

Halket, Jonathan; Nesheim, Lars; Oswald, Florian

Conference Paper

A Selection Model of Hedonic House Values and Rents

55th Congress of the European Regional Science Association: "World Renaissance: Changing roles for people and places", 25-28 August 2015, Lisbon, Portugal

Provided in Cooperation with:

European Regional Science Association (ERSA)

Suggested Citation: Halket, Jonathan; Nesheim, Lars; Oswald, Florian (2015) : A Selection Model of Hedonic House Values and Rents, 55th Congress of the European Regional Science Association: "World Renaissance: Changing roles for people and places", 25-28 August 2015, Lisbon, Portugal, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/124734>

Standard-Nutzungsbedingungen:

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

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

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

Terms of use:

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

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

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

A Selection Model of Hedonic House Values and Rents^{*}

Jonathan Halket[†]

Lars Nesheim[‡]

Florian Oswald[§]

PRELIMINARY: please do not cite or circulate without permission.

February 26, 2015

Abstract

We use a large repeated cross-section of houses to estimate a selection model of the supply of owner-occupied and rental housing. We find that physical characteristics and unobserved heterogeneity and not location are important for selection. We interpret this as strong evidence in favor of contracting frictions in the rental market relating to maintenance and modification of a dwellings physical characteristics.

^{*}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.

[†]University of Essex and CeMMAP, Institute for Fiscal Studies

[‡]University College London and CeMMAP, Institute for Fiscal Studies

[§]University College London

1 Introduction

In the UK and the US approximately 60% of households tie-up nearly their entire 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 three bedroom detached 120 square-meter house in the suburbs while enjoying the dividend flows and potential capital gains of the wider financial market? In a frictionless world with complete markets and complete contracts, this is the efficient outcome one would expect.

At the same time, rent-to-value ratios vary across the urban landscape as do the market shares of owner-occupied housing and rental housing. The economic tradeoffs households and landlords face vary dramatically across the urban landscape. Why are some houses or dwellings more likely to be owned by owner-occupiers while others are more likely to be owned by landlords? How do dwelling values and rents vary with the propensity to be selected into owner-occupancy and what can these facts tell us about contracting frictions in the rental market? We use data on rents, values and market shares to document some of these tradeoffs and to suggest answers to some of the more complex dynamic questions lurking in the housing market.

Owner-occupied housing and rental housing both provide housing services. At the physical level, the technologies are essentially the same. A physical housing unit provides the services. However, from a theoretical perspective, it seems likely that there are differential costs in maintaining and managing the housing services provided by these two types of housing. A literature going back to ? and ? has argued that contracting frictions result in higher maintenance costs in the rental sector. Because of monitoring costs, tenants under invest in maintenance resulting in a higher depreciation rate in the rental sector. In this literature, these differential costs are likely to be correlated with the physical characteristics of the property. A housing unit with higher rental maintenance costs, or for which contracting frictions are greater, is more likely to be found in the owner-occupied sector.

In this paper, we develop a simple model of the supply side of the housing market. As opposed to the extensive tenure-choice literature that seeks to explain why some consumers choose to rent and others choose to own, we look at what aspects of a housing unit 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. In our model, the supply of housing units is fixed (In England, housing supply is severely constrained by planning regulations.). Property owners (investors or developers) who are endowed with a property, take prices and rents as fixed and decide whether to sell their property to the owner-occupied sector or the rental sector. We exploit the tremendous variation in dwelling characteristics, values, rents and ownership rates within a single but large housing market in England over four years. We then analyse the implications of our findings in the context of the simplest possible dynamic housing model to show that:

1. Observable physical characteristics are important determinants of the probability of being in the owner-occupied sector. Housing units with high value structural attributes (large or more detached dwellings) are more likely to be owner-occupied. At the same time, rent-to-value ratios for these “large” properties are higher than for smaller properties. These relationships are stable over time despite large changes in property values during the time period. These facts are consistent with large or detached properties having higher maintenance needs and therefore higher rental maintenance costs due to contracting frictions.
2. Location is unimportant for the likelihood of being owner-occupied once physical characteristics are controlled for. Rent-to-value ratios vary significantly with location but in a unstable manner that may be due to time and location dependent expectations of capital gains. As long as these expectations are common to both landlords and owner-occupiers, this variation in rent-to-value ratios has no implication for selection into owner-occupancy.
3. Housing units with high unmeasured quality (and therefore higher rents) are much less likely to be in the rental sector. Attributes of a dwelling that are unmeasured in our data include features like the layout of the dwelling, the architectural style, whether it is south facing, whether it has a high quality kitchen or if there’s a garden that requires upkeep. Rental dwellings are of lower unobserved quality. This is consistent with unobservable attributes being hard to contract upon for landlords and thus having higher cost in the rental sector. A further implication is that rent-to-value measures based only on observable attributes that do not control for selection are biased.

We are part of a growing literature that exploits the variation over time and within the cross-section in rent-to-value ratios. In our UK data, as in the various data from the U.S.A. in ???, rent-to-value ratios decline with the value of the house: the more expensive a house is, the lower is its predicted rent-to-value ratio. As ownership rates are (unconditionally) increasing in the value of the house, this could lead one to the conclusion that households tend to own homes that have low rent-to-value ratios. Here we show that this unconditional correlation between ownership and rent-to-value ratios breaks down once we look at this variation characteristic by characteristic: households tend to own houses whose structural characteristics imply high rent-to-value ratios.

Several studies find evidence of higher maintenance costs in rentals. ? estimates that owner-occupiers occupy better properties and better maintain them. ? estimates a hedonic model of sales values for rental and owner-occupied single-family property in a single parish in Louisiana and finds that rentals depreciate faster. ? estimates hedonic models of rents and values from the AHS and finds that rents decline evenly with age whereas values decline at a declining rate.

In the corporate finance literature, studies of selection due to various contracting frictions and its effects on the distribution of observed returns are common.¹ In the housing literature, despite numerous theories of contracting frictions and ownership² and the cautionary warning in ?, there are no studies that control for selection in hedonic estimations of rents and values. We are the first to (a) estimate the size of these effects using a representative data set for an entire housing market (in our case, the London housing market), (b) to econometrically account for the fact that frictions imply selection effects on hedonic value estimates, (c) show that contracting frictions have a powerful effect on unobservables.

If we assume that a single stochastic discount factor prices all housing in the market as in ?, then we can say more. In this case, if different houses have different rent-to-values, then these differences must be due either to differing expectations about future

¹Examples include ??????.

²? notes that the “efficiency” of landlord provided housing services is an important part of the determinant of ownership rates. ? examine how a jurisdiction’s legal system and propensity to enforce contracts affects ownership propensities, ? looks at how the mortgage interest tax deduction interacts with home sizes to effect ownership rates in the U.S.A., ? examines neighborhood externality risks in the AHS and finds them negatively correlated with home-ownership. ? studies how size affects a building’s management structure. ? finds that homeowners that are more likely to default on their mortgage may undermaintain their house.

or to differences in the costs of renting out the property. We argue that expectations about the future, given rents and values, should not systematically affect selection into renting. Furthermore, we use repeated cross-sections sampled from periods of both house price decline and boom in England and find that the estimated relationship between structural attributes, selection and rent-to-value remains very stable. We find that recent movements in the rent-to-value of housing are reflected entirely in the value of land.

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 measured characteristics $z \in \mathbf{R}_n$ and unmeasured characteristics $\varepsilon \in \mathbf{R}_2$. Measured characteristics include features like the location, type of dwelling (detached, semi-detached, etc.), size (square meters), number of bedrooms, and age of structure. We assume that the value of unmeasured characteristics is completely captured by a two dimensional vector that we label “unmeasured quality”. This vector has dimension of at least two because selection into the owner-occupied sector is not perfectly correlated with values. A priori, it is theoretically plausible that some characteristics are more valued in the rental sector and some are more valued in the owner-occupied sector.

The value of a property is a function of its characteristics (z, ε) . The value of a property is the value of the property in the market where it is most valuable:

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

where z is a vector of attributes observable to an econometrician and ε is a vector of unobservables. The functions above and below may vary over time. We leave their dependence on t implicit.

If a dwelling unit is in the rental sector we observe its rent. If it is in the owner-occupied sector, we observe its value. 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 hedonic rent function. The parameters $(\alpha, \lambda_1^r, \lambda_2^r)$ measure the percentage impact of measured and unmeasured quality on rental prices. Let the log value 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 value function. The parameters $(\beta, \lambda_1^o, \lambda_2^o)$ capture the percentage impact of measured and unmeasured quality on values in the owner-occupied sector.

