e.g., [Geanakoplos, Magill, and Quinzii \(2004\)](#), [Favero, Gozluklu, and Tamoni \(2011\)](#), [Poterba, Weil, and Shiller \(1991\)](#)). In column 1, we look at the age profile of the district, which is likely to capture slow-moving long-term expectations about the housing market due to life-cycle portfolio effects ([Favero et al. \(2011\)](#)). The ratio of elderly dependent (above 65 years old) to working age population loads positively and significantly on rent-to-price ratios. This points to positive housing valuation effects of people in the prime of their careers coupled with a depressing effect of the elderly relative to active population ([Takáts \(2012\)](#)). At the same time, in column 2 we show that the district's total population correlates negatively with valuation ratios. This is arguably capturing a mere size effect, which could be driven by agglomeration economies (see [Combes and Gobillon \(2015\)](#), and references therein).

Structural features of the local economy, such as disposable income, unemployment, and industry composition may also correlate with housing valuation ratios, both through expected rent growth and expected discount rates. For instance, income per capita ought to co-move with house prices in the long-run equilibrium (e.g., [Abraham and Hendershott \(1996\)](#)). The coefficient estimates in columns 3 and 4 appear to confirm this intuition: rent-to-price ratios are decreasing in both district-level disposable income per capita and GDP per capita.

A long-standing theoretical and empirical literature has uncovered rich interactions between local housing and labor markets (e.g., [Cameron and Muellbauer \(2001\)](#), [Branch, Petrosky-Nadeau, and Rocheteau \(2016\)](#), [Zabel \(2012\)](#)). In this spirit, in column 5 we

find that districts with more unemployment display significantly higher rent-to-price ratios, consistently with [Vermeulen and Van Ommeren \(2009\)](#), who suggest that cheaper housing may compensate for lower wages in such regions and explain the persistence of across-region heterogeneity in unemployment rates.

In column 6, we observe that a higher importance of the manufacturing sector is associated with higher rent-to-price ratios. A possible story is that regions more reliant on traditional—possibly declining—industries exhibit a more pervasive presence of displaced workers ([Case and Mayer \(1996\)](#)), which is in line with the previous finding on unemployment.⁶ The sheer number of registered business in the district correlates instead negatively with rent-to-price ratios. This correlation, like in the case of total population, is probably a manifestation of a size effect and is consistent with higher demand for dwellings (and by higher-paid workers) in metropolitan areas.

Housing Market Characteristics.

Frictions in the housing market are prevalent and could lead to a wide no arbitrage interval. The wider the no arbitrage interval, the higher the price heterogeneity we will observe in the market, which may then transmit to valuation ratios. In other words, the liquidity of the local housing market is likely to be a non-negligible determinant of rent-to-price ratios (e.g. [Krainer \(2001\)](#)). Besides trading frictions, housing valuations should also hinge on the rigidity of local supply. Indeed, whereas standard finance theory relies on the assumption of stocks being close substitutes to each other, so that their demand curves are horizontal and single stocks' supply does not affect their price ([Shleifer \(1986\)](#)), a real estate property is not just a claim on a firm's cash flows, but offers a service to the (homeowner) investor. The consumption value of residential properties makes their market akin to service markets, which are heavily driven by the demand and supply dynamics, and

⁶The relation between house prices and industry structure can go both ways. For instance, [Adelino, Schoar, and Severino \(2015\)](#) show how increasing house prices favors self-employment by providing capital to new ventures through the collateral channel.

not just by investment motives. Property taxes may also affect valuations as a specific form of friction ([Poterba et al. \(1991\)](#)). For example, in the user cost model, the rent-to-price ratio is a function of the property tax rate (e.g., [Hill and Syed \(2016\)](#)).

In Table [OA.7](#), we look at the relation of several measures capturing local housing market liquidity, supply rigidities, and property taxes. Columns 1 to 3 consider three different measures of liquidity. The first one captures the standardization of the properties listed on the online platform in a given calendar quarter: the district-level interquartile range of flat surface. The second measure is the median number of days a rental listing stays online.⁷ The third measure is the number of flats posted on the online platform in a given district-calendar quarter, which also captures the size of the market. In line with intuition, the first and the third measure correlate negatively with rent-to-price ratios, the second positively. Put differently, housing valuations are increasing in the liquidity of the local housing market, which in turn tends to be higher in growing urban areas.

Moving to supply factors, in columns 4 to 6 we study the role of rigidities in the provision of dwellings, which depend both on zoning laws and the geographic conformation of the area (e.g., [Pogodzinski and Sass \(1991\)](#), [Harari \(2020\)](#)). Real estate valuations correlate negatively with the housing stock per capita, and positively with the price of construction land. Rent-to-price ratios, less intuitively, decline with completed living space per capita in a given year. The rationale for such a correlation may stem from reverse causality: property developers respond to higher demand (and prices) in areas where housing is in short supply by building more dwellings.

Finally, column 7 investigates whether property taxes feed into rent-to-price ratios. In Germany, a property tax is levied on all the land used for residential buildings, whereas the transaction tax is levied on traded properties. We focus on the former, for which each district's government can then apply a multiplier to the federal property assessment rate. Such a rate loads insignificantly.

⁷Note that these housing market liquidity measures are computed from RWI-GEO-RED data, not from regional data.

This analysis highlights the importance of numerous local factors. The signs of most correlations align well with intuition, thus enhancing the credibility of our synthetic rent-to-price ratios. To reduce the dimensionality of the analysis, in $\mathbf{p}_d$ we then include those factors that are statistically significant at the 1% level and exhibit low pairwise correlations among each other: the old-to-working age ratio, disposable income per capita, completed living space per capita, and the number of businesses. These factors heuristically condense information on the demographics, the economic fundamentals, the housing market, and the size of the district.

Based on the adjusted R^2 , in the main analysis we show that district-level time-varying factors can potentially explain 14% ($= 44\% - 30\%$) of yield variation, as documented in Table 5. Column 1 of Table OA.8 shows that directly controlling for the covariates in $\mathbf{p}_d$ —rather than including district-quarter fixed effects—produces an adjusted R^2 of 36%, meaning that we can trace back to specific factors only around 42% ($= 6\%/14\%$) of the district-quarter fixed effect. All coefficient estimates display coefficient signs and magnitudes that are consistent with those discussed above.

In Table OA.8, we then decompose the rent-to-price ratio, investigating separately the rent price per sqm (columns 2-5) and the sale price per sqm (columns 6-9). The specifications in columns 2-4 (6-8) illustrate that time-varying district-level factors can potentially explain up to 19% (24%) of the variation in rent (sale) prices, and that the covariates in $\mathbf{p}_d$ capture 63% (63%) of this incremental explanatory power. Moreover, both the rent and the sale price exhibit correlations with the four district-level factors of opposite sign relative to the rent-to-price ratio in column 1. In other words, the rent-to-price ratio co-moves more strongly with the sale price at the denominator than with the rent price at the numerator—which is consistent with the greater variability of the former. The inclusion of postal code-quarter fixed effects in column 5 (9) raises the adjusted R^2 for the rent (sale) price to 83% (79%), as opposed to the 59% found for the rent-to-price ratio.

Note that in our attempt to re-conduct local heterogeneity in yields to specific

regional-level factors—namely to “explain” postal-quarter fixed effects—we have only considered covariates at the district level, the most granular administrative subdivision at which a comprehensive set of statistics is publicly available (including, e.g., local output measures, industry structure, or new building completion). However, a narrower set of statistics exists also at municipality level and, more interestingly, for finely-defined 1-km raster cells ([Breidenbach and Eilers \(2018\)](#)). We leave the study on the determinants of rent-to-price ratios at this level of geographic granularity to future research.

OA.4. Further Tests on Unexplained Variation

This section examines whether the unexplained variation in housing yields documented in the baseline analysis can be accounted for by costs borne by property owners, regulatory distortions, heterogeneous regional conditions, or a less restrictive matching procedure. The aim is to assess if the residual dispersion can be attributed to such so far omitted sources of variation.

OA.4.1. Costs

For robustness, we adjust our baseline measure of gross yields for costs possibly borne by property owners. Because we do not directly observe such costs, we follow several approaches to proxy for them. In the spirit of [Eichholtz et al. \(2021\)](#), we focus on three sources of costs: nontax costs (including expenses for maintenance), vacancy costs, and taxes. By making different assumptions, we obtain four sets of net yield estimates.

- (a) To compute the first measure of net yields, we assume that nontax costs are at 30% of the annual property-specific listed rent as [Eichholtz et al. \(2021\)](#) do. Similar to [Demers and Eisfeldt \(2022\)](#), the vacancy loss is assumed to be given by the product of the local vacancy rate and the annual property-specific listed rent. The vacancy rate is measured at the district level, based on the 2011 German census, which is then assumed to be constant in the previous and subsequent years. Property taxes are computed

as the product of the implicit recurrent property tax rate estimated for Germany by [Barrios, Denis, Ivaškaitė-Tamošiūnė, Reut, and Vázquez Torres \(2019\)](#) and the listed property-specific sale price.

- (b) Relative to the first measure, the second adjusts both vacancy costs and property taxes. The vacancy costs are computed from a time-varying measure of the local vacancy rate. Starting from the 2011 census estimate of the vacancy rate, we link its variation in the surrounding years to the growth rate of property prices (e.g., an increase in residential prices relative to 2011 is assumed to lead to a lower vacancy rate). Then, the property tax component is adjusted for the district-level property assessment rate B .
- (c) The third measure deviates from the second by assuming that nontax costs are equal to 1.5% of the listed property-specific sale price, as suggested by [Barrios et al. \(2019\)](#).
- (d) The fourth measure deviates from the second by linking the sources of costs to average district-level rent and sale prices, rather than to property-specific ones.

In each case, we impose that property-level net yields cannot be lower than 0. Panel A of Table [OA.9](#) displays summary statistics for the four measures of net yields, which are on average 30% lower than gross ones.

Panel B re-estimates the same specification as in column 6 of Table 5 using the net yield as the dependent variable. No matter how we measure net yields, the adjusted R^2 lines up with the baseline result. Admittedly, the cost specifications (a)-(c) amount to implicitly assuming that costs are perfectly correlated with underlying property prices, muting any additional sources of idiosyncratic risk that could drive costs. If we were to observe costs, it is thus theoretically possible that a “diversification effect” would reduce dispersion in net rental income relative to the gross one. However, [Chambers et al. \(2021\)](#) empirically show that costs drive up volatility of rental income. Interestingly, even indexing costs to district- rather than property-level prices, as we do in case (d), does not produce an increase in the explanatory power of geographic fixed effects.

OA.4.2. Diagnostic Tests for PSM

We report diagnostic tests for PSM procedure used in Table 7. Table OA.10 reports summary statistics for rental and sale flats before and after PSM. Although several differences remain statistically significant at conventional levels after matching, most become economically small, indicating that the procedure substantially improves covariate balance.

Figure OA.6 shows clear differences in the distribution of propensity scores between rental and sale flats, though also pointing to a substantial region of common support between the two groups. This suggests that rental and sale listings differ systematically in observable characteristics, but that PSM can still be implemented on a meaningful set of comparable properties.

OA.4.3. Conditioning on Local Housing Market Frictions

We then examine whether the unexplained dispersion in rent-to-price ratios varies systematically with local proxies for information frictions, supply rigidities, and housing-market regulation. In Table OA.11, we provide suggestive evidence on the role of the housing market frictions discussed in Section IV.C by estimating the specification with postal code-quarter fixed effects on subsamples of the baseline matched dataset. We define such subsamples based on the median district in terms of (i) median online listing time (columns 1 and 2), (ii) housing stock per capita (columns 3 and 4), (iii) living space completed per capita (columns 5 and 6), and (iv) the property assessment rate B (column 7 and 8). Case (i) aims at capturing differences in informational frictions as proxied by the liquidity of the local housing market. Cases (ii), (iii), and (iv) gauge the role of supply rigidities and regulations in the form of zoning laws and property taxation. In each case, we do not uncover any economically significant difference in the explanatory power of our regression specification relative to the baseline in Table 5. This is not surprising because postal code-quarter fixed effects arguably absorb all cross-postal code variation in yields due

to time-varying local housing liquidity, supply rigidities, or regulation, making it effectively a test of within-postal code-quarter differential effects of such frictions. Our results point to their limited quantitative importance.

OA.4.4. Rent Controls

To assess the potential role of local regulation in shaping heterogeneity in housing yields, as discussed in Section IV.C, we focus on rent controls—a prominent form of housing-market regulation that is widespread across Europe and also present in Germany (Kholodilin (2020)). Germany has long had rent controls in place. First-generation rent control laws, which simply set rents at a certain level, were introduced already in 1922; in 1971, policy-makers moved to second-generation laws, which link allowed new rents and rent increases to a reference local rent; a nationwide cap on rent increases of 30% (over 3 years) for existing contracts was passed in 1982 and subsequently reduced to 20% in 1993 (Kholodilin (2017)). During our sample period, the federal government passed two major reforms: on rent updating in 2013 (the *capping limit*) and on rent setting in 2015 (the *rental brake*). These reforms allowed local state governments to introduce more stringent controls for specific municipalities characterized by a “tight” rental market (Kholodilin (2017)). The capping limit restricts rent increases for existing contracts at 15% over 3 years for the designated municipalities. The rental brake constrains new rents in the designated municipalities to be at most 10% higher than a reference rent from the area, except for new or modernized properties.

Also in this case, postal code-quarter fixed effects in Table 5 could absorb any increase in dispersion induced by these reforms as long as they are equally binding across the properties in a given municipality (or postal code). Suppose instead that rent controls bind only for certain properties and, at same time, that market participants expect the controls to be lifted in the future—the new municipality-level rent controls always come with an expiry date and their renewal is subject to varying degrees of uncertainty—, then we could

in principle observe a depressing effect on the current rent (the numerator) coupled with a mild-to-null effect on the price (the denominator), artificially driving down rent-to-price ratios for such properties. Although nationwide rent controls were active in Germany throughout our sample period, they have been relatively loose until the 2010s, whereas cross-sectional heterogeneity in housing yields is substantial even in the early years of our sample (see Figure OA.1), suggesting that the recent reforms are hardly its main driver.

