

Bracke, Philippe

Conference Paper

House Prices and Rents: Micro Evidence from a Matched Dataset in Central London

53rd Congress of the European Regional Science Association: "Regional Integration: Europe, the Mediterranean and the World Economy", 27-31 August 2013, Palermo, Italy

Provided in Cooperation with:

European Regional Science Association (ERSA)

Suggested Citation: Bracke, Philippe (2013) : House Prices and Rents: Micro Evidence from a Matched Dataset in Central London, 53rd Congress of the European Regional Science Association: "Regional Integration: Europe, the Mediterranean and the World Economy", 27-31 August 2013, Palermo, Italy, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/123867>

Standard-Nutzungsbedingungen:

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

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

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

Terms of use:

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

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

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

House Prices and Rents: Micro Evidence from a Matched Dataset in Central London*

Philippe Bracke[†]

February 2013

Using the proprietary dataset of a real estate agency, I analyze tens of thousands of housing sale and rental transactions in Central London during the 2005-2011 period. I run hedonic regressions on both prices and rents and show that price-rent ratios are higher for bigger and more central units. Since this result could be driven by differences in unobserved characteristics between properties for sale and properties for rent, I replicate my analysis using only units that were both sold and rented out within 6 months, and get similar results. I discuss several possible explanations for my findings.

*I thank Jonathan Halket, Steve Gibbons, Christian Hilber, Olmo Silva, Silvana Tenreyro, and participants to the LSE Spatial Economics Research Centre and STICERD seminars for useful comments. I am grateful to James Wyatt of John D Wood & Co. for providing access to the data.

[†]Spatial Economics Research Centre (SERC), London School of Economics. Email: p.bracke@lse.ac.uk

The relation between house prices and rents matters both for households and real estate investors. Housing is the largest asset in homeowners' balance sheets,¹ and rents represent the major expenditure for most renters, amounting to 20-30% of monthly payments (Genesove, 2003). For investors, rental yields—the inverse of price-rent ratios—are a fundamental component of housing returns. Differences in rental yields across property types are important for portfolio management (Plazzi, Torous and Valkanov, 2011).

In this paper I study unit-level data on house prices and rents in Central London. I document the existence of systematic differences in price-rent ratios across property types within the same urban area: bigger properties and properties located in more expensive neighborhoods have higher price-rent ratios. My analysis is based on a novel proprietary dataset from a Central London real estate agency. The dataset contains information on achieved prices and rents for tens of thousands of properties, as well as detailed descriptions of property characteristics. The period of analysis, 2005 to 2011, covers the last part of the housing boom, the bust of 2008, and the subsequent recovery.² The area under study hosts a mix of owner-occupied and private-rented properties, which often lie side by side. Observed prices and rents are the result of genuine market forces, because the UK private rental market is essentially unregulated.³

In terms of empirical methodology, I use hedonic regressions to estimate average prices and rents within cells of observationally equivalent properties. Since hedonic regressions cannot control for unobserved characteristics, and these could differ between sold and rented dwellings, I also run a restricted analysis with properties that are both sold and rented out within 6 months. In this way I am able to measure price-rent ratios directly: I have enough observations to focus only on prices and rents observed *on the same property at approximately the same time*. I run the same hedonic regressions on this subset of properties and get coefficients that are very similar to the ones obtained from the whole dataset.

In the last part of the paper I discuss potential explanations for the differences in price-rent ratios. One possibility is that gross price-rent ratio disparities hide differences in maintenance costs or vacancy rates: perhaps price-rent ratios are equalized once these differences are taken into account. Another possibility, related to the dividend pricing model, is that properties with higher price-rent ratios feature higher expected rent growth or lower risk premia. Contrary to

this second view, I find that within Central London the rent growth rates of more expensive properties are not different from those of cheaper properties, but their volatility is significantly higher. This is consistent with the hedging model of Sinai and Souleles (2005): higher price-rent ratios are associated with higher rent volatility, which pushes people to buy in order to lock in future rents. The submarkets where price-rent ratios are highest—big and centrally located properties—are characterized by rental markets that are “thin” in a search-theory sense (Ngai and Tenreyro, 2009), so that finding and maintaining good matches is difficult.

In order to verify the above mechanisms, I use price and rent indexes derived from the hedonic regressions to estimate the growth and aggregate volatility of prices and rents for different property categories. Using data at the individual property level, I also measure idiosyncratic volatilities by restricting attention to properties that were sold or rented at least twice during the sample period. Since the expectations of agents might differ from the actual historical performance of house prices and rents, I complement my analysis with an expectation survey carried out through the mailing list of the real estate agency that provided the property data.

In the literature, price-rent ratios have been repeatedly measured and studied using aggregate data. For instance, Gallin (2008) uses US city-level data to check if changes in price-rent ratios anticipate future price or rent growth, as the dividend discount model would predict. Ambrose, Eichholtz and Lindenthal (2013) analyze the behavior of the aggregate price-rent ratio in Amsterdam from 1650 through 2005.

Due to the lack of data, very few papers have analyzed rents at the micro local level. An exception is Hwang, Quigley and Son (2006), who use micro data on prices and rents from South Korea to test the dividend pricing model. Hwang, Quigley and Son exploit the high homogeneity of apartments in Seoul and the surrounding areas to compute price-rent ratios and see how they evolve over time. By contrast, in this paper I exploit the *heterogeneity* of housing units in Central London to shed light on the cross-sectional variation of price-rent ratios. Another recent exception is Hill and Syed (2012), who use micro data from Sidney, Australia, to compare actual price-rent ratios with housing user costs. Using imputation hedonic methods, they also find that more expensive properties have higher price-rent ratios. However, they do not decompose these differences into size and location effects. Finally, also Garner and Verbrugge

(2009), using answers from the US Consumer Expenditure Survey to compare self-reported rents and house values, find that more expensive properties have higher price-rent ratios.⁴

The rest of the paper is organized as follows. The first section describes the data. Next the empirical methodology of the paper is presented and the main results are shown. I then discuss the theories that can explain the main results, and provide some final considerations.

Data

The main dataset used in this paper comes from John D Wood & Co. (JDW), a real estate agency that operates in London and the surrounding countryside.⁵ The JDW Dataset includes observations from the Central-Western area of London, corresponding to the local authorities of Camden, Westminster, Kensington and Chelsea, Hammersmith and Fulham, and Wandsworth. These local authorities are shown on the left-hand side of Figure 1. They are responsible for running services such as schools, waste collection, and roads. This area is one of the most densely populated in London: most of the housing stock is made of flats rather than single-family houses; approximately one fourth of dwellings are privately rented.⁶

The right-hand side of Figure 1 shows the postcode districts included in the JDW Dataset. In the U.K. postal code system, the postcode district represents the first half of the postcode (one or two letters followed by one or two numbers) and corresponds to 10,000 - 20,000 unique addresses. In the empirical analysis, I use postcode district dummies to capture the effect of location on house prices.

Big agencies like JDW have valuation teams who keep track of market trends. Agents assemble sale and rental data from their own records as well as from other agencies. Before using the JDW data for the analysis, I perform several cleaning and checking procedures. To remove potential duplicates, every sale or rental contract which refers to the same property and occurs within one month is excluded. This operation has the additional advantage of removing short-term rental contracts, which are usually more expensive than other rentals and targeted to specific markets (e.g. business travelers and tourists). Moreover, since London houses and flats can also be sold

on a leasehold—an arrangement by which the property goes back to the original landlord after the lease expires—I drop all sales of properties with a leasehold expiring in less than 80 years.⁷ Finally, to avoid outliers, I trim properties whose price or rent is below the 1st percentile or above the 99th percentile of the price or rent distribution for their transaction year.

Figure 2 plots the sale and rental observations on the London map. There are 20,154 sales and 43,361 rentals in total. Figure 3 shows the quarterly number of transactions for these two groups of observations. The number of sales varies substantially from one period to another. In the 2005-2007 period, when the market was characterized by rising prices, the average number of quarterly transactions was four times as high as the number of transactions during the 2008 bust. The number of rental contracts, by contrast, appears less volatile from one year to another. However, rental contracts display a sharper seasonal pattern: the third quarter always has 50% more transactions than the first quarter. For sales, the first quarter has usually a lower number of transactions, but seasonality is less pronounced.

The different parts of the empirical analysis focus on subsets of the JDW data. The first subset, denoted as the Hedonic Dataset, contains the properties on which the main hedonic regressions are run. These are the properties for which floor area (measured in square feet) is available.⁸ The second subset, denoted as Repeat Transactions, contains the properties that were sold or rented multiple times. The third subset, denoted the Matched Units, contains the properties that were sold and rented out within 6 months.

The first two columns of Table 1 illustrate the summary statistics for the sold properties (Sales) and rented properties (Rentals) in the Hedonic Dataset. Consistently with the composition of housing stock in this part of London, the majority of housing units are flats. There are more flats in Rentals (84%) than in Sales (75%). Moreover, Sales contain a higher number of large flats (3 or more bedrooms) than Rentals. The median floor area is larger for Sales (1,059 square feet against 879 square feet). Other studies report similar differences between owner-occupied and rented units. For instance, Glaeser and Gyourko (2007) use the 2005 American Housing Survey to show that “The median owner occupied unit is nearly double the size of the median rented housing unit,” and that rental units are more likely to be located near the city centre.

Some of the analyses carried out in the following sections require to focus only on properties that appear at least twice in the Sales or Rental datasets. Since the turnover in rental contracts is higher than the turnover in owner occupation, repeat observations in Rentals are more common than repeat observations in Sales. In the 2005-2011 JDW data, 1,049 properties are sold twice and 45 properties are sold three times. By contrast, 5,774 properties appear twice in the rental dataset, 1,594 properties appear three times, and 556 properties appear four or more times. The Appendix reports that the proportion of repeat sales out of all sales in the JDW Sales Dataset and the Land Registry look similar.

Table 1 shows that the average property for sale in the Repeat Transactions Dataset is more expensive and bigger than the average property in the whole JDW Sales Dataset. By contrast, the average property for rent in the Repeat Transactions Dataset is very similar to the average property for rent in the whole JDW Rentals Dataset.

The Matched Dataset

The Matched Dataset contains properties that appear both in the sale and rental datasets, with the sale taking place between 0 and 6 months before the corresponding rental contract. To increase the number of matched observations, I expand the search for matches to the Land Registry, which contains official records of all housing transactions in England and Wales.⁹

In all datasets properties are uniquely identified by their address. For houses, the address is made of the street name and number. For apartments, the address contains additional information such as floor and unit number. For each property in the JDW Rental Dataset, the matching algorithm looks for a sale of the same property either in the JDW Sales Dataset or in the Land Registry. Since every record comes with a transaction date, the distance in days between sales and rentals is measured. As there can be multiple sales and multiple rentals for each property, for every sale the algorithm keeps only the closest rental contract. If a rental contract can be imputed to multiple sales, the algorithm keeps only the closest sale. Since prices and rents can diverge over time, it is necessary to keep only rental contracts that were signed shortly after the sale of the property. I choose 6 months as the cutoff distance between the sale

and the rental. My window around the sale date is asymmetric in the sense that I do not select rental contracts signed a few months before a sale.

Table 1 shows that the average property in the Matched Dataset is slightly more expensive than the average property in the Hedonic Rental Dataset but less expensive than the average property in the Hedonic Sale Dataset. The proportion of flats and the geographical distribution of properties are similar to the ones in the Rentals Dataset.