Let the log value 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 values are related to rents in section 2.1 below.

We observe a housing unit in the owner-occupied sector if

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

That is, if

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

Let $Y = \ln \pi^o$ if a housing unit is in the owner-occupied sector and let $Y = \ln R$ if it is in the rental sector. We observe

$$Y(z, \varepsilon) = \mathbf{I}(\pi^o(z, \varepsilon) \geq \pi^r(z, \varepsilon)) \ln \pi^o(z, \varepsilon) + \mathbf{I}(\pi^o(z, \varepsilon) < \pi^r(z, \varepsilon)) \ln R(z, \varepsilon) \quad (6)$$

If we assume that $\varepsilon \sim N(0, \Sigma)$ and that ε is independent of z , then this is a standard

Heckman selection model (?). 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 (Σ, Λ) are not identified. Instead, we define $\eta = \Lambda\varepsilon$ and seek to estimate the parameters (α, β, γ) and $\Omega = \Lambda\Sigma\Lambda^T$ the covariance matrix of η .

Note that, since $\varepsilon \in \mathbf{R}_2$ by assumption, Ω is not full rank. We seek to understand the implications of this restriction for our analysis. The model has several important features. First, the value of unmeasured characteristics in the owner-occupied sector is not perfectly correlated with the value in the rental sector. Second, the impact of unmeasured characteristics on selection is not perfectly correlated with owner-occupied value nor with rental value. Third, the correlation of η_3 and η_2 may differ from the correlation between η_3 and η_1 . We can identify the variance of η_1 and the variance of η_2 . We can also identify the correlation of η_1 and η_3 . We cannot identify the correlation of η_1 and η_2 . To do so requires either observations on properties that switch sectors or imposing some restrictions on how the parameters change over time.

2.1 User-costs and the rent-to-value ratio

The value of a property in a sector, either $\pi^o(z, \varepsilon)$ or $\pi^r(z, \varepsilon)$, equals the willingness to pay of the marginal buyer in that sector. The willingness to pay of each buyer is determined by his or her utility or rent flow from the property, the cost of maintenance and expectations about future utility and financial value of the property. We assume this can be characterized by two Poterba-like user cost equations (?):

$$\pi^o(z, \varepsilon) (m^o(z, \varepsilon) - g^o(z, \varepsilon) + c^o(z, \varepsilon)) = u(z, \varepsilon) \quad (7)$$

$$\pi^r(z, \varepsilon) (m^r(z, \varepsilon) - g^r(z, \varepsilon) + c^r(z, \varepsilon)) = R(z, \varepsilon), \quad (8)$$

where $m^i(z, \varepsilon)$ is the effective discount rate in sector i , $g^i(z, \varepsilon)$ is expected capital gains in sector i , and $c^i(z, \varepsilon)$ is the cost of maintenance in sector i .³ In general, each of these el-

³Note that this implies that, at least to a first-order approximation, risk aversion, transaction costs, and borrowing constraints are not important for the marginal buyer.

ements may vary across property types defined by (z, ε) .⁴ After estimating (π^o, R, π^r) , we will discuss the extent to which our estimates are consistent with different assumptions about how (m^i, g^i, c^i) vary across the two sectors.

3 Data

We use data from the confidential version 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 conducts a “household survey” and a “physical survey”. 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.⁵

Respondents from this selection (approximately 17,000) comprise the Household Interview sample. The EHS team then chooses a sub-sample of these dwellings (approximately 8000 in 2008/2009), including vacant ones, and performs a physical inspection. This is called the “Physical Survey.” The sub-sample is constructed from the 17,500 by taking a sample of all social housing addresses and a sub-sample of private addresses. Private rental properties are over-sampled. Finally, to construct the final “housing stock sample”, the EHS team combines 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 2 year sampling window.

We focus our discussion on the 2011 wave of the EHS. While we also analyse the 2012, 2013, and 2014 waves, these later waves have some limitations. In the later waves, property values are top-coded at £1,000,000. Also, due to budget cuts, the later waves use smaller samples and collect information on a smaller range of topics. Results are

⁴Mortgage payments are not deductible in England for owner-occupiers but are for landlords. This may be reflected in differences between m^o and m^r . Capital gains are not taxed for owner-occupiers but are taxed for landlords. 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. Assuming a common income tax rate, this can be subsumed into $c^r(z, \varepsilon)$. Costs of vacancies in the rental sector can also be subsumed in c^r .

⁵At each sampled address, one dwelling is sampled. At each dwelling, one household is sampled.

robust across waves.⁶

House values are recorded in one of two ways. For a subset of the owner-occupied properties, the owner self-reports what she thinks the market value of her home is. For the remainder of owner-occupied properties a professional surveyor valued the house on-site.

Much of our analysis focuses on a sub-sample of dwellings within 140 km of Trafalgar Square in London. We call this region "Greater London." We restrict the analysis to this region because, as much as possible, we want to focus on a single economic market and account for important geographic features of the data. Within 140 km we are confident that our non-parametric methods for controlling for the quality of locations robustly capture the shape of the hedonic price surface. For dwellings that are farther away from London, we are concerned that our sample size is not sufficient to accurately estimate the hedonic price function.

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 landlords (RSL).⁷ The Greater London is roughly similar to the entire country with regards to tenure: there are slightly more private rentals and few owner-occupiers in London. The share of private rentals has increased very slightly over the four waves at the expense of owner-occupancy.

⁶Because each wave uses data from a two year span, the samples are not independent. For instance, 2011 and 2012 each use the same data collected in 2009/2010.

⁷RSL's are non-profit organisations that provide low-cost housing. They are regulated by the government and highly subsidised.

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

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

Uses sampling weights for each wave. London refers to the Greater London sample area.

Table 2 shows how some of these markets shares vary when one conditions on dwelling characteristics. Semi-detached and detached houses and bungalows are much more likely to be in the owner-occupied sector and converted flats and multi-unit dwellings are more likely to be in the rental sector. A question we seek to shed light on in this paper, is why is that the case? How much is it due to relative prices, how much to preferences, and how much to the costs of providing housing in the two sectors?

Table 2: Market Share By Dwelling Type: Greater London

Dwelling Type	Owner-occupied	Private rented	LA or RSL	Total share
Semi detached	73.9%	13.0%	13.68%	46.5%
Detached	94.4%	5.0%	0.40%	16.8%
Bungalow	76.8%	5.0%	18.31%	5.3%
Converted flat	39.3%	48.5%	15.15%	6.2%
Low rise	32.2%	26.7%	38.39%	21.5%
High rise	20.7%	19.7%	48.06%	3.6%

Semi-detached includes End Terrace and Mid Terrace. From the 2011 wave of the EHS, using sampling weights.

4 Estimation Procedure

We estimate the model parameters using maximum likelihood. We explore several specifications. In our main specification, we include: dummy variables for dwelling type and

dwelling age⁸, an eighth-order polynomial in dwelling size (square meters), and a set of variables that account for the dwelling’s geographic location.⁹ For owner-occupiers, we also include a dummy for whether the value was self-reported or not. We measure location using polar coordinates (r, θ) where r is distance from Trafalgar Square and θ is the arc distance from due east. That is, $\theta = 0$, is east, $\theta = 0.5\pi$ is south, etc. We model location effects as a nonparametric function of (r, θ) . In our main specification, we use the interaction of an eighth-order polynomial in distance with a 5th order trigonometric expansions in terms of θ . The distance variable captures the impact of distance from London on property values and on selection into either the owner-occupied sector or the rental sector. The arc distance variable θ captures variation in outcomes that depend on direction of travel. In a second specification, we drop the arc distance variable in order to measure the average effect of distance. We also explore including dummy variables for the numbers of bedrooms, kitchens, living rooms and bathrooms as well as using levels rather than logs of values and rents. For all cases, our sample was private rental and owner-occupied houses within Greater London, using the sample weights provided in the EHS.

5 Results

We estimate equations (1), (2), and (6) using maximum likelihood. Parameter estimates for our main specification are detailed in Tables 3 and 4 in the appendix. Most of the parameters are statistically significant and have plausible values. Because the parameters are difficult to interpret we plot predictions for log-rent, log-value and for the market share of the owner-occupied sector as functions of the explanatory variables. We also

⁸Dwelling 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.

