

# The housing stock, housing prices, and user costs: the roles of location, structure and unobserved quality

---

**Jonathan Halket**  
**Lars Nesheim**  
**Florian Oswald**

The Institute for Fiscal Studies  
Department of Economics, UCL

**cemmap** working paper CWP73/15

# The Housing Stock, Housing Prices, and User Costs: The Roles of Location, Structure and Unobserved Quality \*

Jonathan Halket<sup>†</sup>

Lars Nesheim<sup>‡</sup>

Florian Oswald<sup>§</sup>

December 9, 2015

## Abstract

Using the English Housing Survey, we estimate a supply side selection model of the allocation of properties to the owner-occupied and rental sectors. We find that location, structure and unobserved quality are important for understanding housing prices, rents and selection. Structural characteristics and unobserved quality are important for selection. Location is not. Accounting for selection is important for estimates of rent-to-price ratios and can explain some puzzling correlations between rent-to-price ratios and homeownership rates. We interpret this as strong evidence in favor of contracting frictions in the rental market likely related to housing maintenance.

---

\*This research is supported by the UK Economic and Social Research Council through the Centre for Microdata Methods and Practice (CeMMAP grant number ES/I034021/1) and through an ESRC transformative research grant, grant number ES/M000486/1. This work is based on data from the English Housing Survey, produced by the Office for National Statistics (ONS) and supplied by the Secure Data Service at the UK Data Archive. The data are Crown Copyright and reproduced with the permission of the controller of HMSO and Queen's Printer for Scotland. The use of the data in this work does not imply the endorsement of ONS or the Secure Data Service at the UK Data Archive in relation to the interpretation or analysis of the data. This work uses research datasets which may not exactly reproduce National Statistics aggregates.

<sup>†</sup>University of Essex and CeMMAP, Institute for Fiscal Studies

<sup>‡</sup>University College London and CeMMAP, Institute for Fiscal Studies

<sup>§</sup>Sciences Po

# 1 Introduction

Homeownership is puzzling. Why is the market share of the owner-occupied sector in most developed countries so large relative to that of the rental sector? On the demand side, numerous theoretical explanations have been proposed, including insurance motives, tax considerations, and a “warm glow” from housing. On the supply side, theories have focused on a variety of contracting problems between renters and landlords. Most of these theories have implications not only for the aggregate market share of homeownership, but also for which properties are owned and which are rented and how rent-to-price ratios might vary across properties. Yet surprisingly little is known empirically about the supply side choices that govern which housing characteristics are likely to be owned rather than rented and about how rent-to-price ratios vary across the urban landscape.

Casual empiricism suggests that large, detached, expensive, suburban houses are more likely to be owned while smaller, cheaper apartments in urban centers are less likely to be owned. These simple correlations mask a much more nuanced picture. The market share of homeownership, the rent-to-price ratio, and the physical characteristics of the housing stock all vary dramatically across the urban landscape. How do dwelling prices and rents vary with the propensity to be selected into owner-occupancy? What do the empirical patterns imply about user costs and contracting frictions in the rental and owner-occupied markets?

These are crucial questions for understanding how household choices are affected by local geography. They are also crucial because they are intimately tied together with important macroeconomic questions about household savings and insurance against financial risks. In the UK and the US more than 60% of households tie up a large part of their financial portfolio in a single, risky, illiquid asset; housing. Why is this the case? Why don't financial and rental housing markets provide contracts that enable households to enjoy the consumption flow from a rented three bedroom detached 120 square-meter house in the suburbs while enjoying the dividend flows and potential capital gains of the wider financial market?

Our approach is straightforward. We use a rich micro data set to estimate a simple model of the supply of housing to the rental and owner-occupied housing markets, exploiting the tremendous variation in dwelling characteristics, prices, rents and own-

ership rates within London between 2008-2012. We model the economy as endowed with a set of properties. Landlord and owner-occupiers take prices and rents as given. A property ends up in either the rental or owner-occupied part of the housing market depending on which part of the market values the property more. We then analyze the implications of our findings in the context of the simplest possible dynamic setting. We find that:

1. Observable physical characteristics of a property are by far the most important determinants of the probability of being in the owner-occupied sector. Housing units with high value physical attributes (large or more detached dwellings) are more likely to be owner-occupied. At the same time, rent-to-price ratios for these “large” properties are higher than for smaller properties. In other words, despite their relatively high gross yield in the rental sector, properties with high value physical characteristics are less likely to be bought up by landlords and supplied to renters. These relationships are fairly stable over time despite large changes in property prices during the time period. This is consistent with the theory that agency problems in the rental sector increase the cost of upkeep and care<sup>1</sup> for properties relative to the cost in the owner-occupied sector; especially for larger properties and for properties that are more detached. In other words, rental sector user costs of housing capital increase with property size faster than owner-occupied user costs.
2. Location is unimportant for the likelihood of being owner-occupied after controlling for physical characteristics. Rent-to-price ratios vary significantly with location most likely because changing locations leads to variation in the value of the property that does not imply changes in gross maintenance levels. Some features of the rent-to-price patterns are unstable over time which may reflect time and location dependent expectations of capital gains. These facts are all consistent with a theory of selection based on the physical character of the house.
3. Modeling and measuring differences in unobserved quality is essential for understanding which homes become rentals and at what price. The selection of

---

<sup>1</sup>Throughout we refer to these costs as maintenance costs. They are the costs required to keep a property at constant quality.

properties into owner-occupancy has a strong effect on the relative distribution of both observed and unobserved characteristics in the two markets. In particular, we find that the data reject uni-dimensional models of unobserved quality and that selection affects the distributions of the different dimensions of quality differently. Attributes of a dwelling that are unobserved in our data include features like the layout of the dwelling, the architectural style, whether it has South facing windows, whether it has a high quality kitchen or if there's a garden that requires upkeep. Unobserved quality that raises the value of owner-occupied houses is valued equally by landlords; so this dimension of quality does not lead to selection between the two sectors. In contrast, the type of unobserved quality that only raises rents but not values ("rental quality") is on average lower in the rental sector. Thus rentals have lower unobserved rental quality on average. This is consistent with some unobservable attributes being harder to contract upon for landlords and thus having higher maintenance costs in the rental sector. A further implication is that rent-to-price measures based only on observable attributes that do not control for selection are biased.

## 1.1 Relation to the literature

There is an extensive household tenure-choice literature that studies demand side selection into homeownership. For example, see Rosen (1979), Goodman (1988), Kan (2000), or Campbell and Cocco (2007). This literature studies how observable and unobservable household factors affect selection into and welfare from homeownership. For instance, Diaz and Luengo-Prado (2008) and Blow and Nesheim (2009) examine how the flow shadow price of housing for owner-occupiers can differ systematically from the cost of renting a similar housing unit in large part due to differences in the Lagrange multipliers from households' constrained maximization problems. In general, the multipliers are functions of households' current assets, income and other state variables.

In contrast, in this paper we study supply side selection into homeownership. We study what aspects of housing units explain why some units are more likely to end up in the rental sector while others are more likely to end up in the owner-occupied sector. Glaeser and Shapiro (2003) and Amior and Halket (2014) observe that there is a strik-

ing difference in the homeownership rates of single versus multi-family housing units throughout the US. This leads the former to note that “homeownership is particularly correlated with housing structure.” Here we show that the same is true in England and that this correlation is found consistently across many structural characteristics including dwelling type, dwelling size and unobserved quality.

What explains this correlation? The technology that provides housing services is the same regardless of housing sector: in both sectors, a physical housing unit provides the services. However, it has long been suggested that the contracts that govern the provision of housing services in the rental and owner-occupied markets may be constrained by tenure-dependent information frictions. A literature going back to Sweeney (1974) and Henderson and Ioannides (1983) has argued that contracting frictions in the rental sector result in higher maintenance costs and less upkeep and investment. Because of monitoring costs, tenants under-invest in maintenance resulting in a higher depreciation rate in the rental sector. In this literature, it is assumed but often unstated that these differential costs are likely to be correlated with the physical characteristics of the property. As a result, a housing unit with higher rental-specific maintenance costs, or for which rental contracting frictions are greater, is more likely to be found in the owner-occupied sector.

Heretofore, empirical support for these theories is scant.<sup>2</sup> In the corporate finance literature, studies of selection due to various contracting frictions and the effects of selection on the distribution of observed returns are common.<sup>3</sup> In the housing literature, despite the long-standing theories discussed above and the cautionary warning in Glaeser and Gyourko (2007), there are no studies that control for supply side selection when estimating hedonic rent and price functions. The closest perhaps is Heston and Nakamura (2009) which uses a small sample of federal employee data in several small markets to show that owner-occupied housing units are 15 percent more valuable

---

<sup>2</sup>Linneman (1985) notes that the “efficiency” of landlord provided housing services is an important factor determining ownership rates. Casas-Arce and Saiz (2010) examine how different jurisdictions’ legal systems and propensity to enforce contracts affect ownership propensities. Hanson (2012) looks at how the mortgage interest tax deduction interacts with home sizes to affect ownership rates in the US. Hilber (2005) examines neighborhood externality risks in the AHS and finds that they are negatively correlated with homeownership. Coulson and Fisher (2014) study how property size affects a building’s management structure. Harding et al. (2000) find that homeowners that are more likely to default on their mortgage may under-maintain their house.

<sup>3</sup>Among many others, see Prabhala (2008).

than observably equivalent rentals. The result is based on self-reported estimates of rental flows obtained from owner-occupiers. We are the first to (a) estimate a selection model of hedonic housing prices and rents for a large housing market (in our case, the Greater London housing market), (b) to estimate the importance of unobserved quality in this market, (c) to analyze the implications of these models for sector specific user costs. Furthermore, we use repeated cross-sections sampled from periods of both housing price decline and boom in England and find that the estimated relationship between physical attributes, selection and rent-to-price remains very stable. We find that recent movements in the rent-to-price ratios of housing are reflected entirely in the value of land,<sup>4</sup> which is consistent with the methods and findings in Davis and Heathcote (2007), Davis and Palumbo (2008), and Amior and Halket (2014).

Taking this one step further, if we assume that a single risk-adjusted discount rate prices all housing in the market as in Epple et al. (2013), then we can say more. In this case, if different housing units have different rent-to-price ratios, then these differences must be due either to differing expectations about the future or to differences in the costs of renting out the property. We argue that expectations about the future, given rents and prices, should not systematically affect selection into renting. This enables us to use our selection model results to estimate the potential size of the moral hazard problem described above.

Our findings have important implications for macro models of the housing market. As we stated at the outset, homeownership is a puzzle. To explain the high market share of homeownership, many theories of household demand for homeownership incorporate higher maintenance costs for rentals and a limited supply of “high value” rentals into their models (e.g. Chambers et al., 2009b,a; Chen, 2010). Often these two elements are the primary factors that explain why so many households choose to own in these models. We show that selection due to agency problems can simultaneously explain both the higher maintenance costs and the limited supply of *physically* valuable rentals. Our estimate of the size of the agency problem is large but consistent with the less direct estimates from these models.

Finally, accounting for unobserved quality explains two curious features in the raw data. Firstly, direct measures of maintenance expenditures rarely show a meaningful

---

<sup>4</sup>One possible exception to this is in the 2014 wave rental sector findings. In that year, it appears that the rent function, as a function of property size may be flatter than in other years.

difference between rentals and owner-occupied properties.<sup>5</sup> Our findings on bias and selection imply that comparisons across tenure of observationally similar housing units will not necessarily reveal evidence of contracting problems. Properties with more contracting issues select into owner-occupancy. We show that not properly accounting for selection leads to biased measures of the difference in maintenance costs that are large enough to obscure the agency costs. Indeed, under some assumptions about interest rates and capital gains, our estimates suggest (as an upper bound) that contracting problems can increase maintenance costs of large rental properties by as much as 50 percent.