Table 2 shows how many matches are retrieved in each year, and their average rent-price ratios. Most matches come from the Land Registry. Some matches are found both in the JDW Dataset and the Land Registry, so that the sum of the second and third column in the table is in some cases less than the number in the third column. The low number of transactions in 2008 and 2009 causes the number of matches to be low in those years. Moreover, since the available Land Registry data on individual addresses covers only the 2006-2010 period, I concentrate only on these years when analyzing matched properties.¹⁰

In Table 2, the price-rent ratios computed on the JDW rentals-sales matches are slightly lower than the corresponding ratios computed on the JDW rentals-Land Registry matches. The variation over time is similar: price-rent ratios increase from 2006 to 2009 and then decrease in the last year.

As a preliminary glance at the main finding documented in the paper, Figure 4 plots rents and prices for the observations of the Matched Dataset. The plot on the left contains all observations. Since the distribution of prices and rents is skewed to the right, the axes use a log scale. Notice that the dots would cluster around a 45-degree line if price-rent ratios were the same for expensive and cheap properties. The dots appear to have a slope greater than 1, indicating that price-rent ratios are increasing in rents. To see this more formally, consider the following regression:

$$\log \left(\frac{Price}{Rent} \right) = \alpha + \beta \log Rent \tag{1}$$

where a β coefficient significantly greater than zero indicates that price-rent ratios are positively correlated with rents. Table 3 shows the output of this regression. In the first column, the

coefficient is positive and significant: more valuable houses have higher price-rent ratios. The second column displays the regression results when year dummies are added: the coefficient on $\log Rent$ is virtually unchanged. Lastly, by interacting year dummies with the log rents, it is possible to separate the different effects of log Rent for each year—results are displayed in the third column of Table 3. Coefficients are positive and significant in all years except for 2008. This might be due to the exceptional nature of 2008 market conditions and the low number of observations. Moreover, coefficients are larger in later years, indicating an increasing divergence of price-rent ratios across properties.

The right-hand side plot of Figure 4 zooms on properties with a weekly rent of less than £1000. The axes follow now a linear scale, and the plot contains the same line as the first plot. Dots follow a nonlinear pattern, with higher rents corresponding to much higher prices. This nonlinearity is highlighted also by Garner and Verbrugge (2009), using data from the U.S. Consumer Expenditure Survey, and Hill and Syed (2012), using sale and rental data from Sidney. The next section pushes the analysis further by decomposing this pattern into the contribution of size and location on price-rent ratios.

Main findings

Empirical methodology

The log price of a house i at time t can be modeled as the sum of three elements:

$$p_{it} = q_i + \lambda_t + u_{it}, \tag{2}$$

where q_i represents the quantity of housing services that the house provides (the “quality” of the house), λ_t is the quality-adjusted price for one unit of housing services at time t , and u_{it} is an idiosyncratic shock centered around zero. The first term varies across properties but is constant over time; the second term is constant across properties but varies over time; and the third term captures property- and time-specific shocks.

Housing is a composite and heterogeneous good and every property represents a different combination of characteristics. Hence q_i can be decomposed as follows:

$$q_i = X_i\beta^* + Z_i\gamma^*, \quad (3)$$

where X_i is a vector of observed characteristics and Z_i is a vector of unobserved characteristics. This formulation is at the basis of the hedonic method (Court, 1939; Griliches, 1961). In the context of housing, assuming that the market for properties is competitive and property characteristics enter the utility function, the coefficients β^* and γ^* represent the shadow prices of an additional unit of each characteristic (Rosen, 1974).

Similarly, λ_t can be decomposed as:

$$\lambda_t = \sum_{d=1}^D \delta_d Time_d, \quad (4)$$

where $Time_d$'s are dummy variables equal to 1 if $d = t$ and 0 otherwise, and the δ_d 's represent the coefficients on those dummies. By assumption, the prices of characteristics are held fixed over time: all time variation is captured by λ_t . The sequence of estimated δ_d coefficients can be interpreted as an index of log house prices. This regression is commonly referred to as the “time-dummy” hedonic regression (Hill, 2012).

A more general model would include interactions between the characteristics included in X_i and the $Time_d$ dummies. In the section on mechanisms, I allow different categories of houses to have their own price indexes, and define categories according to one or more of the characteristics included in X_i . Pushing this argument forward, one could also allow the price of *every* observed characteristic to change over time, making the aggregate price index λ_t redundant. At the end of this section I briefly explore this formulation, which is equivalent to interacting all the elements of the vector X_i with time dummies. In the main part of the analysis, I keep the prices of characteristics fixed and stick to the time-dummy regression, which conveniently separates cross-sectional and time variation. Since the analyzed dataset covers only 7 years, from 2005 to 2011, changes in the relative prices of characteristics are likely to be limited.

In the empirical work, the vector Z_i is unobservable. The estimated model is therefore:

$$p_{it} = \alpha + X_i\beta + \sum_{d=1}^D \delta_d \text{Time}_d + \varepsilon_{it}, \quad (5)$$

where α is a constant which serves the purpose of normalizing the β and δ coefficients with respect to a base category (e.g. 1-bedroom flats in the first quarter of 2005). The coefficients estimated from equation 5 are affected by omitted variable bias (OVB). For instance, the coefficient β is equal to $\beta^* + \gamma^*\phi_X$, where $\phi_X = (X'X)^{-1}X'Z$.

The dataset used in this paper contains information on both sale prices and rental prices. To distinguish between the two, I use the subscripts s for sales and r for rentals. Equation 5 becomes:

$$p_{hit} = \alpha_h + \beta_h X_i + \lambda_{ht} + u_{hit}, \quad (6)$$

where $h \in \{s, r\}$ and $\lambda_{ht} = \sum_{d=1}^D \delta_{hd} \text{Time}_{hd}$. This formulation allows for quality, quality-adjusted prices, and errors to differ between sales and rentals. It is quite natural to expect the estimated α_s to be significantly higher than the estimated α_r : on average, prices are higher than rents. Similarly, one would expect λ_{st} 's to be different from λ_{rt} 's: aggregate prices and rents might move differently over time, generating changes in the aggregate price-rent ratio (Gallin, 2008; Campbell, Davis, Gallin and Martin, 2009). Indeed, different coefficients in the price and rent hedonic equations imply an effect of the regressors on price-rent ratios, because $E p_s - E p_r = E(p_s - p_r)$, and $p_s - p_r$ is the log price-rent ratio. Hence, obtaining different estimates for the β_s and the β_r is a nontrivial finding: it means that some property characteristics have an effect on price-rent ratios.

Clearly, estimated coefficients are biased. Using the OVB formula, the difference between β_s and β_r computed from the hedonic regressions is:

$$\begin{aligned} \beta_s - \beta_r &= \beta_s^* - \beta_r^* + \phi_{Xs}\gamma_s^* - \phi_{Xr}\gamma_r^* \\ &= \beta_s^* - \beta_r^* + (\phi_{Xs} - \phi_{Xr})\gamma_s^* + (\gamma_s^* - \gamma_r^*)\phi_{Xr} \end{aligned}$$

where the final step is obtained by adding $\phi_{Xr}\gamma_s^* - \phi_{Xr}\gamma_r^* = 0$ to the equation. The difference in the estimated coefficients is equal to the true difference in coefficients plus two terms—the first depending on the different types of houses that belong to the sales and rentals datasets ($\phi_{Xs} - \phi_{Xr}$), and the second depending on the different coefficients that regulate the relation between unobserved characteristics and log prices or rents ($\gamma_s^* - \gamma_r^*$).

The Matched Dataset contains properties that were both sold and rented within a short period of time. These observations can be helpful in reducing the OVB. For these properties, the price-rent ratio can be directly observed and can serve as dependent variable in the following regression:

$$\log\left(\frac{Price}{Rent}\right)_{it} = \alpha_m + X_i\beta_m + \lambda_{mt} + \varepsilon_{mit}, \quad (7)$$

which mimics the hedonic model and where $y_{it} = p_{sit} - p_{rit}$. For these properties $\phi_{Xs} = \phi_{Xr}$: the bias in measuring the effect of property characteristics on price-rent ratios is reduced to $(\gamma_s^* - \gamma_r^*)\phi_{Xr}$.

Results

I start the empirical analysis by estimating equation 6 separately for Sales and Rentals. The vector of characteristics X_{it} contains: a dummy variable to indicate whether the property is a house (as opposed to a flat); three dummy variables indicating the number of bedrooms of the property: 2 bedrooms, 3 bedrooms, and 4 bedrooms or more¹¹ (1-bedroom properties are the baseline category); floor area measured in square feet; floor area squared, to take into account the tendency of prices and rents to rise less than proportionally with size; and postcode district dummies to capture the effects of local amenities. The estimated regression is therefore:

$$\begin{aligned} p_{it} = & \alpha + \beta_1 \text{House} + \beta_2 \text{2-bed} + \beta_3 \text{3-bed} + \beta_4 \text{4-bed} + \beta_5 \text{sqft} + \beta_6 \text{sqft}^2 \\ & + \sum_{q=7}^Q \beta_q \text{Postcode}_q + \sum_{d=1}^D \delta_d \text{Time}_d + \varepsilon_{it}, \end{aligned}$$

where I omit the subscript h to ease notation. I use quarterly dummies to construct a quarter-

by-quarter index of log house prices and rents (δ_{st} 's and δ_{rt} 's). Ferreira and Gyourko (2011) employ a similar hedonic regression—only with sale prices—for their recent neighborhood-level analysis of the start of the US housing boom.

Table 4 shows the output of the hedonic regressions on the complete Sales and Rentals dataset in columns 1 and 2. Column 3 computes the implied effect on price-rent ratios of the characteristics X . The coefficients in column 3 are equal to the difference between the coefficients in column 1 and those in column 2. Equivalently, they can be computed by stacking together the sale and rental prices in the same dataset and running the following regression:

$$\begin{aligned}
p_{it} = & \alpha + \pi_0 \text{Sale} + \beta_1 \text{House} + \pi_1 (\text{House} \times \text{Sale}) \\
& + \beta_2 \text{2-bed} + \pi_2 (\text{2-bed} \times \text{Sale}) + \beta_3 \text{3-bed} + \pi_3 (\text{3-bed} \times \text{Sale}) \\
& + \beta_4 \text{4-bed} + \pi_4 (\text{4-bed} \times \text{Sale}) \\
& + \beta_5 \text{sqft} + \pi_5 (\text{sqft} \times \text{Sale}) + \beta_6 \text{sqft}^2 + \pi_6 (\text{sqft}^2 \times \text{Sale}) \\
& + \sum_{q=7}^Q \beta_q \text{Postcode}_q + \sum_{q=7}^Q \pi_q (\text{Postcode}_q \times \text{Sale}) \\
& + \sum_{d=1}^D \delta_d \text{Time}_d + \sum_{d=1}^D \eta_d (\text{Time}_d \times \text{Sale}) + \varepsilon_{it},
\end{aligned} \tag{8}$$

where Sale is a dummy variable that takes value 1 if the observation refers to a sale, and 0 if the observation refers to a rental. The π and η coefficients measure the effect of certain variables on the price-rent ratio. Finally, column 4 of Table 4 shows the output from estimating equation 7 on the Matched Dataset. It is interesting to compare the coefficients in this column with the ones computed in column 3.