To verify this conjecture, we implement a staggered difference-in-difference design on housing yields dispersion across municipalities around the time at which the new rent controls started to be enforced.⁸ We estimate specifications of the following form:

$$(OA.2) \quad SD(\ln(H/P))_{m,t} = \alpha \times Rental\ brake_{m,t} + \beta \times Capping\ limit_{m,t} + \tau_m + \tau_t + \epsilon_{m,t},$$

where the dependent variable, $SD(\ln(H/P))_{m,t}$ is the standard deviation of rent-to-price ratios in municipality m at quarter t .⁹ To compute the indicators $Rental\ brake_{m,t}$ and $Capping\ limit_{m,t}$, we manually collected from state-level laws the list of municipalities touched by the new rent controls (for an overview, see Kholodilin (2017)). $Rental\ brake_{m,t}$ is an indicator variable equal to 1 in quarters in which a given municipality is subject to the rental brake, and 0 otherwise. Starting from 2015, the rental brake has been introduced at different points in time by 13 out of 16 states. The $Capping\ limit_{m,t}$ indicator is defined in a similar way. Starting from 2013, the capping limit has been adopted in a staggered fashion by 14 states. Because the house listings in our data relate to new contracts, our focus is on the rental brake reform of 2015. Yet, we control for the capping limit on existing contracts introduced in 2013 because of its potential effects on investors' expectations.¹⁰

⁸Mense, Michelsen, and Kholodilin (2019) and Mense, Michelsen, and Kholodilin (2023) use a similar setting to examine the consequences of the rental brake for land values and new rents, respectively. Breidenbach, Eilers, and Fries (2022) perform an evaluation of the reform by means of an event-study approach.

⁹To compute these standard deviations, we remove municipality-quarters with less than 10 matched properties. We then trim the standard deviations at the 99.5% level.

¹⁰We do not study the *rental freeze* introduced in Berlin in 2020 because it is outside our sample period (Hahn, Kholodilin, Walzl, and Fongoni (2024)).

We then account for time-invariant differences across municipalities and for changes in macroeconomic conditions via municipality (τ_m) and time (τ_t) fixed effects. We cluster standard errors at the district level.

In Table OA.12, we report coefficient estimates from equation (OA.2). In column 1, we compute standard deviations of rent-to-price ratios for all properties. In column 2, we focus on properties that are neither new nor recently modernized, i.e., those that are subject to the rental brake in the designated municipalities. In column 3, we look at spillover effects of the rental brake on the dispersion of yields of new or recently modernized properties. To ameliorate overlap between treated and control municipalities, in columns 4 to 6 we re-estimate the same specifications but restricting the control group to non-designated municipalities belonging to districts comprising at least one designated municipality. If anything, the evidence is that the rental brake appears to have reduced—rather than increased—dispersion in housing yields.¹¹

OA.4.5. *Climate Risk and Energy Efficiency*

Motivated by recent evidence on the pricing of climate risk in housing markets (e.g., Giglio, Maggiori, Rao, Stroebel, and Weber (2021)), we examine whether local climate risk exposure helps explain dispersion in housing valuations. In Table OA.13, we incorporate district-level measures of climate risk and assess their association with dispersion in housing valuations. Specifically, we use a set of climate risk indicators based on historical meteorological data over 1989–2018 from Angelova and Lupio (2020), including measures of summer heat, temperature volatility, precipitation level and volatility, and hot–wet stress. We complement these with JRC PESETA estimates of heat-related excess mortality under current climate conditions and alternative warming scenarios (García-León, Masselot, Mistry, Gasparrini, Motta, Feyen, and Ciscar (2024)), which provide an outcome-based

¹¹In unreported analysis, we find that this result is robust to using the CSDID estimator developed by Callaway and Sant’Anna (2021) and Sant’Anna and Zhao (2020), which helps mitigate the inference issues linked to staggered difference-in-differences designs such as ours.

proxy for local climate vulnerability. For both sets of variables, higher values indicate greater relative climate risk exposure.

Panel A of Table [OA.13](#) reports summary statistics about our measures of climate risk. Panel B relates these measures to postal code–quarter-level dispersion in rent-to-price ratios. The evidence is mixed: higher volatility in summer temperatures, as well as higher levels and volatility of precipitation, are associated with lower dispersion, whereas heat-related mortality under current climate conditions is associated with higher dispersion in housing yields. Other proxies for climate risk load insignificantly. Overall, these results suggest that climate risk, as captured by our proxies, is unlikely to be a primary driver of the observed heterogeneity.

As a related channel, we also examine energy efficiency in the spirit of [Lu and Spaenjers \(2025\)](#), who show that it is capitalized into housing prices. To proxy for energy efficiency, we use annual energy consumption in kWh per sqm, available mostly after 2013, which we clean by removing implausibly low values and by adjusting reported consumption depending on whether hot-water consumption is included in the computation, to improve comparability across properties.

Table [OA.14](#) relates postal code-quarter-level dispersion in rent-to-price ratios to dispersion in energy consumption. The results indicate a positive association between dispersion in housing yields and dispersion in energy consumption, although the explanatory power remains limited.

OA.4.6. Ownership Structure

We also consider the role of institutional ownership. This channel is relevant in the German context, where rental housing markets include private households, municipal housing companies, cooperatives, and private housing companies, with large institutional landlords particularly important in some urban areas. Our listings data do not identify seller or owner type at the property level. We therefore rely on municipality-level census information on the

share of dwellings owned by property companies and other institutional investors.

Table OA.15 reports summary statistics and regression results using ownership shares in 2011, as well as changes in these shares between the 2011 and 2022 censuses.¹² We relate these measures to postal code–quarter-level dispersion in rent-to-price ratios. The estimated associations are small and statistically insignificant, both in levels and in changes, and the explanatory power is low. Thus, although institutional ownership and its dynamics may be relevant for housing valuations, we find little evidence that they explain the dispersion in yields in our setting. This may partly reflect the limitations we face in measuring local housing market ownership structure.

OA.5. Housing Returns

The RWI-GEO-RED database is de facto (repeated) cross-sectional in nature. Hence, a relevant drawback of the matching approach to recovering rent-to-price ratios is the impossibility to compute housing returns and conduct time-series tests. To work around this problem, rather than exploiting the panel nature of those (relatively few) listings that appear more than once in the dataset, we resort to a pseudo-panel in the spirit of Deaton (1985) by creating cohorts of similar properties. Whereas our main analysis builds on synthetic rent-to-price ratios, here we implement complementary time-series analyses using such pseudo-panel returns.

OA.5.1. Construction of the Pseudo-Panel

To obtain unbiased estimates from a pseudo-panel analysis, the cohorts must be defined using time-invariant attributes that are observed in all periods for all properties. Therefore, we define cohorts with respect to three traits: location of the property (at the district level), its number of rooms, and its surface.¹³ Despite having a large number of

¹²Municipality-level measures of ownership structure are publicly available only for 2011 and 2022, creating a mismatch with the sample period we cover, 2007–2017.

¹³Whereas the location of a property is unambiguously time-invariant, its number of rooms or surface may admittedly change following a major renovation. In other words, the consistency of our approach rests on

observations, the presence of 402 districts in Germany requires us to be parsimonious in the granularity of rooms-surface combinations to ensure that we have sufficient observations in each cohort to achieve statistically robust asymptotics. We, therefore, form relatively coarse categories of flats. We split them in three groups in terms of number of rooms: studio flats with one or two rooms, medium-sized flat with two and a half or three rooms; and big flats with more than three rooms. Similarly, we discretize the surface of properties in four intervals: small ($\text{sqm} \in (0, 50)$), medium ($\text{sqm} \in [50, 70)$), large ($\text{sqm} \in [70, 90)$), and very large ($\text{sqm} \in [90, +\infty)$). Moreover, again to achieve well-sized cohorts, we construct the pseudo-panel at quarterly (rather than monthly) frequency.

In each district-quarter, we can thus have up to 12 ($= 3 \times 4$) groups of properties based on our rooms-surface categories. However, not all the combinations of the number of rooms and surface are well-populated enough. For instance, it is extremely rare for a small flat to have more than three rooms (0.02% of the sample). By the same token, a large flat is unlikely to have less than three rooms (1.06 % of the sample). Removing rooms-surface combinations with few observations (accounting for less than 2% of total observations) leaves us with at most 8 groups per district-quarter.

However, not only the distribution of observations across rooms-surface combinations is uneven, also the regional distribution is. Sample sizes for large metropolitan areas are large, whereas many of the rural districts do not have a meaningful number of observations even for the most common rooms-surface combinations. Besides the sheer difference in population, another reason is that we focus on flats instead of houses, which naturally tilts the sample towards metropolitan areas. To meet the conditions for Type 1 asymptotics (Verbeek (2008)), we therefore disregard any district-rooms-surface-quarter with fewer than 5 properties for rent or for sale. At the same time, we require each district-rooms-surface cohort to have an average of at least 30 properties for rent and 10 properties for sale over the period for which it is in the dataset.¹⁴ After screening the pseudo-panel according to

the assumption that major renovations are rare enough events.

¹⁴The lower threshold for flats for sale reflects their lower frequency in the sample relative to flats for rent.

such criteria, we end up with 672 cohorts from 175 districts, each of which we observe for up to 44 quarters (corresponding to 5,236,418 micro-level listings).

For each cohort-quarter, we then compute the natural logarithm of the quarterly rent-to-price ratio as

$$(OA.3) \quad \ln \left(\frac{H_{c,t}^q}{P_{c,t}} \right) = \ln \left(\frac{3 \times \bar{H}_{c,t}}{\bar{P}_{c,t}} \right),$$

where c indicates the cohort of properties. $\bar{H}_{c,t}$ and $\bar{P}_{c,t}$ are the average monthly rent and sale price per sqm in a given cohort-quarter respectively. Though the pseudo-panel only allows us to compute rent-to-price ratios at a less granular level than the matching approach, it makes it possible to investigate their evolution through time.

In the same way, for each property cohort we compute its logarithmic total return between quarter $t - 1$ and quarter t :

$$(OA.4) \quad r_{c,t} = \ln \left(\frac{\bar{P}_{c,t} + 3 \times \bar{H}_{c,t}}{\bar{P}_{c,t-1}} \right),$$

which reflects both property price appreciation and rental income (e.g., [Plazzi et al. \(2010\)](#), [Jordà et al. \(2019\)](#)). We then denote the pure price growth component of returns as $r_{c,t}^*$. Analogously, using the same notation as [Plazzi et al. \(2010\)](#), we obtain the quarterly rent growth over the same horizon:

$$(OA.5) \quad \Delta h_{c,t} = \ln \left(\frac{\bar{H}_{c,t}}{\bar{H}_{c,t-1}} \right).$$

Finally, we define the housing premium (i.e., the return in excess of the risk-free rate) as

$$(OA.6) \quad r_{c,t}^e = r_{c,t} - r_t^f,$$

where r_t^f is the 3-month interbank rate for Germany.

OA.5.2. Stylized Facts

By relying on the pseudo-panel dataset, we can now shift our attention to the time-series dimension of house prices while preserving non-trivial cross-sectional heterogeneity, as captured by cohorts of flats formed by district-rooms-surface category.

In Table OA.16, we provide some stylized facts on cohort-level returns, rent growth rates, and rent-to-price ratios within our pseudo-panel. Panel A shows that the average total housing return (r) stands at 1.99% per quarter, with an average price growth rate (r^*) per quarter of 0.79%. Rents grow at an average quarterly rate (Δh) of 0.35%. Total returns exhibit a correlation of 0.99 with price growth, whereas their correlation with rent growth is a mere 0.04. Excess returns (r^e) exceed total returns, because over our sample period risk-free rates are for long periods negative. The mean non-log-transformed rent-to-price ratio (H^q/P) is 1.21%: this is based on quarterly rather than annual rents, so its magnitude is consistent with the analysis above.¹⁵ However, it is worth noting that rent-to-price ratios in the pseudo-panel are substantially less heterogeneous than those from the matching exercise. The aggregation of flats in cohorts (by taking averages) absorbs a hefty fraction of cross-sectional variation. In other words, tests based on our pseudo-panel are most informative about time-series variation. The rent-to-price ratio is persistent with significant first- and fourth-order autocorrelation coefficients of 0.81 and 0.74. Return and rent growth are less persistent but still exhibit significant first-order autocorrelations. In both cases, moreover, such autocorrelations are negative at -0.36 and -0.34, respectively, pointing to the existence of some reversal effect in the pricing of properties.

Following Piazzesi, Schneider, and Tuzel (2007), Panel B studies time-series variation at different levels of geographic aggregation, namely at the cohort-, district-, state-, and national level. In line with Table 2, we find that national variation accounts only for a

¹⁵These results are not far from those of Demers and Eisfeldt (2022), who recover city- and postal code-level returns for single-family rental properties in the US and find that price growth and rental yields constitute about 50% of total returns each. In our case, the split is 39% vs. 61%, respectively. Given that Demers and Eisfeldt (2022) use *net* rental yields while we use *gross* ones, our decomposition of total returns is probably even closer to theirs.

modest fraction of cohort-level variation, especially in the case of returns and yields, whereas this pattern is less pronounced for rent growth. This is even more remarkable because, as pointed out above, the construction of the pseudo-panel per se absorbs a significant part of variation. It is also apparent that sale prices are substantially more volatile over time than rental prices, confirming the stylized fact documented in Section III.

In Figure OA.7, we visualize (cumulative) housing returns. Graph A plots the distribution of quarterly returns and growth by year. We do observe much more pronounced cross-sectional dispersion in r and r^* than in Δh . This may point to a primary role of expected discount rates in generating variation in rent-to-price ratios across properties. In Graph B, we then look at cumulative returns for buy-and-hold strategies of different housing assets against the cumulative return from holding the DAX index, a proxy for the German stock market. The median housing asset delivers a cumulative return of about 90%, in line with the German stock market. Even housing assets in the bottom percentile generate a non-negligible return of about 25% over the sample period. Distinguishing among properties based on rooms-surface combinations, we obtain cumulative returns ranging approximately between 75% and 110% (Graph C). It is worth noting that small housing units tend to consistently outperform large ones.