Secondly, several studies (see Verbrugge, 2008; Landvoigt et al., Forthcoming; Epple et al., 2013; Verbrugge and Poole, 2010; Heston and Nakamura, 2009) using US data have found that rent-to-price ratios decline with property prices. If we estimate rent and price functions using UK data without accounting for selection, we also find that rent-to-price ratios decline with property prices. The more expensive a property is, the lower is its predicted rent-to-price ratio. As ownership rates are (unconditionally) increasing in dwelling prices, this could lead one to the curious conclusion that households tend to own homes that have low rent-to-price ratios. This is a bit puzzling from the household's perspective. Here we show that this unconditional correlation between ownership and rent-to-price ratios breaks down once one controls for selection. Households actually tend to own housing units whose physical characteristics imply high rent-to-price ratios.

Sections 2 and 3 introduce the model and data, respectively. Section 4 explains our estimation procedure and Section 5 discusses the results.

## 2 Model

A property has observed characteristics  $z \in \mathbf{R}_n$  and unobserved characteristics  $\varepsilon \in \mathbf{R}_2$ . Observed characteristics include location, type of dwelling (detached, semi-detached,

---

<sup>5</sup>Differences between rental and owner-occupied maintenance expenditures in national accounts data are small. There are several studies that attempt to measure at the disaggregate level whether rentals have higher maintenance costs. Their findings are mixed. Galster (1983) estimates that owner-occupiers occupy better properties and better maintain them. Shilling et al. (1991) estimates a hedonic model of sales prices for rental and owner-occupied single-family property in a single parish in Louisiana and finds that rentals depreciate faster. Malpezzi et al. (1987) estimates hedonic models of rents and prices from the AHS and finds that rents decline evenly with age whereas prices decline at a declining rate. However, Gatzlaff et al. (1998) finds limited evidence of differential maintenance by comparing appreciation rates of rentals and owner-occupied housing units.



etc.), size (square meters), number of bedrooms, and age of structure. We assume that the value of unobserved characteristics is completely captured by a two dimensional vector that we label “unobserved quality”. This vector has dimension of at least two because selection into the owner-occupied sector is not perfectly correlated with prices. In addition, *a priori*, it seems likely that some characteristics are more valued in the rental sector while others are more valued in the owner-occupied sector.

If a dwelling unit is in the rental sector, its rent is observed. If it is in the owner-occupied sector, its price is observed. Let log annual rent be given by

$$\ln R(z, \varepsilon) = \alpha z + \lambda_1^r \varepsilon_1 + \lambda_2^r \varepsilon_2. \quad (1)$$

This is a log linear approximation to the true hedonic rent function. The parameters  $(\alpha, \lambda_1^r, \lambda_2^r)$  measure the percentage impact of observed and unobserved quality on rental prices. Let the log price in the owner-occupied sector be

$$\ln \pi^o(z, \varepsilon) = \beta z + \lambda_1^o \varepsilon_1 + \lambda_2^o \varepsilon_2. \quad (2)$$

This is a log linear approximation to the hedonic price function. The parameters  $(\beta, \lambda_1^o, \lambda_2^o)$  capture the percentage impact of observed and unobserved quality on prices in the owner-occupied sector.

Let the log price of a dwelling in the rental sector be

$$\ln \pi^r(z, \varepsilon) = (\beta - \gamma)z + (\lambda_1^o - \lambda_1^s) \varepsilon_1 + (\lambda_2^o - \lambda_2^s) \varepsilon_2. \quad (3)$$

The parameters  $(\gamma, \lambda_1^s, \lambda_2^s)$ , when they are positive, capture the reduced form net loss in value of renting out a dwelling relative to selling it in the owner-occupied sector. When they are negative, they capture the net gain from renting the dwelling in the rental sector. We discuss how rental prices are related to rents in Section 2.1 below.

We do not observe the rental sector price  $\pi^r$  for any property in our data. However, assuming investors maximise profits, a housing unit is observed in the owner-occupied sector if

$$\ln \pi^o(z, \varepsilon) \geq \ln \pi^r(z, \varepsilon). \quad (4)$$

That is, if

$$\gamma z \geq -\lambda_1^s \varepsilon_1 - \lambda_2^s \varepsilon_2 \quad (5)$$

The prices  $\pi^o$  and  $\pi^r$  are conditioned on sector. They measure willingness to pay of buyers in each sector. The unconditional price of a property is the price of the property in the market where it is most valuable. That is,

$$P(z, \varepsilon) = \max_{\{own, rent\}} \{\pi^o(z, \varepsilon), \pi^r(z, \varepsilon)\}$$

The parameters of the price and rent functions may vary over time. We leave their dependence on  $t$  implicit.

Under the assumption that  $\varepsilon \sim N(0, \Sigma)$  and that  $\varepsilon$  is independent of  $z$ , this is a standard Heckman selection model (Heckman, 1979). Define

$$\Lambda = \begin{bmatrix} \lambda_1^r & \lambda_2^r \\ \lambda_1^o & \lambda_2^o \\ -\lambda_1^s & -\lambda_2^s \end{bmatrix}.$$

The parameters  $(\Sigma, \Lambda)$  are not identified. Instead, we define  $\eta = \Lambda \varepsilon$  and seek to estimate the parameters  $(\alpha, \beta, \gamma)$  and  $\Omega = \Lambda \Sigma \Lambda^T$  where  $\Omega$  is the covariance matrix of  $\eta$ , with matrix elements  $\omega_{ij}$ . Note that  $\eta \in \mathbf{R}_3$ , but since  $\varepsilon \in \mathbf{R}_2$  by assumption,  $\Omega$  is not full rank.  $\eta_1$  is the error in the rent equation,  $\eta_2$  is the error in the owner-occupied price equation, and  $\eta_3$  is the error in the selection equation.

Our data on prices, rents and allocations do not fully identify  $\Omega$ . As usual with discrete choice models, the variance parameter  $\omega_{33}$  is not identified. In addition,  $\omega_{12}$  is not identified. However, we can identify  $(\alpha, \beta)$  as well as  $(\tilde{\gamma}, \tilde{\Omega})$  (excluding  $\tilde{\omega}_{12}$ ) where the latter parameters are rescaled by the variance of  $\eta_3$  ( $\omega_{33}$ ).

For much of the analysis below, the rescaling plays no role. However, the rescaling does play an important role in the analysis of user costs, yields, and contracting costs in Section 5.4. For these parts of the analysis, we estimate  $\omega_{33}$  by fitting the predicted rental sector yields from the model to data on rental sector yields in Bracke (2015). We discuss this further in Section 5.4.

The model has several important features. First, the value of unobserved character-

istics in the owner-occupied sector is not restricted to be perfectly correlated with the value in the rental sector. Second, the impact of unobserved characteristics on selection is not restricted to be perfectly correlated with owner-occupied price nor with rents. Third, the correlation of  $\eta_3$  and  $\eta_2$  may differ from the correlation between  $\eta_3$  and  $\eta_1$ . We can identify the variance of  $\eta_1$  and the variance of  $\eta_2$ . We can also identify the correlation of  $\eta_1$  and  $\eta_3$ . We cannot identify the correlation of  $\eta_1$  and  $\eta_2$ . However, we show in Appendix A that because  $\tilde{\Omega}$  is rank 2, it can take on only two possible values.

## 2.1 User costs and the rent-to-price ratio

A hypothetical investor's willingness to pay for a property in a sector, either  $\pi^o(z, \varepsilon)$  or  $\pi^r(z, \varepsilon)$ , equals the stream of utility or rent flows from the property discounted by the sector-specific user cost of capital. The user cost of capital in a sector is determined by the effective rate of interest, the cost of maintenance and expectations about future capital gains, taking into account taxes, transactions costs, inflation, and uncertainty. We assume that these relationships can be characterized by two Poterba-like user cost equations (Poterba, 1992):

$$\pi^o(z, \varepsilon) = \frac{u(z, \varepsilon)}{r^o(z, \varepsilon) + c^o(z, \varepsilon) - g^o(z, \varepsilon)} \quad (6)$$

$$\pi^r(z, \varepsilon) = \frac{R(z, \varepsilon)}{r^r(z, \varepsilon) + c^r(z, \varepsilon) - g^r(z, \varepsilon)} \quad (7)$$

where for each sector  $i$ ,  $r^i(z, \varepsilon)$  is the effective discount rate,  $c^i(z, \varepsilon)$  is the cost of management and maintenance (including amortized vacancy costs), and  $g^i(z, \varepsilon)$  is expected capital gains.  $u(z, \varepsilon)$  is the utility flow for an owner-occupier.

Each element in these equations may vary both across property types and sector. Since tax policies and borrowing constraints frequently depend on tenure, there need not be a single discount rate that prices all property in equilibrium. For example, mortgage interest payments are not deductible from taxable income in England for owner-occupiers but are for landlords. This may be reflected in differences between  $r^o$  and  $r^r$ . Capital gains are not taxed for owner-occupiers but are taxed for landlords. We assume this is subsumed in differences between  $g^o$  and  $g^r$ . Lettings are exempt from Value Added Taxes in the UK but *net* rental income may be subject to income taxes. Assum-

ing a common income tax rate, this can be subsumed into  $c^r(z, \varepsilon)$ . Costs of vacancies in either sector can also be subsumed in  $c^i$ .

In general, for a property of type  $(z, \varepsilon)$ , user costs in the two sectors will differ. Properties with relatively high rental sector user costs will be selected into the owner-occupied sector. Owners of inframarginal properties will not be indifferent between the two sectors. For inframarginal owners in the owner-occupied sector,  $P(z, \varepsilon) = \pi^o(z, \varepsilon) > \pi^r(z, \varepsilon)$  while for inframarginal rental sector owners  $P(z, \varepsilon) = \pi^r(z, \varepsilon) > \pi^o(z, \varepsilon)$ . Only for owners of those properties at the margin, where  $P(z, \varepsilon) = \pi^o(z, \varepsilon) = \pi^r(z, \varepsilon)$ , do the two user cost equations (6)-(7) collapse to the more familiar, single equation user cost formula (such as the one in Poterba (1992)).

In the next two sections we discuss the data and estimation results. Then in Section 5, after discussing estimates of  $(\pi^o, R, \pi^r)$  and of selection of properties into the two sectors, we discuss the extent to which the estimates are consistent with different assumptions about how  $(r^i, g^i, c^i)$  vary across properties and across the two sectors.

### 3 Data

We use data from the “secure access” version of waves 2011-2014 of the English Housing Survey (EHS). The EHS uses a complex multistage methodology. Each wave comprises two surveys which are then combined to produce two samples. Each sample is constructed using data from surveys from multiple waves.

In each wave, the EHS team conducted a “household survey” and a “physical survey”. For example, to construct the 2011 wave, the EHS team sampled approximately 17,500 households in the financial year 2008/2009 (April 2008 - March 2009). These households were drawn from the list of addresses held by Royal Mail.<sup>6</sup>

Respondents from this selection (approximately 17,000) comprised the household interview sample. The EHS team then chose a subsample of these dwellings (approximately 8,000 in 2008/2009), including vacant ones, and performed a physical inspection. This is called the “physical survey.” The subsample was constructed from the 17,500 by including all social housing<sup>7</sup> addresses and taking a subsample of private

<sup>6</sup>At each sampled address, one dwelling was sampled. At each dwelling, one household was sampled.

<sup>7</sup>Social housing units are non-market rate rentals where the government either directly owns the properties or offers subsidies to landlords in the sector.

addresses. Private rental properties were over-sampled. Finally, to construct the final “housing stock” sample, the EHS team combined data from two physical surveys. For instance, the housing stock sample in the 2011 wave is comprised of the physical surveys from 2008/2009 and 2009/2010. Weighting for the final sample is based on this two year sampling window.