Table 4 shows that, conditional on number of bedrooms and floor area, houses command a positive premium on sales but a small negative premium on rentals. Therefore, on average, houses have higher price-rent ratios than flats. The effect is consistent with the hedonic regression on the matched dataset. Conditional on floor area, the number of bedrooms has a higher effect on Rentals than Sales. The contribution of floor area is positive, but more for prices than rents. As expected, the coefficient on floor area squared is negative.¹²

In Table 4, I sort neighborhoods from those with the highest price premium (SW3, Chelsea) to

those with the lowest one (SW6, Fulham)—the baseline postcode district is W2 (Paddington). In terms of coefficients, both the complete JDW dataset and the Matched Dataset show that more expensive neighborhoods have higher price-rent ratios. In other words, both prices and rents rise for more expensive neighborhood, but prices rise more than rents. This fact is well known by housing market practitioners.¹³

The fact that coefficients in column (3) and (4) are similar is reassuring: it speaks in favor of the consistency, in terms of results, between the Hedonic Dataset and the sample of matched properties. Table 1 already showed that the two samples are similar in terms of descriptive statistics.

The left-hand side part of Figure 5 plots the coefficients on time dummies from the hedonic regression in Sales (λ_{st}) and Rentals (λ_{rt}). In the boom period, prices grew at a rate approximately double that of rents. After the peak at the end of 2007, the gap between prices and rents has continued growing, albeit more slowly. During the sample period the correlation of the two indexes is very high (90%). The different growth rates of prices and rents produced increasing price-rent ratios—as shown in the right-hand side of Figure 5. The dashed line represents the price-rent ratios implied by the price and rent indexes. The solid line represents an index of actual price-rent ratios computed from the Matched Dataset. The two samples give similar results, although the matched sample is more volatile because of the smaller sample size.

Hedonic regressions with time-varying characteristic prices

Dropping the assumption of constant characteristic prices β_h , the hedonic equation 6 becomes:

$$p_{hit} = \beta_{ht}X_i + u_{hit}. \quad (9)$$

According to this equation house prices are a combination of the time-varying prices of their characteristics. The practical implementation of this approach consists in estimating equation 9 for each period. The JDW Dataset contains 28 quarters. Using the 23 explanatory variables listed in Table 4 (6 variables for physical characteristics and 17 dummies for postcode districts), separately for sale and rental prices, produces $28 \times 23 \times 2 = 1288$ coefficients. Figure 6 plots the

time evolution of some of these coefficients. Some quarters have a limited number of observations and this generates volatile characteristic prices. Despite volatility, however, the main message of these coefficients is consistent over time.

Houses enjoy a positive price premium and a negative rent premium with respect to flats. However, Figure 6 shows that this differential has been declining over time. It is possible that, in the aftermath of the housing bust, the demand for house purchases has declined and has been substituted by an increasing demand for housing rentals. Figure 6 also confirms that, conditional on floor area, rentals enjoy a premium for a high number of bedroom (4+). Moreover, the price of a square foot has been rising over time for sales but has stayed constant for rentals. This pattern is consistent with the general price and rent indexes in Figure 5, which show a higher appreciation of prices in the 2005-2011 sample period. The last four charts show the effect of location on prices and rents. It is clearly the case that, in all periods, properties in prime neighborhoods such as Chelsea (SW3) or Kensington (SW7) command a bigger premium on sales than rentals. When analyzing other neighborhoods, such as Fulham (SW6), the distinction between price and rent coefficients becomes much less clear.

Mechanisms

The previous section has shown that price-rent ratios are higher for bigger and better-located properties. A well-functioning market would arbitrage these asymmetries away unless they correspond to fundamental differences in the investment characteristics of properties. In this section, I explore some possible explanations for these disparities. I distinguish between three groups of arguments: explanations based on hidden costs, explanations based on asset pricing, and explanations based on housing submarkets. Hidden costs create a wedge between gross and net yields. It is possible that, while gross yields differ across property categories, net yields are similar. Asset-pricing explanations take the dividend discount model as starting point and try to rationalize differences in price-rent ratios with differences in expected growth or risk premia. The third group of explanations is based on the fact that urban housing markets are made of several submarkets with different demand-supply conditions, and price-rent ratios could differ

across these submarkets.

My focus is on general explanations rather than explanations based on specific characteristics of Central London—the presence of foreign buyers, for instance, or the importance of corporate lettings in the market for big apartments. This is because the stylized fact illustrated in this paper has been shown elsewhere in the UK (Joseph Rowntree Foundation, 1996; Association of Residential Letting Agents, 2012), the US (Garner and Verbrugge, 2009), and Australia (Hill and Syed, 2012). Thus, any comprehensive explanation cannot be London-specific.¹⁴

Hidden costs: Gross vs. net yields

A first reason for a wedge between gross and net yields relates to transaction costs. If fixed costs do not rise proportionally with prices, it becomes relatively more expensive to trade cheap properties. These properties would then sell for a lower price, which in turn would raise their rental yield.

A second potential explanation refers to the fact that gross yields do not account for periods in which rental properties are vacant. If small and less central properties are vacant for longer, this would explain the difference in yields. However, it is commonly believed that more expensive properties stay on the market for longer—the higher the price, the more important it is to find a good match with a buyer that likes the property. Despite the fact that the JDW Dataset does not contain data on vacancies, it has been possible to retrieve vacancy data for a subset of 2,325 rental properties marketed by John D Wood & Co. in the 2005-2011 period. Table 5 shows the results from regressing the number of days a property is waiting for a new tenant on the log rent and the type of property (flat or house). This subset of the data confirms the belief that more expensive properties stay on the market for longer periods. Moreover, conditional on rent, houses are marketed on average 28 days more than flats.

It could be that, despite having longer vacancies when on the market, big properties are on the market less often, i.e. tenants that use bigger properties stay for longer periods. This possibility, acknowledged by Halket and Pignatti (2012), would increase the net yield of big properties with respect to small properties.

A third possibility is that depreciation and maintenance costs are higher for smaller and less central properties; in equilibrium, then, their gross yields would have to be higher. Even if data on maintenance costs are not readily available, the conceptual decomposition of property prices into structure and land prices can help to infer the general pattern of depreciation across the price distribution. Structure costs vary less than house prices (Davis and Heathcote, 2007), which implies that expensive properties have a higher land value share. Since depreciation and maintenance costs refer to structures, they should be lower, in percentage terms, for expensive houses. Hill and Syed (2012) find that this explanation could justify at least part of the cross-sectional differences in price-rent ratios.

It has to be noted that professional landlord or real estate investor associations often publish estimates of net yields, which take into account maintenance costs and vacancies. The report by Association of Residential Letting Agents (2012), for instance, shows that (1) prime Central London in particular has a lower net yields than the rest of London and other UK regions, and (2) houses have lower net yields than flats. Thus, it is difficult to *completely* explain the pattern in price-rent ratios highlighted on the basis of transaction costs, maintenance costs, or vacancies.

Asset pricing: Expected appreciation and risk premia

According to the dividend discount model, the price of an asset corresponds to the present discounted value of its future dividends. In a real estate context, this equivalence translates into a relation between the sale price of a property and its future rents; or, alternatively, a link between the current price-rent ratio and future rent expectations. Properties with higher price-rent ratios should feature higher expected rent growth, higher risk premia, or both (Campbell, Davis, Gallin and Martin, 2009). The previous section has shown that size and location are positively correlated with price-rent ratios. Hence, big and better-located properties should display higher rent growth or be associated with lower volatility.

Aggregate evidence To check whether the empirical findings are consistent with the dividend discount model, I create two housing category classifications. In terms of size, I divide

observations into: 1-bedroom flats, 2-bedroom flats, 3-or-more-bedroom flats, and houses. The summary statistics in Table 1 show the dimension of these groups with respect to the overall dataset. In terms of location, I divide observations into prime neighborhoods and other neighborhoods. Prime neighborhoods are the most expensive six postcode districts in the hedonic regression of Table 4: SW3 (Chelsea), SW7 (South Kensington), W8 (Holland Park), W1 (Mayfair), W11 (Notting Hill), SW1 (Belgravia and Pimlico). In the JDW Dataset, prime neighborhoods correspond to 53% of Sales and 54% of Rentals. I take equation 6 and allow the coefficients on property characteristics to differ across categories:

$$p_{hct} = \alpha_{hc} + X_{ic}\beta_{hc} + \lambda_{hct} + \varepsilon_{hct}, \quad (10)$$

where c denotes a category of properties. This is equivalent to interacting the category dummy with all property characteristics. The average growth rate for a given property category c is $E(\lambda_{hct+1} - \lambda_{hct})$ and the corresponding aggregate risk is $\text{Var}(\lambda_{hct+1} - \lambda_{hct})$.

Figure 7 plots the λ_{hct} 's over the different quarters t , estimated using equation 10. The upper part of the figure shows results according to the first category classification, based on physical characteristics. Consistently with the housing-ladder model of Ortalo-Magné and Rady (2006), the prices of bigger houses grew more in the 2005-2007 boom period. This trend was partially reversed during the brief bust of 2008 but restarted immediately after. In terms of rents, the pattern is the same but more pronounced: the rent volatility of bigger properties is clearly higher. A similar impression is given by the prime vs. non-prime neighborhood sale price comparison at the bottom left of Figure 7. Sale prices in prime neighborhoods have grown more but are also more volatile. However, rents have behaved very similarly in prime and other neighborhoods, both in terms of growth and volatility.

Table 6 lists the average growth and volatilities of the different property categories and confirms the impressions gathered from Figure 7. In particular, the standard deviation of rents for houses is twice that for 2-bed flats. The numbers in Table 6 imply that it is difficult to rationalize the observed differences in price-rent ratios using the dividend pricing model. First, rent growth in the 2005–2011 period was not substantially higher for bigger or better-located properties. Second, for these properties aggregate rent volatility was significantly more pronounced, which

is inconsistent with their risk premium being lower.

Evidence from repeat sales and rentals Table 6 shows results on the volatility of *aggregate* indexes, not individual properties. The balance sheet of most homeowners contains just one property (Flavin and Yamashita, 2002), and most renters are obviously subject to just one rental contract (Genesove, 2003).¹⁵ Idiosyncratic volatility might be more relevant than its aggregate counterpart. To measure idiosyncratic volatility, I restrict attention to properties that sold or rented at least twice during the sample period. Suppose we observe the price of one of these properties (i) at time T and t . Differencing Equation 6 gives the log appreciation of property i :

$$p_{hiT} - p_{hit} = \underbrace{\lambda_{hT} - \lambda_{ht}}_{\text{aggregate movement}} + \underbrace{u_{hiT} - u_{hit}}_{\text{idiosyncratic movement}}. \quad (11)$$

Equation 11 constitutes the basis of the repeat sales method (Bailey, Muth and Nourse, 1963; Case and Shiller, 1989), which allows for the estimation of the term $u_{hiT} - u_{hit}$. Similarly to aggregate risk, idiosyncratic risk is defined as $\text{Var}(u_{it+1} - u_{it})$. Case and Shiller (1989) assume that $u_{it} = v_{it} + h_{it}$, where v_{it} is a white noise with mean zero and variance σ_v^2 , and h_{it} is a random walk with mean zero and variance $t\sigma_h^2$. Under these assumptions, $\text{Var}(u_{iT} - u_{it}) = 2\sigma_v^2 + \sigma_h^2(T - t)$ and $\text{Var}(u_{it+1} - u_{it}) = 2\sigma_v^2 + \sigma_h^2$. Case and Shiller employ these volatility estimates to improve the efficiency of the repeat sales regression and call their approach the weighted repeat sales estimator (WRS).¹⁶

Figure 8 compares the WRS indices estimated on the JDW Dataset with the hedonic indexes estimated before. The WRS index for sale prices displays a significantly steeper appreciation than the corresponding hedonic index. This difference might be due to that fact that, to be included in the repeat sales regression, a property must have sold twice between 2005 and 2011, a relatively short period. Property that resell quickly have usually undergone substantial improvements, or belong to a seller who has received a particularly good offer—indeed, Table 1 shows that the median price in the Repeat Transactions Dataset is higher than the one in the Hedonic Dataset.¹⁷

By contrast, when measuring the index for rental prices, the series computed through repeat transactions is smoother and shows a lower appreciation rate than the one measured through the hedonic method. The different behavior of the repeat rent index with respect to the repeat sales index is consistent with the findings of Genesove (2003) who uses data from the American Housing Survey and shows that rents on the same units are sticky, especially when tenants do not change. Moreover, since landlords tend to postpone maintenance works, repeat rents on the same unit suffer from unaccounted depreciation.

