SERC DISCUSSION PAPER 127

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

Philippe Bracke (SERC, LSE)

February 2013

This work is part of the research programme of the independent UK Spatial Economics Research Centre funded by a grant from the Economic and Social Research Council (ESRC), Department for Business, Innovation & Skills (BIS) and the Welsh Assembly Government. The support of the funders is acknowledged. The views expressed are those of the authors and do not represent the views of the funders.

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

Philippe Bracke*

February 2013

* SERC, Department of Economics, London School of Economics and Political Science

Acknowledgements
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.

Abstract
Using the proprietary dataset of a real estate agency, I analyse 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.

JEL Classifications: G10, R21, R31
Keywords: House prices, housing rents, price index

1 Introduction

The value of the entire stock of housing stands at 16 trillion dollars in the US and 3.9 trillion pounds in the UK,[1] making housing the biggest item among households' assets. Similarly, rents represent the major expenditure item for many households, amounting to 20-30% of monthly expenses (Genesove, 2003).[2] The booms and busts that affect house prices are not equally pronounced for rents. In fact, the recent housing boom was characterised by a significant rise in the price-rent ratio (Campbell et al., 2009), and historically rents are less volatile than house prices (Gallin, 2008), as dividends are less volatile than stock prices (Shiller, 1981). Understanding and modeling price-rent ratios is therefore crucial to improve our knowledge of the housing market.

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 neighbourhoods 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.[3] The area under study contains 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.[4]

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 exactly: 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.

The empirical analysis shows that price-rent ratios are higher for bigger properties and properties located in more expensive neighbourhoods. The finding that more expensive properties have higher price-rent ratios is consistent with Garner and Verbrugge (2009) who use answers from the US Consumer Expenditure Survey to compare self-reported rents and house values. In the UK, 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). The present paper contributes to the literature by presenting a detailed micro dataset

[1] U.S. data from the Federal Reserve Board's Flow of Funds Accounts, (table B100, number 49). U.K. data from `http://www.lloydsbankinggroup.com/media/pdfs/halifax/2012/1102_value.pdf`

[2] 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). According to the UK Office for National Statistics (`http://www.ons.gov.uk/ons/rel/family-spending/family-spending/family-spending-2012-edition/rft---table-2-8.xls`), "Expenditure on rent by renters" corresponds to 28% of income. In the US, the Bureau of Labour Statistics (`http://www.bls.gov/cex/2010/share/tenure.pdf`) puts this number at 25%.

[3] Differently 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).

[4] The 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`

and implementing a unique empirical methodology to uncover these patterns. Despite the geographical focus of the dataset, the findings are not London-specific and time-specific; they are consistent with previous studies which concentrated on other geographical areas and periods.

Using micro local data to infer general features of housing markets is a common approach in housing research. For instance, Guerrieri et al. (2010) analyse house prices at the zip-code level in a group of US cities to propose a model of neighbourhood gentrification. Ferreira and Gyourko (2011) use US individual-level transaction data to produce local price indices and study the start of the recent house price boom. Due to the lack of reliable data, the analysis of rental prices at the micro local level has been so far limited, despite the importance of rents in determining housing market conditions. Price-rent ratios have been repeatedly measured and studied using aggregate data. For instance, Gallin (2008) uses city-level data to check if changes in price-rent ratios anticipate future price or rent growth, as the dividend discount model would predict. Verbrugge (2008) analyses aggregate price and rent indices to show a persistent divergence of rents and ex ante user costs.

A notable exception is Hwang et al. (2006), who use micro data on prices and rents from South Korea to test the dividend pricing model. Hwang et al. (2006) 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. While the study of Hwang et al. (2006) relies on the very specific features of the Korean housing market, my analysis is based on a market that functions similarly to most housing markets in the US and Europe.

In the last part of the paper I discuss the potential explanations for the findings on price-rent ratios. One possibility is that gross price-rent ratio differences hide differences in maintenance costs or vacancy rates: once these differences are taken into account, it could be that "net" price-rent ratios are actually quite homogeneous. I don't find evidence consistent with this view. 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 in Central London rent growth rates of bigger properties are not different from those of smaller properties, but their volatility is significantly higher. Similarly, rents are not growing faster in more expensive neighbourhoods, but are more volatile. 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 characterised by rental markets that are "thin" in a search-theory sense (Ngai and Tenreyro, 2009), so that finding and maintaining good matches is difficult.

I use price and rent indices 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 allows me also to measure idiosyncratic volatilities by restricting attention to properties that were sold or rented at least twice during the sample period and applying the weighted repeat sales estimator of Case and Shiller (1989). Since most people own or rent only one property, idiosyncratic volatilities of prices and rents might be a better risk measure than aggregate indices, but previous studies only measure rent risk at the aggregate level. 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 that is regularly carried out by the same real estate agency that has provided the property dataset.

The results of this paper are relevant both for consumers and investors. Returns to housing are given by the sum of capital gains and rental yields, where rental yields are defined as the inverse of price-rent ratios. The finding of different rental yields across property types is useful for real estate investors' portfolio management (Plazzi et al., 2011). Moreover, the recent crisis has interrupted the upward trend in homeownership in many countries such as the U.S. (Gabriel and Rosenthal, 2011) and the U.K. (Department for Communities and Local Government, 2012). The prospect of more concentration in the ownership of the housing stock makes these portfolio considerations quite relevant.

The rest of the paper is organised as follows. Section 2 describes the data. Section 3 presents the empirical methodology of the paper and shows the main results. Section 4 discusses the theories that can explain the main results. Section 5 concludes.

2 Data

The main dataset used in this paper comes from John D Wood & Co., a real estate agency that operates in London and the surrounding countryside.[5] I refer to these data as the JDW Dataset. Sometimes the empirical analysis is restricted to subsets of the JDW Dataset: the Matched Dataset and the Repeat Transactions Dataset. To match rented properties with sale transactions and to assess the representativeness of the JDW Dataset, I also use data from the Land Registry, which contains official records of all housing transactions in England and Wales. These datasets are described in the remaining of this section.

2.1 The JDW Dataset

The JDW Dataset includes observations from the Central-Western area of London. London is divided in 33 local authorities, which are responsible for running services such as schools, waste collection, and roads. The local authorities covered by the JDW dataset are Camden, Westminster, Kensington and Chelsea, Hammersmith and Fulham, and Wandsworth. These local authorities are shown on the left-hand side of Figure 1.