We focus discussion on the 2011 wave of the EHS. While we also analyze the 2012, 2013, and 2014 waves, these later waves have some limitations. In the later waves, property prices were top-coded at £1,000,000. Also, due to budget cuts, the later waves used smaller samples and collected information on a smaller range of topics. Despite these limitations, our results are robust across waves.<sup>8</sup>

Property prices recorded in the survey were obtained in one of two ways. For a subset of the owner-occupied properties, owners self-reported what they thought the market value of their home was. For the remainder of owner-occupied properties, a professional surveyor valued the property on-site. We obtain results similar to those reported below even if we restrict our sample of owner-occupied properties to either subset. Rental sector rents were self-reported by tenants.

Much of our analysis focuses on a subsample of dwellings within 140 km of Trafalgar Square in London. We call this region "Greater London." We restrict the analysis to this region because we want to focus on a single economic market.

We present summary statistics for the owner-occupied, private rental and social housing sectors. However, when we estimate the model, we restrict the analysis to private sector housing. Determinants of housing supply in the private sector are very different from those in the social housing sector. In the private sector, investors may buy and sell properties freely and prices are determined by the market. In the social sector, supply is largely determined by political forces, not by choices of investors. In addition, prices in this sector are subsidised and highly regulated.

Table 1 displays the overall market shares of owner-occupied housing, private rentals, and publicly assisted housing in England and Greater London. In England in 2008-2010, 67.9% of housing units were owner-occupied units while 14.3% were private rentals and 17.8% were publicly assisted units. In England, publicly assisted housing consists of Local Authority provided housing (LA) and housing provided by registered social

---

<sup>8</sup>Because the samples in each wave use data from a two year span, the samples overlap. For instance, the samples in the 2011 and 2012 waves each use the same data collected in 2009/2010.

landlords (RSL).<sup>9</sup>The Greater London area is roughly similar to the entire country with regards to tenure: there are slightly more private rentals and fewer owner-occupiers in London. The share of private rentals has increased by three percentage points over the four waves at the expense of owner-occupancy.

Table 1: Market shares: Greater London and England (%)

Region	EHS Wave	Owner-occupied	Private rented	LA or RSL
London	2011	66.6	15.7	17.7
	2012	65.3	17.0	17.8
	2013	63.1	18.7	18.1
	2014	62.4	19.5	18.2
England	2011	67.9	14.3	17.8
	2012	67.0	15.1	17.9
	2013	65.3	16.4	18.3
	2014	65.0	17.1	17.9

Note: Market shares are computed using sampling weights for each wave. London refers to the Greater London sample area. The 2011 wave uses data from April 2008 - March 2010. The 2012 wave uses data from April 2009 - March 2011. The 2013 wave uses data from April 2010 - March 2012. The 2014 wave uses data from April 2011 - March 2013.

Table 2 shows how market shares vary with distance from the center of London. Within 10 km of the center, the owner-occupied share is 37.9% while the private rented and social housing shares are 23.7% and 38.4% respectively. The owner-occupied share increases with distance. More than 50 km from the center, the owner-occupied share is 72.9% while the rental and social housing shares decline to 13.4% and 13.7% respectively. We will see that these patterns with respect to distance do not persist after controlling for structural characteristics.

---

<sup>9</sup>RSL's are non-profit organizations that provide low-cost housing. They are regulated by the government and highly subsidized.

Table 2: Market share by distance: Greater London 2011 wave (%)

Distance	Owner-occupied	Private rented	LA or RSL
Less than 10 km	37.9	23.7	38.4
10 - 20 km	61.6	19.8	18.6
20 - 30 km	69.8	13.5	16.8
30 - 50 km	71.4	13.1	15.5
More than 50 km	72.9	13.4	13.7

Note: Market shares are calculated using data and sampling weights from the 2011 wave of the EHS. The 2011 wave uses data from April 2008 - March 2010.

Tables 3 and 4 show how market shares vary with size and dwelling type. Large properties, semi-detached and detached houses and bungalows are much more likely to be in the owner-occupied sector while small properties, converted flats and dwellings in multi-unit structures are more likely to be in the rental sector. We will see that these patterns hold up even after controlling for location and other property characteristics. However, they cannot be explained by relative prices in the two sectors.

Table 3: Market share by dwelling size: Greater London 2011 wave (%)

Dwelling size	Owner-occupied	Private rented	LA or RSL
Less than 50 sq. m.	33.1	27.4	39.5
50 - 60 sq. m.	47.5	25.4	27.2
60 - 80 sq. m.	60.3	17.1	22.6
80 - 100 sq. m.	74.6	12.6	12.8
More than 100 sq. m.	90.1	7.24	2.63

Note: Market shares are calculated using data and sampling weights from the 2011 wave of the EHS. The 2011 wave uses data from April 2008 - March 2010.

Table 4: Market share by dwelling type: Greater London 2011 wave (%)

Dwelling Type	Owner-occupied	Private rented	LA or RSL
Semi detached	73.9	13.0	13.7
Detached	94.4	5.0	0.40
Bungalow	76.8	5.0	18.3
Converted flat	39.3	48.5	15.2
Low rise	32.2	26.7	38.4
High rise	20.7	19.7	48.1

Note: Market shares are calculated using data and sampling weights from the 2011 wave of the EHS. The 2011 wave uses data from April 2008 - March 2010. The semi-detached category includes "End Terrace" and "Mid Terrace".

## 4 Estimation procedure

We estimate the model parameters from equations (1), (2), and (5) using maximum likelihood. As noted above, the parameters  $\tilde{\gamma}$  and  $\tilde{\Omega}$  are normalized so that  $\tilde{\omega}_{33} = 1$ . For much of the analysis below, this normalization does not matter. However, in Section 5.4 where we discuss estimation of user costs, we do require an estimate of  $\omega_{33}$ . Hence, in Section 5.4 we use additional information on rental sector yields from Bracke (2013) to estimate  $\omega_{33}$ .

We explore several specifications. In the main specification, we include indicator variables for dwelling type and dwelling age<sup>10</sup>, an eighth-order polynomial in dwelling size (square meters), and a set of variables that account for the dwelling's geographic location. For owner-occupied properties, we also include an indicator for whether the property price is self-reported or not.

The confidential version of the EHS reports each dwelling's full postcode. We match each postcode with its geographic coordinates using the Office for National Statistics's Postcode Directory for 2013. Because postcodes can change over time, there are a few unmatched postcodes. In 2014, there is 1 unmatched owner-occupied property out of

<sup>10</sup>Dwelling types are detailed in Table 3. Dwelling age categories include: 1) pre-1919, 2) 1919 to 1944, 3) 1945 to 1964, 4) 1965 to 1980 or 5) post 1980.



5,184 and 2 unmatched private rentals out of 2,683 for all of England. In 2011, there are none. The numbers of unmatched properties for other waves are similar. For these unmatched properties, we imputed coordinates using the mean geographic coordinates of all postcodes sharing the same postcode district (postcodes are grouped geographically and the first three to four characters of a 7-8 character postcode are its postcode district). For each property, we then convert its geographic coordinates to polar coordinates  $(r, \theta)$  centered around Trafalgar square. For each property we compute  $r$ , the Euclidean distance from Trafalgar Square and  $\theta$ , the angular distance from due east measured in radians. That is,  $\theta = 0$ , is east,  $\theta = 0.5\pi$  is south, etc.

In our empirical model, we model location effects as a nonparametric function of  $(r, \theta)$ . In the main specification, we include the interaction of an eighth-order polynomial in distance with a fifth order trigonometric expansion in terms of  $\theta$ . The distance variable captures the impact of distance from London on property prices and on selection into the owner-occupied sector. The angular distance variable  $\theta$  captures variation in outcomes that depends on direction of travel. For example, the rate of decline of prices with distance is higher heading west than east. In a second specification, we drop the angular distance variable in order to measure the average effect of distance. We also explore including indicator variables for the numbers of bedrooms, kitchens, living rooms and bathrooms as well as using levels rather than logs of prices and rents. Results for these additional specifications are available from the authors upon request. For all cases, the sample is private rental and owner-occupied housing units within Greater London, using the sample weights provided in the EHS.

## 5 Results

Parameter estimates for the main specification are detailed in Table 5. Most of the parameters are statistically significant and have plausible values. Because the parameters are difficult to interpret we plot predicted values of log rent, log price and owner-occupied market share as functions of individual explanatory variables while holding other property characteristics fixed at baseline values. The baseline house is a semi-detached house with 75 square meters of floor space, 10 km Northwest of Trafalgar Square, built between 1919-1944. We also plot point-wise confidence bands for the predictions. The graphs are discussed below.

## 5.1 Location

The upper left panel of Figure 1 shows how prices and rents decline with distance conditional on moving Northwest from the centre. Pointwise confidence bands are illustrated with shaded areas. The decline of property prices and rents with distance is dramatic. The panel shows that an owner-occupied property 20 km from the center of London is worth only 30% of an identical property at the center. For the same change in distance, rent in the rental sector declines to 37% of the rent at the center. For the first 10 km, owner-occupied properties decline in price about 10% per km. Rental properties rents decline by about 8% per km. In both sectors, the hedonic functions flatten out significantly at distances greater than 10 km. In the owner-occupied sector, moving from 10 km to 20 km reduces property prices by about 2.0% per km and from 20 km to 40 km by about 1.0% per km. For rental properties, moving from 10 km to 20 km reduces rents by 1.0% per km and from 20 km to 40 km reduces rents by 0.5% per km.

The other panels in Figure 1 show a very similar pattern for three other directions, Northeast, Southwest and Southeast. Regardless of direction, the qualitative pattern is the same. Both prices and rents fall dramatically, prices fall faster than rents, and the functions flatten out after about 20 km and even more after 40 km. The function is flatter in the other directions and is almost completely flat after 40 km. In some directions (e.g. Southeast and Northeast), 140 km from Trafalgar Square is a point in the middle of the sea. Thus the confidence intervals in those directions blow up. In these cases, the estimated hedonic price function should be interpreted as a prediction for what the price of a property would be if it was feasible to fill in the sea and build new housing.

These results are obtained without controlling for lot size. Our data do not include information on lot size. This likely biases upward (towards zero) the estimated slopes of the hedonic rent and price functions with respect to distance. Lot sizes are probably larger further away from the city center where land is cheaper.

Figure 2 shows the estimated relationship between distance and the owner-occupancy rate moving Northwest from London. The “unconditional” line plots the estimated relationship between distance and ownership when no other correlates are included. The curve reproduces the numbers in Table 2<sup>11</sup>. The “conditional” line plots the relationship

---

<sup>11</sup>The “unconditional” line differs from the those in Table 2 only because the former is the homeownership rate for houses not in the LA/RSL.

with distance holding other characteristics fixed at their baseline values. Unsurprisingly, the unconditional line shows that owner-occupancy is far more prevalent 20 km outside of London than inside the city center. However, the conditional line shows that, once one controls for other housing unit characteristics, distance essentially plays no role in selection into owner-occupancy. Owner-occupancy as a function of distance is essentially flat at around 80%.

In summary, both rents and prices fall with distance. Rents relative to prices rise with distance but housing units are not more likely to be found in the rental sector. Why then don't investors in properties far from the centre convert more properties into rental units? We use equations (5) and (7) to frame an answer.

One possible explanation is that maintenance costs relative to rents rise as the value of the location falls. This point is best illustrated by thinking of the value of a property as being composed of the value of land and the value of built structure. In general, structure requires far more maintenance than land. So, maintenance costs are increasing in the proportion of the value that is structural. Given that the value of land decreases quickly as distance increases, the proportion of value that is structural increases with distance. A second possible explanation is that vacancy costs in the rental sector are higher further out because the rental markets there are thinner resulting in longer expected vacancy durations. A third possible explanation is that properties close to the city center had higher expected capital gains during the period of our study. In any case, no matter the cause, our results indicate that the *relative* value of a house on the two sectors ( $\pi_r$  versus  $\pi_o$ ) is not affected; the higher costs or lower expected capital gains of suburban properties is capitalized into rents.