One way to limit this problem is to exclude properties whose “holding period” (the time between two sales or two rentals) is below a certain threshold. I choose an arbitrary threshold of 1000 days (corresponding to approximately 3 years). The result is displayed in Figure 8: the WRS indexes with no close transitions are more similar to the hedonic indexes, both for prices and rents.

To mirror the analysis above with hedonic regressions, Table 7 shows the outcome of running the regression described by equation 11, using both the unadjusted WRS procedure (upper panel) and the WRS with minimum holding period of 1000 days (lower panel). The WRS with a minimum holding period relies on smaller samples; nevertheless, coefficients under both approaches are similar. The output of Table 7 follows the results on aggregate volatility. Bigger houses and expensive neighborhoods have higher idiosyncratic volatility, although the distinction between prime and non-prime neighborhoods is less sharp than the distinction between big and small properties. This is in contrast with the dividend discount model, which would predict a *lower* idiosyncratic volatilities for those properties that display higher price-rent ratios.

Survey evidence The difficulty of reconciling these results with the dividend discount model might be due to the way in which expectations are measured. The analysis so far has followed the common practice of studying the historical trend of economic variables and then assume that expectations reflect this trend. An alternative approach would be to measure people’s expectations directly. John D Wood & Co., whose Sale and Rental Dataset is used in the present analysis, conducts every six months an online survey of the members of its mailing list.¹⁸ The January 2012 survey contains a couple of questions on local price and rent expectations,

formulated as follows:

“The next few questions are about nominal house prices in the area where you live.

Please enter the first part of your postcode: ----

- In terms of nominal value, what do you think will happen to *house prices* in your area after 1 year?

- In terms of nominal value, what do you think will happen to *rents* in your area after 1 year?”

Both expectation questions are followed by a drop-down menu where the respondents can choose a percentage. Figure 9 shows the frequency of each answer.

For the analysis, I exclude respondents who live outside London and I divide the remaining respondents in two groups: those living in London but not in a prime neighborhood, and those living in a London prime neighborhood.¹⁹ The definition of prime neighborhood is the same as the one in the rest of the paper, namely an address belonging to the following six postcode districts: SW1, SW3, SW7, W1, W8, W11. With this information, I can check whether the high price-rent ratios of prime neighborhoods are correlated with high price or rent growth expectations, in accordance with the dividend pricing model.²⁰ Since there are no explicit questions on rent risk, I take the dispersion of rent expectations as a measure of rent uncertainty. This approach is consistent with the empirical literature that looks at disagreement about inflation (Mankiw, Reis and Wolfers, 2003) or the stock market (Vissing-Jorgensen, 2003). Figure 9 shows that disagreement about house prices and rents can be substantial: taken together, respondents fill almost the entire range of possible answers, with round numbers (“-10% or more”, “-5%”, “0%”, “+5%”, “+10% or more”) being chosen more often.²¹

The upper half of Table 8 shows the differences in price and rent expectations between prime and other neighborhoods. In terms of price expectations, respondents in prime neighborhoods are slightly more optimistic than other Londoners, but the difference is not significant. In terms of rent expectations, people living in London non-prime neighborhoods expect slightly higher growth (the difference is significant at the 10% level). The standard deviation of rent expectations is significantly higher for prime London than other parts of London.

Therefore, according to the survey, rent uncertainty is higher for London prime neighborhoods but expected rent growth is not. This is again inconsistent with the dividend discount model. Clearly, the evidence presented here is only suggestive. Nevertheless, the respondents to this survey are likely to be representative of the house buyers and sellers of this particular area.

Housing submarkets

As noted by Sinai and Souleles (2005), the traditional dividend pricing model ignores that housing is a necessary consumption good and all households must either rent or own. From this perspective, higher rent volatility might *increase* the demand for housing, because it induces people to choose homeownership as a way to insure against future rent changes. In places with inelastic housing supply, such as London, this insurance motive results in higher price-rent ratios rather than higher homeownership rates. Sinai and Souleles (2005) show that, consistently with this model, higher rent volatility is associated with higher price-rent ratios across US cities.

In this paper I use data from London to show that the finding of Sinai and Souleles (2005) also holds *within* cities, at the submarket level. In Central London, households looking for small properties face thick markets both in sales and rentals. By contrast, households looking for big properties face a thin rental market and are pushed toward buying. Thin markets are more volatile and, as in Ngai and Tenreyro (2009), are less likely to generate good matches between property characteristics and people’s tastes. While Ngai and Tenreyro look at the thick vs thin market distinction over time, I look at it over the cross-section of property types.

In a recent paper, Bar-Isaac and Gavazza (2013) look at the Manhattan rental market and classify properties according to their “uniqueness” or “atypicality”—i.e. how much their features (e.g. size, location, number of rooms) depart from the average characteristics of properties in that area. They find that owners of more atypical properties are more likely to sign exclusive agreements with real estate agents, because renting out atypical properties requires more marketing effort. In the context of the present paper, the most expensive properties in the dataset might be characterized as atypical: they are in neighborhoods where it is very difficult to build new housing units; as the Appendix shows, when new properties are built, it is usually small

flats, not expensive houses. Atypicality might provide a premium in the sale market but, at the same time, might induce a discount in the rental market, leading to lower rental yields.

Conclusion

This paper presents novel findings on house prices and rents at the individual-property level. Price-rent ratios are shown to be higher for bigger properties and properties located in more expensive neighborhoods.

The main contribution of this paper is the empirical methodology, which consists of two steps: first, I run hedonic regressions on both price and rent data and check if there are statistically significant differences in the coefficients. Second, to avoid any bias caused by unobserved heterogeneity between sale and rental properties, I restrict the analysis on those properties for which it is possible to observe a sale and a rental during a short time span (6 months). By measuring prices and rents *on the same property at approximately the same time*, I can regress price-rent ratios on the same characteristics used in the hedonic regression and compare the estimated coefficients. Reassuringly, the coefficients obtained under the two methods are very similar.

While the main objective of this work is empirical, in the last part of the paper I explore several possible theoretical explanations for the stylized fact that I uncover. I divide the potential mechanisms in three groups: explanations based on the difference between gross and net yields, explanations based on asset pricing, and explanations based on housing submarkets. I measure rent risk at both the aggregate and individual level, and find that, in accordance with the hedging model of Sinai and Souleles, 2005, the properties with higher price-rent ratios are those with higher rent risk.

Consistently with the finance literature, I measure risk as price volatility, which is also the approach of Sinai and Souleles. However, the hedging model leaves open the possibility that other kinds of risk play a role in the renting vs. buying decision. For instance, search costs: a household looking for a 4-bedroom house to rent is not only worried about changes in rental

prices, but also about *finding* a 4-bedroom house to rent. Moreover, households might differ in their risk preferences. Workers whose income covaries positively with rents are less sensitive to rent volatility (Ortalo-Magné and Rady, 2002). Families with children—who usually demand bigger properties—are more risk averse (Banks, Blundell, Oldfield and Smith, 2010). Future research should expand on these different aspects of rent risk and housing market liquidity.

References

- Ambrose, B. W., P. Eichholtz and T. Lindenthal. 2013. House Prices and Fundamentals: 355 Years of Evidence. *Journal of Money, Credit and Banking* 45 (2): forthcoming.
- Association of Residential Letting Agents. 2012. Review and Index for Residential Investment. Quarterly report.
- Bailey, M. J., R. F. Muth and H. O. Nourse. 1963. A Regression Method for Real Estate Price Index Construction. *Journal of the American Statistical Association* 58 (304): 933–942.
- Banks, J., R. Blundell, Z. Oldfield and J. P. Smith. 2010. House Price Volatility and the Housing Ladder. IZA Discussion Papers 5173. Institute for the Study of Labor (IZA).
- Bar-Isaac, H. and A. Gavazza. 2013. Brokers' Contractual Arrangements in the Manhattan Residential Rental Market. MPRA Paper 43967. University Library of Munich, Germany.
- Bracke, P. and E. W. Pinchbeck. 2013. Embedded Options in the Housing Market: Evidence from the 1993 Leasehold Reform. Mimeo. London School of Economics.
- Calhoun, C. 1996. OFHEO House Price Indexes: HPI Technical Description. Tech. rep. Office of Federal Housing Enterprise Oversight.
- Campbell, S., M. Davis, J. Gallin and R. Martin. 2009. What Moves Housing Markets: A Variance Decomposition of the Rent–Price Ratio. *Journal of Urban Economics* 66 (2): 90–102.
- Case, K. E. and R. J. Shiller. 1989. The Efficiency of the Market for Single-Family Homes. *American Economic Review* 79 (1): 125–37.
- Clapp, J. M. and C. Giaccotto. 1992. Estimating Price Indices for Residential Property: a Comparison of Repeat Sales and Assessed Value Methods. *Journal of the American Statistical Association* 87 (418): 300–306.
- Court, A. 1939. Hedonic Price Indexes with Automotive Examples. In: *The Dynamics of Automobile Demand*. New York: General Motors, 98–119.

Davis, M. A. and J. Heathcote. 2007. The Price and Quantity of Residential Land in the United States. *Journal of Monetary Economics* 54 (8): 2595–2620.

Ferreira, F. and J. Gyourko. 2011. Anatomy of the Beginning of the Housing Boom: U.S. Neighborhoods and Metropolitan Areas, 1993-2009. Working Paper 17374. National Bureau of Economic Research.

Flavin, M. and T. Yamashita. 2002. Owner-Occupied Housing and the Composition of the Household Portfolio. *The American Economic Review* 92 (1): 345–362.

Gallin, J. 2008. The Long-Run Relationship Between House Prices and Rents. *Real Estate Economics* 36 (4): 635–658.

Garner, T. and R. Verbrugge. 2009. Reconciling User Costs and Rental Equivalence: Evidence from the US Consumer Expenditure Survey. *Journal of Housing Economics* 18 (3): 172–192.

Genesove, D. 2003. The Nominal Rigidity of Apartment Rents. *The Review of Economics and Statistics* 85 (4): 844–853.