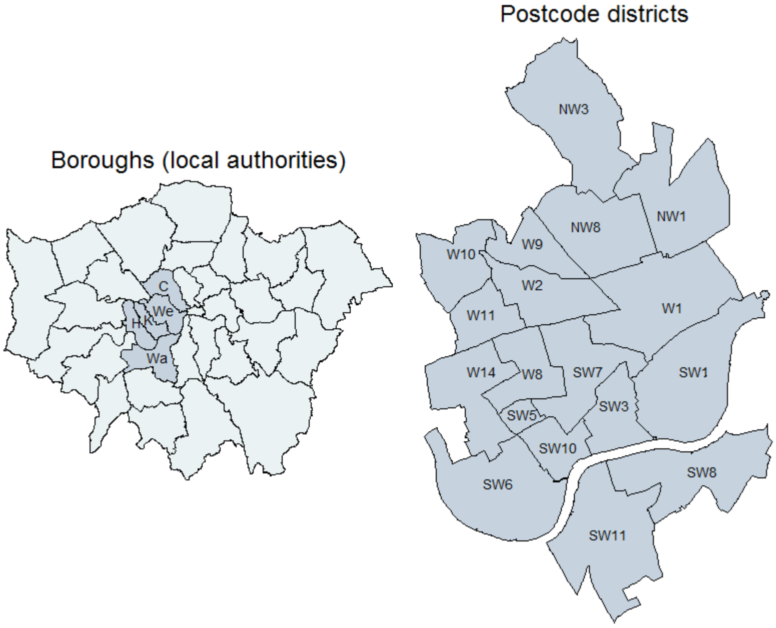
This area is one of the most densely populated in London. Most of the housing stock is made of flats rather than single-family houses. Approximately one fourth of dwellings are privately rented.[6] Appendix Table A1 shows detailed statistics on the area, gathered from public sources. A more detailed partition of this area can be obtained using postcode districts. In the U.K. postal code, 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. The right-hand side of Figure 1 shows the postcode districts included in the JDW Dataset. In the empirical analysis, I use postcode district dummies to capture the effect of location on house prices.

[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. Big agencies have valuation teams whose duties include keeping track of market trends. Agents assemble sale and rental data from their own records as well as from other agencies.

[6] In addition to the private-rented sector, 30% of the housing stock is rented at subsidised prices by local authorities or housing associations. This part of the market is not included in the JDW Dataset.

Figure 1: Geographical coverage of the JDW Dataset

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



To remove potential duplicate observations in the dataset, 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 travellers 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 of their transaction year. Figure 2 plots the sale observations on the London map.

The JDW Dataset contains only a fraction of the housing units present in the whole Central London area. In Appendix C, I compare some of the features of the JDW Sales Dataset with the Land Registry for the 2005-2010 period. Compared to the JDW Sales dataset, the Land Registry does not contain important information on housing characteristics, such as floor area.

The first two columns of Table 1 contain the summary statistics for the sold properties (Sales) and rented properties (Rentals). Consistently with the composition of housing stock in this part of London, the majority of housing units in the JDW Dataset are flats. There are more flats in Rentals (88%) than in Sales (76%). 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.

[7] It is commonly believed that the price difference between a freehold (i.e., not subject to leasehold) and leasehold property is negligible for leaseholds longer than 80 years.

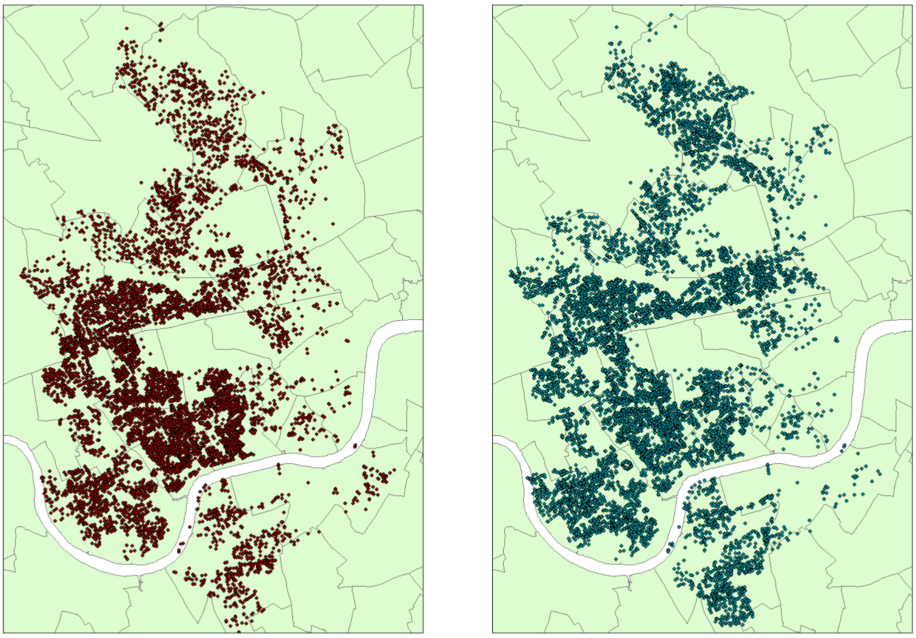
Table 1: JDW Datasets: Summary statistics

	Complete dataset		**Matched units**	**Repeat transactions**	
	Sales	Rentals	Sales & Rentals	Sales	Rentals
	(1)	(2)	(3)	(4)	(5)
Observations	20,154	43,361	1,661	1,233	18,710
Median price	694,323		532,746	920,000	.
Median rent		460	524		460
(in 2005 £; rent per week)					
Floor area (sqft)	1059	879	781	1245	863
Property type (%)					
1-bed flat	0.20	0.34	0.33	0.17	0.36
2-bed flat	0.35	0.39	0.41	0.29	0.38
3-bed+ flat	0.21	0.14	0.12	0.20	0.12
House	0.24	0.12	0.14	0.35	0.14
Postocde districts (%)					
NW1	0.03	0.02	0.03	0.03	0.02
NW3	0.03	0.05	0.05	0.02	0.04
NW8	0.05	0.03	0.04	0.04	0.03
SW1	0.15	0.13	0.15	0.15	0.14
SW10	0.07	0.04	0.06	0.09	0.05
SW11	0.03	0.04	0.03	0.01	0.03
SW3	0.09	0.10	0.10	0.14	0.12
SW5	0.03	0.04	0.06	0.04	0.04
SW6	0.06	0.07	0.08	0.03	0.07
SW7	0.09	0.08	0.08	0.12	0.09
SW8	0.02	0.03	0.02	0.01	0.03
W1	0.07	0.12	0.08	0.09	0.13
W10	0.01	0.01	0.01	0.00	0.01
W11	0.04	0.05	0.04	0.03	0.05
W14	0.03	0.03	0.03	0.02	0.03
W2	0.09	0.07	0.08	0.08	0.06
W8	0.08	0.06	0.07	0.08	0.06
W9	0.03	0.02	0.02	0.02	0.01

Figure 2: Observations in the JDW Dataset

Notes: Property addresses were geocoded using Google Maps.