OA.5.3. Further Tests on Predictability

Finally, we complement the main evidence on predictability with two further tests. First, in Table OA.17 we perform predictive regressions augmented with calendar quarter fixed effects, which absorb macroeconomic fluctuations in the spirit of Fama and French (2025). The inclusion of time fixed effects increases the R^2 across the board, in particular for returns, pointing to the relevance of aggregate shocks. Yet, the rent-to-price ratio retains its predictive ability. Put differently, the part of yield variation driven by local factors significantly contributes to the predictability of returns and rent growth rates. This in turn lines up well with the decomposition in Panel B of Table OA.16, where we highlight how

time-series variation in yields largely originates at a granular level.

Second, Table OA.18 shows that the relation between excess returns and rent growth, and the rent-to-price ratio remains qualitatively unchanged if we directly control for selected local factors. But the within R^2 substantially increases (up to 31% and 16%, respectively), i.e., current local factors contain information useful to predict future return and rent dynamics that is not fully incorporated in the rent-to-price ratio, even though the micro-level evidence of Section III unambiguously shows that these quantities co-move.

OA.6. Persistence in Dispersion of Housing Yields

This section provides additional evidence on the persistence and spatial structure of local dispersion in rent-to-price ratios. The analysis complements Section VI.A by studying whether unexplained heterogeneity in housing yields is persistent over time and spatially clustered across nearby postal codes.

We first examine the time-series persistence of local yield dispersion. In Table OA.19, we estimate postal code-quarter-level panel regressions of the standard deviation of filtered rent-to-price ratios on its one-period lag. We report specifications both without and with postal code fixed effects. The former provides an upper bound for the autoregressive coefficient, while the latter provides a lower bound (Bond (2002)). Across specifications, the estimated AR(1) coefficient remains below 0.5, pointing to a low-to-moderate degree of persistence in local yield dispersion. This suggests that dispersion in rent-to-price ratios is not a highly persistent postal code characteristic, but instead varies substantially across quarters, possibly reflecting the composition of buyers, sellers, tenants, and properties entering the market at each point in time. This evidence is consistent with Table 2, which shows that the time-series standard deviation of rent-to-price ratios at the postal-code level is relatively small. Taken together, these results suggest that local dispersion in housing yields is stationary and only weakly persistent.

We then study whether local yield dispersion is spatially correlated across nearby

areas. For each postal code, we compute the cross-sectional standard deviation of (filtered) rent-to-price ratios and calculate pairwise geographic distances between all postal code centroids in Germany using the haversine formula. We then group postal code pairs by distance bands, ranging from very close pairs to pairs more than 150 km apart, and compute the correlation in yield dispersion within each distance band. Figure [OA.8](#) shows that the correlation declines with distance. Postal codes located very close to one another display a positive correlation in rent-to-price ratio dispersion, whereas the correlation approaches zero for distant postal code pairs.

This spatial pattern indicates that (unexplained) yield dispersion is not randomly distributed across space, but tends to cluster across nearby areas. At the same time, the decline in correlation with distance suggests that this spatial dependence is highly localized. These results complement our baseline evidence of substantial within-postal code-quarter heterogeneity: while the main analysis documents dispersion within narrowly defined local markets, the evidence here shows that areas with high or low dispersion also tend to be geographically clustered across neighboring areas.

References

- Abraham, J. M., and P. H. Hendershott. “Bubbles in Metropolitan Housing Markets.” *Journal of Housing Research*, 7 (1996), 191–207.
- Adelino, M.; A. Schoar; and F. Severino. “House Prices, Collateral, and Self-Employment.” *Journal of Financial Economics*, 117 (2015), 288–306.
- Ahlfeldt, G. M., and A. Feddersen. “From Periphery to Core: Measuring Agglomeration Effects Using High-Speed Rail.” *Journal of Economic Geography*, 18 (2018), 355–390.
- Ahlfeldt, G. M.; S. Heblich; and T. Seidel. “Micro-Geographic Property Price and Rent Indices.” *Regional Science and Urban Economics* (2022), 103836.
- Ahlfeldt, G. M.; S. J. Redding; D. M. Sturm; and N. Wolf. “The Economics of Density: Evidence from the Berlin Wall.” *Econometrica*, 83 (2015), 2127–2189.
- Amaral, F.; M. Dohmen; S. Kohl; and M. Schularick. “Superstar Returns? Spatial Heterogeneity in Returns to Housing.” *Journal of Finance*, 80 (2025), 3057–3094.
- Angelova, D., and N. B. Lupio. “Constructing a Meteorological Indicator Dataset for Selected European NUTS 3 Regions.” *Data in Brief*, 31 (2020), 105786.
- Barrios, S.; C. Denis; V. Ivaškaitė-Tamošiūnė; A. Reut; and E. Vázquez Torres. “Housing Taxation: a New Database for Europe.” Working Paper (2019).
- Bednarek, P.; D. M. t. Kaat; C. Ma; and A. Rebucci. “Capital Flows, Real Estate, and Local Cycles: Evidence from German Cities, Banks, and Firms.” *Review of Financial Studies*, 34 (2021), 5077–5134.
- Bond, S. R. “Dynamic Panel Data Models: a Guide to Micro Data Methods and Practice.” *Portuguese Economic Journal*, 1 (2002), 141–162.
- Bracke, P. “House Prices and Rents: Microevidence from a Matched Data Set in Central London.” *Real Estate Economics*, 43 (2015), 403–431.

- Branch, W. A.; N. Petrosky-Nadeau; and G. Rocheteau. “Financial Frictions, the Housing Market, and Unemployment.” *Journal of Economic Theory*, 164 (2016), 101–135.
- Breidenbach, P., and L. Eilers. “RWI-GEO-GRID: Socio-Economic Data on Grid Level.” *Jahrbücher für Nationalökonomie und Statistik*, 238 (2018), 609–616.
- Breidenbach, P.; L. Eilers; and J. Fries. “Temporal Dynamics of Rent regulations—The Case of the German Rent Control.” *Regional Science and Urban Economics*, 92 (2022), 103737.
- Callaway, B., and P. H. Sant’Anna. “Difference-in-Differences with Multiple Time Periods.” *Journal of Econometrics*, 225 (2021), 200–230.
- Cameron, G., and J. Muellbauer. “Earnings, Unemployment, and Housing in Britain.” *Journal of Applied Econometrics*, 16 (2001), 203–220.
- Campbell, J. Y., and R. J. Shiller. “The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors.” *Review of Financial Studies*, 1 (1988), 195–228.
- Campbell, S. D.; M. A. Davis; J. Gallin; and R. F. Martin. “What Moves Housing Markets: a Variance Decomposition of the Rent-Price Ratio.” *Journal of Urban Economics*, 66 (2009), 90–102.
- Case, K. E., and C. J. Mayer. “Housing Price Dynamics Within a Metropolitan Area.” *Regional Science and Urban Economics*, 26 (1996), 387–407.
- Case, K. E., and R. J. Shiller. “The Efficiency of the Market for Single-Family Homes.” *American Economic Review* (1989), 125–137.
- Chambers, D.; C. Spaenjers; and E. Steiner. “The Rate of Return on Real Estate: Long-Run Micro-Level Evidence.” *Review of Financial Studies*, 34 (2021), 3572–3607.
- Chen, J.; Y. Chen; R. J. Hill; and P. Hu. “The User Cost of Housing and the Price-Rent Ratio in Shanghai.” *Regional Science and Urban Economics*, 92 (2022), 103738.
- Clark, S., and N. Lomax. “Rent/price Ratio for English Housing Sub-Markets Using Matched Sales and Rental Data.” *Area*, 52 (2020), 136–147.

- Cohn, J. B.; Z. Liu; and M. I. Wardlaw. “Count (and Count-Like) Data in Finance.” *Journal of Financial Economics*, 146 (2022), 529–551.
- Combes, P.-P., and L. Gobillon. “The Empirics of Agglomeration Economies.” In *Handbook of Regional and Urban Economics*, Vol. 5. Elsevier (2015), 247–348.
- Couture, V., and J. Handbury. “Urban Revival in America.” *Journal of Urban Economics*, 119 (2020), 103267.
- Deaton, A. “Panel Data from Time Series of Cross-Sections.” *Journal of Econometrics*, 30 (1985), 109–126.
- Demers, A., and A. L. Einfeldt. “Total Returns to Single-Family Rentals.” *Real Estate Economics*, 50 (2022), 7–32.
- Duca, J. V.; J. Muellbauer; and A. Murphy. “House Prices and Credit Constraints: Making Sense of the US Experience.” *Economic Journal*, 121 (2011), 533–551.
- Duca, J. V.; J. Muellbauer; and A. Murphy. “What Drives House Price Cycles? International Experience and Policy Issues.” *Journal of Economic Literature*, 59 (2021), 773–864.
- Eichholtz, P.; M. Korevaar; and T. Lindenthal. “The Housing Affordability Revolution.” Working Paper (2025).
- Eichholtz, P.; M. Korevaar; T. Lindenthal; and R. Tallec. “The Total Return and Risk to Residential Real Estate.” *Review of Financial Studies*, 34 (2021), 3608–3646.
- Engsted, T., and T. Q. Pedersen. “Predicting Returns and Rent Growth in the Housing Market Using the Rent-Price Ratio: Evidence from the OECD Countries.” *Journal of International Money and Finance*, 53 (2015), 257–275.
- Fama, E. F., and K. R. French. “House Prices and Rents.” *Review of Financial Studies*, 38 (2025), 547–563.
- Favero, C. A.; A. E. Gozluklu; and A. Tamoni. “Demographic Trends, the Dividend-Price Ratio, and the Predictability of Long-Run Stock Market Returns.” *Journal of Financial*

and Quantitative Analysis, 46 (2011), 1493–1520.

García-León, D.; P. Masselot; M. N. Mistry; A. Gasparrini; C. Motta; L. Feyen; and J.-C. Ciscar. “Temperature-Related Mortality Burden and Projected Change in 1368 European Regions: a Modelling Study.” *The Lancet Public Health*, 9 (2024), e644–e653.

Geanakoplos, J.; M. Magill; and M. Quinzii. “Demography and the Long-Run Predictability of the Stock Market.” *Brookings Papers on Economic Activity*, 2004 (2004), 241–307.

Giglio, S.; M. Maggiori; K. Rao; J. Stroebel; and A. Weber. “Climate Change and Long-Run Discount Rates: Evidence from Real Estate.” *Review of Financial Studies*, 34 (2021), 3527–3571.

Gyourko, J.; C. Mayer; and T. Sinai. “Superstar Cities.” *American Economic Journal: Economic Policy*, 5 (2013), 167–99.

Hahn, A. M.; K. A. Kholodilin; S. R. Walzl; and M. Fongoni. “Forward to the Past: Short-Term Effects of the Rent Freeze in Berlin.” *Management Science*, 70 (2024), 1901–1923.

Halket, J.; L. Nesheim; and F. Oswald. “The Housing Stock, Housing Prices, and User Costs: the Roles of Location, Structure, and Unobserved Quality.” *International Economic Review*, 61 (2020), 1777–1814.

Harari, M. “Cities in Bad Shape: Urban Geometry in India.” *American Economic Review*, 110 (2020), 2377–2421.

Hill, R. J. “Hedonic Price Indexes for Residential Housing: a Survey, Evaluation and Taxonomy.” *Journal of Economic Surveys*, 27 (2013), 879–914.

Hill, R. J., and I. A. Syed. “Hedonic Price-Rent Ratios, User Cost, and Departures from Equilibrium in the Housing Market.” *Regional Science and Urban Economics*, 56 (2016), 60–72.

Himmelberg, C.; C. Mayer; and T. Sinai. “Assessing High House Prices: Bubbles,

- Fundamentals and Misperceptions.” *Journal of Economic Perspectives*, 19 (2005), 67–92.
- Jordà, Ò.; K. Knoll; D. Kuvshinov; M. Schularick; and A. M. Taylor. “The Rate of Return on Everything, 1870–2015.” *Quarterly Journal of Economics*, 134 (2019), 1225–1298.
- Kholodilin, K. “Long-Term, Multicountry Perspective on Rental Market Regulations.” *Housing Policy Debate*, 30 (2020), 994–1015.
- Kholodilin, K. A. “Quantifying a Century of State Intervention in Rental Housing in Germany.” *Urban Research & Practice*, 10 (2017), 267–328.
- Krainer, J. “A Theory of Liquidity in Residential Real Estate Markets.” *Journal of Urban Economics*, 49 (2001), 32–53.
- Lu, X., and C. Spaenjers. “Energy Labels, Misreporting, and Heuristic Hedonic Models.” *Real Estate Economics*, 53 (2025), 1200–1222.
- Melser, D., and R. J. Hill. “Residential Real Estate, Risk, Return and Diversification: Some Empirical Evidence.” *Journal of Real Estate Finance and Economics*, 59 (2019), 111–146.
- Mense, A.; C. Michelsen; and K. A. Kholodilin. “The Effects of Second-Generation Rent Control on Land Values.” *AEA Papers and Proceedings*, 109 (2019), 385–88.
- Mense, A.; C. Michelsen; and K. A. Kholodilin. “Rent Control, Market Segmentation, and Misallocation: Causal Evidence from a Large-Scale Policy Intervention.” *Journal of Urban Economics*, 134 (2023), 103513.
- Mian, A., and A. Sufi. “The Consequences of Mortgage Credit Expansion: Evidence from the US Mortgage Default Crisis.” *Quarterly Journal of Economics*, 124 (2009), 1449–1496.
- Piazzesi, M.; M. Schneider; and S. Tuzel. “Housing, Consumption and Asset Pricing.” *Journal of Financial Economics*, 83 (2007), 531–569.
- Plazzi, A.; W. Torous; and R. Valkanov. “Expected Returns and Expected Growth in Rents of Commercial Real Estate.” *Review of Financial Studies*, 23 (2010), 3469–3519.
- Pogodzinski, J. M., and T. R. Sass. “Measuring the Effects of Municipal Zoning Regulations:

- a Survey.” *Urban Studies*, 28 (1991), 597–621.
- Poterba, J. M.; D. N. Weil; and R. Shiller. “House Price Dynamics: the Role of Tax Policy and Demography.” *Brookings Papers on Economic Activity*, 1991 (1991), 143–203.
- Reher, M. “Finance and the Supply of Housing Quality.” *Journal of Financial Economics*, 142 (2021), 357–376.
- Saadi, V. “Role of the Community Reinvestment Act in Mortgage Supply and the US Housing Boom.” *Review of Financial Studies*, 33 (2020), 5288–5332.
- Sant’Anna, P. H., and J. Zhao. “Doubly Robust Difference-in-Differences Estimators.” *Journal of Econometrics*, 219 (2020), 101–122.
- Shleifer, A. “Do Demand Curves for Stocks Slope Down?” *Journal of Finance*, 41 (1986), 579–590.
- Smith, M. H., and G. Smith. “Bubble, Bubble, Where’s the Housing Bubble?” *Brookings Papers on Economic Activity*, 2006 (2006), 1–67.
- Takáts, E. “Aging and House Prices.” *Journal of Housing Economics*, 21 (2012), 131–141.
- Van Nieuwerburgh, S., and P. O. Weill. “Why Has House Price Dispersion Gone Up?” *Review of Economic Studies*, 77 (2010), 1567–1606.
- Verbeek, M. “Pseudo-Panels and Repeated Cross-Sections.” In *The Econometrics of Panel Data*. Springer (2008), 369–383.
- Vermeulen, W., and J. Van Ommeren. “Compensation of Regional Unemployment in Housing Markets.” *Economica*, 76 (2009), 71–88.
- Waltl, S. R. “Estimating Quantile-Specific Rental Yields for Residential Housing in Sydney.” *Regional Science and Urban Economics*, 68 (2018), 204–225.
- Zabel, J. E. “Migration, Housing Market, and Labor Market Responses to Employment Shocks.” *Journal of Urban Economics*, 72 (2012), 267–284.