⁹The confidential version of the EHS reports each dwelling’s full post code. We match that postcode with its geographic coordinates (eastings and northings) 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 was 1 unmatched owner-occupied house out of 5,184 and 2 unmatched private rentals out of 2,683 for all of England. In 2011, there were none though. The numbers of unmatched for other waves are similar. For these unmatched properties, we used the mean easting and northings of all postcodes with 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) as our measure of the location for their particular postcode. Our measure of distance from Trafalgar Square is the Pythagorean distance in kilometers from eastings 529.997 and northings 180.559.

plot point-wise confidence bands for these predictions. The graphs are discussed below.

5.1 Location

The effect of location on house values is dramatic. Figure 1 shows that an otherwise identical house 40km away from center of London is 100 log points lower in value. In contrast, rents fall only about 75 log points. The average slope in the owner-occupied sector is much steeper. In both sectors, the hedonic functions flatten out significantly at distances greater than 40km. Moving from 40 km to 50 km only reduces dwelling values by about 10 log points.

Figure 1 also shows the estimated impact of distance on owner-occupancy rates. The “unconditional” line plots the estimated effect of distance on selection into owner-occupancy when no other correlates are included. The “conditional” line plots the effect of distance after controlling for dwelling characteristics. This figure plots the owner-occupancy rate for a semi-detached 100 square meter dwelling built after 1960. Unsurprisingly, the unconditional line shows that owner-occupancy is far more prevalent 20km outside of London than inside the city center. However, the conditional line shows that, once we control for other characteristics of the houses, distance essentially plays no role in selection. The owner-occupancy rate is essentially flat at around 80% for this house type. It is essentially flat for all house types.

Figures 2 and 3 show how the above results vary if we analyse the impact of distance moving in different directions. The figures show the impact of distance along the four points of the compass, East, South, West and North. The graphs show that, regardless of direction, the qualitative pattern is the same. Both values and rents fall dramatically, values fall faster than rents, and the functions flatten out after about 40km. The function is flatter in the Eastern direction and is completely flat after 40km. In contrast, in all other directions, the function is steeper than East and does not completely flatten out after 40 km. In some directions (e.g. East and South), 140 km from Trafalgar Square is no longer dry land. Thus the confidence intervals in those directions blow up. Also, we are not able to control for lot size in our estimation. However this likely biases upward (towards zero) the slopes of the hedonic rent and value functions with respect to distance: lot sizes are probably larger further away from the city center where land is cheaper.

In summary, rents relative to values rise dramatically with distance but housing units are not more likely to be found in the rental sector. Why don't investors in properties far from the centre convert more properties into rental units? The answer has to be in equations (8) and (6). One possibility is that maintenance costs relative to rents fall as the value of the location rises. Though we do not directly observe maintenance in the EHS, this seems likely. The value of a location is essentially the raw value of land and land, unlike structures, does not depreciate. Therefore two otherwise comparable properties in locations with different values will have different rents but similar maintenance costs. This alone means that the rent-to-value ratio of the property in the more valuable location should be lower.

A second possible explanation is that properties close to the city center had higher expected capital gains during this period. Figure 5 shows the estimated hedonic values and rents with respect to distance for the all waves, 2011-2014.¹⁰ In the most recent wave, 2014, the rental function steeper with respect to distance than the value function. Figure 8 shows a value index for houses close to city center and further away.¹¹ values closer to London have risen much more than values further away, particularly since the end of 2011. This is also reflected in 7, relative values in London fall in 2009-2010 (which is reflected in the 2012 wave) but then rise again in the 2013 and 2014 waves. To the extent that these ex-post high relative capital gains in London were expected ex-ante, this would help explain the low rent-to-value ratios in London.

5.1.1 Economic Interpretation

The rent-to-value ratio increases with distance in several waves but decreases with distance in 2014. However, location is unimportant for selection. Using equations 7 and 8, and letting Δx and Δz denote the changes in variable x and distance respectively, this

¹⁰Note that due to disclosure requirements of the Secure Data Service, we had to normalize values each year. So all values and rents were normalized to be equal across waves at Trafalgar Square.

¹¹We use data from the Land Registry to form this index. The data include all house sales in England from January 1, 2009 to June 30, 2014. As above we restrict our sample to house sales within 100km of Trafalgar Square using the reported postcode and only use non-newly built houses. We form two sub-samples: houses within 10km of Trafalgar Square and houses further than 10km away. For each subsample, we regress log values on a 10th order spline of time at the daily frequency. The plotted lines are the fitted values, normalizing log values to be 0 for each series on January 1, 2010.

suggests that $\frac{\Delta m^o}{\Delta z} = \frac{\Delta m^r}{\Delta z}$, $\frac{\Delta g^o}{\Delta z} = \frac{\Delta g^r}{\Delta z}$ and $\frac{\Delta c^o}{\Delta z} = \frac{\Delta c^r}{\Delta z}$ ¹². As discussed above, it is likely that $\frac{\Delta c^o}{\Delta z} = \frac{\Delta c^r}{\Delta z} < 0$, as maintenance costs likely do not change much when the land value changes¹³ and therefore maintenance as a proportion of the value of the house, c^i , falls when land value increases. Apparently though, in some time periods like those covered in the 2014 wave, this maintenance effect on the rent-to-value ratio is dominated by other effects. Apparently during this period, either $\frac{\Delta m^o}{\Delta z} = \frac{\Delta m^r}{\Delta z} > 0$, or $\frac{\Delta g^o}{\Delta z} = \frac{\Delta g^r}{\Delta z} < 0$. That is, either the discount rate in London went up relative to outside London or relative expected capital gains fell in London.

5.2 Structure

5.2.1 Dwelling type

Figure 4 shows how values, rents and ownership vary with structure type. House values for detached houses and bungalows are about 20 log points higher than for semi-detached houses whereas converted flats and dwelling in lo-rise units are about 20 log points cheaper. Rents follow a similar pattern. Rents for detached houses and bungalows are about 20 log points more expensive. So, the rent-to-value ratio appears to be approximately constant within the category of semi-detached, detached and bungalows. For dwellings in multi-unit structures, rents are 40 log points lower. So, the rent-to-value ratio for multi-unit structures is much lower. These results are stable over waves.

In contrast to location, the conditional effects of dwelling type on predicted ownership are qualitatively similar to the unconditional ownership rates. The unconditional ownership rates are detailed in Table 2: 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%. Figure 4 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 dwelling in multi-unit structures. Structural features

¹²It is unlikely that instead changes in one factor in one sector happen to offset changes in another factor in the other (e.g. $\frac{\Delta m^o}{\Delta z} = \frac{\Delta c^r}{\Delta z}$).

¹³However, more expensive locations may have slightly higher maintenance costs if labor costs are higher.

are important determinants of selection into the owner-occupied sector.

The pattern is similar to the stylized fact documented in ? that in the US, houses 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). Ownership rates do not vary quite as much in the U.K. across structure types (this is true even in the full sample).

5.2.2 Size

Figure 9 (left panel) shows how rents and values change with respect to the total floor space of the house. House values increase approximately 40 log points per 50 square meters. Rents increase approximately 50 log points per 100 square meters. The rent-to-value ratios increases with respect to size. The different is not as big as the changes in rent-to-value ratio with respect to structure nor with respect to location. Figure 10 shows that this pattern remains stable over time. Most of the variation in values and rents over time is captured by changes in the value of location (i.e. land) and not in the valuation of structure.

The center panel of figure 9 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.

It is also interesting to see what this selection effects implies for unmeasured quality. The right panel in Figure 9 shows how the predicted average unmeasured quality for an average housing unit conditional on sector varies with size. In the owner-occupied sector, average unmeasured quality does not vary with size. However, this is not true in the rental sector. In the rental sector, bigger dwellings have much lower unmeasured quality. The average quality difference between a 50 square meter rental property and a 100 square meter rental property is almost 20 log points. Large housing units in the rental sector are likely to be of much lower unobserved quality.

5.2.3 Economic Interpretation

The more detached and/or the larger a property is the higher is its rent-to-value ratio but the lower is its likelihood of being a rental. Detachedness and size are each positively valued. Using equations 7 and 8, this implies that either $\frac{m^o}{m^r}$ or $\frac{c^o}{c^r}$ decreases or $\frac{g^o}{g^r}$ increases with detachedness or with size.