Glaeser, E. L. and J. Gyourko. 2007. Arbitrage in Housing Markets. Working Paper 13704. National Bureau of Economic Research.

Goetzmann, W. N. and M. Spiegel. 1995. Non-Temporal Components of Residential Real Estate Appreciation. *The Review of Economics and Statistics* 77 (1): 199–206.

Griliches, Z. 1961. Hedonic Price Indexes for Automobiles: An Econometric of Quality Change. Tech. rep.

Halket, J. and M. Pignatti. 2012. Housing Tenure Choices with Private Information. Discussion Paper 717. University of Essex.

Hill, R. 2012. Hedonic Price Indexes for Residential Housing: A Survey, Evaluation and Taxonomy. *Journal of Economic Surveys*.

- Hill, R. J. and I. A. Syed. 2012. Hedonic Price-Rent Ratios, User Cost, and Departures from Equilibrium in the Housing Market. Graz Economics Papers 2012-08. Karl-Franzens University Graz, Department of Economics.
- Hudomiet, P., G. Kézdi and R. Willis. 2011. Stock Market Crash and Expectations of American Households. *Journal of Applied Econometrics* 26 (3): 393–415.
- Hwang, M., J. Quigley and J. Son. 2006. The Dividend Pricing Model: New Evidence from the Korean Housing Market. *The Journal of Real Estate Finance and Economics* 32 (3): 205–228.
- Joseph Rowntree Foundation. 1996. Index of Private Rents and Yields. Housing Research 194.
- Mankiw, N., R. Reis and J. Wolfers. 2003. Disagreement about Inflation Expectations. *NBER Macroeconomics Annual*, 209–248.
- Ngai, R. and S. Tenreyro. 2009. Hot and Cold Seasons in the Housing Market. CEP Discussion Paper 922. Centre for Economic Performance, LSE.
- Ortalo-Magné, F. and S. Rady. 2002. Tenure Choice and the Riskiness of Non-Housing Consumption. *Journal of Housing Economics* 11 (3): 266–279.
- Ortalo-Magné, F. and S. Rady. 2006. Housing Market Dynamics: On the Contribution of Income Shocks and Credit Constraints. *The Review of Economic Studies* 73 (2): 459–485.
- Plazzi, A., W. Torous and R. Valkanov. 2011. Exploiting Property Characteristics in Commercial Real Estate Portfolio Allocation. *The Journal of Portfolio Management* 35 (5): 39–50.
- Rosen, S. 1974. Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *The Journal of Political Economy* 82 (1): 34–55.
- Sinai, T. and N. Souleles. 2005. Owner-Occupied Housing as a Hedge against Rent Risk. *The Quarterly Journal of Economics* 120 (2): 763–789.
- Vissing-Jorgensen, A. 2003. Perspectives on Behavioral Finance: Does ‘Irrationality’ Disappear with Wealth? Evidence from Expectations and Actions. *NBER Macroeconomics Annual*, 139–208.

Notes

¹ In 2008 residential real estate constituted 39% of households' assets in the U.K. (Survey of Assets and Wealth) and 29% of households' assets in the U.S. (Flows of Funds).

²Differently from many advanced economies and the rest of the United Kingdom, nominal house prices in Central London are currently higher than in 2007 (the previous peak).

³The most common form of rental contract, the “assured shorthold tenancy”, leaves landlords and renters free to renegotiate any rental increase or decrease at the end of the rental period (usually one year). See http://www.direct.gov.uk/en/HomeAndCommunity/PrivateRenting/Tenancies/DG_189101

⁴The results are also consistent with non-academic researches. In the UK, for instance, reports from several sources, such as the Joseph Rowntree Foundation (1996) and the Association of Residential Letting Agents (2012), also show that price-rent ratios are higher for bigger properties (houses against flats) and expensive regions (London against the rest of the UK).

⁵<http://www.johndwood.co.uk/>. John D Wood & Co. was established in 1872 and has now 20 offices: 14 in London and 6 in the countryside. UK real estate agencies provide several services ranging from assistance in selling properties to management of rental units.

⁶Appendix Table 9 shows detailed statistics on the area, gathered from public sources. In addition to the private-rented sector, 30% of the housing stock is rented at subsidized prices by local authorities or housing associations. This part of the market is not included in the JDW Dataset.

⁷It is commonly believed among real estate practitioners that the price difference between a freehold (i.e., not subject to leasehold) and leasehold property is negligible for leaseholds longer than 80 years. Preliminary work by Bracke and Pinchbeck (2013) confirms this intuition.

⁸There is no clear pattern to model which observations have or don't have floor area information. Missing values are distributed randomly.

⁹Compared to the JDW Sales dataset, the Land Registry does not contain important information on housing characteristics, such as floor area. This is why the JDW Sales dataset is the preferred data source for the main part of the analysis. Appendix contains the descriptive statistics for the Land Registry; the type of properties transacted and the evolution of prices are similar to those of the JDW Sales Dataset.

¹⁰The 2005 file of the Land Registry does not contain individual addresses but only postcodes (corresponding to 10-20 properties).

¹¹Only 2.5% of the properties in the sample have more than 4 bedrooms. Properties with more than 10 bedrooms are discarded as outliers

¹²The data allow me to measure *gross* price-rent ratios, i.e. price-rent ratios which do not take into account maintenance expenses and, for rented properties, vacancies. If these were higher for smaller properties, *net* rent yields (rent-price ratio net of costs) could be more similar than what suggested by their gross counterparts. I will return to these issues in section .

¹³See for instance “London buyers find streets paved with gold”, Financial Times, 13 March 2011.

¹⁴Likewise, I do not expand on behavioral explanations—the fact that there might be a homeownership premium paid on top of the price for housing services, and this premium might be positively correlated with size and location prestige. These behavioral mechanisms might well play a role in Central London, but the facts presented here are so persistent and general that a convincing explanation cannot rely on behavioral arguments only.

¹⁵According to the U.K. Wealth and Assets Survey, only 10% of households own property other than their main residence. Similarly, the English Private Landlord Survey of 2010 reveals that 78% of landlords owns just one property for rent.

¹⁶In practice, the Case and Shiller (1989)’s procedure involves three steps: first, running an OLS regression to estimate Equation 11; second, regressing the resulting residuals on a constant (which will provide an estimate $2\sigma_u^2$) and the $(T - t)$ term (which will provide an estimate for σ_h^2); third, estimating Equation 11 again running a GLS regression where observations are weighted by the inverse of the square root of the predicted residuals.

¹⁷A substantial literature (e.g. Clapp and Giaccotto, 1992; Goetzmann and Spiegel, 1995) addresses the issue of sample selection bias in repeat sales indexes—an issue that is especially important when indexes are estimated over short periods of time.

¹⁸House price expectations are rarely surveyed. This is contrast with inflation expectations, which are regularly surveyed by Central Banks and other institutions (Mankiw et al., 2004).

¹⁹The question on the postcode appears at the very beginning of the survey and 95% of people that clicked on the survey link filled that question. The last part of the questionnaire contains questions on the socio-demographic characteristics of respondents. A table with summary statistics is shows in Appendix .

²⁰The survey makes no distinction between properties with different physical characteristics, e.g. flats vs houses. Hence, I can only test the part of results that relates to differences between neighborhoods, not the one regarding differences between properties of different sizes.

²¹This is a common feature of expectation surveys (Hudomiet, Kézdi and Willis, 2011).

Appendices

Housing Statistics for Central-Western London

The first two columns of Table 9 refer to the London local authorities covered by the JDW Dataset (Camden, Westminster, Kensington and Chelsea, Hammersmith and Fulham, and Wandsworth), the third and the fourth columns refer to the whole London area, and the fifth and sixth columns refer to England.

The upper panel takes data on sales from the 2011 Land Registry. In England as a whole, houses constitute 81% of sales, whereas they are only half of sales in London, and only one quarter of sales in Central-Western London. The median sale price in Central-Western London is more than two times and a half the median English price.

The middle panel takes data on housing tenure from the 2001 Census. Going from England to London and then to Central-Western London, the percentage of owner occupied properties goes down, and the percentage of privately rented properties goes up. A quarter of properties in Central-Western London belong to the privately rented market. The percentage of properties rented by a social landlord (either a local authority or a registered housing association) is also higher in London and Central-Western London.

The bottom panel takes data on house building from the U.K. Communities and Local Government Department.²² (These data are not available at the local authority level). The figures show that, both in England and London, house building tend to focus more on flats than houses, compared to the composition of the existing stock. Within flats, most of the building activity is centered on 2-bedroom flats.

The other two figures of the Appendix focus on the subset of observations in the Land Registry that belong to the postcode districts listed in Table 1. Figure 10 shows the quantity of registered sales in the area during the 2005–2011 period; figure 11 shows the hedonic and repeat sales house price indexes computed for the same period. Both figures display patterns very similar to the ones seen in the JDW Dataset.

Summary statistics for expectation survey respondents

The January 2012 John D Wood & Co. online survey of expectations asked respondents for many demographic information, which are summarized below. Respondents are mostly males, married with children, graduated, and homeowners. The sample is not representative of the general UK population, but is reasonably consistent with the expected profile of a home buyer in Central London.

The characteristics in Panel A were asked at the beginning of the online questionnaire, while the characteristics in Panel B were asked at the end. It is common for a percentage of respondents of online questionnaires to drop out of the survey before the end. This explains the lower number of observations for characteristics listed in Panel B.

Figure 1: Geographical coverage of the JDW Dataset

Notes: The figure shows the London local authorities covered by the JDW Dataset: Camden (C), Westminster (We), Kensington and Chelsea (K), Hammersmith and Fulham (H), and Wandsworth (W).

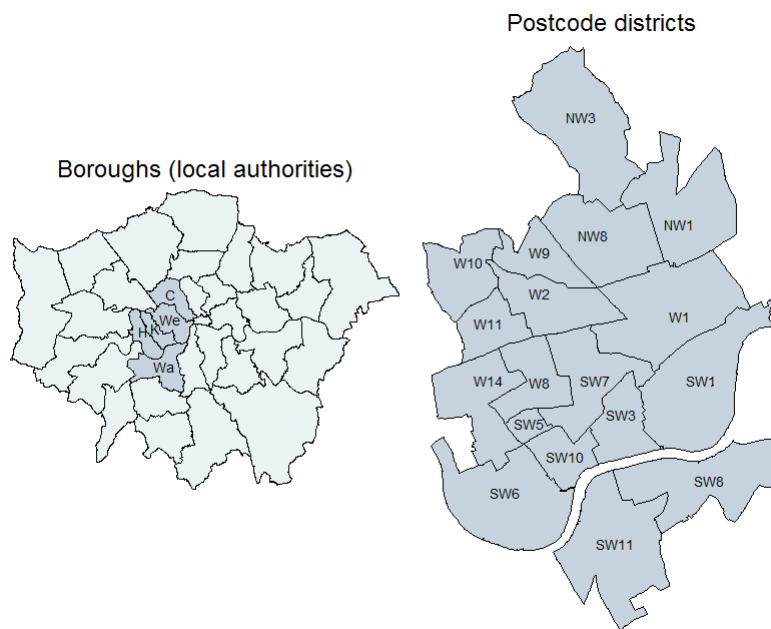
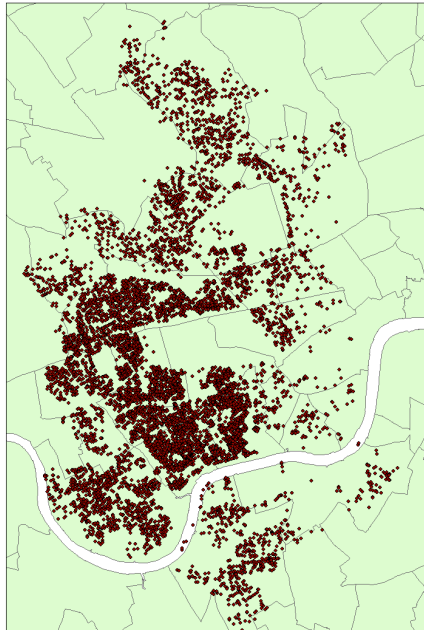


Figure 2: Observations in the JDW Dataset

Notes: Property addresses from the JDW Dataset were geocoded using Google Maps and then overlapped with a map of London postcode districts. The map is restricted to the area covered by the data.