(a) Sales (b) Rentals



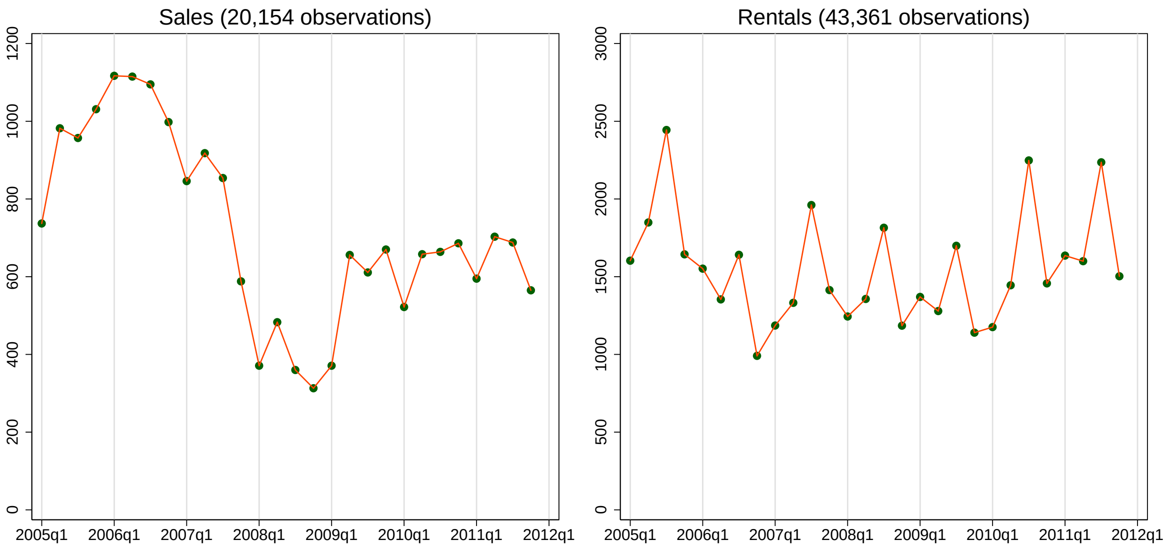
These facts are consistent with Linneman (1985)'s production-efficiency argument: smaller units demand less management costs and both landlords and households prefer them for renting.

Before proceeding to the main analysis, it is useful to measure the evolution of the number of transactions in the sale and rental market. Prices are not the only margin of adjustment in the housing market: volumes and liquidity are also important (Novy-Marx, 2009; Ngai and Tenreyro, 2009). Figure 3 shows the quarterly number of transactions in the JDW Sales and Rentals datasets. The number of sales varies a lot 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 time 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 much clearer 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 than for rentals.

2.2 The Matched Dataset

The Matched Dataset contains properties that appear both in the Sales and Rentals datasets, with the sale taking place between 0 and 6 months before the corresponding rental contract. To increase the number of matched observations, I also add properties that appear both in the JDW Rentals dataset and the Land Registry—again, with a maximum distance of 6 months between the sale and the subsequent rental. 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.

Figure 3: Sale and rental contracts per quarter



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. Since there can be multiple sales and multiple rents 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 2: Properties sold and rented out within 6 months

	JDW rent - JDW sales	JDW rent - Land Registry	All
2006	98	165	259
2007	132	347	475
2008	56	214	270
2009	96	109	203
2010	163	224	384

Table 2 shows how many matches are retrieved in each year, and the average rent-price ratios. Most matches come from the Land Registry. Some JDW matches are also found in 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 those years when analysing matched properties.[8]

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

Figure 4: Observations in the Matched Dataset

Notes: The continuous line represents the linear relation Price = 1000 × Weekly rent.

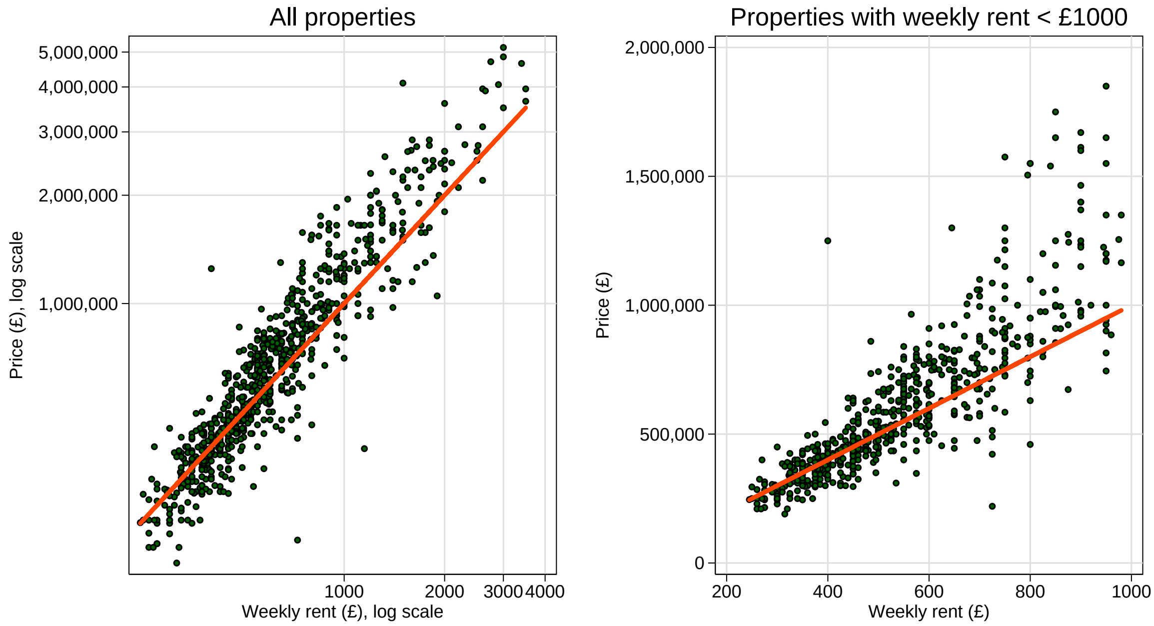


Figure 4 plots rents and prices for the observations of the Matched Dataset. The plot on the left contains all observations. Since the distribution of prices and rents is skewed to the right, the axes use a log scale. Notice that, if price-rent ratios were the same for expensive and cheap properties,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