Is it likely that $\frac{g^o}{g^r}$ is dramatically different for detached houses than for dwelling in multi-unit structures? Is this ratio likely to be dramatically different for 100 square meter flats versus 50 square meter flats. This ratio may vary slightly with these structural features due to sectoral differences in taxation of capital gains. However, any variation due to this factor should also be reflected in our findings with respect to location. In the location dimension, consistent lack of a selection effect suggests that $\frac{g^o}{g^r}$ is constant with respect to location.

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 we compare detached houses to dwelling in multi-unit structures. This is the direct or indirect implication of ?, ? and ?.

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 structurally more valuable houses (but not locationally more valuable houses). 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 house characteristics, then m^r/m^o would increase relative to structural 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 mortgage costs vary across owner-occupiers, across landlords, and across different types of housing units.

5.3 Unobserved quality

Table 4 shows the effect of unobserved quality on selection. Properties with unmeasured characteristics that would imply a 1 percent higher rent are 1 percent less likely to be in the rental sector. However, characteristics that lead to high unmeasured house values do

not affect selection.

The results of our analysis are subtle here. One way to explain these results would be if $\lambda_1^r, \lambda_2^o, \lambda_1^s > 0$ and $\lambda_1^o, \lambda_2^s = 0$. In this case, we could think of ε_1 as an amenity which affects the user-enjoyment of a house but comes at a high maintenance cost such that net value of the characteristic is zero. ε_2 on the other hand is an amenity that affects both values and rents but does not affect selection. In other words, perhaps ε_1 is a jacuzzi - nice to use but a nightmare to maintain, and ε_2 is a nice view from the bedroom.

Characteristics that are unobservable to econometricians may also suffer more acutely from a third-party verification problem. Enforcing contracts to invest in and/or maintain these characteristics may be particularly costly if even possible. Landlords may therefore choose properties with fewer of these characteristics.¹⁴

6 Conclusion

Models of households' homeownership decisions, such as ????, largely have abstracted away from explicit considerations of the multi-characteristic nature of a house. Our findings here point to a need to examine households' desire to own jointly with their desire to live in certain houses. Households that have a higher demand for larger houses or detached houses or houses with high maintenance amenities are more likely to save for a downpayment everything else equal.

¹⁴?? builds a theory of underinvestment in housing and security of tenure that has a similar mechanism.

A Figures

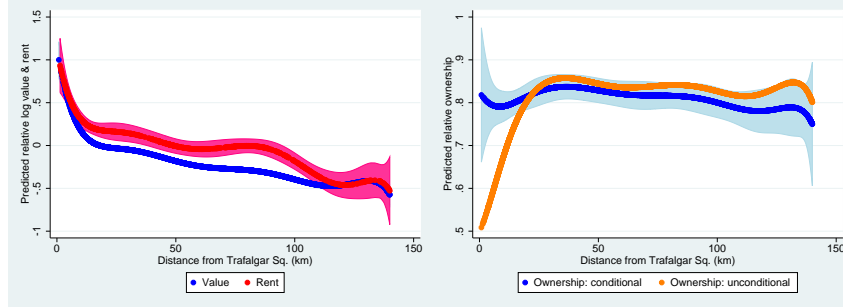


Figure 1: Location: Estimated values and selection

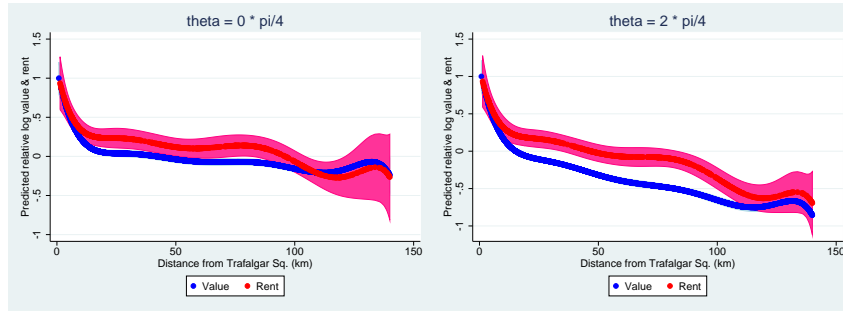


Figure 2: Location : Estimated values (East and South)

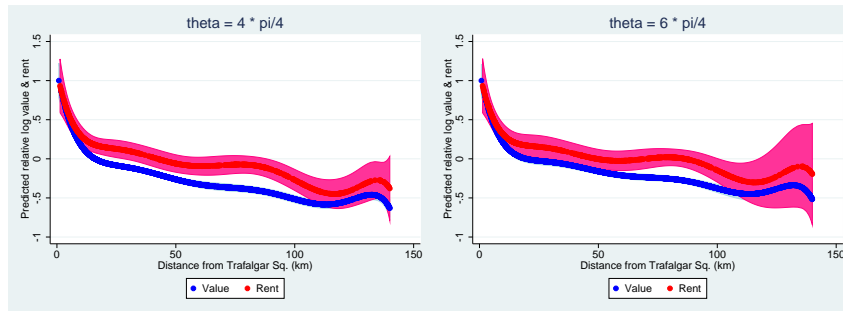


Figure 3: Location: Estimated values (West and North)