OA.7. Figures and Tables

Figure OA.1. Level and Dispersion of Rent-to-Price Ratios Over Time

This figure shows the evolution of the level and dispersion of rent-to-price ratios (obtained via the baseline matching) between 2007 and 2017. Graph A reports the mean and the standard deviation dynamics (blue and red line, respectively). Graph B reports the median and the interquartile range trends (blue and red line, respectively).

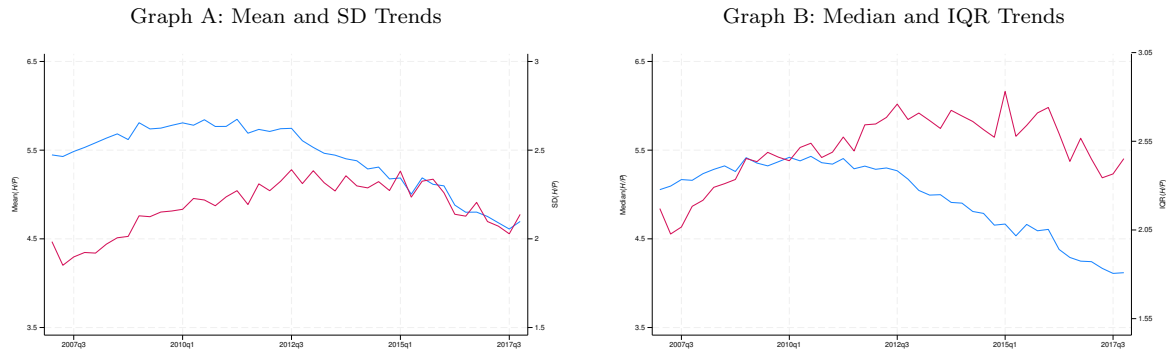


Figure OA.2. Median Rent-to-Price Ratio at the Postal Code Level Within Other Major Cities

This figure visualizes the median rent-to-price ratio (obtained via the baseline matching) at the five-digit postal code level across selected German major cities, pooling all periods between 2007 and 2017. Postal codes with fewer than 10 observations are excluded (white colored). Reported cities are those assigned a “sufficiency” rating in the 2020 ranking by the Globalization and World Cities Research Network. Each of the panels from A to H corresponds to a different city.

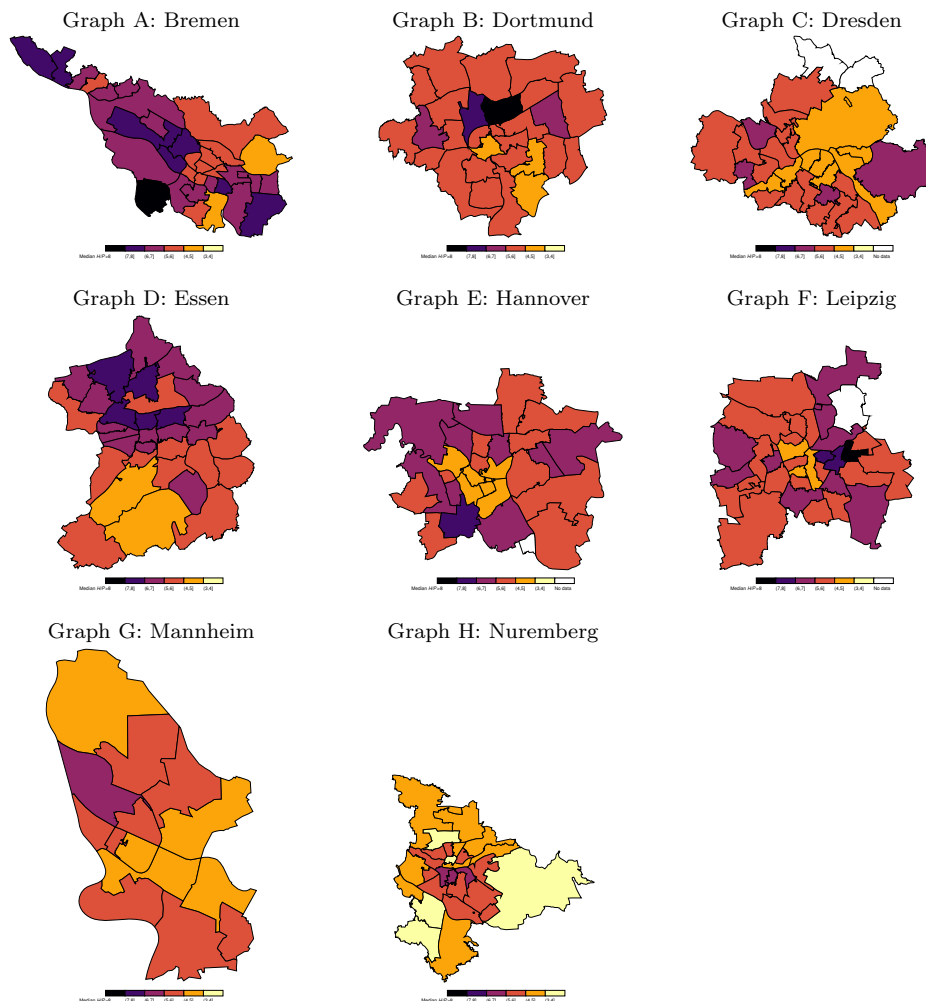


Figure OA.3. Validation of Rent-to-Price Ratios Obtained via the Baseline Matching

This figure compares matched rent-to-price ratios against their actual counterparts, for properties for which both are available. Actual rent-to-price ratios refer to properties on sale for which a rental income is reported. Graph A shows a scatterplot of the two rent-to-price ratio measures (all in natural logarithm). A line fitted with a linear regression is also plotted (together with 95% confidence bands). Graph B shows empirical densities for matched and actual rent-to-price ratios, whereas Graph C visualizes their empirical cumulative distribution functions (CDFs). In all graphs, both matched and actual rent-to-price ratios exclude the top 1% of observations to favor readability of the graphs.

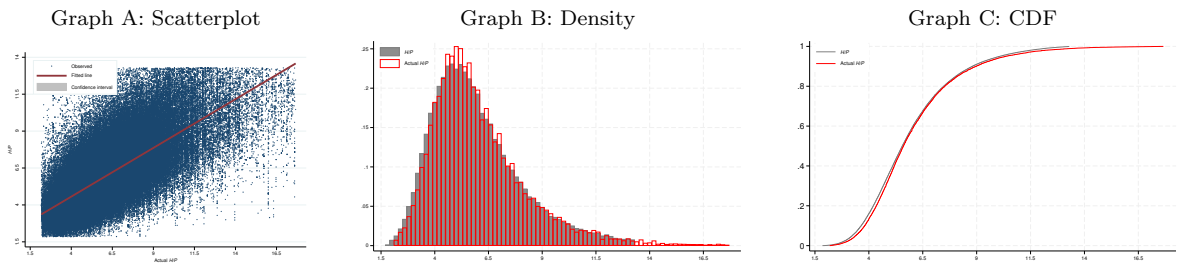


Figure OA.4. The Role of Outliers

This figure shows densities for quantities obtained from the baseline specification in column 6 of Table 5 against those from a similar specification not including property-specific control variables. Graph A compares observed log-transformed rent-to-price ratios against fitted values (excluding estimated postal code-quarter fixed effects) from the baseline specification. Graphs B, C, and D compare, respectively, estimated fixed effects, fitted values (including estimated fixed effects), and residuals from the two specifications.

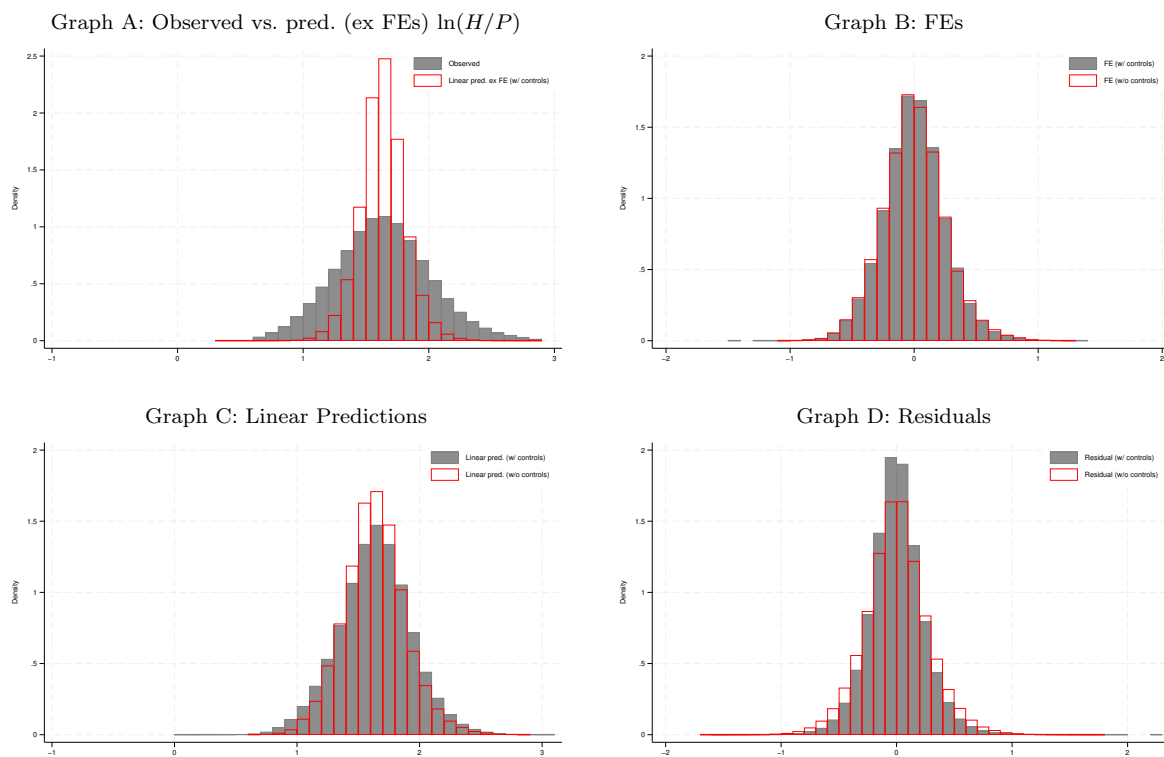


Figure OA.5. Restricting the Regression Sample

This figure shows densities for quantities obtained from the baseline specification in column 6 of Table 5 against those from the same specification estimated over a sample that excludes observations linked to outliers. Such observations are identified as those for which either the estimated fixed effect or the residual are more than five standard deviations away from their mean. Graphs A and B compare, respectively, estimated fixed effects and residuals from the two specifications.

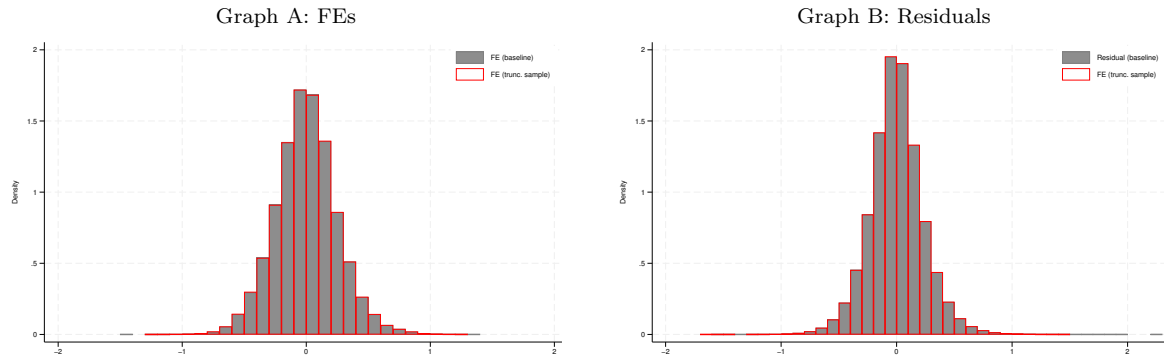


Figure OA.6. Propensity Scores

This figure shows densities for the propensity scores for properties for sale and for rent used to implement the PSM algorithm in column 2 of Table 7.