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

where β represents $\frac{\mathrm{d}\ (\log \text{Price} - \log \text{Rent})}{\mathrm{d}\ \log \text{Rent}} = \frac{\mathrm{d}\ \log \text{Price}}{\mathrm{d}\ \log \text{Rent}} - 1$. Hence, a coefficient significantly greater than zero indicates that the price-rent ratio is positively correlated with rents. Table 3 shows the output of the regression in Equation 1. In the first column, the coefficient is positive and significant, meaning that more valuable houses have lower rent-price ratios. The second column displays the regression results when year dummies are added. Without these dummies, one might suspect that years with a lower aggregate rent-price ratio also display higher rents, and drive the results. As the table shows, however, adding year dummies leaves the coefficient on $\log \text{Rent}$ virtually unchanged. Even controlling for year dummies, the regression in column 2 forces the coefficients on $\log \text{Rent}$ to be the same in all years. By interacting year dummies with the log rents, it is possible to separate the different effects of $\log \text{Rent}$ for each year:

$$\log\left(\frac{\text{Price}}{\text{Rent}}\right) = \alpha_t + \beta_t \log \text{Rent}$$

where β_t is allowed to change from year to year. Results for the whole sample are displayed in the fourth 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 and the low number of observations.

Table 3: Regression of price-rent ratios on rents (matched dataset)

	$\log\left(\frac{\text{Price}}{\text{Rent}}\right) = \alpha + \beta \log \text{Rent}$		
	(1)	(2)	(3)
	Baseline	Year dummies	Year interactions
β	0.084^{***}	0.081^{***}	
	(0.015)	(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

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 seem to cluster around an increasing 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.

2.3 The Repeat Transactions Dataset

Some of the analyses carried out in Section 4 require to focus only on properties that appear at least twice in the Sales or Rental datasets. Table 4 shows how many repeat observations are contained in the JDW Dataset. 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. Appendix C shows that the proportion of repeat sales out of all sales in the JDW Sales Dataset and the Land Registry look similar.

Table 4: Repeat transactions in the JDW Dataset

JDW Sales		JDW Rentals	
# Transactions	Properties	# Transactions	Properties
1	17,921	1	24,651
2	1,049	2	5,774
3	45	3	1,594
		4	430
		5	102
		6	18
		7	6

3 Main findings

3.1 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:

$$q_i = X_i\beta^* + Z_i\gamma^*, \tag{3}$$

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:

$$\lambda_t = \sum_{d=1}^{D} \delta_d \text{ Time}_d, \tag{4}$$

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_n 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 section 4 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 analysed dataset covers only 7 years, from 2005 to 2011, changes in the relative prices of characteristics are likely to be limited.

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

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

where α is a constant which serves the purpose of normalising 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:

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

where $h \in \{s, r\}$ and $\lambda_{ht} = \sum_{d=1}^{D} \delta_{hd}\ \mathrm{Time}_{hd}$. This formulation allows for quality, quality-adjusted prices, and errors to differ between sales and rentals.[9] 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 et al., 2009). Indeed, different coefficients in the price and rent hedonic equations imply an effect of the regressors on price-rent ratios, because $\mathrm{E}p_s - \mathrm{E}p_r = \mathrm{E}(p_s - p_r)$, 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 $\phi_{Xr}\gamma_s^* - \phi_{Xr}\gamma_s^* = 0$ 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 dataset used for the empirical analysis 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:

$$y_{it} = \alpha_m + X_i\beta_m + \lambda_{mt} + \varepsilon_{mit}, \tag{7}$$

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

3.2 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 more[10] (1-bedroom properties are the

[9] In what follows, I often refer to sale prices as "prices" and rental prices as "rents".

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

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:

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

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 neighbourhood-level analysis of the start of the US housing boom.

Table 5 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:

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

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 5 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 5 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.[11]

In Table 5 I sort neighbourhoods from those with the highest price premium (SW3, Chelsea) to those with the lowest one (SW6, Fulham)—the baseline postcode district is W2 (Paddington).

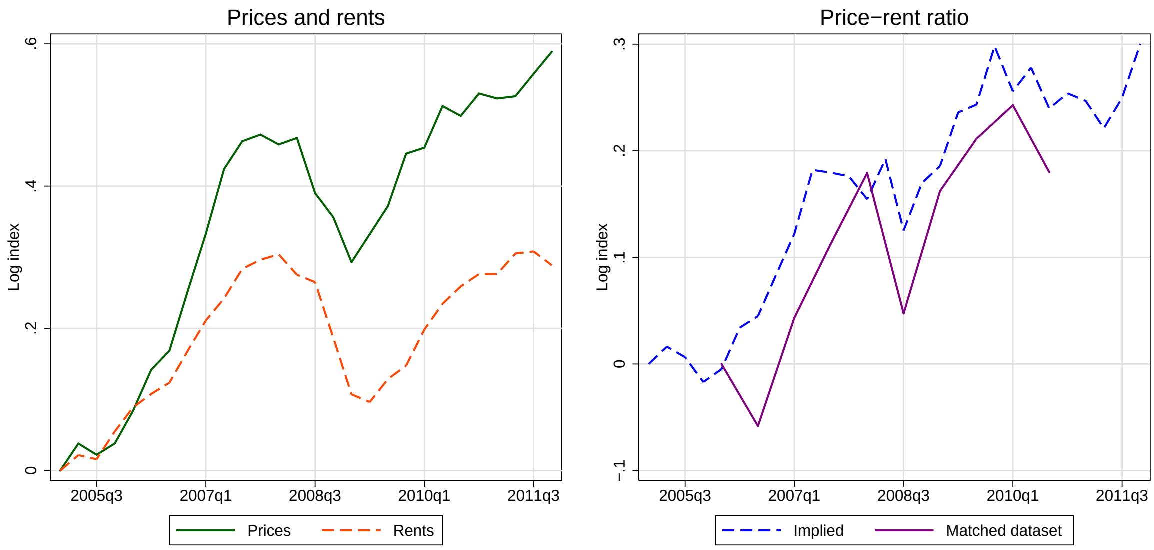
[11] The data allow me to measure *gross* price-rent ratios, i.e. price-rent ratios which do not take into account maintenance expenses and, for rented properties, vacancies. If these were higher for smaller properties, *net* rent yields (rent-price ratio net of costs) could be more similar than what suggested by their gross counterparts. However, maintenance is commonly thought to be proportionally cheaper for smaller properties (Linneman, 1985). Moreover, anecdotal evidence suggests that more expensive properties stay vacant for longer.

Table 5: Hedonic regressions

Notes: Quarterly time dummies used for the complete dataset and half-year dummies for the matched dataset. The baseline property is a 1-bedroom flat in W2.