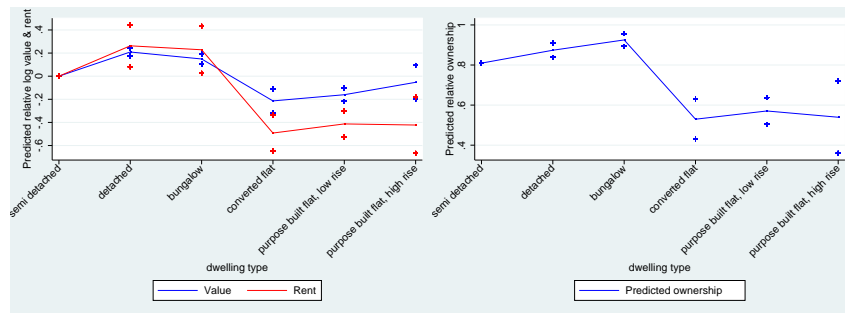


Figure 4: Dwelling Type: Estimated values and selection

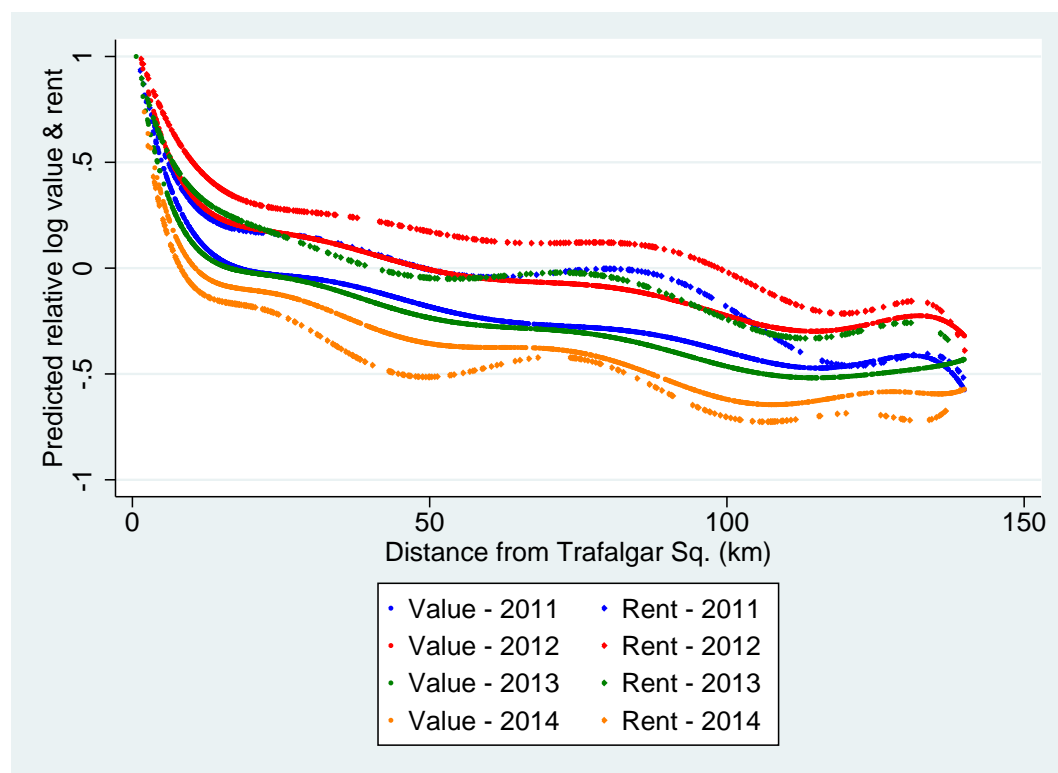


Figure 5: Location: variation over time

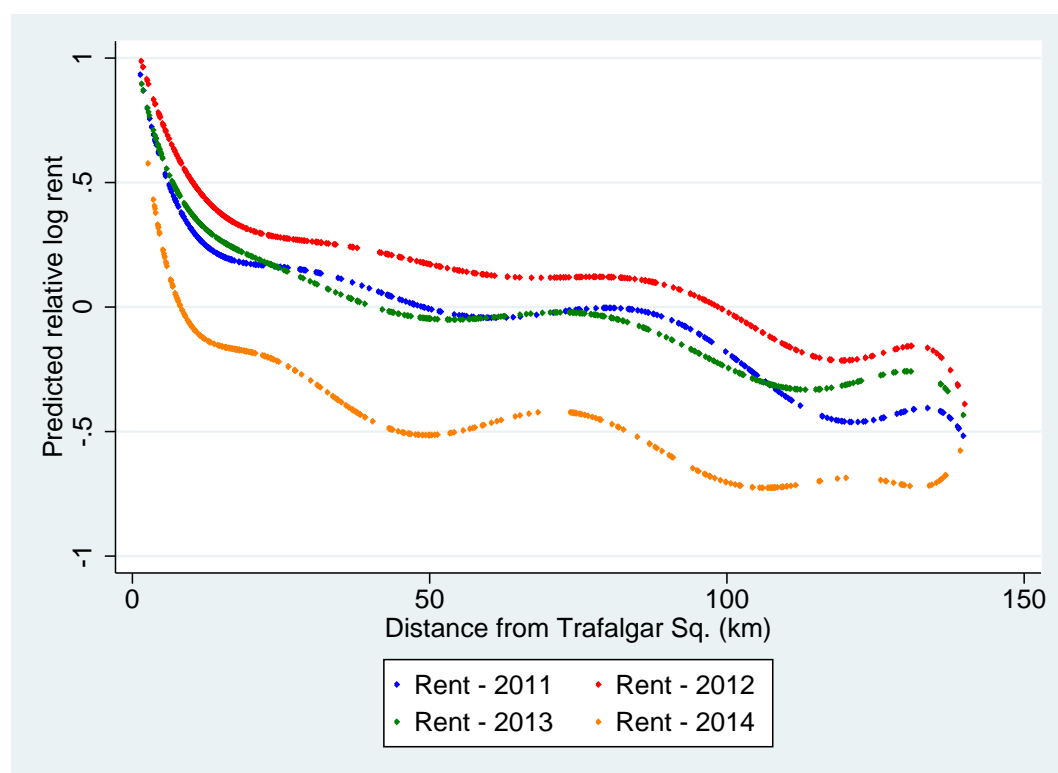


Figure 6: Location: variation over time (2)

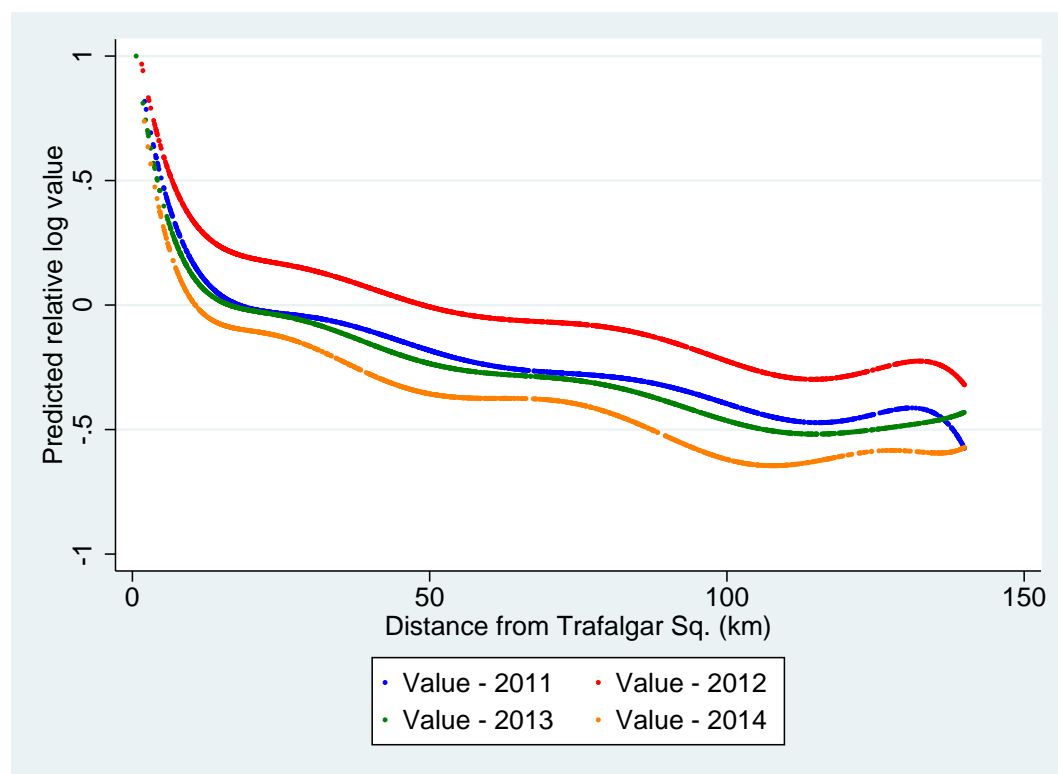


Figure 7: Location: time-variation (3)