Some limited evidence on these points can be obtained by studying changes in the hedonic functions over time. Figure 3 shows estimated hedonic prices and rents with respect to distance for all four waves of the EHS, 2011-2014. The functions become steeper over time and the rent-to-price ratio is unstable. In the most recent wave, 2014, the rental function is steeper with respect to distance than the price function, the opposite of the earlier waves. These facts suggest that perhaps (a) expected capital gains are location-dependent and (b) these relative difference across locations may change at a high frequency. Though we do not have good data on household expectations by location, it is at least true that ex-post capital gains have varied widely over this same time period. For example, UK Land Registry data that show that, from 2009 to 2014, prop-

erty prices within 10km of London increased 42% while prices further out increased only by 22%.<sup>12</sup>

## 5.2 Structure

### 5.2.1 Size

Figure 4 (left panel) shows how rents and prices change with respect to the total floor space of the property. Property prices increase 7.3% per 10 square meters. Rents increase 10.5% per 10 square meters. As a result, rent-to-price ratios increase with size. Figure 6 shows that this pattern remains stable over time. Other than for rents in the 2014 wave, most of parameter estimates measuring the impact of dwelling type and size are stable over time, whereas the impact of location on prices and rents (but not selection) is less stable. Therefore most of the medium to high frequency variation in prices and rents over time in our data is captured by changes in the value of location (i.e. land) and not in the valuation of dwelling type or size, consistent with findings throughout the housing literature (Davis and Heathcote (2007); Davis and Palumbo (2008); Amior and Halket (2014)).

The top right panel of Figure 4 shows how size affects the probability of being owner-occupied. Again, we compare the results from the selection model to an “unconditional” probit of ownership on size. The effects are dramatic. Unlike location and like dwelling type, size is hugely important for explaining variation in selection, even after controlling for other covariates. The market share of the owner-occupied sector increases from less than 60% for properties less than 50 square meters to more than 90% for properties larger than 150 square meters.

### 5.2.2 Dwelling type

Figure 5 shows how prices, rents and ownership vary with structure type. Property prices for detached houses and bungalows are about 22% higher than for semi-detached houses whereas converted flats and dwellings in low-rise units are about 20-22% cheaper. Rents follow a similar pattern. Rents for detached houses and bungalows are about 24% more expensive than semi-detached properties. The rent-to-price ratio is approximately

---

<sup>12</sup>Further details on these calculations are available from the authors upon request.

constant across the categories of semi-detached, detached and bungalows. For dwellings in multi-unit structures, rents are 33% lower than semi-detached properties. So, the rent-to-price ratio for multi-unit structures is much lower than for detached or semi-detached. These results are stable over waves.<sup>13</sup>

In contrast to location, the conditional relationship between dwelling type and predicted ownership is qualitatively similar to the unconditional relationship. The unconditional relationship is detailed in Table 4. Excluding the social housing sector, 95% of detached, 85% of semi-detached properties and 93.9% of bungalows are in the owner-occupied sector while various types of dwellings in multi-unit structures (converted flats, low rise and high rise) have ownership rates that vary between 44.8% and 54.7%. Physical features are important determinants of selection into the owner-occupied sector. Figure 5 shows that conditional on location and other characteristics, the average predicted ownership rate is between 80% and 90% for semi-detached, detached and bungalows and falls to around 60% for dwellings in multi-unit structures.

The pattern is similar to the stylized fact documented in Glaeser and Shapiro (2003) that in the US, housing units in multi-unit structures are extremely likely to be rented (85.9% in their study) whereas single-unit housing is very likely to be owned (85.5% in their study). Unconditional ownership rates do not vary quite as much in England across structure types (this is true even in the full sample). Conditioning narrows the difference still further.

In England property ownership predominantly takes one of two forms, freehold or leasehold. Freehold ownership is ownership in perpetuity. Leasehold ownership is ownership of a long lease (for example 75 years or 99 years).<sup>14</sup> It is clear that property prices should depend on the freehold or leasehold status of the property. Unfortunately the EHS only records information on the type of holding for owner-occupied properties. For the Greater London subsample, leaseholds comprise only 10 percent of owner-occupied properties. In addition, holding type is highly correlated with dwelling type. In the EHS sample, fewer than 23% of flats are freeholds while nearly 94% of detached houses are freeholds. Giglio et al. (2015) finds that leasehold flats sell for a noticeable duration-dependent discount compared to otherwise identical freeholds. So, some of

---

<sup>13</sup>Results from other waves on dwelling type as well as other unreported results are available upon request.

<sup>14</sup>A third form of ownership, “commonhold”, exists but is almost never used due to legal uncertainties.

the decline in prices that is captured by dwelling type may in fact be due to the higher prevalence of leaseholds for flats. However they also find that the type of holding does not affect rents. So the decline in rent-to-price ratios for flats is likely *understated* by omitting holding type. When an indicator for ownership type is included in the property price equation, the parameter estimate is 0.1198 (with a standard error of 0.0447). Freehold status increases property prices nearly 12% relative to leasehold status. The estimates in Figure 5 are robust to including an indicator for leasehold in our estimation of equation (2).

### 5.3 Unobserved quality

The final two columns of Table 5 report estimates of the error correlations from the selection model. For the rental sector results, the correlation between  $\eta_1$  and  $\eta_3$  is  $-0.9759$  (0.0048). Properties that are likely to be in the rental sector (high  $\eta_3$ ) have much lower unobserved rental quality ( $\eta_1$ ). For the owner-occupied sector, the correlation between  $\eta_2$  and  $\eta_3$  is only 0.1040 (0.0533). It is not statistically distinct from zero at the 5% level. Selection into the owner-occupied sector has zero or low correlation with unobserved owner-occupied quality ( $\eta_2$ ).

The implication of this is illustrated in the bottom panel of Figure 4. The figure shows, for an average housing unit, how the predicted average unobserved owner-occupied quality ( $\eta_2$ ) and rental quality ( $\eta_1$ ) vary with size. Conditional on being in the owner-occupied sector, average unobserved owner-occupied quality does not vary with size. However, this is not true in the rental sector. Conditional on being in the rental sector, bigger dwellings have much lower unobserved rental quality. The average rental quality difference between a 50 square meter rental property and a 100 square meter rental property is almost 22%. Large housing units in the rental sector are likely to be of much lower unobserved rental quality.

Characteristics that are unobservable to econometricians may suffer more acutely from third-party verification problems. Enforcing contracts to invest in and/or maintain these characteristics may be particularly costly, if possible at all. Landlords may therefore choose properties with fewer of these characteristics.<sup>15</sup> With this in mind, one way

---

<sup>15</sup>Kanemoto (1990) builds a theory of underinvestment in housing and security of tenure that has a similar mechanism.

to explain the results is as follows. Suppose  $\lambda_1^r, \lambda_2^o, \lambda_1^s > 0$  and  $\lambda_1^o, \lambda_2^s = 0$ . In this case, one could think of  $\varepsilon_1$  as an amenity that affects users' enjoyment of a property but that comes at a high maintenance cost. Such an amenity raises rents, is negatively correlated with selection into the rental sector, and has zero net impact on property prices. The high maintenance cost exactly offsets the use value.  $\varepsilon_2$  on the other hand could be an amenity that affects both prices and rents but does not affect selection, perhaps because there are no maintenance concerns associated with it. In other words,  $\varepsilon_1$  could be a jacuzzi - nice to use but a nightmare to maintain, while  $\varepsilon_2$  could be the unimpeded light from South facing windows.

Finally, some models of homeownership and housing demand use a preference specification which includes a preference for owning (a.k.a a "warm glow" from owning).<sup>16</sup> In calibration exercises, such a preference for owning is often required to generate high homeownership rates. Our results show that an econometrician measuring demand for homeownership using only observable housing characteristics would indeed find a preference for owning. This is because rentals, on average, have lower unobserved rental quality.

## 5.4 Implications for user cost and yields

To estimate the model, we rescaled  $\gamma$  and  $\Omega$  by  $\sqrt{\omega_{33}}$ . Let  $\tilde{\eta}_3$  be the normalized error term. We now derive formula for user costs and yields based on these parameters.

Let  $u^o = \frac{u(z, \varepsilon)}{\pi^o(z, \varepsilon)}$  and  $u^r = \frac{R(z, \varepsilon)}{\pi^r(z, \varepsilon)}$  and recall that  $\eta = \Lambda\varepsilon$ . The variable  $u^r$  also is the yield on a rental property (gross of maintenance expenses). If we assume that the service flows from dwelling  $(z, \varepsilon)$  are the same regardless of which sector the dwelling is in, then  $R(z, \varepsilon) = u(z, \varepsilon)$ . Combining this with equations (1) - (3) implies that

$$\ln u^o = (\alpha - \beta)z + \eta_1 - \eta_2 \quad (8)$$

$$\ln u^r = (\alpha - \beta + \sqrt{\omega_{33}}\tilde{\gamma})z + \eta_1 - \eta_2 - \sqrt{\omega_{33}}\tilde{\eta}_3. \quad (9)$$

Both user costs and yields depend on the unknown parameter  $\omega_{33}$ . Since  $\eta$  is normally

---

<sup>16</sup>For example, see Iacoviello and Pavan (2013) or Kiyotaki et al. (2011).

distributed, we can calculate the conditional mean rental sector yield as

$$\begin{aligned}
E(u^r | rental) &= \int_z E(u^r | z, \tilde{\eta}_3 \geq \tilde{\gamma}z) f(z | rental) dz \\
&= \int_z \frac{1 - \Phi(\tilde{\gamma}z - \Psi_{12})}{1 - \Phi(\tilde{\gamma}z)} e^{(\alpha - \beta + \sqrt{\omega_{33}}\tilde{\gamma})z + \Psi_{11}} f(z | rental) dz
\end{aligned} \tag{10}$$

where  $\Psi_{12} = cov(\tilde{\eta}_3, \eta_1 - \eta_2 + \sqrt{\omega_{33}}\tilde{\eta}_3)$  and  $\Psi_{11} = var(\eta_1 - \eta_2 + \sqrt{\omega_{33}}\tilde{\eta}_3)$ .

Figure 3 in Bracke (2015), based on data on rents and prices for the UK buy-to-let market, shows that the average rental sector yield in London in 2009-2014 was 0.062. Matching this moment to equation (10), we find  $\omega_{33} \in \{0.011, 0.215\}$ . There are two possible values corresponding to the two possible values of  $\tilde{\omega}_{12}$  derived in Appendix A.

Finally, equations (8) and (9) also imply that the ratio of user costs satisfies:

$$\ln u^r - \ln u^o = \gamma z - \eta_3. \tag{11}$$

In the following sections we discuss estimates of user costs for the same baseline semi-detached house as above. We present results for two cases,  $\omega_{33} = 0.215$  and  $\omega_{33} = 0.011$ .