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

FIGURE 5: PRICE AND RENT INDEXES, PRICE-RENT RATIOS



In terms of coefficients, both the complete JDW dataset and the Matched Dataset show that more expensive neighbourhoods have higher price-rent ratios. In other words, both prices and rents rise for more expensive neighbourhood, but prices rise more than rents. This fact is well known by housing market practitioners.[12]

The results presented here demonstrate the consistency between the whole JDW dataset and the small subsample of matched properties. Results (non-tabulated) show that the percentages of 1-bed, 2-bed, 3-bed+ and houses among matched properties are similar to those in the JDW rental dataset.

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 at a lower rate. 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.

3.3 Hedonic regressions with time-varying characteristic prices

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

$$p_{hit} = \beta_{ht} X_i + u_{hit}. \tag{8}$$

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

[12] See for instance "London buyers find streets paved with gold", Financial Times, 13 March 2011.

Figure 6: Time-varying hedonic prices

8 for each period. The JDW Dataset contains 28 quarters. Using the 23 explanatory variables listed in Table 5 (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 increase in 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, in all periods, that properties in prime neighbourhoods such as Chelsea (SW3) or Kensington (SW7) command a bigger premium on sales than rentals. When analysing other neighbourhoods, such as Fulham (SW6), the distinction between price and rent coefficients become much less clear or is reversed.

4 Mechanisms

The previous section has shown that price-rent ratios are higher for bigger and better-located properties. Higher price-rent ratios mean lower rental yields. In a well-functioning market these differences in yields would be arbitraged away, unless they correspond to fundamental differences in the investment characteristics of properties. In this section, I explore some possible explanations for these differences. The goal is not to provide an exhaustive theory but to discuss the plausibility of potential mechanisms.

In what follows, I distinguish between two groups of arguments: explanations based on hidden costs and explanations based on asset pricing. Hidden costs create a wedge between gross yields (which are measured in this paper) and net yields. It is possible that, while gross yields differ across property categories, net yields are actually quite similar. Asset-pricing explanations take the dividend discount model as starting point and try to rationalise differences in price-rent ratios with differences in expected growth or discount factors. Neither hidden cost arguments nor the traditional dividend pricing model are able to explain yield differences in a satisfactory way. An alternative asset pricing model, where ownership is a hedge against rent risk (Sinai and Souleles, 2005), seems more able to provide a framework consistent with the data.

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 stylised fact illustrated in this paper has been shown elsewhere in the UK (Joseph Rowntree Foundation, 1996; Association of Residential Letting Agents, 2012) and in the US (Garner and Verbrugge, 2009). Thus, any satisfying explanation cannot be London-specific. Likewise, I do not expand on behavioural 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 behavioural 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 behavioural arguments only.

4.1 Hidden costs: Gross vs. net yields

A first reason why there could be a wedge between gross and net yields is related to transaction costs. There are fixed costs to be paid each time properties are bought or sold. If these fixed cost do not rise proportionally with prices, flats would be subject to higher transaction costs, in relative terms. They should then sell for a lower price, as a compensation, and this would raise their yield.

A second possibility is that maintenance costs are higher for smaller and less central properties—in equilibrium, then, their gross yields would have to be higher. Unfortunately, data on maintenance costs are not readily available. However, common sense suggests that maintenance costs are proportionally higher for bigger properties. Linneman (1985) made the famous point that managers of small properties can better exploit economies of scale. Moreover, in most European cities, London included, properties in more expensive neighbourhoods are older, and hence need more maintenance. Thus, it doesn't seem likely that maintenance costs can explain the differences in gross yields measured in the paper.

A third potential explanations 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

Table 6: 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.

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

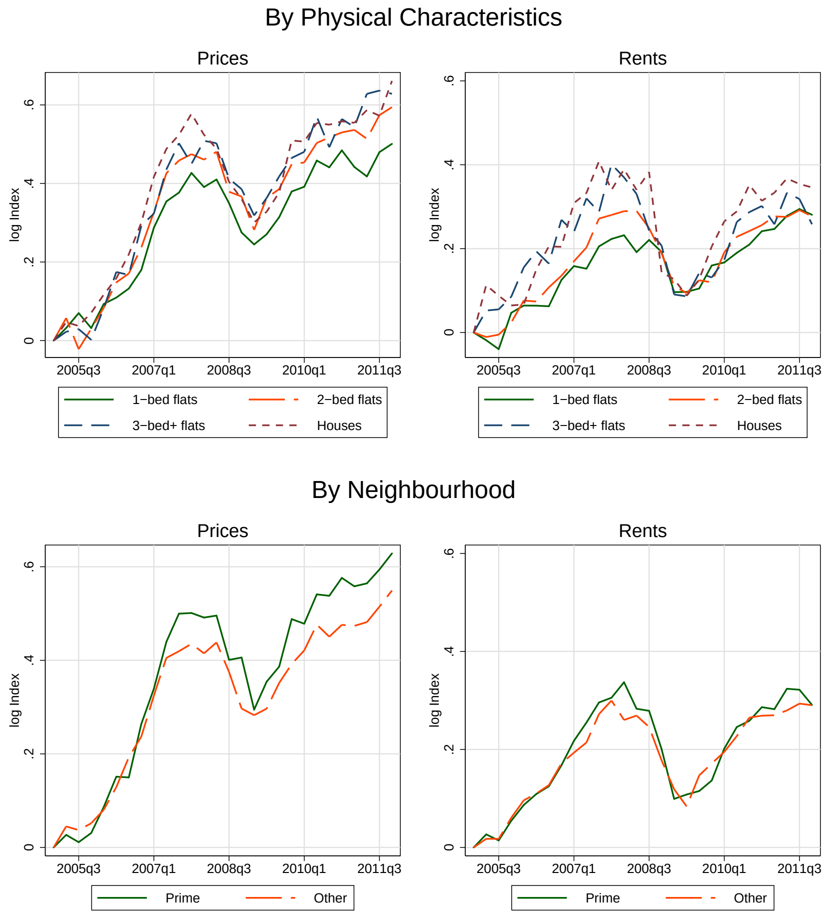
could explain the differences in yields. However, it is commonly believed that more expensive properties stay on the market for longer—the higher the price, the more important it is to find a good match with a buyer that likes the property. Despite the fact that the JDW Dataset does not contain data 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 6 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 for an average of 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. Again, it is extremely difficult to find data to prove this claim. However, 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) 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 explain the pattern in price-rent ratios highlighted on the basis of transaction costs, maintenance costs, or vacancies.

4.2 Asset pricing: Expected appreciation and risk premia

The most basic asset pricing model is the dividend discount model, according to which the price of an asset corresponds to the present discounted value of all its future dividends. In a real estate context, this equivalence translates into a relation between the sale price of a property and all its future rents, which in turn implies 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 expected returns (i.e. risk premia), or both (Campbell et al., 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.