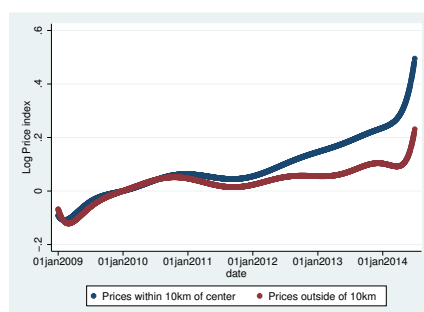


Figure 8: Land registry: Estimated values

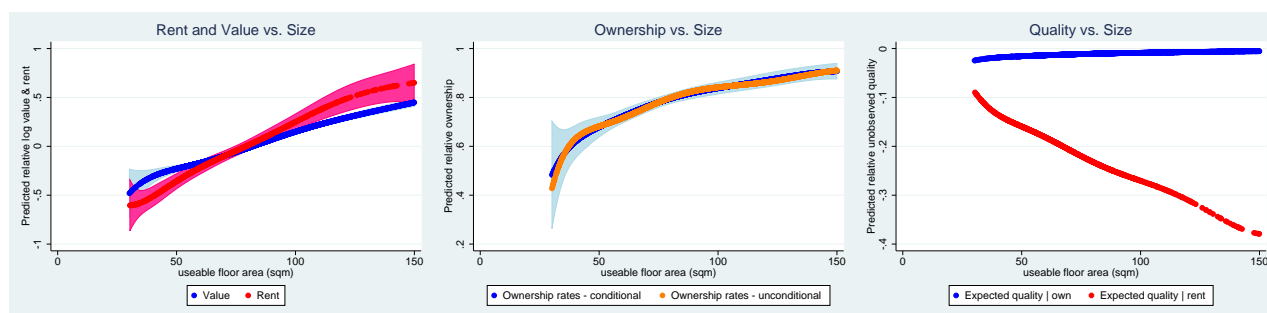


Figure 9: Size: Estimated values, selection and unobserved quality

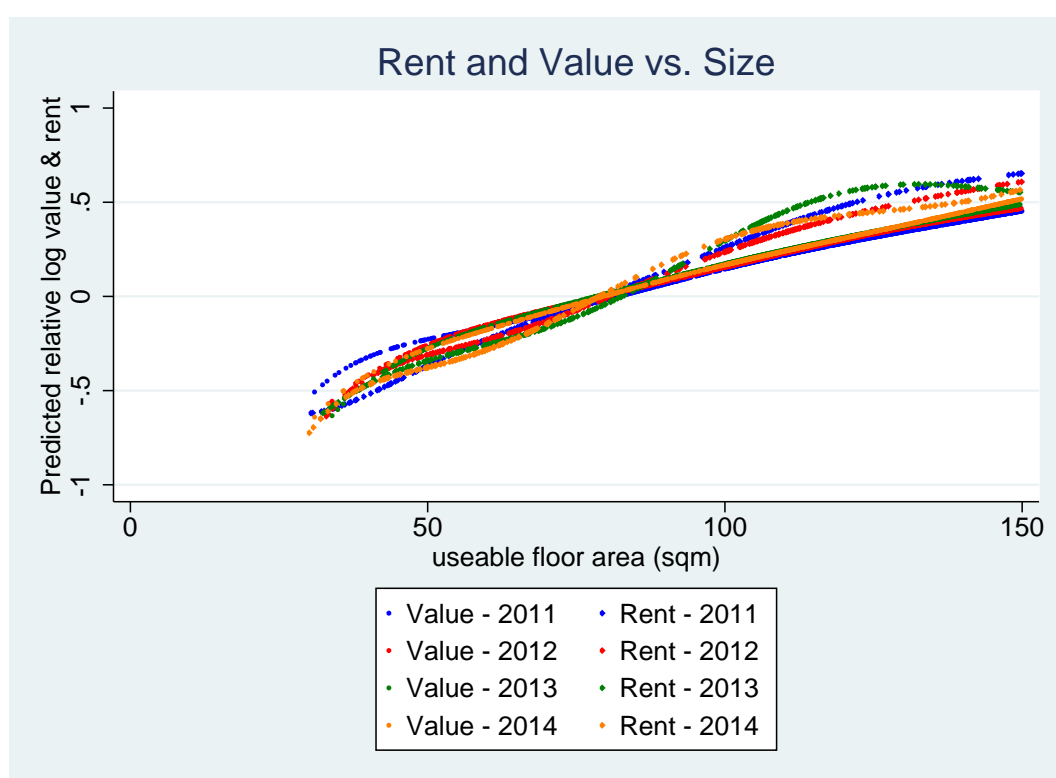


Figure 10: Time variation in value of size

B Tables

Table 3: Estimation Results - Hedonics

	MLE - Own	Two Step - Own	MLE - Rent	Two Step - Rent	Probit
dist1	-0.2513	-0.2515	-0.2388	-0.225	0.0005
	0.0107	0.0108	0.0278	0.0243	0.04
dist2	0.0895	0.0924	0.0352	0.0323	-0.0359
	0.0096	0.0102	0.025	0.0286	0.0373
dist3	-0.0486	-0.0516	-0.0388	-0.0543	0.0453
	0.0086	0.0089	0.0228	0.0324	0.0312
dist4	0.0623	0.0627	0.0699	0.0651	-0.0048
	0.0078	0.0078	0.0197	0.0171	0.0276
dist5	-0.04	-0.0394	-0.004	0.0025	-0.0126
	0.0072	0.0073	0.0196	0.0176	0.0267
dist6	0.0174	0.0168	0.0156	0.0196	0.0128
	0.0066	0.0067	0.0204	0.0174	0.0257
dist7	-0.035	-0.0337	-0.0457	-0.047	-0.0214
	0.0063	0.0065	0.019	0.0176	0.0255
dist8	-0.0028	-0.0027	-0.0033	0.0001	-0.0021
	0.0059	0.0059	0.0186	0.0162	0.025
fsin1	-0.0225	-0.0179	-0.0539	-0.0617	-0.077
	0.0101	0.0111	0.0318	0.0438	0.0447
fsin2	0.0106	0.0139	0.0413	0.0356	-0.044
	0.0084	0.0087	0.0286	0.0327	0.0384
fsin3	0.0078	0.0081	0.0153	0.0322	0.0006
	0.0084	0.0085	0.0274	0.0224	0.0382
fsin4	-0.0109	-0.0103	-0.0122	-0.0022	-0.0123
	0.0082	0.0083	0.0282	0.0237	0.0374
fsin5	0.0301	0.0299	0.0128	-0.0012	0.002
	0.0086	0.0087	0.0276	0.0214	0.0374
fcos1	-0.1076	-0.1111	-0.0444	-0.0359	0.0569
	0.0098	0.0106	0.0274	0.0376	0.0388
fcos2	0.005	0.0102	-0.0363	-0.0409	-0.0786
	0.0091	0.0101	0.0281	0.051	0.0384
fcos3	-0.0251	-0.0256	0.0138	0.0107	0.0011
	0.0086	0.0087	0.0283	0.0212	0.0395
fcos4	-0.0306	-0.0315	0.016	0.0243	0.0097

	0.0083	0.0084	0.0266	0.0215	0.038
fcos5	0.0014	0.0025	-0.0278	-0.0244	-0.0182
	0.0083	0.0083	0.0278	0.0245	0.0374
distsin1_1	-0.0515	-0.0502	-0.0635	-0.0563	-0.0384
	0.0116	0.0117	0.037	0.0405	0.0523
distsin2_1	0.0062	0.0084	-0.0246	-0.04	-0.0366
	0.0118	0.0119	0.0363	0.0385	0.0523
distsin1_2	-0.0037	-0.0053	-0.0046	-0.0159	0.0102
	0.0102	0.0104	0.0303	0.0237	0.0431
distcos1_1	0.0621	0.0621	0.0247	0.031	0.0028
	0.0114	0.0115	0.0319	0.0266	0.0434
distcos2_1	-0.0137	-0.0127	-0.0218	-0.0318	-0.0159
	0.0124	0.0126	0.0325	0.0286	0.0437
distcos1_2	0.0367	0.0344	0.0179	0.0003	0.0171
	0.0111	0.0115	0.0326	0.0285	0.0454
Quarter					
2	-0.0677	-0.0716	-0.0068	-0.0337	0.0597
	0.0275	0.0273	0.0876	0.0831	0.1116
3	-0.0925	-0.0903	0.0486	0.065	-0.0265
	0.0233	0.0235	0.0743	0.0604	0.1019
4	-0.0808	-0.0766	-0.0776	-0.0993	-0.0555
	0.0248	0.0253	0.0813	0.0678	0.1049
5	-0.1083	-0.1038	-0.0313	0.0066	-0.0712
	0.0229	0.0238	0.0761	0.0712	0.1048
6	-0.1181	-0.1168	-0.0202	-0.0069	-0.0169
	0.0239	0.0241	0.0783	0.0642	0.105
7	-0.1111	-0.098	-0.053	-0.012	-0.181
	0.0228	0.0269	0.0724	0.1052	0.1004
8	-0.0628	-0.0636	0.0732	0.116	0.0257
	0.0241	0.0243	0.0757	0.0612	0.1078
Dwell. type					
detached	0.2094	0.1981	0.2628	0.2244	0.2681
	0.0174	0.0222	0.0922	0.2004	0.0882
bungalow	0.149	0.1154	0.2286	0.1938	0.572
	0.0225	0.0391	0.1038	0.3549	0.1157

converted	-0.2139	-0.1201	-0.4918	-0.4276	-0.8019
	0.053	0.1178	0.0796	0.4044	0.1284
low rise	-0.1603	-0.0898	-0.413	-0.3197	-0.6955
	0.0295	0.0787	0.0567	0.3801	0.0875
high rise	-0.053	0.0265	-0.4226	-0.3247	-0.773
	0.0749	0.1083	0.1248	0.4183	0.2315
size	0.2503	0.2294	0.3081	0.2685	0.3129
	0.0095	0.0252	0.0477	0.198	0.0335
size2	-0.0289	-0.0183	-0.1455	-0.0866	-0.0568
	0.0093	0.016	0.0536	0.0861	0.0308
size3	-0.0027	-0.0079	-0.0554	0.0027	0.0668
	0.0088	0.0118	0.0464	0.0844	0.0267
size4	0.0054	0.0066	-0.0229	0.0017	0.0127
	0.0078	0.0083	0.0267	0.0443	0.025
size5	0.0021	0.0013	-0.0034	0.0009	-0.0079
	0.0074	0.0077	0.0242	0.0298	0.0271
size6	-0.0078	-0.0061	0.0044	-0.0147	-0.0585
	0.007	0.0075	0.0378	0.0678	0.0259
size7	0.0065	0.0076	0.0167	-0.0172	-0.0558
	0.0066	0.0068	0.0392	0.0706	0.0255
size8	-0.0055	-0.0046	0.0212	-0.0067	-0.0395
	0.0061	0.0063	0.0276	0.0439	0.0272
Dwell age					
1919 - 1944	-0.0175	-0.0355	0.1884	0.1938	0.2557
	0.0211	0.0296	0.0602	0.1448	0.0806
1945 - 1964	-0.0909	-0.1085	0.0481	0.0066	0.2575
	0.0205	0.0283	0.0679	0.158	0.0869
1965 - 1980	-0.1068	-0.1316	0.2064	0.1599	0.3613
	0.02	0.0325	0.0685	0.201	0.0841
Post 1980	-0.0498	-0.0702	0.1694	0.1891	0.287
	0.021	0.0292	0.0654	0.1639	0.0878
selfReport					
1	-0.0878	-0.0879			
	0.0158	0.016			