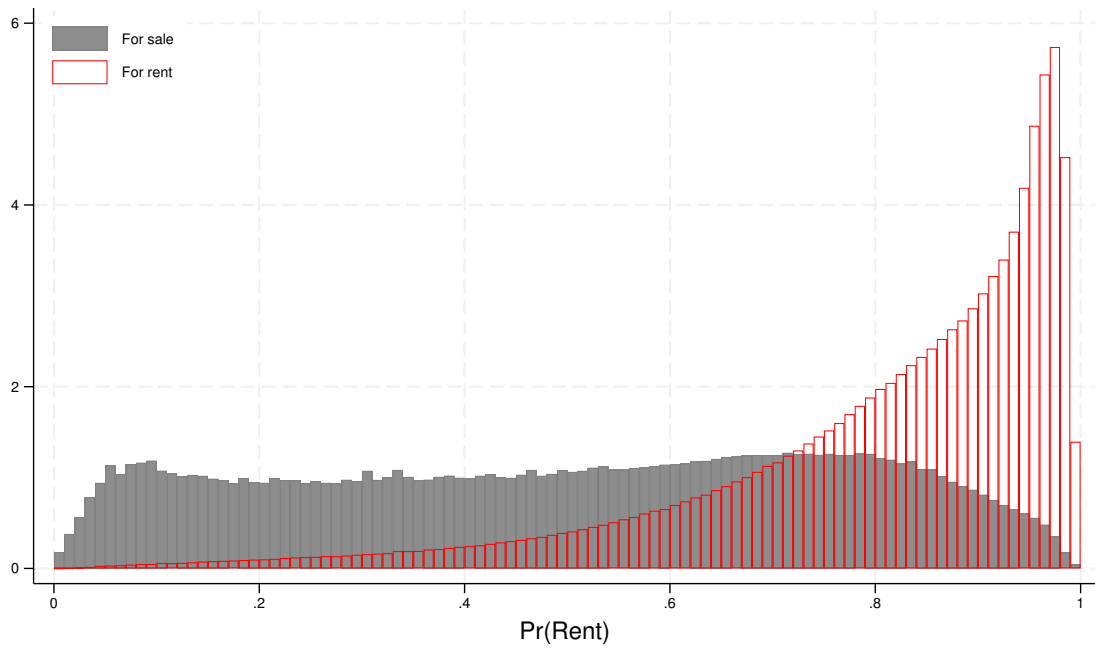
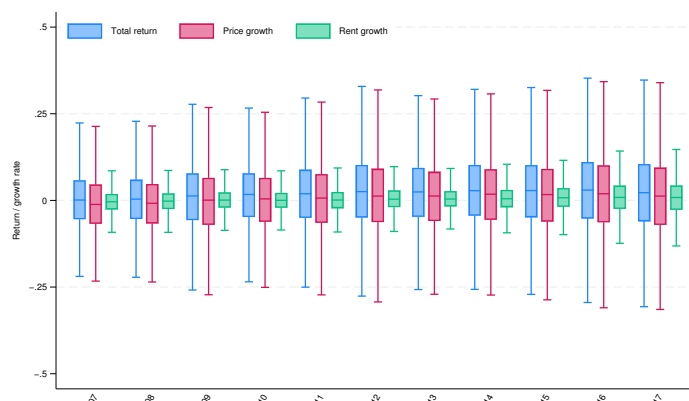


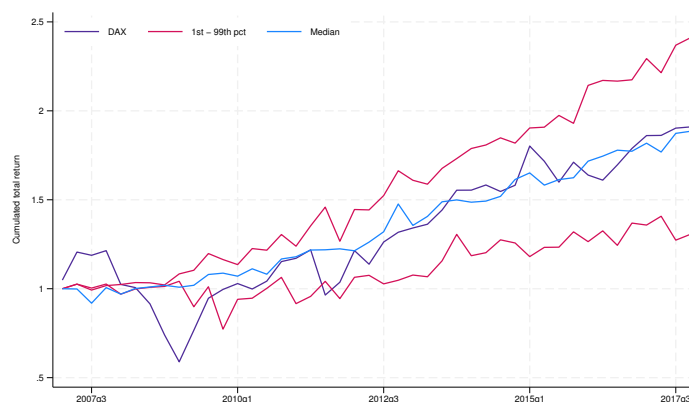
Figure OA.7. Housing Returns

This figure shows the distribution of housing returns obtained via the pseudo-panel approach for the period 2007-2017. Graph A reports the empirical distribution of quarterly total returns (r), price growth rates (r^*), and rent growth rates (Δh) by means of year-by-year box plots. Graph B illustrates the evolution of the median as well as the 1st and 99th percentile of cumulative returns (focusing on cohorts with consecutive non-missing observations over the entire sample period). These time-series are plotted together with the cumulative return on the DAX stock index. Graph C illustrates the evolution of the median of cumulative returns conditional on the rooms-surface category of properties. In each case, the quantiles of cumulative returns are computed as of 2017Q3.

Graph A: Empirical Distribution by Year



Graph B: Cumulative Returns



Graph C: Cumulative Returns by Flat Type

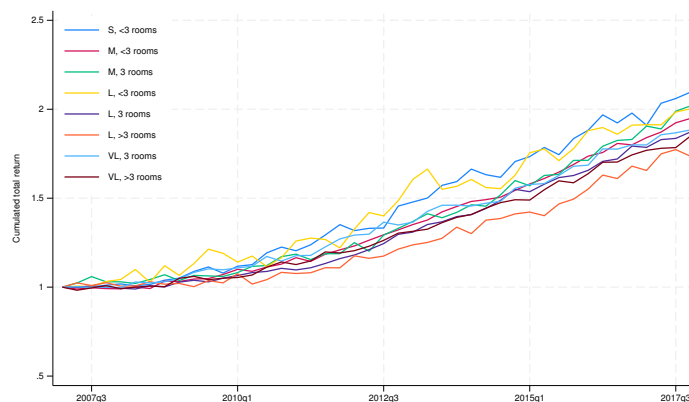


Figure OA.8. Spatial Correlation in Local Rent-to-Price Ratio Dispersion

This figure shows average correlation in postal code-level (filtered) rent-to-price ratio standard deviation across postal codes as a function of their distance. The filtered rent-to-price ratio is the residual from a regression of non-log-transformed rent-to-price ratios specified as in column 6 of Table 5. Suspect duplicate property listings as well as postal codes with fewer than 30 observations are removed from the sample.

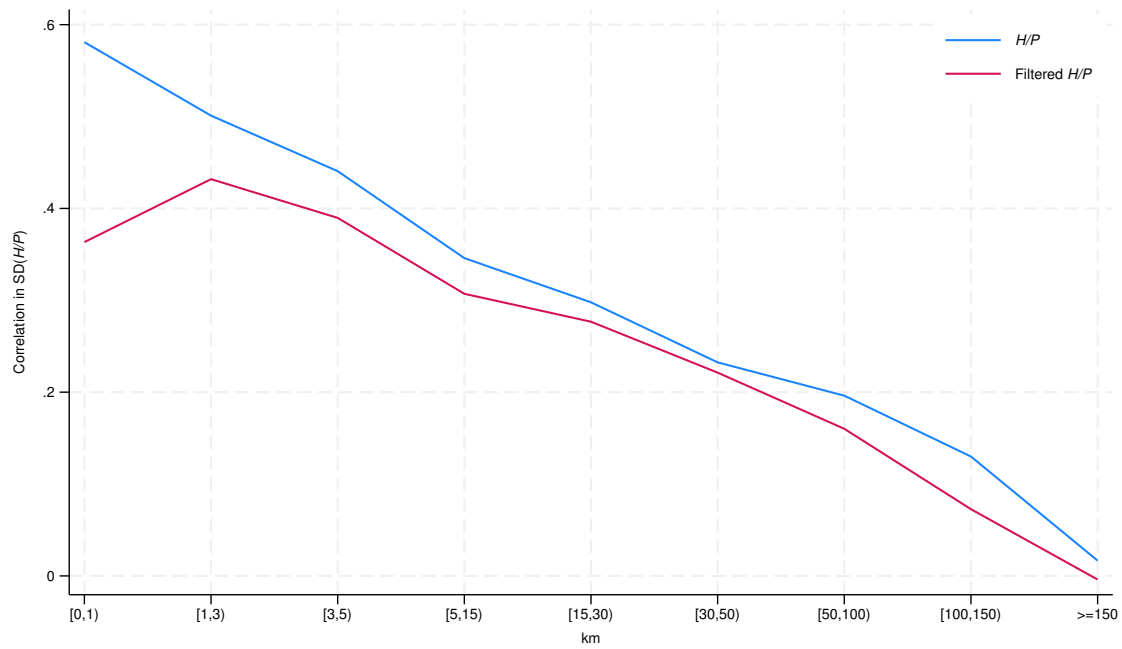


Table OA.1. Definition of Further Variables Used in the Online Appendix

Variable	Databases		Definition
<i>National-level variables</i>			
10 YTM (Bund)	FRED		Annualized yield to maturity on the 10-year Bund.
GDP per capita growth	FRED		Quarterly annualized logarithmic growth rate of GDP per capita.
Housing stock per capita growth	Federal Office	Statistical	Year-on-year logarithmic growth rate of housing stock per capita. Note that nationwide data on housing stock are available at annual frequency, at year end, so it is assumed that it stays constant between the fourth quarter of year y and the fourth quarter of year $y + 1$.
<i>District-level variables</i>			
Old-to-working age ratio	Federal Office	Statistical	Ratio of population older than 65 years to working age population (20-65 years) in a given district-year.
Population	Federal Office	Statistical	Total population in a given district-year.
Disposable income per capita	Federal Office	Statistical	Disposable income per capita in thousand EUR in a given district-year.
GDP per capita	Federal Office	Statistical	GDP per capita in thousand EUR in a given district-year.
Unemployment rate	Federal Office	Statistical	Unemployment rate in a given district-year.
Manufacturing industry share	Federal Office	Statistical	Ratio of the number of manufacturing firms to the number of registered businesses across all industries in a given district-year.
No. businesses	Federal Office	Statistical	Number of registered businesses across all industries in a given district-year.
IQR of surface	RWI-GEO-RED		Interquartile range of the surface of the flats listed on the online platform in a given district-calendar quarter.
Median online listing time	RWI-GEO-RED		Median number of days online of the rental listings on the platform in a given district-calendar quarter.
No. posted flats	RWI-GEO-RED		Number of flats listed on the online platform in a given district-calendar quarter.
Housing stock per capita	Federal Office	Statistical	Residential housing stock per capita in sqm in a given district-year.
Living space completed per capita	Federal Office	Statistical	Living space completed per capita in sqm in a given district-year.
Land price	Federal Office	Statistical	Average ready-for-building land price per sqm in a given district-year.
Property assessment rate B	Federal Office	Statistical	Property tax multiplier (type B) in a given district-year.
Summer max. temperature (mean)	Historical meteorological data, 1989–2018		Average daily maximum temperature during summer months, based on the climate-risk indicators constructed by Angelova and Lupio (2020) .
Summer max. temperature (volatility)	Historical meteorological data, 1989–2018		Volatility of daily maximum temperature during summer months, based on the climate-risk indicators constructed by Angelova and Lupio (2020) .
Daily precipitation (mean)	Historical meteorological data, 1989–2018		Average daily precipitation, based on the climate-risk indicators constructed by Angelova and Lupio (2020) .
Daily precipitation (volatility)	Historical meteorological data, 1989–2018		Volatility of daily precipitation, based on the climate-risk indicators constructed by Angelova and Lupio (2020) .

(Continued)

Table OA.1. – Continued

Summer temperature-precipitation corr.	Historical meteorological data, 1989–2018		Correlation between daily maximum temperature and precipitation during summer months, based on the climate-risk indicators constructed by Angelova and Lupio (2020) .
Excess heat-related mortality (present)	JRC PESETA		Heat-related excess mortality under current climate conditions, based on JRC PESETA estimates used by García-León et al. (2024) .
Excess heat-related mortality (+4°)	JRC PESETA		Projected heat-related excess mortality under a +4°C warming scenario, based on JRC PESETA estimates used by García-León et al. (2024) .
<i>Municipality-level variables</i>			
Rental brake	State-level laws		Indicator equal to one in quarters in which a municipality is subject to the rental brake, which constrains new rents in designated municipalities to be at most 10% above the local reference rent, except for new or modernized properties. The list of treated municipalities is manually collected from state-level laws, following the institutional overview in Kholodilin (2017) .
Capping limit	State-level laws		Indicator equal to one in quarters in which a municipality is subject to the capping limit, which restricts rent increases for existing contracts in designated municipalities to 15% over three years. The list of treated municipalities is manually collected from state-level laws, following the institutional overview in Kholodilin (2017) .
Property companies (% , 2011)	German 2011	census	Share of dwellings in a municipality owned by property companies in the 2011 census.
Other institutions (% , 2011)	German 2011	census	Share of dwellings in a municipality owned by other institutional investors in the 2011 census.
Δ Property companies (% , 2022-2011)	German 2011 and 2022	census	Change between the 2022 and 2011 censuses in the municipal share of dwellings owned by property companies.
Δ Other institutions (% , 2022-2011)	German 2011 and 2022	census	Change between the 2022 and 2011 censuses in the municipal share of dwellings owned by other institutional investors.
<i>Housing returns</i>			
r	RWI-GEO-RED		Quarterly logarithmic total return (i.e., reflecting also rental income) for the cohort of properties of the pseudo-panel dataset.
r^*	RWI-GEO-RED		Quarterly logarithmic ex-rent return for the cohort of properties of the pseudo-panel dataset.

Table OA.2. Rent-to-Price Ratio, Property Characteristics, and Macroeconomic Conditions

This table re-estimates the specification from column 4 of Table 4, augmenting it with interactions with measures of macroeconomic conditions in Germany. Column 1 interacts each of the matching covariates (surface (squared), no. of rooms, no. of bedrooms, no. of bathrooms, floor number) with the yield-to-maturity on the 10-year Bund. Column 2 adds similar interactions with GDP per capita growth in Germany. Column 3 adds similar interactions with housing stock per capita growth in Germany. The t -statistics (in parentheses) are based on standard errors clustered by district. Significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively. Refer to Tables A.1 and OA.1 for variable definitions.