Figure 7: Growth and Volatility by Housing Categories



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 neighbourhoods and other neighbourhoods. Prime neighbourhoods are the most expensive six postcode districts in the hedonic regression of Table 5: SW3 (Chelsea), SW7 (South Kensington), W8 (Holland Park), W1 (Mayfair), W11 (Notting Hill), SW1 (Belgravia and Pimlico). In the JDW Dataset, prime neighbourhoods correspond to 53% of Sales and 54% of Rentals. I take equation 6 and allow the coefficients on property characteristics to differ across categories:

$$p_{hct} = \alpha_{hc} + X_{ic}\beta_{hc} + \lambda_{hct} + \varepsilon_{hct}, \tag{9}$$

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 $\mathrm{E}(\lambda_{hct+1} - \lambda_{hct})$ and the corresponding aggregate risk is $\mathrm{Var}(\lambda_{hct+1} - \lambda_{hct})$.

Figure 7 plots the λ_{hct}'s over the different quarters t, estimated using equation 9. 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),

Table 7: Prices and rents: Growth and systemic risk

	$p_{hct} = \alpha_{hc} + X_{ic}\beta_{hc} + \lambda_{hct} + \varepsilon_{hct}$,			
	JDW Sales Dataset		JDW Rentals Dataset	
	$\mathrm{E}(\lambda_{ct+1} - \lambda_{ct})$	St.Dev.$(\lambda_{ct+1} - \lambda_{ct})$	$\mathrm{E}(\lambda_{ct+1} - \lambda_{ct})$	St.Dev.$(\lambda_{ct+1} - \lambda_{ct})$
All	*0.022*	*0.042*	*0.011*	*0.033*
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

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 neighbourhood sale price comparison at the bottom left of Figure 7. Sale prices in prime neighbourhoods have grown more but are also more volatile. However, rents have behaved very similarly in prime and other neighbourhoods, both in terms of growth and volatility.

Table 7 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 7 imply that it is difficult to rationalise 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.

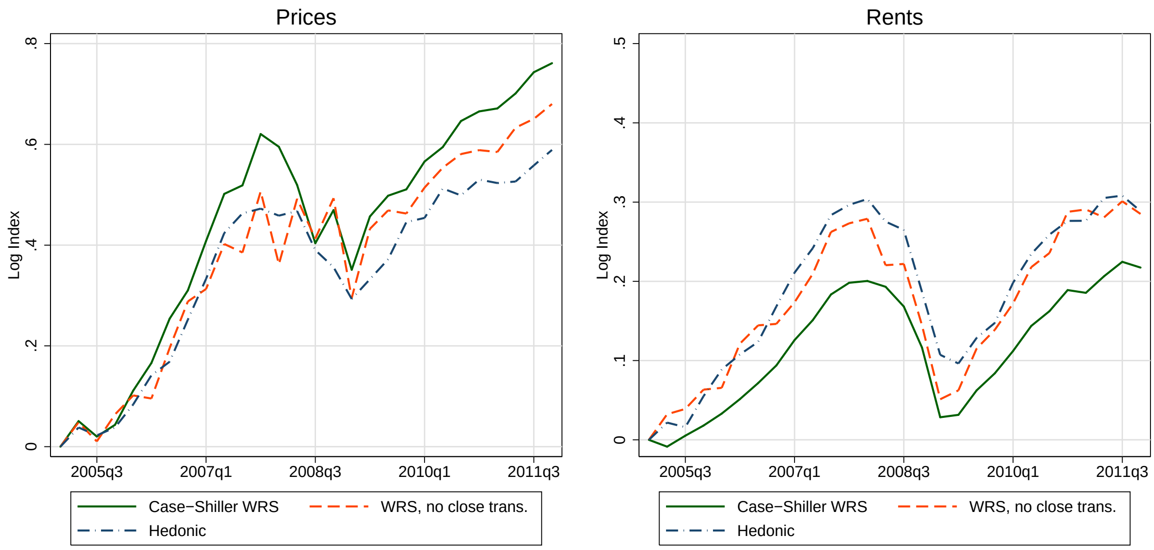
Evidence from repeat sales and rentals Table 7 shows results on the volatiltiy of *aggregate* indices, not individual properties. The balance sheet of most homeowners contains just one property (Flavin and Yamashita, 2002), and most renters are obviously subject to just one rental contract (Genesove, 2003).[13] Idiosyncratic volatility might be more relevant than its aggregate counterpart. To measure idiosyncratic volatility, I restrict my 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:

$$p_{hiT} - p_{hit} = \underbrace{\lambda_{hT} - \lambda_{ht}}_{\substack{\text{aggregate}\\ \text{movement}}} + \underbrace{u_{hiT} - u_{hit}}_{\substack{\text{idiosyncratic}\\ \text{movement}}}. \tag{10}$$

Equation 10 constitutes the basis of the repeat sales method (Bailey et al., 1963; Case and Shiller, 1989), which allows for the estimation of the term $u_{hiT} - u_{hit}$. Simlarly to aggregate risk, idiosyncratic risk is defined as $\mathrm{Var}(u_{it+1} - u_{it})$. 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, $\mathrm{Var}(u_{iT} - u_{it}) = 2\sigma_v^2 + \sigma_h^2(T - t)$

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

Figure 8: Hedonic and repeat sales indices



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).[14]

Figure 8 compares the WRS indices estimated on the JDW Dataset with the hedonic indices 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.[15] One way to limit this problem is to exclude properties whose "holding period" (the time between two sales) is below a certain threshold. I choose a threshold of 1000 days (corresponding to approximately 3 years). The result is displayed in the left-hand side of Figure 8: the WRS index with no close transitions appreciates less than the unadjusted WRS index, but still more than the original hedonic index.

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 behaviour 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. Again, one can limit the selection bias by

[14] In practice, the Case and Shiller (1989)'s procedure involves three steps: first, running an OLS regression to estimate Equation 10; 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 10 again running a GLS regression where observations are weighted by the inverse of the square root of the predicted residuals.

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

excluding from the sample rental contracts that are too close from each other. To be consistent with the procedure adopted for the price index, I choose the same threshold of 1000 days. The resulting WRS index with minimum holding period is closer to the hedonic index than the unadjusted WRS index.

I compute separate WRS regressions to estimate idiosyncratic volatilities for different property categories. For some categories the number of repeat sales is low, but this is less of a problem for rents, because repeat rents are more common than repeat sales. Table 8 shows the outcome of running the regression described by equation 10, 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 significantly smaller samples; nevertheless, coefficients under both approaches are similar. The output of Table 8 mirrors the results on aggregate volatility. Bigger houses and expensive neighbourhoods have higher idiosyncratic volatility, although the distinction between prime and non-prime neighbourhoods is less sharp than the distinction between big and small properties.