mills2		-0.2368			
		0.2208			
mills1				-0.6888	
				0.7491	

Table 4: Estimation Results - Selection

	Own	Rent
dist1	0.0010	
	0.0400	
dist2	-0.0351	
	0.0373	
dist3	0.0456	
	0.0312	
dist4	-0.0028	
	0.0278	
dist5	-0.0120	
	0.0267	
dist6	0.0110	
	0.0258	
dist7	-0.0210	
	0.0255	
dist8	-0.0032	
	0.0250	
fsin1	-0.0764	
	0.0447	
fsin2	-0.0447	
	0.0385	
fsin3	-0.0008	
	0.0381	
fsin4	-0.0099	
	0.0376	
fsin5	0.0025	
	0.0373	
fcos1	0.0583	
	0.0389	

fcos2	-0.0775	
	0.0385	
fcos3	0.0031	
	0.0397	
fcos4	0.0102	
	0.0380	
fcos5	-0.0166	
	0.0374	
distsin1_1	-0.0381	
	0.0524	
distsin2_1	-0.0384	
	0.0524	
distsin1_2	0.0097	
	0.0431	
distcos1_1	0.0058	
	0.0436	
distcos2_1	-0.0140	
	0.0440	
distcos1_2	0.0167	
	0.0453	
Quarter		
2	0.0649	
	0.1114	
3	-0.0243	
	0.1017	
4	-0.0523	
	0.1045	
5	-0.0672	
	0.1046	
6	-0.0103	
	0.105	
7	-0.1797	
	0.0998	
8	0.0329	
	0.1084	

Dwell. type		
detached	0.2690	
	0.0881	
bungalow	0.5648	
	0.1155	
converted	-0.8012	
	0.1285	
low rise	-0.6975	
	0.0873	
high rise	-0.7765	
	0.2313	
size	0.3141	
	0.0335	
size2	-0.0549	
	0.0309	
size3	0.0658	
	0.0267	
size4	0.0126	
	0.0249	
size5	-0.0085	
	0.0273	
size6	-0.0599	
	0.0261	
size7	-0.0564	
	0.0255	
size8	-0.0393	
	0.027	
Dwell age		
1919 - 1944	0.2583	
	0.0806	
1945 - 1964	0.2617	
	0.0875	
1965 - 1980	0.3643	
	0.0842	
Post 1980	0.2908	

	0.0878	
$\rho_{23} = \text{corr}(\eta_2, \eta_3)$	-0.104	
	0.0553	
$\rho_{13} = \text{corr}(\eta_1, \eta_3)$		-0.9759
		0.0048
Σ_{22}	0.2855	
	0.0096	
Σ_{11}		0.7533
		0.0415
$\rho\Sigma$	-0.0297	-0.7351
	0.0156	0.0429

C AHS results

In this section we present results from a similar analysis to the preceding but using American Housing Survey (AHS) Data. We have data from the 2007 edition of the dataset, which covers 7 statistical metropolitan areas: Boston, Tampa, Baltimore, Houston, Miami, Minneapolis and Washington DC. The statistical model we estimate is identical to the one used before. The main noteworthy difference is that in the AHS we can control for location only at the pseudo census tract level, rather than at the postcode level as in the EHS. [add more data description]

The AHS table corresponding to table 3 is table 5.

The AHS table corresponding to table 4 is table 6.

There are associated graphs in figures 11, 12 and 13.

	Own 2step	Own MLE	Own linreg	Renter 2step	Renter MLE	Renter linreg
(Intercept)	12.64*** (0.17)	12.06*** (0.09)	11.97*** (0.07)	8.74*** (0.12)	8.78*** (0.11)	8.79*** (0.11)
ns(unitsqm, df = 3)1	0.28** (0.10)	0.63*** (0.05)	0.69*** (0.04)	0.40* (0.19)	0.52*** (0.09)	0.56*** (0.07)
ns(unitsqm, df = 3)2	0.28 (0.25)	1.15*** (0.15)	1.28*** (0.12)	1.52*** (0.27)	1.69*** (0.15)	1.75*** (0.13)
ns(unitsqm, df = 3)3	0.64*** (0.08)	0.91*** (0.04)	0.95*** (0.03)	0.63** (0.20)	0.74*** (0.13)	0.78*** (0.12)
built<1950	−0.04* (0.02)	−0.06*** (0.02)	−0.06*** (0.02)	0.03 (0.03)	0.02 (0.03)	0.01 (0.03)
built1950-1970	−0.04* (0.02)	−0.06*** (0.01)	−0.07*** (0.01)	−0.05 (0.03)	−0.06* (0.02)	−0.06** (0.02)
built1991-2000	0.11*** (0.02)	0.10*** (0.02)	0.10*** (0.02)	0.16*** (0.04)	0.15*** (0.04)	0.15*** (0.04)
built2001-2010	0.16*** (0.02)	0.15*** (0.02)	0.14*** (0.02)	0.29*** (0.04)	0.28*** (0.04)	0.28*** (0.04)
nunit2detached	0.08** (0.03)	0.16*** (0.02)	0.17*** (0.02)	−0.03 (0.09)	0.02 (0.05)	0.04 (0.04)
nunit22+ apartments	0.12* (0.05)	−0.07* (0.03)	−0.09*** (0.02)	0.06 (0.07)	0.01 (0.04)	0.00 (0.03)
invMillsRatio	−0.38*** (0.08)			0.13 (0.14)		
sigma	0.47 (999.00)	0.42*** (0.00)		0.53 (999.00)	0.53*** (0.01)	
rho	−0.81 (999.00)	−0.12 (0.08)		0.24 (999.00)	0.06 (0.10)	
Num. obs.	11678	11678	7813	11678	11678	3865
Not selected	3865	3865		7813	7813	
Selected	7813	7813		3865	3865	
Log Likelihood		−8506.44			−7132.01	
R ²			0.59			0.24
Adj. R ²			0.58			0.21

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

Table 5: AHS estimation results: outcome equation estimated with heckit two step estimator, maximum likelihood of the heckman model, and a simple linear regression (disregarding selection).

	Own 2step	Own MLE	Renter 2step	Renter MLE
(Intercept)	−1.07*** (0.12)	−1.06*** (0.12)	1.07*** (0.12)	1.07*** (0.13)
ns(unitsqm, df = 3)1	2.06*** (0.08)	2.06*** (0.08)	−2.06*** (0.08)	−2.06*** (0.08)
ns(unitsqm, df = 3)2	4.10*** (0.24)	4.08*** (0.24)	−4.10*** (0.24)	−4.10*** (0.24)
ns(unitsqm, df = 3)3	1.78*** (0.11)	1.78*** (0.11)	−1.78*** (0.11)	−1.78*** (0.11)
built<1950	−0.25*** (0.04)	−0.25*** (0.04)	0.25*** (0.04)	0.25*** (0.04)
built1950-1970	−0.23*** (0.04)	−0.23*** (0.04)	0.23*** (0.04)	0.23*** (0.04)
built1991-2000	−0.09 (0.06)	−0.09 (0.06)	0.09 (0.06)	0.09 (0.06)
built2001-2010	−0.19** (0.07)	−0.18** (0.07)	0.19** (0.07)	0.18** (0.07)
nunit2detached	0.72*** (0.05)	0.72*** (0.05)	−0.72*** (0.05)	−0.72*** (0.05)
nunit22+ apartments	−0.80*** (0.05)	−0.80*** (0.05)	0.80*** (0.05)	0.80*** (0.05)
Num. obs.	11678	11678	11678	11678
Not selected	3865	3865	7813	7813
Selected	7813	7813	3865	3865
Log Likelihood		−8506.44		−7132.01

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

Table 6: AHS estimation results: selection equations for the Heckman selection models. These are probit regression coefficients.