#### 5.4.1 Variation with respect to location

As discussed above, the rent-to-price ratio increases with distance in several waves but decreases with distance in 2014. At the same time, location is unimportant for selection. Using equation (11), these facts imply that differential users costs between the two sectors do not vary much with distance. This can be seen in the top panels of Figures 8-9 which show average log user costs in the two sectors as a function of distance for the two estimates of  $\omega_{33}$ . The two curves increase with distance but are nearly parallel. For the high estimate of  $\omega_{33}$ , the gap between the two log user costs curves is 0.40 at a distance of zero and 0.45 at a distance of 40 km.<sup>17</sup> In other words, while effective interest rates, maintenance costs, and expected capital gains, may vary with distance from London, they do display significant differential variation across the two housing

<sup>17</sup>For the case  $\omega_{33} = 0.011$ , the results are qualitatively similar. The difference in log user costs increases from 0.09 to 0.10.



sectors.<sup>18</sup> As discussed in Section 5.1, it is likely that both rental and owner-occupied maintenance costs as a proportion of value rise with distance. Apparently though, in some time periods like those covered in the 2014 wave, this maintenance effect on the rent-to-price ratio is dominated by other effects. That is, during this period either the discount rate in London went up relative to outside London or relative expected capital gains fell in London.

#### 5.4.2 Variation with respect to structure

Above we found that the more detached and/or the larger a property is the higher is its rent-to-price ratio but the lower is its likelihood of being a rental. Detachedness and size are each positively valued and are each negatively correlated with being in the rental sector. Considering equations (6) and (7), this implies that either  $\frac{r^o}{r^r}$  or  $\frac{c^o}{c^r}$  decreases or  $\frac{g^o}{g^r}$  increases with detachedness or with size. The lower two panels in Figure 8 show the resulting predictions for average log user costs as functions of size and dwelling type in the two sectors. For the high estimate of  $\omega_{33}$ , user costs in the owner-occupied sector increase from about 6.7% to 9.1% when moving from a 50 square meter property to 110 square meters. For the same change in property size, rental sector user costs rise from 8.2% to 14.2% of the price of the property. Similarly owner-occupied user costs are 7.7% for detached dwellings vs. 5.3% for dwellings in high rise multi-unit structures. Detached property rental sector user costs are 11.1% vs. 5.5% for flats in high rise structures. There are large increases in relative rental sector user costs for bigger properties and for more detached properties.<sup>19</sup>

What explains these increases? Is it likely that  $\frac{g^o}{g^r}$  is dramatically different for detached houses than for dwellings in multi-unit structures? Is this ratio likely to be dramatically different for 110 square meter flats versus 50 square meter flats? This ratio may vary slightly with these physical features due to sectoral differences in taxation of capital gains. For instance, capital gains below a certain threshold are tax exempt for owner-occupiers. However, any variation due to differential tax treatments should also

<sup>18</sup>It seems unlikely that the individual factors vary differentially in such a way to cancel one another out.

<sup>19</sup>If  $\omega_{33} = 0.011$ , user-costs increase from 6.7% to 8.8% and 7% to 9.8% for owner-occupied and rental houses, respectively, when size increases from 50 square meters to 110 square meters and they decrease from 7.7% to 5.3% and 8.3% to 5.3% for owner-occupied and rental houses, respectively, when the houses changes from a detached property to a high rise.

be reflected in our findings with respect to location. This is not the case. Location is uncorrelated with ownership after controlling for observable and unobservable structural characteristics. As discussed above, this suggests that even if expected capital gains varies with location,  $\frac{g^o}{g^r}$  is roughly constant.

In contrast, in the case of costs, it is theoretically plausible that rental costs  $c^r$  increase faster than  $c^o$  when size increases or when one compares detached houses to dwellings in multi-unit structures. This is the direct or indirect implication of Galster (1983), Henderson and Ioannides (1983) and Coulson and Fisher (2014). Our findings suggest that this differential increase in costs is large. We compute an estimate of an upper bound for this magnitude in Section 5.4.3.

In the case of interest costs, it is also theoretically plausible that interest costs in the rental sector,  $m^r$ , increase faster than  $m^o$ . For instance, some property owners (either landlords or owner-occupiers) may face tighter borrowing constraints than others. As a result, they may face higher interest rates. It is possible that less constrained owner-occupiers tend to live in physically more valuable housing units (but not locationally more valuable housing units). In this case, the marginal owner-occupier's discount rate would decline with structure value. At the same time, if the marginal landlord's discount rate was independent of property characteristics, then  $m^o/m^r$  would decrease relative to physical value. Our study does not provide direct evidence that can shed further light on the extent to which this theoretical possibility is empirically relevant. Further investigation is required to determine whether and to what degree mortgage costs vary across owner-occupiers, across landlords, and across different types of housing units.

### 5.4.3 Implication for maintenance costs

Based on our above estimates, we can back out an estimate of how tenure affects maintenance costs for a property if we make additional assumptions about the discount rates and expected capital gains. For instance, if we make the extreme assumption that  $r^r - g^r = r^o - g^o$ , equation (11) becomes:

$$\frac{1 + \frac{c^r}{r-g}}{1 + \frac{c^o}{r-g}} = e^{\sqrt{\omega_{33}}\tilde{\gamma}z - \sqrt{\omega_{33}}\tilde{\eta}_3}.$$

In Figures 10-11, we calibrate  $c^o = .017$  (consistent with measures of owner-occupied depreciation in Gatzlaff et al. (1998); Malpezzi et al. (1987); Amior and Halket (2014), among others) and set  $r - g = 0.01$  for both sectors. For each estimate of  $\omega_{33}$ , we then plot the conditional and unconditional mean values of  $c^r$  as functions of location, size, and dwelling type. The conditional mean functions show the means conditional on being either in the rental or the owner-occupied sector. The unconditional average,  $E(c^r)$ , which measures the average cost of maintenance if all properties were rented, is higher than  $c^o$  and increases faster with  $z$ . The conditional mean,  $E(c^r | owner)$ , which measures the counterfactual cost of maintaining owner-occupied properties were they instead rented, is still higher. In contrast, the conditional mean,  $E(c^r | rent)$  is lower. As one would expect, properties in the rental sector have lower average renter-specific maintenance costs. This stems from the fact that these properties have lower unobserved rental quality. These findings suggest why estimates of differential maintenance costs that do not control for selection fail to find higher costs in the rental sector. They are not comparing counterfactual maintenance costs. Conditional on the extreme assumptions in this section, this shows that the moral hazard problem for landlords and tenants can be very large.

## 5.5 Bias in imputed rents and/or prices

Beyond simply biasing estimates of maintenance costs, the estimation results imply that hedonic estimates of rents and prices that do not control for selection are also biased. For rents, these biases are statistically and economically significant. Moreover the bias creates striking patterns in the data.

To illustrate the bias, we re-estimate  $\alpha$  and  $\beta$  without controlling for selection. We then use these biased estimates to predict rents and prices for all housing units in the sample. Figure 12 plots the predicted rent-to-price ratios against the predicted log price from the biased regressions for both rentals and owner-occupied properties. The biased estimates imply that homeownership rates are increasing in the predicted price of a home even though the predicted rent-to-price ratio declines in the price of the home. This is a common finding: Verbrugge (2008); Heston and Nakamura (2009); Verbrugge and Poole (2010); Bracke (2013); Epple et al. (2013) all find that rent-to-price ratios decline with prices while, similarly, Landvoigt et al. (Forthcoming) estimates that housing

service flows rise less than one-for-one with property prices in the cross-section.<sup>20</sup> The relationships shown in Figure 12 and in these studies are puzzling from a certain angle and are a challenge for models which attempt to explain the distribution of household homeownership choices: why do so many households choose to buy expensive properties when seemingly equivalent rental properties are relatively cheap? Estimates from our selection corrected model provide the answer: The rentals are not equivalent. More expensive properties in the rental sector on average have lower unobserved quality.

Using parameter values from the biased regression results, Figure 13 (left, center and right) plots the predicted prices and rents as functions of distance, size, and dwelling type respectively. Comparing these figures to their counterparts from section 4 (Figures 1, 4, and 5 respectively), several facts are apparent. There is no bias in the hedonic estimates with respect to location. However, the predicted relationships between rents and physical characteristics - dwelling type and size - are biased. In fact, the relationship between rent-to-price and price predicted by the biased estimates is negative whereas the relationship predicted by the selection corrected model is positive. Using the biased estimates, one would conclude that larger, detached homes (which are more expensive) have lower rent-to-price ratios. The selection corrected model results show that the opposite is true. This dichotomy between location and structure was already evident from our findings discussed above. These figures show how large the bias is: it qualitatively reverses patterns. Correcting for the bias eliminates the puzzle raised by Figure 12.

## 6 Conclusion

Housing units are not randomly selected into a housing sector. Physical attributes including some that are unobservable in our data are important for selection. Location is not. These findings are consistent with theories of contracting frictions over maintenance and upkeep of the property. Most existing models of households' homeownership decisions, such as Landvoigt et al. (Forthcoming); Cocco (2005); Diaz and Luengo-Prado (2008); Henderson and Ioannides (1983), largely have abstracted away from explicit considerations of the multi-characteristic nature of housing units. To understand the puzzles of homeownership, our findings point to a need to examine both sector-

---

<sup>20</sup>Halket and Pignatti (2014) shows that some of this co-variation can be explained by differences in rental vacancy rates.

specific housing costs as well as, on the demand side, to model household choices of both ownership and housing characteristics. Perhaps, households that have a higher demand for larger housing units or detached houses or housing units with high maintenance amenities are more likely to save for a downpayment everything else equal.

The results also imply that properly accounting for the bias that selection imparts may encourage refinements in the construction of price indices both for housing and for consumer prices as well as national accounts. It may also help to better understand some of the relative movements of rents and prices over time such as those documented in Campbell et al. (2009).

## References

- AMIOR, M. AND J. HALKET (2014): “Do Households Use Homeownership To Insure Themselves? Evidence Across U.S. Cities,” *Quantitative Economics*, 5.
- BLOW, L. AND L. NESHEIM (2009): “Dynamic housing expenditures and household welfare,” .
- BRACKE, P. (2013): “House Prices and Rents: Micro Evidence from a Matched Dataset in Central London,” SERC Discussion Papers 0127, Spatial Economics Research Centre, LSE.
- (2015): “How Much Do Investors Pay for Houses?” .
- CAMPBELL, J. Y. AND J. F. COCCO (2007): “How do house prices affect consumption? Evidence from micro data,” *Journal of Monetary Economics*, 54, 591–621.
- CAMPBELL, S. D., M. A. DAVIS, J. GALLIN, AND R. F. MARTIN (2009): “What moves housing markets: A variance decomposition of the rent-price ratio,” *Journal of Urban Economics*, 66, 90–102.
- CASAS-ARCE, P. AND A. SAIZ (2010): “Owning versus Renting: Do Courts Matter?” *Journal of Law and Economics*, 53, 137–165.

- CHAMBERS, M., C. GARRIGA, AND D. SCHLAGENHAUF (2009a): “The Loan Structure and Housing Tenure Decisions in an Equilibrium Model of Mortgage Choice,” *Review of Economic Dynamics*, 12, 444–468.
- CHAMBERS, M., C. GARRIGA, AND D. E. SCHLAGENHAUF (2009b): “Accounting For Changes In The Homeownership Rate,” *International Economic Review*, 50, 677–726.
- CHEN, K. (2010): “A Life-Cycle Analysis of Social Security with Housing,” *Review of Economic Dynamics*, 13, 597–615.
- COCCO, J. (2005): “Portfolio choice in the presence of housing,” *Review of Financial Studies*, 18, 535–567.
- COULSON, N. E. AND L. M. FISHER (2014): “Houses, apartments and condos: The governance of multifamily housing,” Tech. rep., Penn State University.
- DAVIS, M. A. AND J. HEATHCOTE (2007): “The Price and Quantity of Residential Land in the United States,” *Journal of Monetary Economics*, 54, 2595–2620.
- DAVIS, M. A. AND M. G. PALUMBO (2008): “The price of residential land in large US cities,” *Journal of Urban Economics*, 63, 352–384.
- DIAZ, A. AND M. J. LUENGO-PRADO (2008): “On the user cost and homeownership,” *Review of Economic Dynamics*, 11, 584–613.
- EPPLE, D., L. QUINTERO, AND H. SIEG (2013): “Estimating Hedonic Functions for Rents and Values in the Presence of Unobserved Heterogeneity in the Quality for Housing,” Tech. rep.
- GALSTER, G. C. (1983): “Empirical Evidence on Cross-Tenure Differences in Home Maintenance and Conditions,” *Land Economics*, 59, pp. 107–113.
- GATZLAFF, D. H., R. K. GREEN, AND D. C. LING (1998): “Cross-Tenure Differences in Home Maintenance and Appreciation,” *Land Economics*, 74, pp. 328–342.
- GIGLIO, S., M. MAGGIORI, AND J. STROEBEL (2015): “Very Long-Run Discount Rates,” *The Quarterly Journal of Economics*, 130, 1–53.