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.[16] The last survey (January 2012) contains a couple of questions on local price and rent expectations:

> 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.

The question on postcodes aims at identifying the postcode district of the respondent. I divide respondents into three groups: those living in the UK outside London, those living in London but not in a prime neighbourhood, and those living in a London prime neighbourhood.[17] The definition of prime neighbourhood 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 neighbourhoods are correlated with high price or rent growth expectations, in accordance with the dividend pricing model.[18] Since there are no explicit questions on rent risk, I take the dispersion of rent

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

[17]The 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. A few respondents live outside the UK, and are excluded from the present statistics. The last part of the questionnaire contains questions on the socio-demographic characteristics of respondents. A table with summary statistics is shows in Appendix B.

[18]The 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 neighbourhoods, not the one

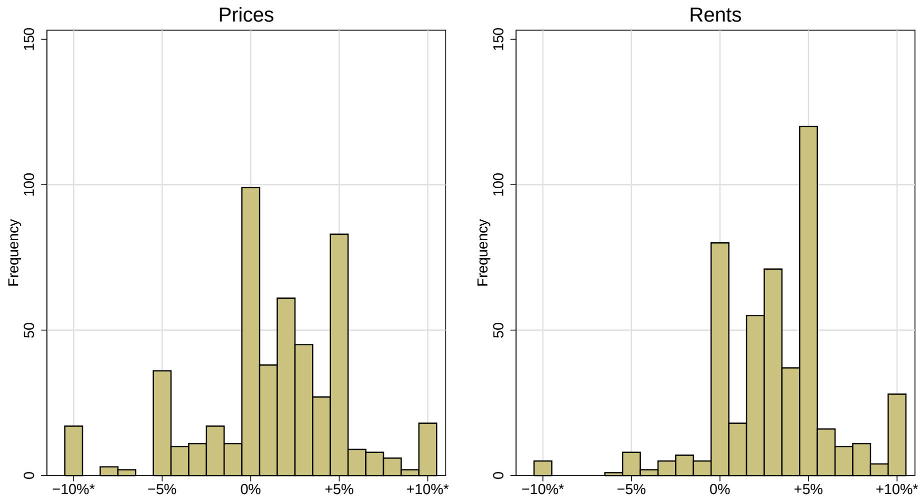
Table 8: Prices and rents: Idiosyncratic risk

Notes: The table shows the coefficients obtained by estimating the second stage of the WRS method of Case and Shiller (1989). 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.

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

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".



expectations as a measure of rent uncertainty. This approach is consistent with the empirical literature that looks at disagreement about inflation (Mankiw et al., 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.[19]

The upper half of Table 9 shows the differences in price and rent expectations between prime and other neighbourhoods. To provide another relevant comparison, the lower half of the table shows the same differences between London and other parts of the UK. In terms of price expectations, respondents in prime neighbourhoods are slightly more optimistic than other Londoners, but the difference is not significant. By contrast, Londoners are significantly more optimistic than other respondents in the UK, consistently with the different performances of the UK housing market inside and outside London in the last years.[20] In terms of rent expectations, people living in the non-prime neighbourhoods of London expect slightly higher growth (the difference is significant at the 10% level). Londoners in general expect higher rent growth than non Londoners. The standard deviation of rent expectations is significantly higher for prime London than other parts of London and the same is true for the London vs Outside London comparison.

Therefore, according to the survey, rent uncertainty is higher for London prime neighbourhoods 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 people on the mailing list of a Central London real estate agency: their opinions are likely to be representative of the buyers and sellers of this particular housing market.

regarding differences between properties of different sizes.

[19]This is a common feature of expectation surveys (Hudomiet et al., 2011).

[20]See "How did London get away with it?", CentrePiece, Winter 2010/2011 (`http://cep.lse.ac.uk/pubs/download/cp333.pdf`).

Table 9: 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?"

	Price Expectations ($E_t p_{st+1}$)			Rent Expectations ($E_t p_{rt+1}$)		
	Mean	(St. Dev.)	Obs.	Coeff.	(St. Dev.)	Obs.
	(1)	(2)	(3)	(4)	(5)	(6)
London, prime neighbourhood	2.25	(4.37)	79	2.84	(3.81)	74
London, other neighbourhood	2.10	(4.14)	189	3.78	(3.38)	183
Mean Diff. (T-test)	*0.15*	*(0.58)*		*-0.94**	*(0.51)*	
StDev Ratio (F-test)		*1.12*			*1.28**	
London	2.15	(4.20)	268	3.51	(3.53)	257
Outside London	0.20	(3.58)	200	2.67	(2.97)	191
Mean Diff. (T-test)	*1.95****	*(0.36)*		*0.84****	*(0.31)*	
StDev Ratio (F-test)		*1.37****			*1.41****	

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 their finding 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.

5 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 neighbourhoods.

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 stylised fact that I uncover. I divide the potential mechanisms in two groups: explanations based on the difference between gross and net yields, and explanations based on asset pricing. 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 et al., 2010). Future research should expand on these different aspects of rent risk and housing market liquidity.

References

Association of Residential Letting Agents (2012). Review and index for residential investment. Technical report.

Bailey, M., Muth, R., and Nourse, H. (1963). A regression method for real estate price index construction. *Journal of the American Statistical Association*, pages 933–942.

Banks, J., Blundell, R., Oldfield, Z., and Smith, J. P. (2010). House price volatility and the housing ladder. IZA Discussion Papers 5173, Institute for the Study of Labor (IZA).

Calhoun, C. (1996). Ofheo house price indexes: Hpi technical description. Technical report, Office of Federal Housing Enterprise Oversight.

Campbell, S., Davis, M., Gallin, J., and Martin, R. (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 Shiller, R. J. (1989). The efficiency of the market for single-family homes. *American Economic Review*, 79(1):125–37.

Clapp, J. and Giaccotto, C. (1992). Estimating price indices for residential property: a comparison of repeat sales and assessed value methods. *Journal of the American Statistical Association*, pages 300–306.

Court, A. (1939). Hedonic price indexes with automotive examples. In *The Dynamics of Automobile Demand*, pages 98–119, New York. General Motors.

Department for Communities and Local Government (2012). English housing survey (ehs) household report 2010-11. Technical report.

Ferreira, F. and Gyourko, J. (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 Yamashita, T. (2002). Owner-occupied housing and the composition of the household portfolio. *The American Economic Review*, 92(1):345–362.