	$\ln(H/P)$		
	1	2	3
Missing indicators	Yes	Yes	Yes
Covariate distances	Yes	Yes	Yes
Matching covariates $\times$ 10Y YTM (Bund)	Yes	Yes	Yes
Matching covariates $\times$ GDP per capita growth	No	Yes	Yes
Matching covariates $\times$ Housing stock per capita growth	No	No	Yes
Time FE	Yes	Yes	Yes
Unit of observation	Match. flat	Match. flat	Match. flat
Mean(y)	1.63	1.63	1.63
SD(y)	0.38	0.38	0.38
Observations	1,625,131	1,625,131	1,625,131
Adjusted R^2	0.30	0.30	0.30

Table OA.3. Re-Estimating the Main Specifications with Non-Transformed Rent-to-Price Ratios

This table re-estimates the specifications from column 4 of Table 4 and column 6 of Table 5, using non-log-transformed property-level rent-to-price ratio as the dependent variable as well non-log-transformed flat-level covariates. Estimated coefficients are multiplied by 100 to favor readability. The t -statistics (in parentheses) are based on standard errors clustered by district. Significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively. Refer to Table A.1 for variable definitions.

	H/P	
	1	2
Surface	-7.101*** (-9.57)	-9.436*** (-27.21)
Surface ²	0.026*** (8.88)	0.038*** (28.95)
No. rooms	55.966*** (11.17)	12.547*** (4.16)
No. bedrooms	-0.937 (-0.27)	12.989*** (6.51)
No. bathrooms	-3.710 (-0.67)	-6.992 (-1.42)
Floor no.	5.092*** (2.75)	4.979*** (4.28)
Expenses	-0.138*** (-3.06)	-0.031 (-1.29)
Heating expenses	-0.235*** (-4.69)	-0.037 (-1.33)
Heating included	5.929 (1.00)	10.905*** (7.45)
Deposit	-0.049*** (-4.95)	0.056*** (10.09)
Community charge	0.074** (2.32)	0.122*** (4.46)
Holiday property	-25.088*** (-4.68)	-15.503*** (-4.54)
Rented	34.043*** (5.23)	31.137*** (5.93)
Missing indicators	Yes	Yes
Covariate distances	Yes	Yes
Time FE	Yes	No
Postal code×Time FE	No	Yes
Unit of observation	Match. flat	Match. flat
Postal code×Time FE	No	Yes
Mean(y)	5.51	5.51
SD(y)	2.22	2.22
Observations	1,625,131	1,606,359
Adjusted R^2	0.24	0.54
Adjusted within R^2	0.22	0.22

Table OA.4. Local Variation of Rent-to-Price Ratios with Sample Restriction

This table reports coefficients of determination from regressions for property-level rent-to-price ratios on different fixed effect structures for a German sample between 2007 and 2017. The log-transformed rent-to-price ratio is obtained via the baseline matching procedure of flats for rent to flats for sale. Column 1 reports the same estimates in column 6 of Table 5. Column 2 re-estimates the same specification over a samples excluding observations that generate outliers in estimated postal code-quarter fixed effects or residuals. Specifically, the sample excludes observations that are more than five standard deviations from the mean of the estimated fixed effect or of the residual. Standard errors are clustered by district. Refer to Table A.1 for variable definitions.

	$\ln(H/P)$	
	1	2
Flat covariates	Yes	Yes
Missing indicators	Yes	Yes
Covariate distances	Yes	Yes
Postal code×Time FE	Yes	Yes
Unit of observation	Match. flat	Match. flat
Tail restricted	None	5 SD
Mean(y)	1.63	1.63
SD(y)	0.38	0.38
Observations	1,606,359	1,605,639
Adjusted R^2	0.59	0.59
Adjusted within R^2	0.29	0.30

Table OA.5. Summary Statistics at the District Level

This table shows summary statistics on socioeconomic and housing market conditions at the district level between 2007 and 2017. Refer to Table OA.1 for variable definitions.

	Obs.	Mean	SD	Min	Median	Max
<i>Demographic and economic conditions</i>						
Old-to-working age ratio (%)	3,660	35.009	4.520	22.704	34.715	52.192
Population (thousand)	3,660	220.045	249.950	33.944	164.963	3,613.495
Disposable income PC (EUR, thousand)	3,629	18.692	2.207	14.050	18.648	28.394
GDP PC (EUR, thousand)	3,629	30.038	11.855	14.134	26.926	88.045
Unemployment rate (%)	3,660	6.672	3.240	1.200	6.000	21.200
Manufacturing industry share (%)	3,648	7.363	2.418	2.518	7.132	21.360
No. businesses	3,660	10,415.504	13,411.856	1,537.000	7,575.500	189,177.000
<i>Housing market conditions</i>						
IQR of surface (sqm)	3,660	31.197	7.417	2.590	30.750	92.000
Median online listing time	3,660	22.792	13.115	1.000	20.000	160.000
Living space completed PC (sqm)	3,635	0.266	0.138	0.045	0.246	0.747
No. posted flats	3,660	424.770	1,323.217	4.000	153.000	29,504.000
Housing stock PC (sqm)	3,631	44.533	4.059	35.457	44.452	57.357
Land price (sqm)	3,603	140.726	117.271	6.961	105.196	831.342
Property assessment rate B (basis points)	3,660	393.466	82.574	244.000	374.000	855.000

Table OA.6. Rent-to-Price Ratio and Local Demographic and Economic Factors

This table reports estimates from regressions for property-level rent-to-price ratios on selected district-level measures of demographic and economic conditions for a German sample between 2007 and 2017. The log-transformed rent-to-price ratio is obtained via the baseline matching procedure of flats for rent to flats for sale. Each column augments the specification of column 4 of Table 4 with one district-level explanatory variable. Estimated coefficients are multiplied by 100 to favor readability. The t -statistics (in parentheses) are based on standard errors clustered by district. Significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively. Refer to Tables A.1 and OA.1 for variable definitions.

	$\ln(H/P)$						
	1	2	3	4	5	6	7
Old-to-working age ratio	1.935*** (7.60)						
$\ln(\text{Population})$		-2.693*** (-2.94)					
Disposable income PC			-1.949*** (-3.15)				
GDP PC				-0.309*** (-4.22)			
Unemployment rate					1.273** (2.42)		
Manufacturing industry share						1.714*** (4.33)	
$\ln(\text{No. businesses})$							-3.871*** (-3.57)
Flat covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Missing indicators	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Covariate distances	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Unit of observation	Match. flat	Match. flat	Match. flat	Match. flat	Match. flat	Match. flat	Match. flat
Mean(y)	1.63	1.63	1.63	1.64	1.63	1.63	1.63
SD(y)	0.38	0.38	0.38	0.38	0.38	0.38	0.38
Observations	1,625,131	1,625,131	1,614,691	1,611,986	1,624,788	1,607,106	1,625,113
Adjusted R^2	0.34	0.31	0.32	0.32	0.31	0.31	0.31

Table OA.7. Rent-to-Price Ratio and Local Housing Market Conditions

This table reports estimates from regressions for property-level rent-to-price ratios on selected district-level measures of housing market conditions for a German sample between 2007 and 2017. The log-transformed rent-to-price ratio is obtained via the baseline matching procedure of flats for rent to flats for sale. Each column augments the specification of column 4 of Table 4 with one district-level explanatory variable. Estimated coefficients are multiplied by 100 to favor readability. The t -statistics (in parentheses) are based on standard errors clustered by district. Significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively. Refer to Tables A.1 and OA.1 for variable definitions.

	$\ln(H/P)$						
	1	2	3	4	5	6	7
IQR of surface	-0.860*** (-4.96)						
Median online listing time		0.767*** (6.33)					
$\ln(\text{No. posted flats})$			-2.026*** (-3.90)				
Housing stock PC				0.782** (2.44)			
Living space completed PC					-48.033*** (-5.26)		
Land price						-0.044*** (-9.14)	
Property assessment rate (B)							-0.002 (-0.37)
Flat covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Missing indicators	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Covariate distances	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Unit of observation	Match. flat	Match. flat	Match. flat	Match. flat	Match. flat	Match. flat	Match. flat
Mean(y)	1.63	1.63	1.63	1.63	1.64	1.65	1.63
SD(y)	0.38	0.38	0.38	0.38	0.38	0.38	0.38
Observations	1,625,131	1,625,131	1,625,131	1,614,746	1,610,357	1,491,880	1,624,365
Adjusted R^2	0.32	0.32	0.31	0.31	0.32	0.33	0.30

Table OA.8. Decomposition of the Rent-to-Price Ratio and Local Conditions

This table reports estimates from regressions for property-level rent-to-price ratios (and the components thereof) on selected district-level characteristics and fixed effects for a German sample between 2007 and 2017. Column 1 uses as dependent variable the log-transformed rent-to-price ratio is obtained via the baseline matching procedure of flats for rent to flats for sale. Columns 2-5 (6-9) use as dependent variable the log-transformed annual rental (sale) price per sqm listed for the matched flat for rent (for sale). Estimated coefficients are multiplied by 100 to favor readability. The t -statistics (in parentheses) are based on standard errors clustered by district. Significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively. Refer to Tables A.1 and OA.1 for variable definitions.

	$\ln(H/P)$	$\ln(H)$				$\ln(P)$			
	1	2	3	4	5	6	7	8	9
Old-to-working age ratio	1.557*** (7.69)		-1.772*** (-7.82)				-3.359*** (-10.06)		
Disposable income PC	-1.733*** (-5.69)		3.025*** (10.30)				4.845*** (9.56)		
Living space completed PC	-29.881*** (-5.02)		21.568*** (3.39)				50.982*** (5.03)		
$\ln(\text{No. businesses})$	-3.072*** (-3.62)		2.820*** (4.62)				5.915*** (5.48)		
Flat covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Missing indicators	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Covariate distances	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	No	No	Yes	Yes	No	No
District $\times$ Time FE	No	No	No	Yes	No	No	No	Yes	No
Postal code $\times$ Time FE	No	No	No	No	Yes	No	No	No	Yes
Unit of observation	Match. flat	Match. flat	Match. flat	Match. flat	Match. flat	Match. flat	Match. flat	Match. flat	Match. flat
Mean(y)	1.64	4.41	4.41	4.41	4.41	7.38	7.38	7.38	7.39
SD(y)	0.38	0.31	0.31	0.31	0.31	0.52	0.51	0.52	0.52
Observations	1,599,899	1,625,131	1,599,899	1,624,137	1,606,359	1,625,131	1,599,899	1,624,137	1,606,359
Adjusted R^2	0.36	0.56	0.68	0.75	0.83	0.41	0.56	0.65	0.79
Adjusted within R^2	0.34	0.55	0.68	0.31	0.29	0.40	0.55	0.21	0.19

Table OA.9. The Role of Costs in the Variation of Rent-to-Price Ratios

This table reports tests on net rent-to-price ratios for a German sample between 2007 and 2017. The log-transformed gross rent-to-price ratio is obtained via the baseline matching procedure of flats for rent to flats for sale. The net rent-to-price ratio is then computed following the four different cost specifications ((a)-(d)) described in the Online Appendix. Panel A shows summary statistics on gross and net rent-to-price ratios. Panel B reports coefficients of determination from regressions for rent-to-price ratios. Each column re-estimates the specification of column 6 of Table 5 using a different measure of the net rent-to-price ratio as the dependent variable. Standard errors are clustered by district. Refer to Table A.1 for variable definitions.

Panel A: Summary Statistics

	Obs.	Mean	SD	Min	P25	Median	P75	Max
H/P	1,625,131	5.509	2.218	1.847	3.979	5.094	6.528	17.573
Net H/P (a)	1,621,566	3.501	1.454	0.735	2.498	3.235	4.183	16.132
Net H/P (b)	1,616,375	3.472	1.453	0.669	2.470	3.209	4.153	16.117
Net H/P (c)	1,616,375	3.627	2.130	0.000	2.159	3.235	4.617	21.997
Net H/P (d)	1,616,375	3.455	1.546	0.000	2.398	3.171	4.166	18.447

Panel B: Regression Analysis

	$\ln(\text{Net } H/P)$			
	1	2	3	4
Flat covariates	Yes	Yes	Yes	Yes
Missing indicators	Yes	Yes	Yes	Yes
Covariate distances	Yes	Yes	Yes	Yes
Postal code $\times$ Time FE	Yes	Yes	Yes	Yes
Unit of observation	Match. flat	Match. flat	Match. flat	Match. flat
Measure of costs	(a)	(b)	(c)	(d)
Mean(y)	1.18	1.17	1.11	1.11
SD(y)	0.39	0.40	0.64	0.64
Observations	1,603,196	1,598,772	1,598,196	1,598,196
Adjusted R^2	0.56	0.56	0.55	0.55
Adjusted within R^2	0.29	0.29	0.29	0.29

Table OA.10. Summary Statistics on Flats Before and After PSM

This table shows summary statistics on property characteristics distinguishing between flats for rent and matched flats for sale, before (Panel A) and after PSM (Panel B) between 2007 and 2017. The t -statistics for mean-comparison tests are based on standard errors clustered by district. Refer to Table A.1 for variable definitions.