(a) Sales



(b) Rentals

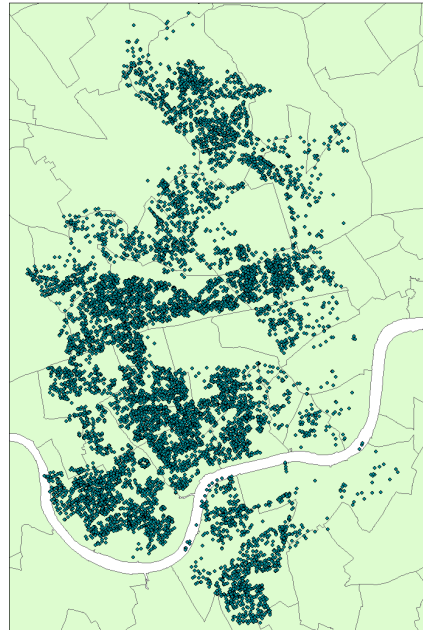


Figure 3: Sale and rental contracts per quarter

Notes: The figure shows the number of sale and rental observations in the JDW Dataset grouped by quarter.

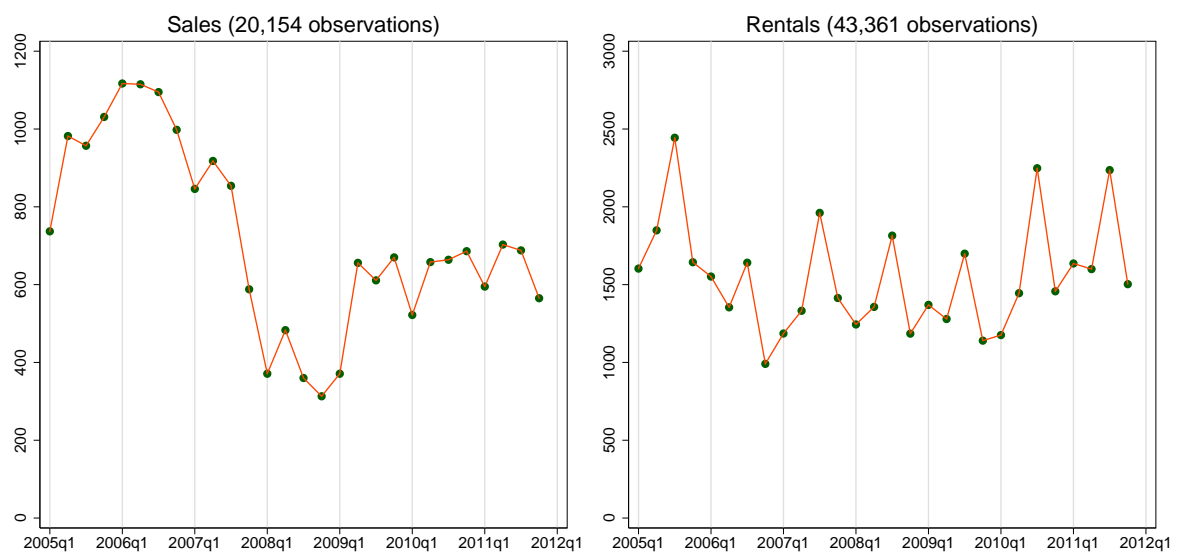


Figure 4: Observations in the Matched Dataset

Notes: The figure plots the price and rent for properties which were sold and rented out within a 6-month time window. The continuous line represents the linear relation $\text{Price} = 1000 \times \text{Weekly rent}$.

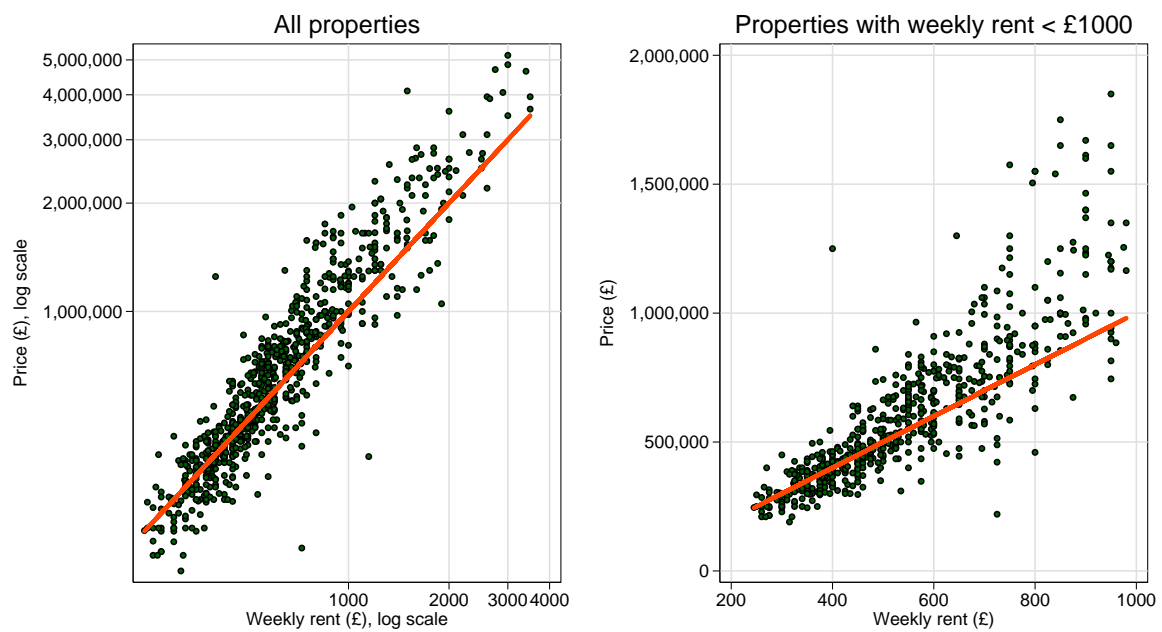


Figure 5: Price and rent indexes, price-rent ratios

Notes: The figure on the left plots the price and rent indexes computed using the main hedonic regression. The figure on the right plots two indexes for the price-rent ratio: one represents “implied” price-rent ratios computed using all the sale and rental observation as in equation 8, the other one comes from the price-rent ratios estimated in the Matched Dataset.



Figure 6: Time-varying hedonic prices

Notes: The panels show some of the estimates for the time-varying hedonic coefficients computed estimating equation 9 on prices and rents.

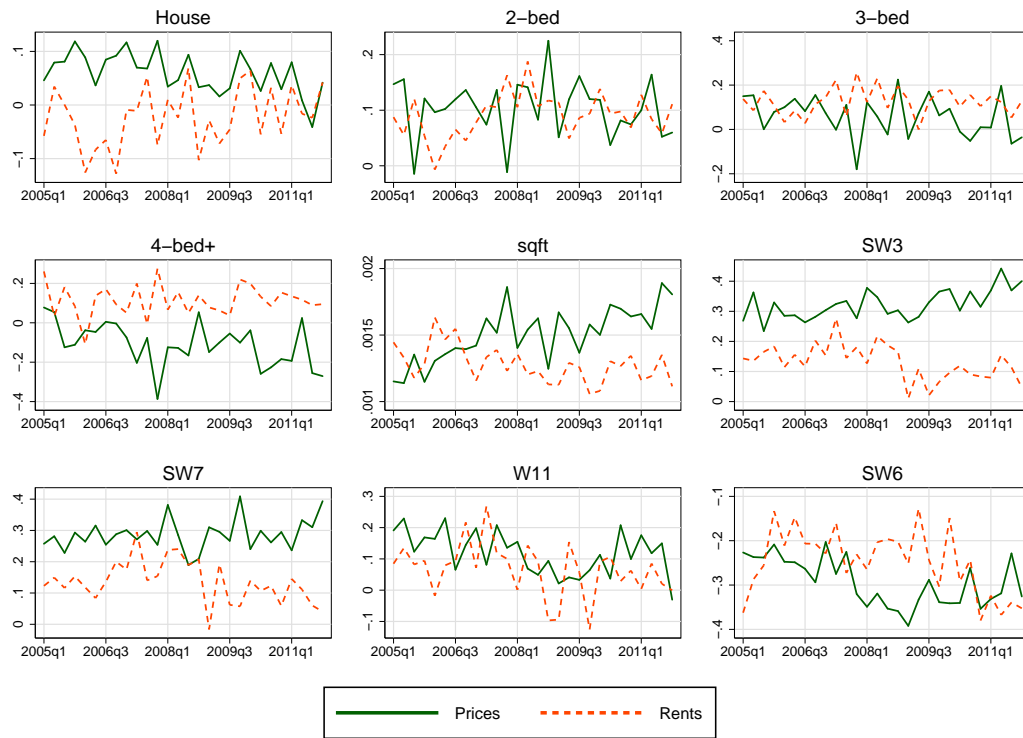


Figure 7: Growth and Volatility by Housing Categories

Notes: The charts represent house price and rent indexes computed separately for different property categories, as in equation 10.

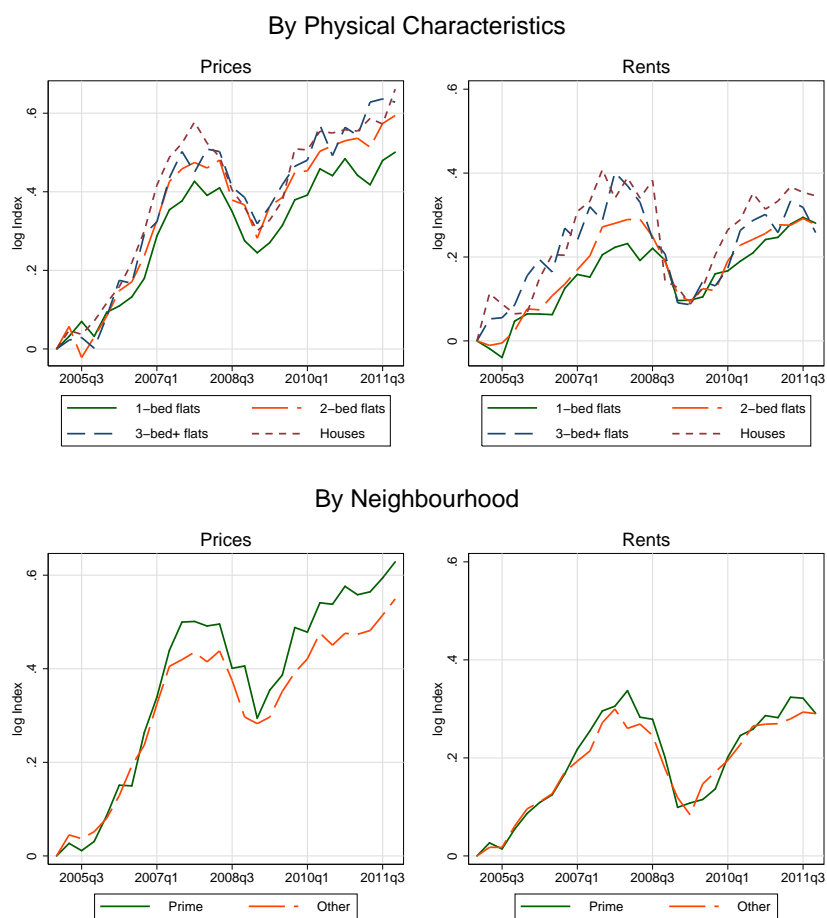


Figure 8: Hedonic and repeat sales indices

Notes: The figure shows the repeat sales and repeat rentals indexes computed using the weighted repeat sales (WRS) methodology of Case and Shiller (1989). The indexes are also computed excluding repeat transactions which happen within less than 1000 days one from another, and compared with the hedonic indexes.