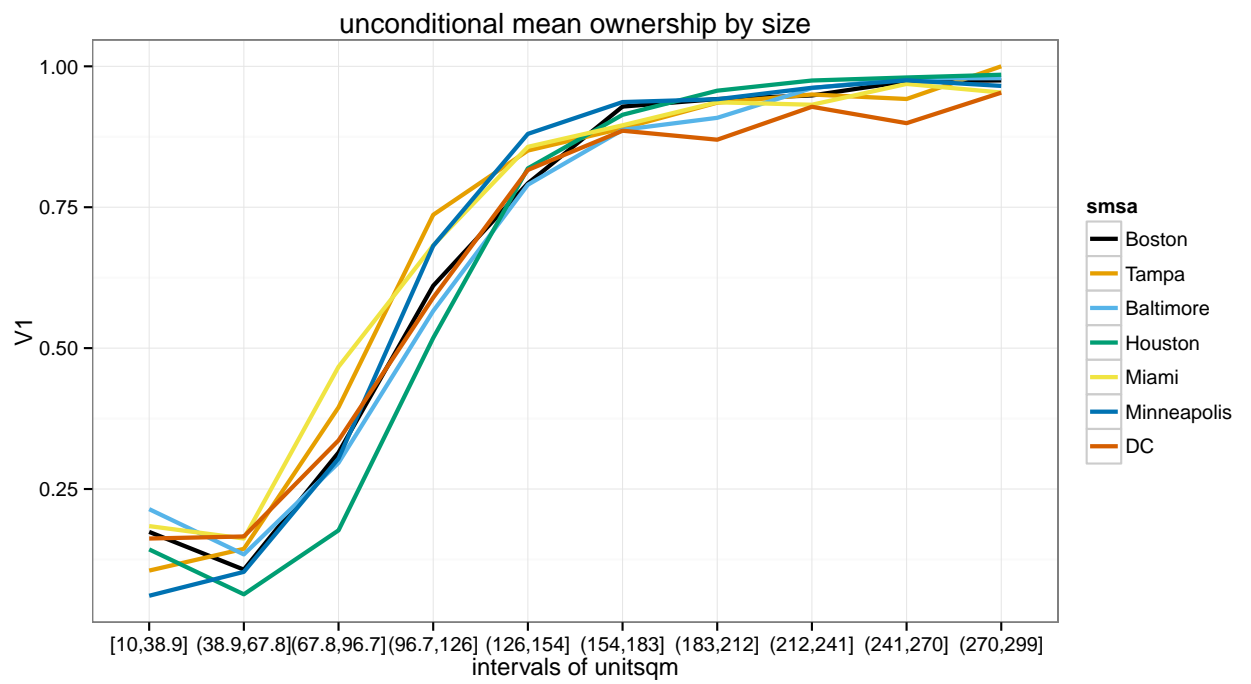


Figure 11: AHS results: non parametric ownership estimate.

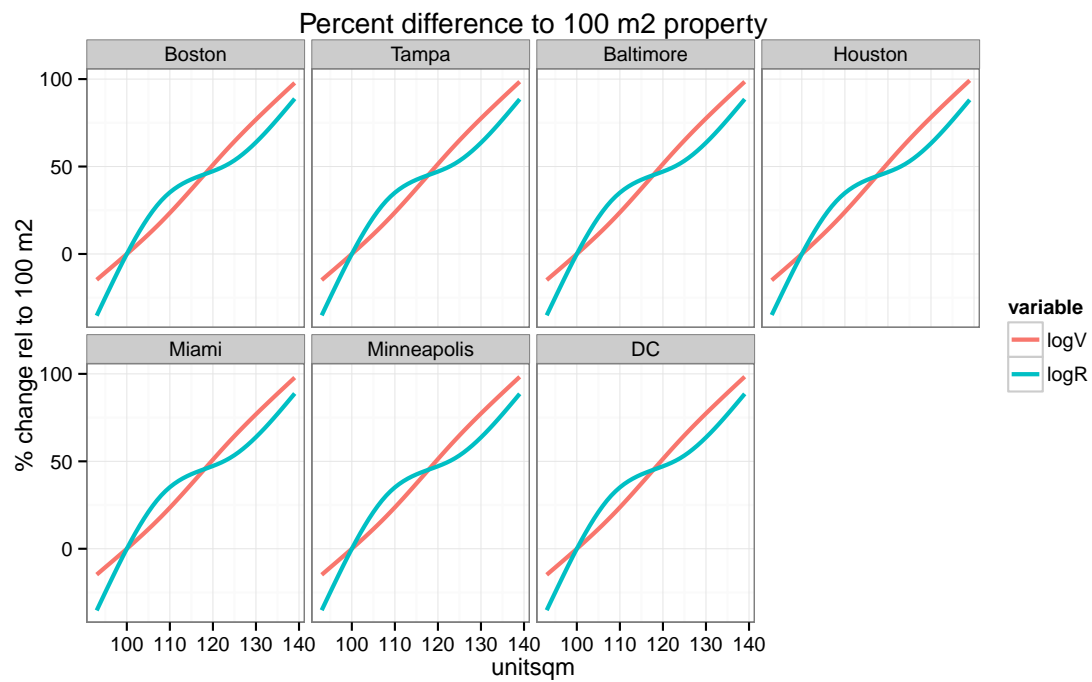


Figure 12: AHS results: predicted percentage changes in property value and rents as a function of size (unit square meters). This uses the MLE estimates from table 5 and predicts value or rent in each city for a property located in the zone with the largest number of observations, of type semi-detached and that was built in 1970-1990.

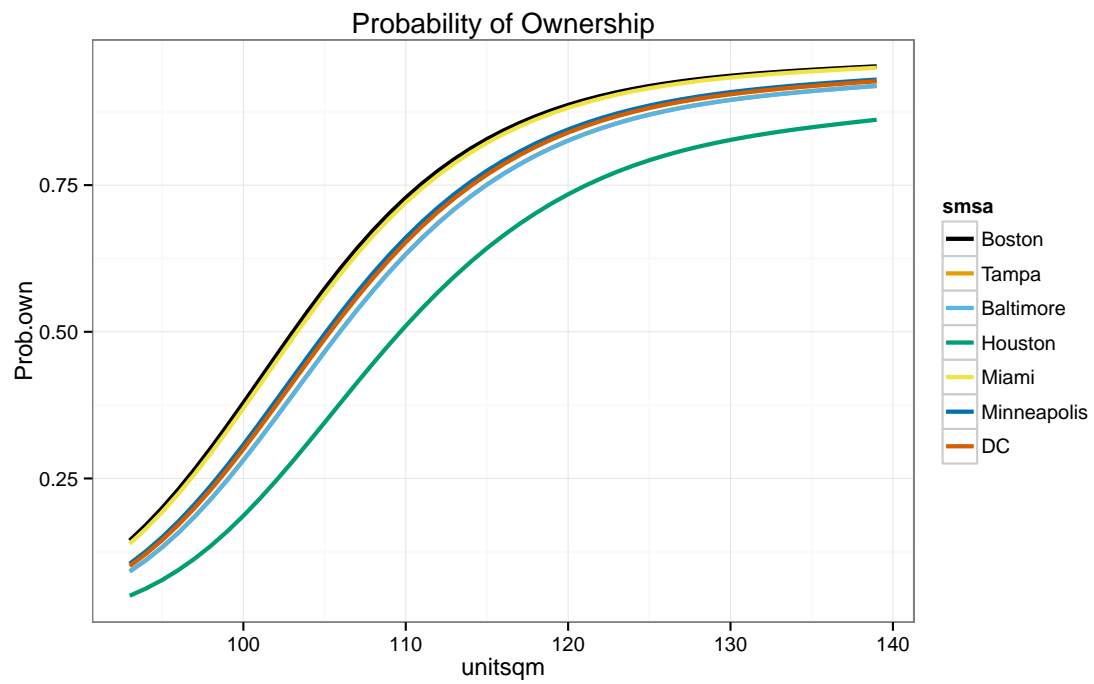


Figure 13: AHS results: predicted probability of ownership as a function of size (unit square meters). This uses the MLE estimates from table 5 and predicts the same property as in figure 12.