Panel A: Flat Characteristics Before PSM

	For rent			For sale			Mean-comparison test	
	Obs.	Mean	SD	Obs.	Mean	SD	Diff.	t -stat
<i>Rental listing</i>								
Annual rent (EUR)	5,100,753	6,748.543	3,778.684					
Annual inclusive rent (EUR)	3,168,689	8,556.069	4,144.971					
Expenses (EUR)	4,813,021	140.172	62.704					
Heating expenses (EUR)	1,172,014	77.165	32.204					
Heating included	4,471,388	0.683	0.465					
Deposit (EUR)	4,440,116	1,419.935	853.491					
<i>Sale listing</i>								
Sale price (EUR)				1,982,461	222,945.074	186,164.069		
Community charge (EUR)				988,087	222.827	92.910		
Rented				1,141,339	0.242	0.428		
Annual rent income (EUR)				246,002	6,570.241	4,604.669		
Holiday property				1,243,362	0.132	0.338		
<i>Matching covariates</i>								
Surface (sqm)	5,100,753	74.893	26.731	1,982,461	87.145	34.687	12.252	5.654
No. rooms	5,100,753	2.665	0.894	1,982,461	2.983	1.010	0.318	9.754
No. bedrooms	5,100,753	1.604	0.734	1,982,461	1.848	0.856	0.244	9.772
No. bathrooms	5,100,753	1.069	0.253	1,982,461	1.182	0.401	0.113	5.780
Floor no.	5,100,753	1.918	1.375	1,982,461	2.053	1.593	0.135	3.231

Panel A: Flat Characteristics Before PSM (continued)

	For rent			For sale			Mean-comparison test	
	Obs.	Mean	SD	Obs.	Mean	SD	Diff.	t-stat
<i>Other characteristics</i>								
No. building floors	3,928,914	3.460	1.390	1,557,308	3.775	1.580	0.315	3.221
Usable surface (sqm)	1,922,739	81.081	30.196	765,901	96.219	36.710	15.138	7.976
Construction year	3,200,502	1,966.745	41.249	1,750,550	1,977.446	38.844	10.701	13.220
Year last modernization	1,474,034	2,009.523	5.607	490,645	2,008.281	6.504	-1.242	-4.793
Property conditions	4,285,838	2.689	0.691	1,723,268	2.902	0.904	0.213	2.544
Quality of facilities	2,920,808	2.479	0.608	1,160,718	2.710	0.656	0.231	3.880
Energy consumption (kWh/(sqm× y))	1,185,051	123.023	49.415	598,024	109.832	49.822	-13.191	-5.136
Energy rating	104,088	4.429	1.984	77,529	3.554	1.789	-0.875	-16.105
Warm water in energy exp.	2,880,925	0.149	0.356	991,109	0.203	0.402	0.054	3.403
Balcony	4,457,683	0.775	0.417	1,844,451	0.907	0.291	0.132	22.849
Available parking	479,364	0.940	0.237	325,428	0.970	0.171	0.030	5.090
Accessible with wheelchair	56,532	0.840	0.367	56,010	0.945	0.228	0.105	5.094
Guest WC	3,653,437	0.240	0.427	1,392,020	0.426	0.494	0.185	10.105
Garden	3,360,448	0.334	0.472	1,231,472	0.413	0.492	0.078	5.961
Cellar	4,128,981	0.778	0.416	1,567,621	0.838	0.369	0.060	4.934
Kitchen	3,848,070	0.547	0.498	1,423,383	0.578	0.494	0.031	1.682
Elevator	3,503,453	0.286	0.452	1,429,294	0.577	0.494	0.291	12.656
Assisted living	1,680,997	0.058	0.235	516,779	0.157	0.364	0.099	7.653
Immediate availability	4,802,079	0.324	0.468	1,385,736	0.282	0.450	-0.042	-1.380
No. of days online	4,831,426	31.104	39.725	1,890,412	45.553	57.232	14.449	3.694
No. of hits	5,016,478	920.315	1,028.774	1,948,593	690.368	950.423	-229.948	-2.819
No. of clicks (contact button)	3,236,726	15.685	22.452	949,749	8.796	13.659	-6.889	-9.261
No. of clicks (customer profile)	800,805	6.360	6.628	275,960	4.163	4.977	-2.196	-17.875
No. of clicks (share button)	1,891,663	2.152	1.700	484,304	1.975	1.581	-0.176	-4.029
No. of clicks (customer URL)	2,454,790	6.747	7.491	946,495	5.984	7.210	-0.763	-2.107

Panel B: Flat Characteristics After PSM

	For rent			For sale (matched)			Mean-comparison test	
	Obs.	Mean	SD	Obs.	Mean	SD	Diff.	t-stat
<i>Rental listing</i>								
Annual rent (EUR)	3,683,096	6,643.953	3,630.195					
Annual inclusive rent (EUR)	3,683,096	8,046.362	3,202.142					
Expenses (EUR)	3,683,096	133.526	57.642					
Heating expenses (EUR)	3,683,096	66.578	15.142					
Heating included	3,683,096	0.717	0.451					
Deposit (EUR)	3,683,096	1,394.455	784.051					
<i>Sale listing</i>								
Sale price (EUR)				3,683,096	148,627.503	126,277.009		
Community charge (EUR)				3,683,096	191.119	60.176		
Rented				3,683,096	0.122	0.327		
Annual rent income (EUR)				3,683,096	5,104.733	1,518.651		
Holiday property				3,683,096	0.053	0.224		
<i>Matching covariates</i>								
Surface (sqm)	3,683,096	75.467	27.109	3,683,096	77.791	29.837	2.325	3.740
No. rooms	3,683,096	2.654	0.895	3,683,096	2.740	0.927	0.085	4.890
No. bedrooms	3,683,096	1.606	0.734	3,683,096	1.638	0.756	0.033	2.100
No. bathrooms	3,683,096	1.077	0.267	3,683,096	1.089	0.290	0.012	3.309
Floor no.	3,683,096	1.983	1.410	3,683,096	2.011	1.457	0.029	1.148

Panel B: Flat Characteristics After PSM (continued)

	For rent			For sale (matched)			Mean-comparison test	
	Obs.	Mean	SD	Obs.	Mean	SD	Diff.	t-stat
<i>Other characteristics</i>								
No. building floors	2,282,814	3.581	1.365	2,282,814	3.603	1.451	0.022	0.916
Usable surface (sqm)	594,434	82.212	29.775	594,434	84.167	31.163	1.955	4.329
Construction year	2,028,651	1,968.605	38.632	2,028,651	1,967.278	39.874	-1.327	-2.222
Year last modernization	412,112	2,009.871	5.487	412,112	2,007.972	6.412	-1.898	-16.559
Property conditions	2,641,776	2.711	0.700	2,641,776	2.593	0.788	-0.118	-17.754
Quality of facilities	1,347,222	2.534	0.620	1,347,222	2.482	0.639	-0.052	-4.231
Energy consumption (kWh/(sqm× y))	424,874	122.514	51.084	424,874	124.875	48.401	2.361	3.619
Energy rating	7,495	3.702	1.959	7,495	4.029	1.925	0.326	7.760
Warm water in energy exp.	1,615,062	0.137	0.344	1,615,062	0.162	0.369	0.025	5.251
Balcony	2,958,359	0.787	0.410	2,958,359	0.800	0.400	0.013	2.275
Available parking	207,335	0.957	0.203	207,335	0.961	0.194	0.004	0.792
Accessible with wheelchair	4,789	0.901	0.299	4,789	0.921	0.270	0.020	2.880
Guest WC	2,376,802	0.240	0.427	2,376,802	0.249	0.432	0.009	1.572
Garden	1,854,364	0.268	0.443	1,854,364	0.295	0.456	0.027	4.575
Cellar	2,759,921	0.794	0.405	2,759,921	0.804	0.397	0.010	1.419
Kitchen	2,311,722	0.538	0.499	2,311,722	0.538	0.499	0.000	0.013
Elevator	2,111,793	0.295	0.456	2,111,793	0.279	0.448	-0.016	-1.943
Assisted living	750,796	0.050	0.218	750,796	0.058	0.233	0.008	5.063
Immediate availability	3,270,688	0.314	0.464	3,270,688	0.379	0.485	0.065	4.745
No. of days online	3,338,882	28.627	34.593	3,338,882	53.737	86.871	25.109	24.871
No. of hits	3,552,574	950.744	1,053.453	3,552,574	1,062.554	1,327.038	111.811	2.203
No. of clicks (contact button)	1,872,688	18.371	26.797	1,872,688	13.648	17.435	-4.723	-8.566
No. of clicks (customer profile)	288,387	7.293	7.834	288,387	4.239	4.720	-3.054	-10.547
No. of clicks (share button)	736,639	2.368	1.966	736,639	2.354	2.058	-0.014	-0.258
No. of clicks (customer URL)	937,158	6.857	7.536	937,158	7.989	9.627	1.132	3.657

Table OA.11. Unexplained Variation of Rent-to-Price Ratios and Local Housing Market Characteristics

This table reports coefficients of determination from regressions for property-level rent-to-price ratios for a German sample between 2007 and 2017, exploring the role of local housing market characteristics. The dependent variable is the log-transformed rent-to-price ratio, as obtained via the baseline matching procedure of flats for rent to flats for sale. Each column estimates specification of column 6 of Table 5 over different subsamples defined with respect to local housing market characteristics. The considered housing market characteristics (indicated below) comprise the median online listing time (i), housing stock per capita (ii), living space completed per capita (iii), and the property assessment rate B (iv). Odd (even) columns restrict the sample to districts that are above (below) the median of the considered housing market characteristic. Standard errors are clustered by district. Refer to Tables A.1 and OA.1 for variable definitions.

	$\ln(H/P)$							
	1	2	3	4	5	6	7	8
	High	Low	High	Low	High	Low	High	Low
Flat covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Missing indicators	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Covariate distances	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Postal code×Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Unit of observation	Match. flat	Match. flat	Match. flat	Match. flat	Match. flat	Match. flat	Match. flat	Match. flat
Conditioning variable	(i)	(i)	(ii)	(ii)	(iii)	(iii)	(iv)	(iv)
Mean(y)	1.72	1.60	1.61	1.64	1.55	1.69	1.64	1.61
SD(y)	0.37	0.37	0.38	0.38	0.35	0.38	0.39	0.33
Observations	424,791	1,181,329	248,767	1,346,937	618,522	972,808	1,264,971	340,202
Adjusted R^2	0.54	0.60	0.62	0.58	0.58	0.58	0.59	0.58
Adjusted within R^2	0.27	0.31	0.34	0.29	0.31	0.29	0.29	0.34

Table OA.12. Dispersion of Rent-to-Price Ratios and Rent Controls

This table reports estimates from regressions for the dispersion in rent-to-price ratios on measures of stringency of local rent controls for a German sample between 2007 and 2017, focusing on the impact of the so-called rental brake of 2015. The dependent variable is municipality-quarter-level standard deviation of log-transformed rent-to-price ratios obtained via the baseline matching procedure of flats for rent to flats for sale. Municipality-quarters with fewer than 30 observations are removed from the sample on which the standard deviation of rent-to-price ratios are computed. Columns 1 and 4 consider all properties. Column 2 and 5 consider only non-new properties. Columns 3 and 6 consider only new properties. Columns 1 to 3 use as control group all municipality-quarters not subject to the rental brake. Columns 4 to 6 use as control group municipality-quarters not subject to the rental brake but within districts comprising at least one municipality-quarter subject to the rental brake. All specifications include municipality and calendar quarter fixed effects. Estimated coefficients are multiplied by 100 to favor readability. The t -statistics (in parentheses) are based on standard errors clustered by district. Significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively. Refer to Tables A.1 and OA.1 for variable definitions.

	SD(ln(H/P))					
	1	2	3	4	5	6
Rental brake	-0.577 (-0.95)	-0.420 (-0.63)	-2.487** (-2.02)	-1.028 (-1.63)	-0.812 (-1.16)	-0.886 (-0.61)
Capping limit	-0.472 (-0.85)	-0.465 (-0.82)	-0.488 (-0.54)	-0.928 (-1.58)	-1.009* (-1.74)	0.168 (0.16)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes
Unit of observation	Municip.-time	Municip.-time	Municip.-time	Municip.-time	Municip.-time	Municip.-time
Properties	All	Non-new	New	All	Non-new	New
Control group	All	All	All	Treated distr.	Treated distr.	Treated distr.
Mean(y)	0.28	0.28	0.30	0.27	0.27	0.29
SD(y)	0.10	0.10	0.10	0.09	0.09	0.09
Observations	13,010	12,112	1,772	6,122	5,730	1,167
Adjusted R^2	0.35	0.35	0.38	0.42	0.41	0.40

Table OA.13. Dispersion in Local Rent-to-Price Ratios and Climate Risk

This table reports tests on local dispersion in rent-to-price ratios and climate risk for a German sample between 2007 and 2017. Panel A shows summary statistics on district-level measures of climate risk from [Angelova and Lupio \(2020\)](#) and [García-León et al. \(2024\)](#). Panel B reports estimates from panel regressions of local dispersion in rent-to-price ratios on such measures. The dependent variable is the postal code-quarter-level standard deviation of log-transformed rent-to-price ratios obtained via the baseline matching procedure of flats for rent to flats for sale. The explanatory variables are the time-invariant climate risk measures of Panel A. Postal code-quarters with fewer than 30 observations are removed from the sample on which the standard deviations of rent-to-price ratios are computed. All specifications include calendar quarter fixed effects. Estimated coefficients are multiplied by 100 to favor readability. The *t*-statistics (in parentheses) are based on standard errors clustered by district. Significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively. Refer to Tables [A.1](#) and [OA.1](#) for variable definitions.

Panel A: Summary Statistics								
	Obs.	Mean	SD	Min	P25	Median	P75	Max
Summer max. temperature (mean)	401	23.234	0.968	20.451	22.657	23.223	23.867	25.694
Summer max. temperature (volatility)	401	4.108	0.242	2.984	4.042	4.128	4.255	4.710
Daily precipitations (mean)	401	1.988	0.395	1.350	1.727	1.909	2.171	3.620
Daily precipitation (volatility)	401	3.979	0.635	2.982	3.592	3.852	4.173	6.976
Summer temperature-precipitation corr.	401	-0.010	0.037	-0.158	-0.030	-0.004	0.016	0.066
Excess heat-related mortality (present)	401	10.538	2.097	5.443	9.022	10.476	12.079	17.680
Excess heat-related mortality (+4°)	401	37.209	5.480	22.217	34.118	37.815	41.202	52.400
Panel B: Regression Analysis								
	SD(ln(<i>H</i> / <i>P</i>))							
	1	2	3	4	5	6	7	8
Summer max. temperature (mean)	0.406 (0.42)							1.839 (1.44)
Summer max. temperature (volatility)		-3.218** (-2.59)						-1.550 (-1.14)
Daily precipitation (mean)			-3.041*** (-3.16)					-3.591** (-2.17)
Daily precipitation (volatility)				-3.043*** (-2.91)				3.209* (1.70)
Summer temperature-precipitation corr.					1.568 (1.16)			1.161 (1.50)
Excess heat-related mortality (present)						3.447*** (2.94)		4.224*** (3.43)
Excess heat-related mortality (+4°)							0.840 (0.66)	-2.102 (-1.54)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Unit of observation	Postal code-time	Postal code-time	Postal code-time	Postal code-time	Postal code-time	Postal code-time	Postal code-time	Postal code-time
Mean(<i>y</i>)	0.28	0.28	0.28	0.28	0.28	0.28	0.28	0.28
SD(<i>y</i>)	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
Observations	15,553	15,553	15,553	15,553	15,553	15,553	15,553	15,553
Adjusted <i>R</i> ²	0.01	0.02	0.02	0.02	0.02	0.03	0.01	0.04

Table OA.14. Dispersion in Local Rent-to-Price Ratios and Energy Consumption

This table reports estimates from panel regressions of local dispersion in rent-to-price ratios on dispersion in energy consumption of properties. The dependent variable is the postal code-quarter-level standard deviation of log-transformed rent-to-price ratios obtained via the baseline matching procedure of flats for rent to flats for sale. The explanatory variables is postal code-quarter-level standard deviation in properties' annual energy consumption in kWh per sqm. Postal code-quarters with fewer than 10 observations—in terms of non-missing rent-to-price ratios or energy consumption proxy—are removed from the sample on which the standard deviations of rent-to-price ratios and energy consumption are computed. Columns 1-2 (3-4) columns include (exclude) matches for which the absolute difference in terms of annual energy consumption between the flat for rent and the flat(s) for sale is above 50 kWh per sqm. Included fixed effects are indicated below. Estimated coefficients are multiplied by 100 to favor readability. The *t*-statistics (in parentheses) are based on standard errors clustered by district. Significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively. Refer to Tables A.1 and OA.1 for variable definitions.