Gabriel, S. and Rosenthal, S. (2011). Homeownership boom and bust 2000 to 2009: Where will the homeownership rate go from here? Research paper no. 11-03, Research Institute for Housing America.

Gallin, J. (2008). The long-run relationship between house prices and rents. *Real Estate Economics*, 36(4):635–658.

Garner, T. and Verbrugge, R. (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 Gyourko, J. (2007). Arbitrage in housing markets. Working Paper 13704, National Bureau of Economic Research.

Goetzmann, W. and Spiegel, M. (1995). Non-temporal components of residential real estate appreciation. *The Review of Economics and Statistics*, pages 199–206.

Griliches, Z. (1961). Hedonic price indexes for automobiles: An econometric of quality change. Technical report.

Guerrieri, V., Hartley, D., and Hurst, E. (2010). Endogenous gentrification and housing price dynamics. Working Paper 16237, National Bureau of Economic Research.

Halket, J. and Pignatti, M. (2012). Housing tenure choices with private information. mimeo.

Hill, R. (2012). Hedonic price indexes for residential housing: A survey, evaluation and taxonomy. *Journal of Economic Surveys*.

Hudomiet, P., Kézdi, G., and Willis, R. (2011). Stock market crash and expectations of American households. *Journal of Applied Econometrics*, 26(3):393–415.

Hwang, M., Quigley, J., and Son, J. (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.

Linneman, P. (1985). An economic analysis of the homeownership decision. *Journal of Urban Economics*, 17(2):230–246.

Mankiw, N., Reis, R., and Wolfers, J. (2003). Disagreement about inflation expectations. *NBER Macroeconomics Annual*, pages 209–248.

Ngai, R. and Tenreyro, S. (2009). Hot and cold seasons in the housing market. *Centre for Economic Performance Discussion Paper 922.*

Novy-Marx, R. (2009). Hot and cold markets. *Real Estate Economics*, 37(1):1–22.

Ortalo-Magné, F. and Rady, S. (2002). Tenure choice and the riskiness of non-housing consumption. *Journal of Housing Economics*, 11(3):266–279.

Ortalo-Magné, F. and Rady, S. (2006). Housing market dynamics: On the contribution of income shocks and credit constraints. *The Review of Economic Studies*, 73(2):459–485.

Plazzi, A., Torous, W., and Valkanov, R. (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.

Shiller, R. (1981). Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review*, 71(3):421–436.

Sinai, T. and Souleles, N. (2005). Owner-occupied housing as a hedge against rent risk. *The Quarterly Journal of Economics*, 120(2):763–789.

Verbrugge, R. (2008). The puzzling divergence of rents and user costs, 1980–2004. *Review of Income and Wealth*, 54(4):671–699.

Vissing-Jorgensen, A. (2003). Perspectives on behavioral finance: Does 'irrationality' disappear with wealth? Evidence from expectations and actions. *NBER Macroeconomics Annual*, pages 139–208.

Appendices

A Housing Statistics for Central-Western London

The first two columns of Table A1 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.[21] (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.

Table A1: General Housing Statistics

	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	*No statistics*		106,506	*0.54*	413,902	*0.29*
3-bedroom+ flats	*at Borough level*		10,433	*0.05*	14,421	*0.01*
Houses			35,237	*0.18*	879,721	*0.61*

[21] `http://www.communities.gov.uk/housing/housingresearch/housingstatistics/housingstatisticsby/housebuilding/livetables/`.

B 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 summarised 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.

Table B1: Summary statistics

Variable	%	Obs.	Variable	%	Obs.
Panel A: Housing					
Residence			*Housing tenure*		
Outside UK	0.07	510	Homeowner (mortgage)	0.50	451
UK outside London	0.47	510	Homeowner (outright)	0.33	451
London, prime neighbourhood	0.16	510	Renting	0.01	451
London, other neighbourhood	0.30	510	Other	0.16	451
Panel B: Socio-demographic characteristics					
Gender			*Age*		
Male	0.69	293	Less than 31	0.06	294
			31-40	0.32	294
			41-50	0.25	294
Marital status			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	*Income*		
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
Education			*Occupation*		
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

C Comparing the JDW Sales Dataset with the U.K. Land Registry

This part of the Appendix studies the subset of observations in the Land Registry that belong to the postcode districts listed in Table 1.

Figure C1: Quarterly registered sales (Land Registry)

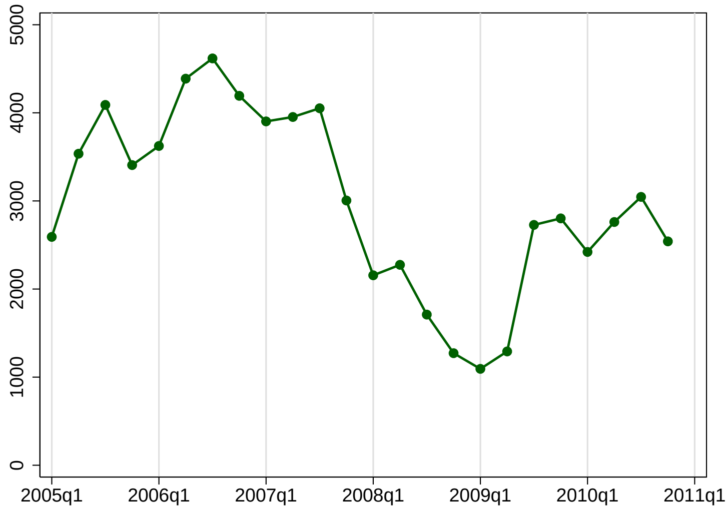


Table C1: Property characteristics (Land Registry)

	2005-2010
Median price (in 2005 £)	427,948
Flat (%)	0.81
Terraced house (%)	0.16
Semi-detached house (%)	0.02
Detached house (%)	0.01
Newly built (%)	0.06
Total observations	71,459

Table C2: Repeat sales (Land Registry, 2006–2010)

Dataset	Units that appear ... times	
Land Registry sales	1	52,167
	2	2,650
	3	117
	4	4

Figure C2: Price indices (Land Registry)

Notes: Indices are normalised to zero in 2006Q1, starting date of the repeat sales index. 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.

BIS Department for Business Innovation & Skills

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Llywodraeth Cynulliad Cymru
Welsh Assembly Government

Spatial Economics Research Centre (SERC)
London School of Economics
Houghton Street
London WC2A 2AE

Tel: 020 7852 3565
Fax: 020 7955 6848
Web: www.spatialeconomics.ac.uk

SERC is an independent research centre funded by the Economic and Social Research Council (ESRC), Department for Business Innovation and Skills (BIS) and the Welsh Assembly Government.