- GLAESER, E. L. AND J. GYOURKO (2007): “Arbitrage in Housing Markets,” NBER Working Papers 13704, National Bureau of Economic Research, Inc.
- GLAESER, E. L. AND J. M. SHAPIRO (2003): “The Benefits of the Home Mortgage Interest Deduction,” in *Tax Policy and the Economy, Volume 17*, National Bureau of Economic Research, Inc, NBER Chapters, 37–82.
- GOODMAN, A. C. (1988): “An econometric model of housing price, permanent income, tenure choice, and housing demand,” *Journal of Urban Economics*, 23, 327 – 353.
- HALKET, J. AND M. PIGNATTI (2014): “Homeownership and the Scarcity of Rentals,” Tech. rep., University of Essex, Department of Economics.
- HANSON, A. (2012): “Size of home, homeownership, and the mortgage interest deduction,” *Journal of Housing Economics*, 21, 195–210.
- HARDING, J., T. J. MICELI, AND C. SIRMANS (2000): “Do Owners Take Better Care of Their Housing Than Renters?” *Real Estate Economics*, 28, 663–681.
- HECKMAN, J. J. (1979): “Sample selection bias as a specification error,” *Econometrica: Journal of the econometric society*, 153–161.
- HENDERSON, J. V. AND Y. M. IOANNIDES (1983): “A model of housing tenure choice,” *American Economic Review*, 73, 98–113.
- HESTON, A. AND A. O. NAKAMURA (2009): “Questions about the equivalence of market rents and user costs for owner occupied housing,” *Journal of Housing Economics*, 18, 273 – 279.
- HILBER, C. A. (2005): “Neighborhood externality risk and the homeownership status of properties,” *Journal of Urban Economics*, 57, 213–241.
- IACOVIELLO, M. AND M. PAVAN (2013): “Housing and debt over the life cycle and over the business cycle,” *Journal of Monetary Economics*, 60, 221–238.
- KAN, K. (2000): “Dynamic modeling of housing tenure choice,” *Journal of Urban Economics*.

- KANEMOTO, Y. (1990): “Contract types in the property market,” *Regional Science and Urban Economics*, 20, 5–22.
- KIYOTAKI, N., A. MICHAELIDES, AND K. NIKOLOV (2011): “Winners and Losers in Housing Markets,” *Journal of Money, Credit and Banking*, 43, 255–296.
- LANDVOIGT, T., M. PIAZZESI, AND M. SCHNEIDER (Forthcoming): “The Housing Market(s) of San Diego,” *American Economic Review*.
- LINNEMAN, P. (1985): “An economic analysis of the homeownership decision,” *Journal of Urban Economics*, 17, 230–246.
- MALPEZZI, S., L. OZANNE, AND T. G. THIBODEAU (1987): “Microeconomic Estimates of Housing Depreciation,” *Land Economics*, 63, 372–385.
- POTERBA, J. M. (1992): “Taxation and Housing: Old Questions, New Answers,” *American Economic Review*, 82, 237–42.
- PRABHALA, N. R. (2008): “Self-selection models in corporate finance,” *Handbook of Empirical Corporate Finance SET*, 1, 37.
- ROSEN, H. (1979): “Housing Decisions and the US Income Tax: An Econometric Analysis,” *Journal of Public Economics*, 11, 1–24.
- SHILLING, J. D., C. SIRMANS, AND J. F. DOMBROW (1991): “Measuring depreciation in single-family rental and owner-occupied housing,” *Journal of Housing Economics*, 1, 368 – 383.
- SWEENEY, J. L. (1974): “Housing unit maintenance and the mode of tenure,” *Journal of Economic Theory*, 8, 111 – 138.
- VERBRUGGE, R. (2008): “The Puzzling Divergence Of Rents And User Costs, 1980-2004,” *Review of Income and Wealth*, 54, 671–699.
- VERBRUGGE, R. AND R. POOLE (2010): “Explaining the Rent-OER Inflation Divergence, 1999-2007,” *Real Estate Economics*, 38, 633–657.



## A Identification of $\Omega$

All the elements of  $\Omega$  are point-identified except  $\omega_{12}$ . The identified set for  $\omega_{12}$  consists of two points. To see this consider the following. Since  $\Omega$  is rank 2, one of its columns can be written as a linear combination of the others. This implies

$$\begin{aligned} \omega_{11} &= \delta_1 \omega_{12} + \delta_2 \omega_{13} \\ \begin{bmatrix} \omega_{21} \\ \omega_{31} \end{bmatrix} &= \begin{bmatrix} \omega_{22} & \omega_{23} \\ \omega_{32} & \omega_{33} \end{bmatrix} \delta \end{aligned}$$

where  $\delta = (\delta_1, \delta_2)$  is a vector of weights. Solving the second two equations for  $\delta$  yields

$$\begin{aligned} \delta_1 &= \frac{\omega_{33} \omega_{12} - \omega_{23} \omega_{13}}{D} \\ \delta_2 &= \frac{-\omega_{23} \omega_{12} + \omega_{22} \omega_{13}}{D} \end{aligned}$$

where  $D = \omega_{22} \omega_{33} - \omega_{23}^2$ . Substituting this into the first equation yields

$$\omega_{11} = \omega_{12} \left( \frac{\omega_{33} \omega_{12} - \omega_{23} \omega_{13}}{D} \right) + \omega_{13} \left( \frac{-\omega_{23} \omega_{12} + \omega_{22} \omega_{13}}{D} \right)$$

which is a quadratic equation in the unknown  $\omega_{12}$ . This equation can be written as

$$\frac{\omega_{33}}{D} \omega_{12}^2 - 2 \left( \frac{\omega_{23} \omega_{13}}{D} \right) \omega_{12} + \left( \frac{\omega_{22} \omega_{13}^2}{D} - \omega_{11} \right) = 0.$$

This equation has solutions

$$\omega_{12} = \frac{\omega_{23} \omega_{13} \pm \sqrt{\omega_{23}^2 \omega_{13}^2 - \omega_{33} (\omega_{22} \omega_{13}^2 - D \omega_{11})}}{\omega_{33}}.$$