	SD(ln(H/P))			
	1	2	3	4
SD(Energy consumption)	0.023*** (4.55)	0.006** (2.02)	0.231*** (6.71)	0.110** (2.47)
Time FE	Yes	Yes	Yes	Yes
Postal code FE	No	Yes	No	Yes
Matching restriction	No	No	Yes	Yes
Unit of observation	Postal code-time	Postal code-time	Postal code-time	Postal code-time
Mean(y)	0.27	0.27	0.21	0.21
SD(y)	0.10	0.10	0.10	0.10
Observations	28,905	28,626	1,054	812
Adjusted R^2	0.01	0.20	0.10	0.29

Table OA.15. Dispersion in Local Rent-to-Price Ratios and Housing Ownership Structure

This table reports tests on local dispersion in rent-to-price ratios and housing ownership structure for a German sample between 2007 and 2017. Panel A shows summary statistics on municipality-level measures of housing ownership by institutional investors. Panel B reports estimates from regressions of local dispersion in rent-to-price ratios on such measures. The dependent variable is the postal code-quarter-level standard deviation of log-transformed rent-to-price ratios obtained via the baseline matching procedure of flats for rent to flats for sale. Postal code-quarters with fewer than 30 observations are removed from the sample on which the standard deviations of rent-to-price ratios are computed. Columns 1-2 (3-4) show specifications in levels (first differences). Specifications in levels include calendar quarter fixed effects. Specifications in first differences consider the change between 2022 and 2011 for the explanatory variables, and between the available year (as long as it is after 2014) and 2011 for the dependent variable. Estimated coefficients are multiplied by 100 to favor readability. The t -statistics (in parentheses) are based on standard errors clustered by district. Significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively. Refer to Tables A.1 and OA.1 for variable definitions.

Panel A: Summary Statistics

	Obs.	Mean	SD	Min	P25	Median	P75	Max
Property companies (% , 2011)	5,830	2.161	3.625	0.100	0.400	0.900	2.200	47.200
Other institutions (% , 2011)	7,382	1.343	1.998	0.100	0.500	0.800	1.500	47.400
Δ Property companies (% , 2022-2011)	3,452	0.308	3.766	-36.600	-0.300	0.100	0.800	56.400
Δ Other institutions (% , 2022-2011)	5,140	0.207	1.904	-28.200	-0.200	0.100	0.600	23.000

Panel B: Regression Analysis

	SD(ln(H/P))		Δ SD(ln(H/P)) (2017-2011)	
	1	2	3	4
Property companies (% , 2011)	0.046 (0.89)			
Other institutions (% , 2011)		0.046 (0.89)		
Δ Property companies (% , 2022-2011)			-0.031 (-0.33)	
Δ Other institutions (% , 2022-2011)				-0.196 (-1.40)
Time FE	Yes	Yes	No	No
Unit of observation	Postal code-time	Postal code-time	Postal code-time	Postal code-time
Mean(y)	0.28	0.28	0.28	0.28
SD(y)	0.09	0.09	0.08	0.08
Observations	15,535	15,535	6,382	6,392
Adjusted R^2	0.01	0.01	-0.00	0.00

Table OA.16. Summary Statistics on the Pseudo-Panel

This table shows summary statistics for the quarterly pseudo-panel between 2007 and 2017. Panel A provides basic statistics. Panel B focuses on time-series variation at cohort-, district-, state-, and national level. The average standard deviation, coefficient of variation (not computed for returns and growth rates), and interquartile range across geographic areas are reported. These averages are computed either giving equal weight to the geographic areas or by weighting them by the number of listed flats. Refer to Table A.1 for variable definitions.

Panel A: Basic Statistics

	Obs.	Mean	SD	Min	P25	Median	P75	Max
r (%)	24,863	1.988	12.353	-41.015	-5.087	1.886	8.846	47.310
r^* (%)	24,863	0.786	12.428	-42.407	-6.350	0.695	7.698	46.575
r^e (%)	24,863	2.025	12.374	-40.688	-5.094	1.898	8.913	47.651
Δh (%)	24,863	0.348	4.283	-14.320	-2.145	0.260	2.806	16.341
H^q/P (%)	24,863	1.210	0.277	0.630	1.012	1.182	1.369	2.269
H^q (EUR)	24,863	20.696	4.886	11.296	16.834	19.870	23.719	38.609
P (EUR)	24,863	1,828.823	708.974	569.200	1,316.466	1,675.489	2,177.365	5,082.414

Panel B: Time-Series Variation

Level of aggregation:	Cohort	District	State	National
<i>Equally weighted</i>				
SD (r)	12.293	10.068	5.890	1.283
SD (r^*)	12.370	10.136	5.927	1.324
SD (r^e)	12.318	10.099	5.972	1.469
SD (Δh)	4.258	3.271	1.673	0.561
SD (H^q/P)	0.175	0.158	0.128	0.087
SD (H^q)	1.499	1.318	1.492	1.230
SD (P)	357.005	324.334	331.430	281.745
CV (r)	.	.	.	.
CV (r^*)	.	.	.	.
CV (r^e)	.	.	.	.
CV (Δh)	.	.	.	.
CV (H^q/P)	0.145	0.132	0.106	0.073
CV (H^q)	0.071	0.067	0.072	0.059
CV (P)	0.192	0.183	0.176	0.153
P75-P25 (r)	15.938	12.966	7.806	1.756
P75-P25 (r^*)	16.066	13.064	7.892	1.849
P75-P25 (r^e)	15.978	13.050	7.976	2.490
P75-P25 (Δh)	5.470	4.189	2.276	0.800
P75-P25 (H^q/P)	0.245	0.223	0.192	0.155
P75-P25 (H^q)	1.888	1.626	1.962	1.195
P75-P25 (P)	494.902	460.144	492.221	423.491
<i>Weighted by no. properties</i>				
SD (r)	9.211	5.447	2.432	1.283
SD (r^*)	9.261	5.482	2.462	1.324
SD (r^e)	9.243	5.495	2.554	1.469
SD (Δh)	3.300	1.913	0.924	0.561
SD (H^q/P)	0.156	0.126	0.095	0.087
SD (H^q)	2.002	1.915	1.705	1.230
SD (P)	449.757	420.299	353.470	281.745
CV (r)	.	.	.	.
CV (r^*)	.	.	.	.
CV (r^e)	.	.	.	.
CV (Δh)	.	.	.	.
CV (H^q/P)	0.134	0.111	0.084	0.073
CV (H^q)	0.087	0.084	0.078	0.059
CV (P)	0.205	0.189	0.170	0.153
P75-P25 (r)	12.045	7.258	3.225	1.756
P75-P25 (r^*)	12.139	7.329	3.310	1.849
P75-P25 (r^e)	12.134	7.426	3.679	2.490
P75-P25 (Δh)	4.210	2.567	1.366	0.800
P75-P25 (H^q/P)	0.235	0.207	0.167	0.155
P75-P25 (H^q)	2.905	2.827	2.295	1.195
P75-P25 (P)	710.960	696.131	582.566	423.491

Table OA.17. Predictive Regressions Including Time Fixed Effects

This table reports estimates from predictive regressions for the housing premium and rent growth on the log-transformed quarterly rent-to-price ratio, including time fixed effects. The regressions are estimated on a pseudo-panel constructed from a sample of German flats listed between 2007 and 2017. The unit of observation is at the cohort-calendar quarter level, where cohorts are defined by the district in which the flat is located, its number of rooms category, and its size category. The dependent variable in columns 1 to 3 (4 to 6) is the k -quarter ahead housing premium $r_{t+1 \rightarrow t+k}^e$ (rent growth $\Delta h_{t+1 \rightarrow t+k}$), with $k = 1, 4, 12$. All specifications include cohort and time fixed effects. Estimated coefficients are multiplied by 100 to favor readability. The t -statistics (in parentheses) are based on Driscoll-Kraay standard errors (number of lags equal to k). Significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively. Refer to Table A.1 for variable definitions.

	$r_{t+1 \rightarrow t+k}^e$			$\Delta h_{t+1 \rightarrow t+k}$		
	1 $k = 1$	2 $k = 4$	3 $k = 12$	4 $k = 1$	5 $k = 4$	6 $k = 12$
$\ln(H^q/P)$	50.219*** (30.40)	73.495*** (18.94)	87.855*** (20.37)	-5.002*** (-16.10)	-7.238*** (-12.20)	-8.562*** (-11.42)
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Unit of observation	Cohort-time	Cohort-time	Cohort-time	Cohort-time	Cohort-time	Cohort-time
Mean(y)	0.02	0.09	0.28	0.00	0.02	0.05
SD(y)	0.12	0.15	0.19	0.04	0.05	0.07
Observations	23,996	20,897	14,866	23,970	20,857	14,860
Within R^2	0.25	0.42	0.59	0.04	0.13	0.35

Table OA.18. Predictive Regressions Controlling for Local Conditions

This table reports estimates from predictive regressions for the housing premium and rent growth on the log-transformed quarterly rent-to-price ratio, controlling for selected district-level characteristics. The regressions are estimated on a pseudo-panel constructed from a sample of German flats listed between 2007 and 2017. The unit of observation is at the cohort-calendar quarter level, where cohorts are defined by the district in which the flat is located, its number of rooms category, and its size category. The dependent variable in columns 1 to 3 (4 to 6) is the k -quarter ahead housing premium $r_{t+1 \rightarrow t+k}^e$ (rent growth $\Delta h_{t+1 \rightarrow t+k}$), with $k = 1, 4, 12$. All specifications include cohort fixed effects. Estimated coefficients are multiplied by 100 to favor readability. The t -statistics (in parentheses) are based on Driscoll-Kraay standard errors (number of lags equal to k). Significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively. Refer to Tables A.1 and OA.1 for variable definitions.

	$r_{t+1 \rightarrow t+k}^e$			$\Delta h_{t+1 \rightarrow t+k}$		
	1 $k = 1$	2 $k = 4$	3 $k = 12$	4 $k = 1$	5 $k = 4$	6 $k = 12$
$\ln(H^q/P)$	37.243*** (17.22)	51.575*** (7.48)	69.340*** (6.72)	-4.909*** (-17.55)	-9.362*** (-10.13)	-15.258*** (-4.81)
Old-to-working age ratio (%)	1.577*** (6.45)	4.017*** (6.48)	8.145*** (12.06)	0.134*** (3.16)	0.655*** (3.98)	2.456*** (10.40)
Disposable income p.c.	1.598** (2.09)	1.906 (1.13)	-3.536** (-2.23)	-0.046 (-0.52)	0.612** (2.68)	0.179 (0.28)
Living space completed p.c.	16.453*** (4.83)	23.569** (2.57)	3.774 (0.33)	-1.738*** (-2.80)	0.338 (0.15)	-0.058 (-0.01)
$\ln(\text{No. businesses})$	3.889 (0.23)	48.422 (1.03)	190.841*** (3.89)	0.753 (0.33)	13.626* (1.96)	25.421* (1.80)
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Unit of observation	Cohort-time	Cohort-time	Cohort-time	Cohort-time	Cohort-time	Cohort-time
Mean(y)	0.02	0.09	0.28	0.00	0.02	0.05
SD(y)	0.12	0.15	0.19	0.04	0.05	0.07
Observations	23,547	20,522	14,634	23,516	20,472	14,611
Within R^2	0.17	0.23	0.31	0.03	0.09	0.16

Table OA.19. Persistence of Local Rent-to-Price Ratio Dispersion

This table reports estimates from panel regressions of local dispersion in rent-to-price ratios on its lag for a German sample between 2007 and 2017. The dependent variable is the postal code-quarter-level standard deviation of (filtered) log-transformed rent-to-price ratios obtained via the baseline matching procedure of flats for rent to flats for sale. The filtered log-transformed rent-to-price ratio is the residual from the specification in column 6 of Table 5. Suspect duplicate property listings as well as postal code-quarters with fewer than 30 observations are removed from the sample on which the measures of local dispersion in rent-to-price ratios are computed. Specification in odd columns include no fixed effects. Specifications in even columns include postal code fixed effects. Estimated coefficients are multiplied by 100 to favor readability. The t -statistics (in parentheses) are based on standard errors clustered by district. Significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively. Refer to Table A.1 for variable definitions.

	SD($\ln(H/P)$)		SD(Filterd $\ln(H/P)$)	
	1	2	3	4
SD($\ln(H/P)$) (lag)	45.590*** (16.00)	8.187*** (6.24)		
SD(Filterd $\ln(H/P)$) (lag)			46.393*** (16.28)	8.849*** (5.12)
Postal code FE	No	Yes	No	Yes
Unit of observation	Postal code-time	Postal code-time	Postal code-time	Postal code-time
Mean(y)	1.62	1.62	1.38	1.38
SD(y)	0.66	0.65	0.59	0.59
Observations	8,659	8,557	8,659	8,557
Adjusted R^2	0.20	0.38	0.21	0.38