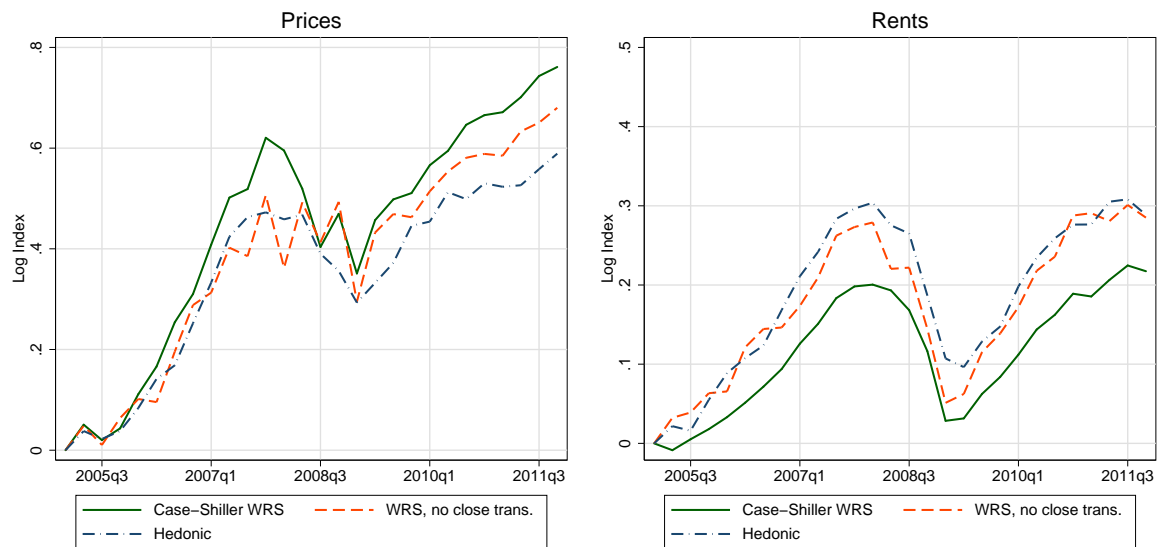


Figure 9: Survey expectations

Notes: The questions are “In terms of nominal value, what do you think will happen to *house prices* in your area after 1 year?” and “In terms of nominal value, what do you think will happen to *rents* in your area after 1 year?” The answers are the bottom and top of the range are “-10% or more” and “+10% or more”.

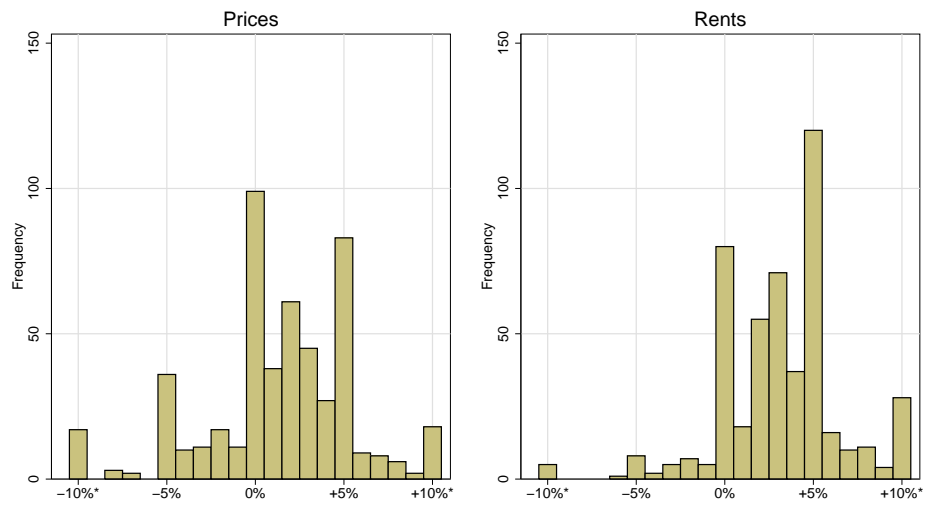


Figure 10: Quarterly registered sales (Land Registry)

Notes: The figure shows the total number of sales in the Land Registry for the postcodes included in the JDW Dataset. Sales are grouped by quarter and follow the same pattern as in the JDW Dataset (displayed in Figure 3).

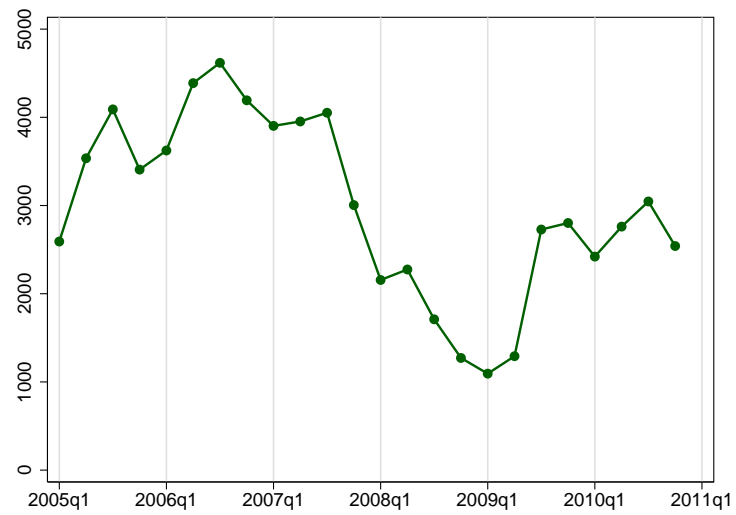


Figure 11: Price Indexes (Land Registry)

Notes: Indexes are normalized to zero in 2006Q1, starting date of the repeat sales index (the data for 2005 contain individual properties, but no individual addresses). In the 2006–2010 period, 2,650 properties in the Land Registry sold twice, 117 sold three times and 4 sold four times. The hedonic index is computed using the two variables available in the Land Registry: property type (flat, terraced house, semi-detached house, and detached house) and whether the property is new.



Table 1: JDW Datasets: Summary statistics

Notes: The Hedonic Dataset contains the properties for which the main hedonic regressions are run, the Matched Dataset contains the properties that were sold and rented out within 6 months, and the Repeat Transactions Dataset contains the properties that were sold or rented multiple times. Observations in the Hedonic dataset have no missing variables. 65% of Matched units, 56% of observations in the Repeat transactions Dataset (Rentals), and 4% of observations in the Repeat transactions Dataset (Sales) have missing floor areas.

	Hedonic Dataset		Repeat Transactions		Matched Units
	Sales	Rentals	Sales	Rentals	Sales & Rentals
	(1)	(2)	(3)	(4)	(5)
Observations	18,864	15,811	2,231	18,710	1,407
Median price	718,475		950,000		539,637
Median rent (in 2005 £; rent per week)		542		507	546
Floor area (sqft)	1059	879	1244	850	791
<i>Property type (%)</i>					
1-bed flat	0.19	0.29	0.17	0.37	0.33
2-bed flat	0.35	0.39	0.29	0.38	0.41
3-bed+ flat	0.22	0.16	0.20	0.11	0.13
House	0.25	0.16	0.35	0.14	0.13
<i>Postcode districts (%)</i>					
NW1	0.03	0.02	0.03	0.02	0.03
NW3	0.03	0.03	0.02	0.04	0.01
NW8	0.04	0.02	0.04	0.03	0.02
SW1	0.16	0.13	0.15	0.14	0.19
SW10	0.07	0.06	0.09	0.05	0.09
SW11	0.03	0.03	0.01	0.03	0.01
SW3	0.10	0.13	0.14	0.12	0.11
SW5	0.03	0.05	0.04	0.03	0.06
SW6	0.06	0.06	0.03	0.07	0.03
SW7	0.09	0.10	0.13	0.09	0.10
SW8	0.02	0.06	0.01	0.03	0.03
W1	0.07	0.11	0.09	0.14	0.08
W10	0.01	0.00	0.00	0.01	0.01
W11	0.04	0.04	0.03	0.05	0.03
W14	0.03	0.03	0.02	0.03	0.03
W2	0.09	0.05	0.08	0.06	0.08
W8	0.08	0.06	0.08	0.06	0.08
W9	0.02	0.01	0.02	0.01	0.01

Table 2: Properties sold and rented out within 6 months

Notes: The table shows the number of properties for which a sale-rental match was found. The first two columns refer to matches that were found within the JDW Dataset. The third and fourth columns refer to matches between JDW rentals and sales in the Land Registry. The last two columns contain all the matches; the numbers in the fifth column are less than the sum of the first and third columns because some matches are duplicates and are removed. For each year and each subsample the average price-rent ratio is computed.

	JDW rentals - JDW sales		JDW rentals - Land Registry		All	
	# Properties	Average price-rent ratio	# Properties	Average price-rent ratio	# Properties	Average price-rent ratio
2006	98	18.81	165	19.14	259	19.02
2007	132	20.39	347	19.40	475	19.67
2008	56	21.15	214	20.20	270	20.40
2009	96	23.34	109	24.10	203	23.74
2010	163	23.16	224	21.99	384	22.48

Table 3: Regression of price-rent ratios on rents (Matched Dataset)

Notes: Regressions are estimated on properties that were both sold and rented out within 6 months. Standard

	$\log\left(\frac{\text{Price}}{\text{Rent}}\right) = \alpha + \beta \log \text{Rent}$		
	(1)	(2)	(3)
	Baseline	Year dummies	Year interactions
β	0.084*** (0.015)	0.081*** (0.014)	
β_{2006}			0.060* (0.034)
β_{2007}			0.060** (0.026)
β_{2008}			0.015 (0.039)
β_{2009}			0.14*** (0.039)
β_{2010}			0.13*** (0.030)
Year dummies		✓	✓
N	1,407	1,407	1,407

errors in parentheses.

Table 4: Hedonic regressions

Notes: Quarterly time dummies are used for the complete dataset and half-year dummies for the matched dataset. The baseline property is a 1-bedroom flat in W2. The coefficients in the third column are computed appending together the Sales and Rentals dataset, as shown in equation 8. The estimation of column 4 does not use all the observations of the Matched dataset, but only those where floor area is present. Standard errors in parentheses.