## B Figures and Tables

Figure 1: Rent and price vs. location

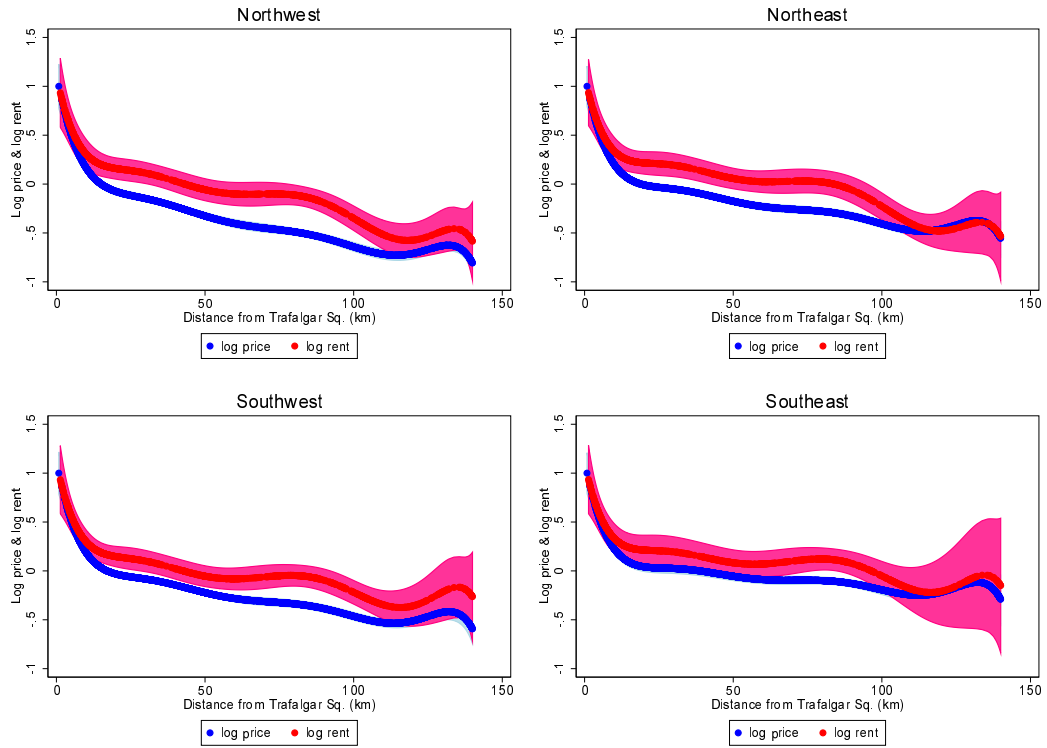


Figure 2: Ownership vs. location

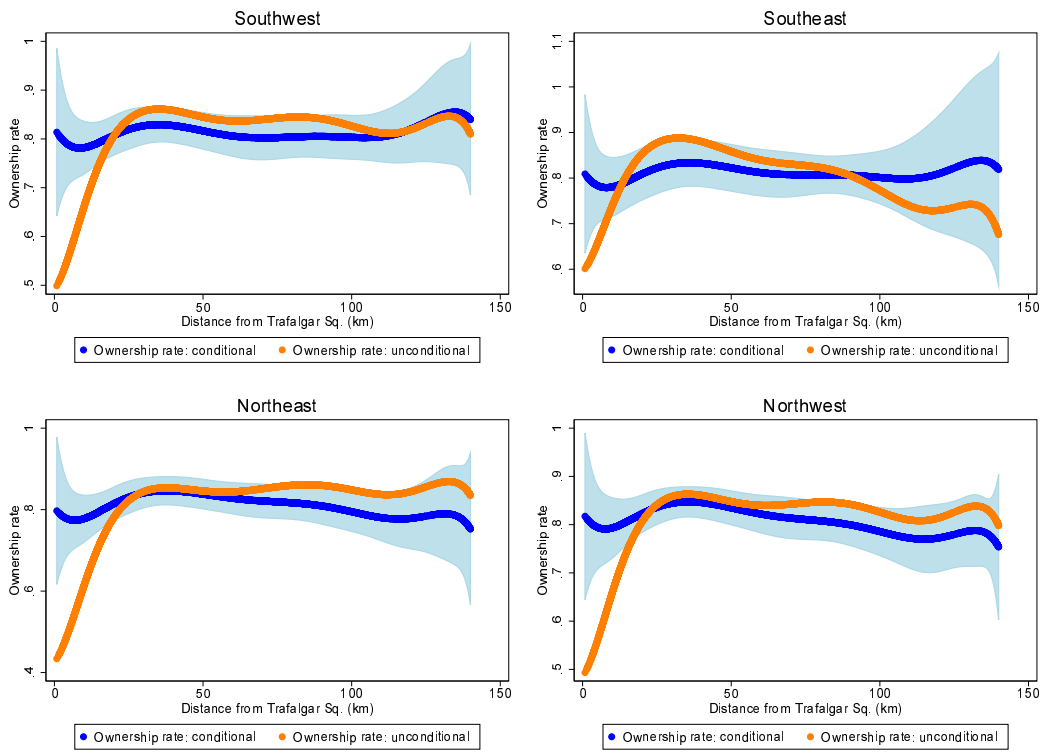


Figure 3: Rent and price vs. location: time variation

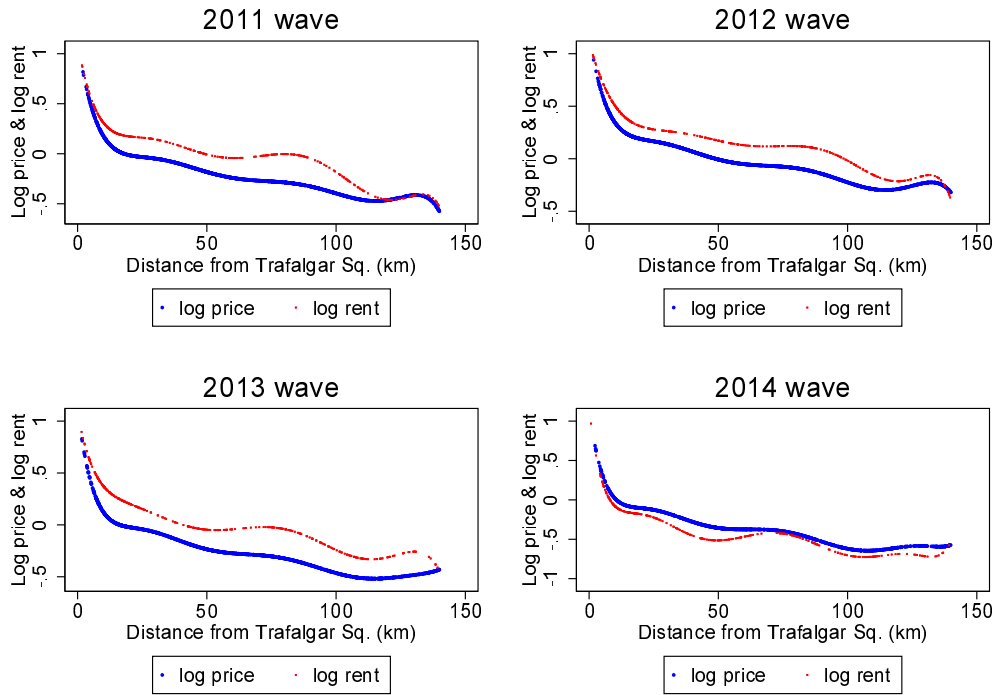


Figure 4: Rent, price, ownership vs. size

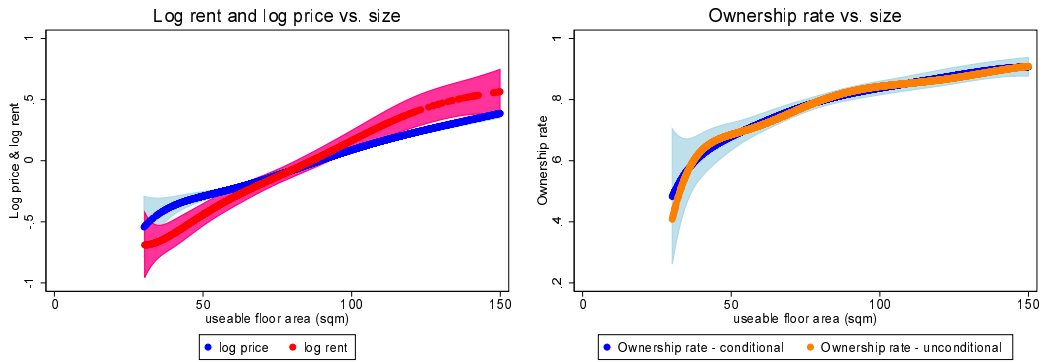


Figure 5: Rent, price and ownership vs. dwelling type

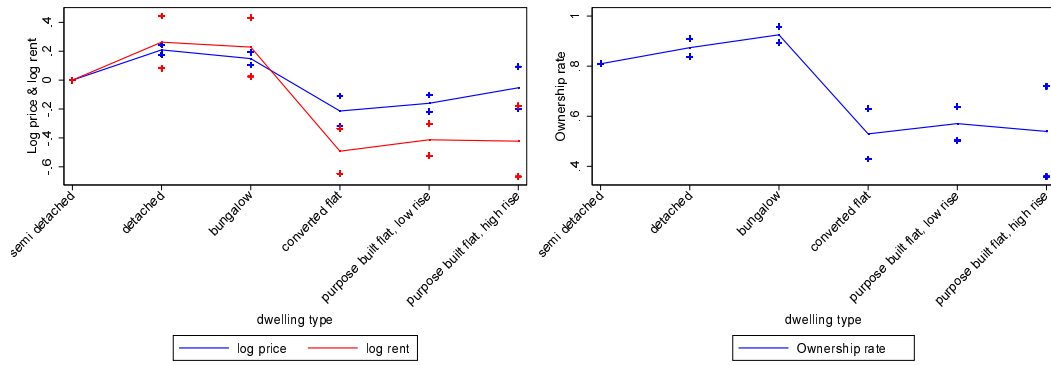


Figure 6: Rent and price vs. size: time variation



Figure 7: Unobserved quality vs. size

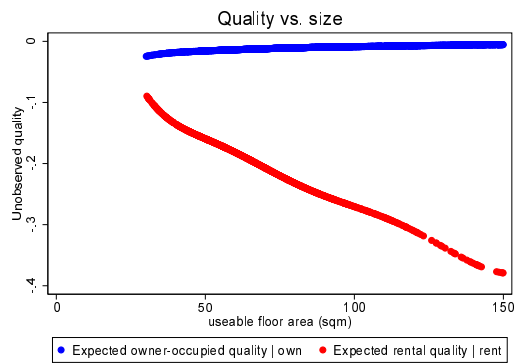


Figure 8: Log user costs vs. characteristics  $z$  :  $\omega_{33} = 0.215$

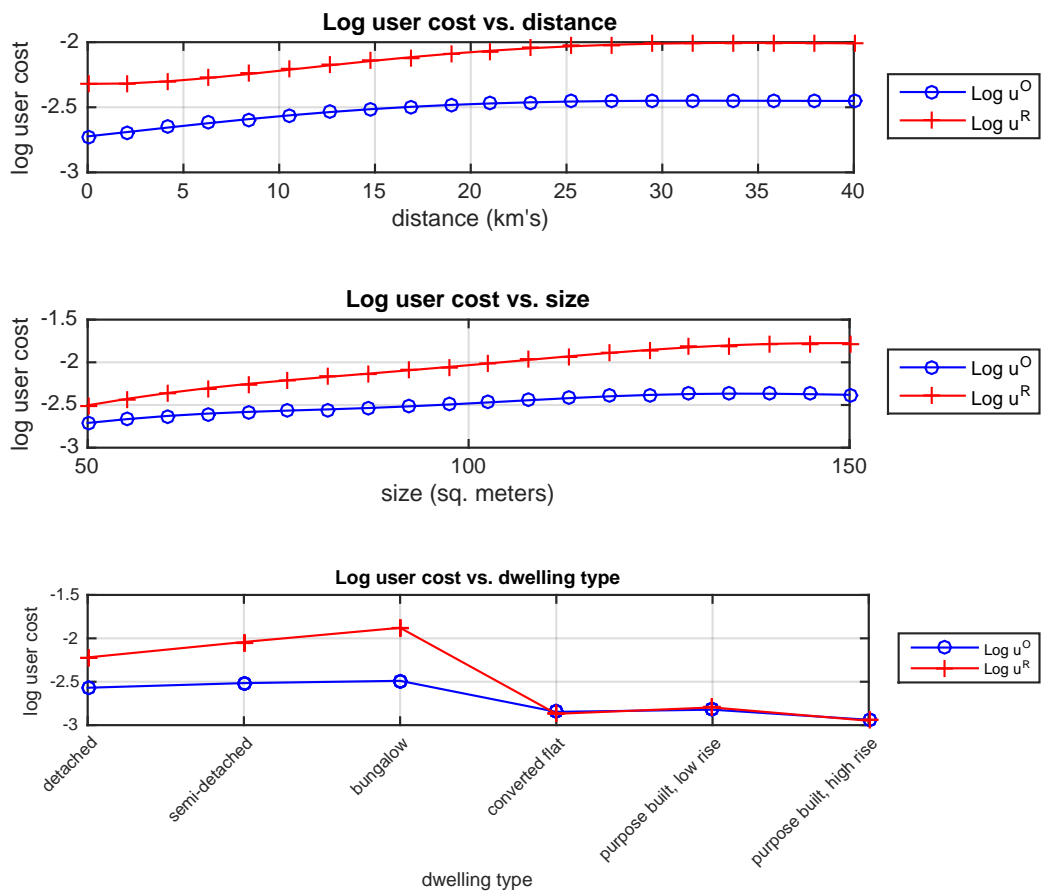


Figure 9: Log user costs vs. characteristics  $z$ :  $\omega_{33} = 0.011$

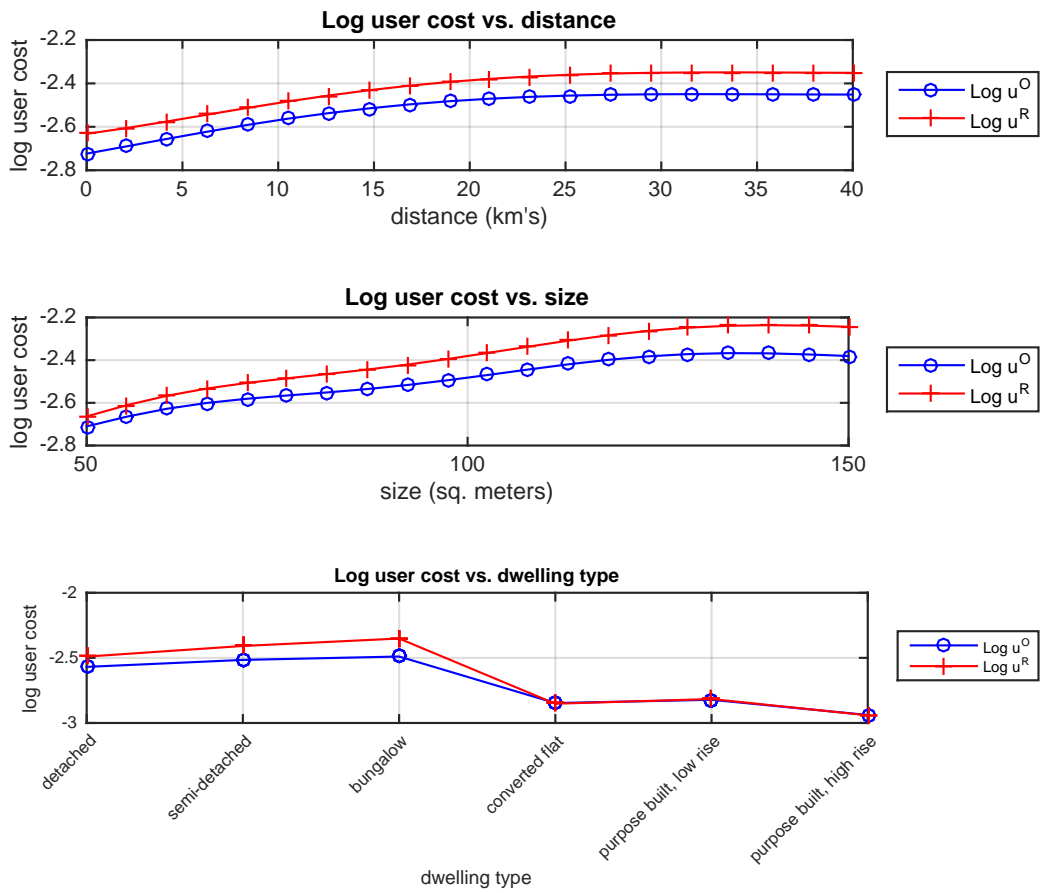


Figure 10: Contracting costs vs. characteristics  $z$ :  $\omega_{33} = 0.215$

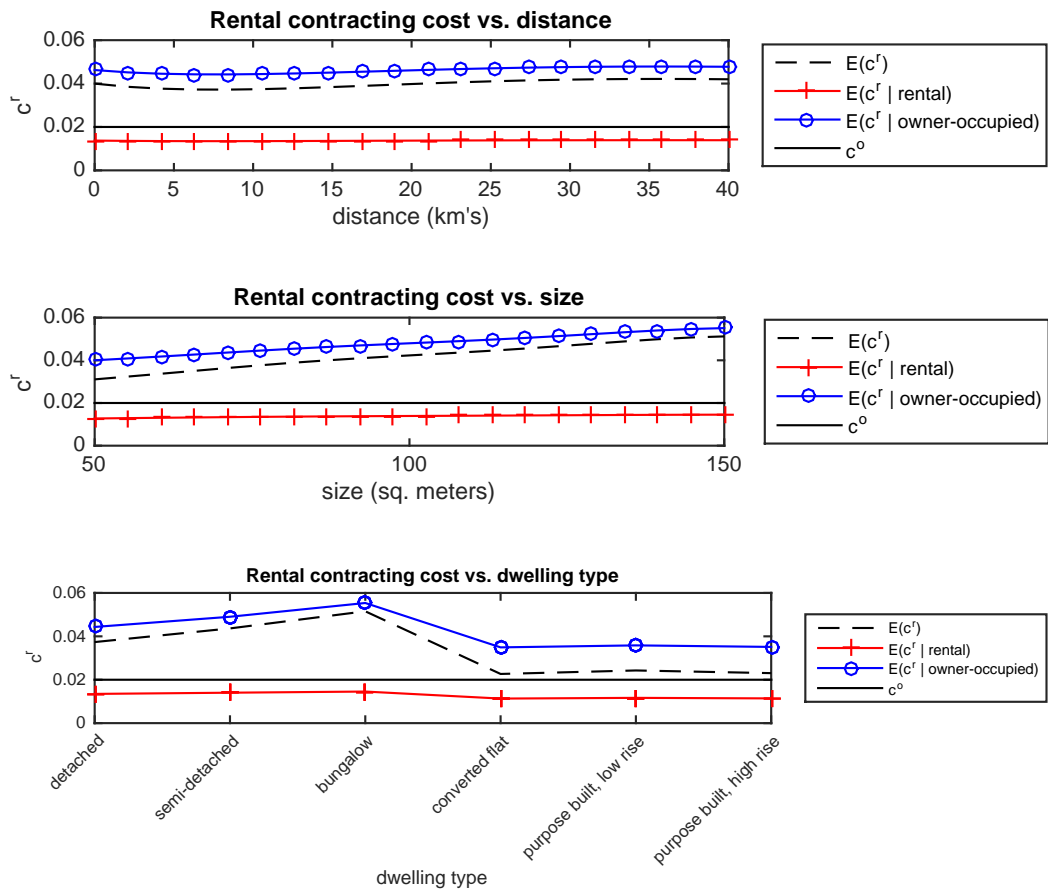




Figure 11: Contracting costs vs. characteristics  $z$ :  $\omega_{33} = 0.011$

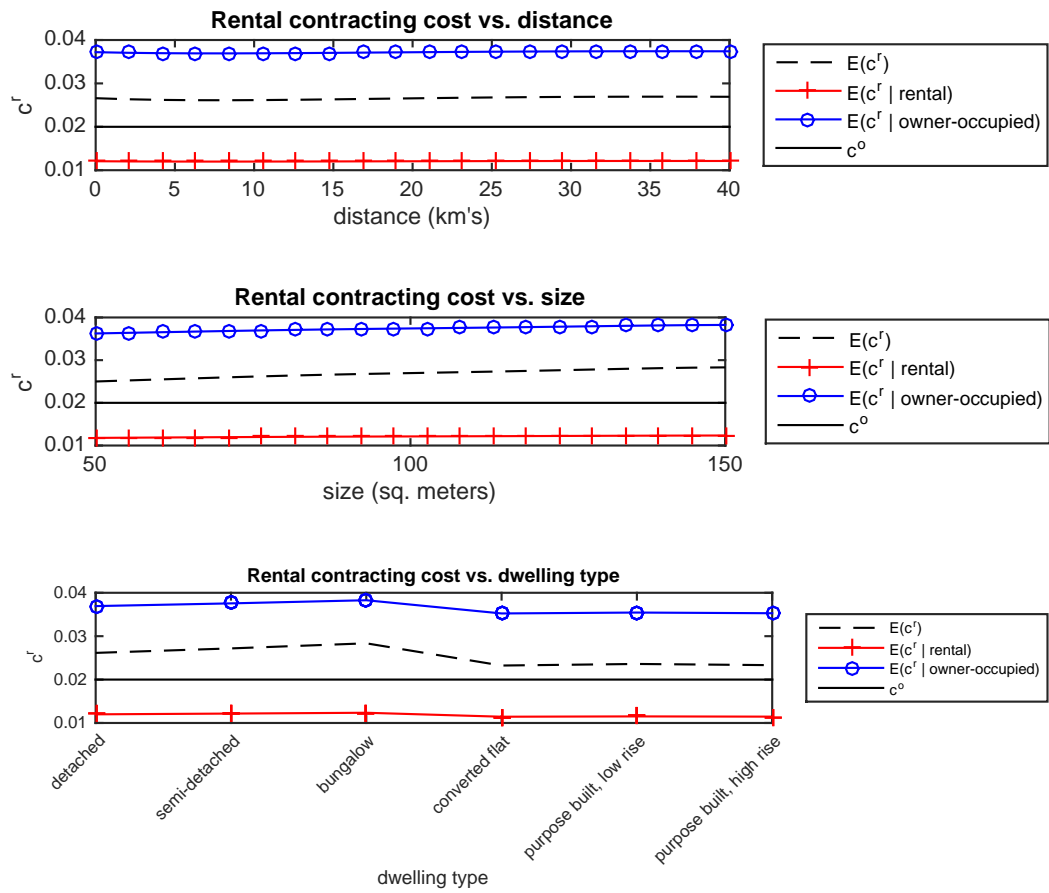


Figure 12: Raw correlation between rent-to-price ratios and ownership (not controlling for selection)

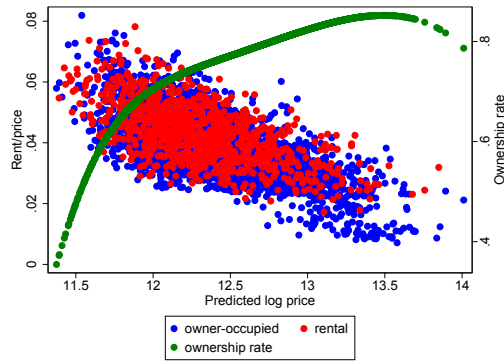


Figure 13: Rents and prices vs. characteristics z (ignoring selection bias)

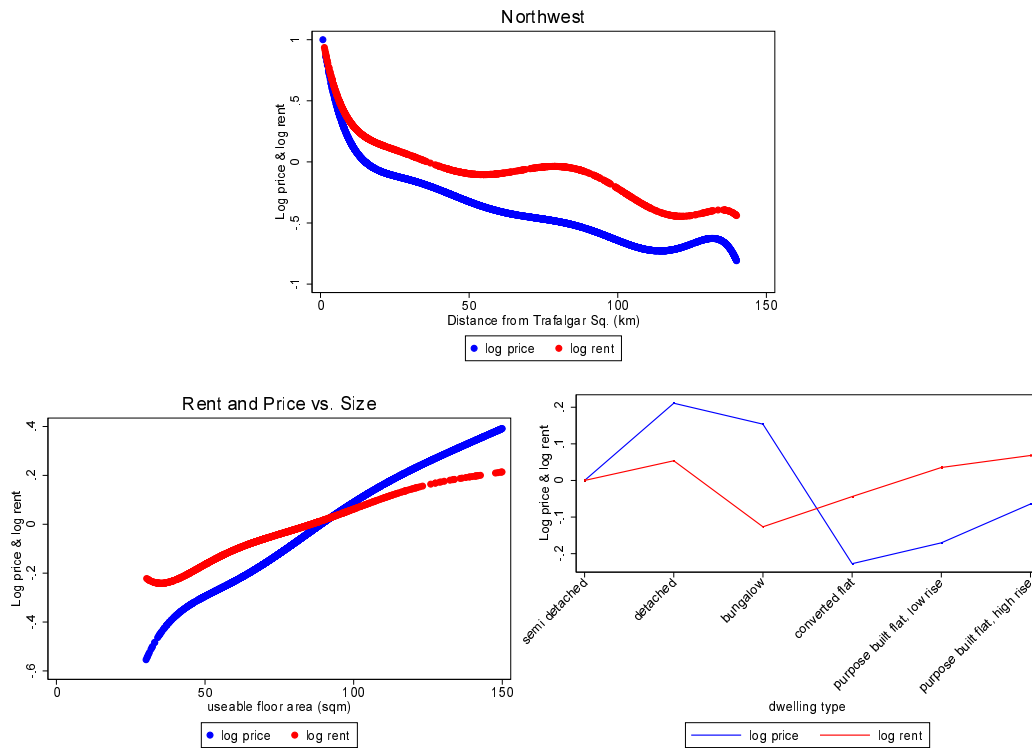


Table 5: Estimation results - hedonics

	MLE - Own	MLE - Rent	Probit	Selection-Own	Selection-Rent
Dwell. type					
detached	0.2094 (0.0174)	0.2628 (0.0922)	0.2681 (0.0882)	0.2690 (0.0881)	
bungalow	0.1490 (0.0225)	0.2286 (0.1038)	0.5720 (0.1157)	0.5648 (0.1155)	
converted	-0.2139 (0.053)	-0.4918 (0.0796)	-0.8019 (0.1284)	-0.8012 (0.1285)	
low rise	-0.1603 (0.0295)	-0.4130 (0.0567)	-0.6955 (0.0875)	-0.6975 (0.0873)	
high rise	-0.0530 (0.0749)	-0.4226 (0.1248)	-0.7730 (0.2315)	-0.7765 (0.2313)	
Dwell age					
1919 - 1944	-0.0175 (0.0211)	0.1884 (0.0602)	0.2557 (0.0806)	0.2583 (0.0806)	
1945 - 1964	-0.0909 (0.0205)	0.0481 (0.0679)	0.2575 (0.0869)	0.2617 (0.0875)	
1965 - 1980	-0.1068 (0.0200)	0.2064 (0.0685)	0.3613 (0.0841)	0.3643 (0.0842)	
Post 1980	-0.0498 (0.0210)	0.1694 (0.0654)	0.287 (0.0878)	0.2908 (0.0878)	
selfReport					
1	-0.0878 (0.0158)				
$\rho_{23} = corr(\eta_2, \eta_3)$				0.1040 (0.0553)	
$\rho_{13} = corr(\eta_1, \eta_3)$					-0.9759 (0.0048)
$\Sigma_{22}$				0.2855 (0.0096)	
$\Sigma_{11}$					0.7533 (0.0415)
$\rho\Sigma$				-0.0297 (0.0156)	-0.7351 (0.0429)

From the 2011 wave of the EHS, using sampling weights. Omitted from the table but included in the estimation are the polynomials in size and the location variables. *Probit* are coefficients from a probit regression of all variables on a dummy for owner-occupancy.