	$y_{hit} = \alpha_h + \beta_h X_{it} + \lambda_{ht} + \varepsilon_{hit}$			2006–2010
	2005–2011			
	(1)	(2)	(3)	(4)
	JDW Sales $y = p_s$	JDW Rentals $y = p_r$	<i>Implied</i> (1) - (2)	Matched $y = p_s - p_r$
House	0.065*** (0.007)	-0.014* (0.008)	0.079*** (0.011)	0.147*** (0.033)
2-bed	0.118*** (0.008)	0.116*** (0.007)	0.002 (0.010)	0.027 (0.027)
3-bed	0.090*** (0.010)	0.158*** (0.011)	-0.068*** (0.015)	-0.033 (0.045)
4-bed	-0.083*** (0.014)	0.124*** (0.016)	-0.206*** (0.021)	-0.150** (0.066)
Floor area (sqft*10 ⁻³)	1.454*** (0.012)	1.160*** (0.013)	0.293*** (0.018)	0.197*** (0.067)
Floor area squared	-0.156*** (0.002)	-0.128*** (0.002)	-0.028*** (0.003)	-0.015 (0.017)
<i>Postcode:</i>				
SW3	0.309*** (0.011)	0.122*** (0.012)	0.187*** (0.017)	0.107*** (0.040)
SW7	0.288*** (0.011)	0.142*** (0.013)	0.146*** (0.017)	0.158*** (0.041)
W8	0.245*** (0.012)	0.104*** (0.014)	0.140*** (0.019)	0.127*** (0.044)
W1	0.189*** (0.012)	0.124*** (0.013)	0.065*** (0.018)	0.081* (0.044)
W11	0.123*** (0.014)	0.071*** (0.016)	0.052*** (0.022)	-0.091 (0.060)
SW1	0.123*** (0.010)	0.122*** (0.012)	0.000 (0.016)	0.119*** (0.036)
SW10	0.098*** (0.012)	-0.029** (0.014)	0.126*** (0.019)	0.089** (0.042)
SW5	0.074*** (0.015)	-0.011 (0.015)	0.086*** (0.021)	0.080* (0.047)
NW8	-0.004 (0.015)	-0.023 (0.020)	0.019 (0.025)	0.135** (0.064)
SW8	-0.022 (0.019)	-0.004 (0.014)	-0.018 (0.023)	0.059 (0.058)
NW1	-0.076*** (0.016)	-0.029 (0.020)	-0.047** (0.025)	-0.054 (0.060)
NW3	-0.067*** (0.016)	-0.108*** (0.017)	0.041** (0.024)	0.092 (0.096)
W14	-0.121*** (0.015)	-0.175*** (0.018)	0.054** (0.024)	0.054 (0.062)
W9	-0.162*** (0.017)	-0.160*** (0.029)	-0.002 (0.034)	-0.050 (0.084)
W10	-0.247*** (0.031)	-0.172*** (0.047)	-0.075* (0.056)	0.085 (0.116)
SW11	-0.275*** (0.017)	-0.306*** (0.020)	0.031 (0.026)	0.060 (0.092)
SW6	-0.285*** (0.014)	-0.226*** (0.014)	-0.059*** (0.020)	-0.061 (0.061)
Time dummies	✓	✓	✓	✓
<i>N</i>	18,864	15,811	34,675	494

Table 5: Time on the market for rental properties

Notes: The table displays the output of regressing the days a property has been on the market (waiting for a new tenant) on the log rent and the type of property (flat or house). Flats are the baseline category. All properties are in the Chelsea-Fulham area in Central London. Standard errors in parentheses.

	(1)	(2)	(3)
	Days	Days	logDays
Constant	89.920*** (3.550)	5.357 (44.497)	1.623*** (0.320)
log Rent		12.525* (6.908)	0.334*** (0.050)
House		28.692** (11.854)	0.128 (0.085)
Year dummies		✓	✓
<i>N</i>	2325	2325	2325

Table 6: Prices and rents: Growth and systemic risk

Notes: Results are from hedonic regressions with quarterly dummies. “Average growth” columns compute the average log difference between two consecutive quarters in the price or rent index. “Volatility” columns report the standard deviations of these log differences.

	Prices		Rents	
	Average growth	Volatility	Average growth	Volatility
<i>All</i>	<i>0.022</i>	<i>0.042</i>	<i>0.011</i>	<i>0.033</i>
1-bed Flats	0.019	0.045	0.010	0.036
2-bed Flats	0.022	0.051	0.010	0.034
3-bed+ Flats	0.023	0.059	0.010	0.059
Houses	0.024	0.053	0.013	0.070
Prime neighbourhood	0.023	0.052	0.011	0.039
Other neighbourhood	0.020	0.039	0.011	0.032

Table 7: Prices and rents: Idiosyncratic risk

Notes: The table shows the coefficients obtained by estimating the second stage of the WRS method of Case and Shiller (1989). Standard errors are in parentheses. The coefficient θ_1 is positive in the rent equations but negative in the price equation. While at odds with the Case and Shiller (1989)'s model, it is not uncommon to estimate negative θ_1 's in empirical work (Calhoun, 1996). These negative coefficients imply that very close transactions have high idiosyncratic volatility.

	$\nu_{cht}^2 = \theta_{c0} + \theta_{c1}(T - t)$					
	JDW Sales Dataset			JDW Rentals Dataset		
	θ_{c0}	θ_{c1}	Obs.	θ_{c0}	θ_{c1}	Obs.
Panel A: WRS						
<i>All</i>	0.077*** (0.008)	-0.021*** (0.007)	1,139	0.010*** (0.001)	0.011*** (0.001)	10,786
1-bed Flats	0.048*** (0.009)	-0.017** (0.008)	176	0.010*** (0.002)	0.005* (0.003)	3,965
2-bed Flats	0.077*** (0.022)	-0.033 (0.020)	286	0.008*** (0.001)	0.009*** (0.002)	4,007
3-bed+ Flats	0.056*** (0.011)	-0.006 (0.010)	201	0.007*** (0.002)	0.016*** (0.002)	1,147
Houses	0.066*** (0.010)	-0.011 (0.008)	395	0.010*** (0.004)	0.017*** (0.004)	1,450
Prime neighbourhood	0.084*** (0.013)	-0.018 (0.012)	703	0.009*** (0.001)	0.012*** (0.002)	6,465
Other neighbourhood	0.057*** (0.006)	-0.018*** (0.005)	435	0.011*** (0.002)	0.009*** (0.003)	4,321
Panel B: WRS with minimum holding period of 1000 days						
<i>All</i>	0.053*** (0.018)	-0.008 (0.012)	512	0.010*** (0.006)	0.009** (0.004)	2,390
1-bed Flats	0.022 (0.014)	-0.005 (0.008)	77	0.004 (0.005)	0.008** (0.004)	771
2-bed Flats	0.0087 (0.016)	0.008 (0.011)	133	0.014* (0.008)	0.004 (0.005)	879
3-bed+ Flats	0.073* (0.038)	-0.024 (0.024)	80	0.006 (0.013)	0.015* (0.009)	288
Houses	0.049 (0.033)	-0.004 (0.021)	197	0.024 (0.021)	0.006 (0.014)	381
Prime neighbourhood	0.056** (0.027)	-0.005 (0.017)	296	0.008 (0.007)	0.011** (0.005)	1,470
Other neighbourhood	0.032** (0.013)	-0.004 (0.008)	214	0.014 (0.010)	0.006 (0.007)	916

Table 8: Survey results

Notes: The questions are “In terms of nominal value, what do you think will happen to *house prices* in your area after 1 year?” and “In terms of nominal value, what do you think will happen to *rents* in your area after 1 year?” A prime neighborhood is one of the following six postcode districts: SW1, SW3, SW7, W1, W8, W11.

	Price Expectations ($E_{tp_{st+1}}$)			Rent Expectations ($E_{tp_{rt+1}}$)		
	Mean	(St. Dev.)	Obs.	Coeff.	(St. Dev.)	Obs.
	(1)	(2)	(3)	(4)	(5)	(6)
London, prime neighborhood	2.25	(4.37)	79	2.84	(3.81)	74
London, other neighborhood	2.10	(4.14)	189	3.78	(3.38)	183
<i>Mean Diff. (T-test)</i>	<i>0.15</i>	<i>(0.58)</i>		<i>-0.94*</i>	<i>(0.51)</i>	
<i>StDev Ratio (F-test)</i>		<i>1.12</i>			<i>1.28*</i>	

Table 9: General Housing Statistics

Notes: The table shows statistics on the housing stock for the part of London covered by the JDW Dataset, for the whole London area, and for England. All data sources are public.

	Cent.-West London		London		England	
	#	%	#	%	#	%
Sales (Land Registry 2011)						
Flats	12,318	0.75	46,832	0.51	121,092	0.19
Houses	4,148	0.25	44,891	0.49	504,909	0.81
(Median price)	(£480,000)		(£287,000)		(£185,000)	
Stock (Census 2001)						
Owner occupied	188,191	0.44	1,675,690	0.58	13,920,429	0.71
Rented from private landlord	108,084	0.25	432,482	0.15	1,798,864	0.09
Rented from social landlord	132,352	0.31	790,371	0.27	3,940,728	0.20
New supply (Local statistics 2001–2011)						
1-bedroom flats			46,658	0.24	137,006	0.09
2-bedroom flats	<i>No statistics</i>		106,506	0.54	413,902	0.29
3-bedroom+ flats	<i>at Borough level</i>		10,433	0.05	14,421	0.01
Houses			35,237	0.18	879,721	0.61

Table 10: Summary statistics

Notes: The table shows the characteristics of respondents in the expectation survey. Questions in Panel A of the table were at the beginning of the questionnaire, while socio-demographic questions (Panel B) were at the end.

Variable	%	Obs.	Variable	%	Obs.
Panel A: Housing					
<i>Residence</i>			<i>Housing tenure</i>		
Outside UK	0.07	510	Homeowner (mortgage)	0.50	451
UK outside London	0.47	510	Homeowner (outright)	0.33	451
London, prime neighborhood	0.16	510	Renting	0.01	451
London, other neighborhood	0.30	510	Other	0.16	451
Panel B: Socio-demographic characteristics					
<i>Gender</i>			<i>Age</i>		
Male	0.69	293	Less than 31	0.06	294
			31-40	0.32	294
			41-50	0.25	294
<i>Marital status</i>			51-60	0.20	294
Single	0.15	294	61-70	0.15	294
Cohab. (child)	0.03	294	Over 70	0.02	294
Cohab. (no child)	0.06	294			
Married (child)	0.50	294	<i>Income</i>		
Married (no child)	0.16	294	Hhold income < £50,000	0.11	278
Separated/divorced	0.07	294	£50,000-100,000	0.29	278
Widowed	0.02	294	£100,000-200,000	0.30	278
			Over £200,000	0.29	278
<i>Education</i>			<i>Occupation</i>		
GSCE	0.05	295	Student	0.01	297
A-level / Bacc.	0.09	295	Employed	0.56	297
University degree	0.43	295	Self-employed	0.23	297
Masters	0.32	295	Looking for a job	0.01	297
PhD	0.04	295	Retired	0.12	297
Other	0.06	295	Other	0.08	297