History Dependence in the Housing Market: Facts and Explanations*

Philippe Bracke[a,b]   Silvana Tenreyro[a,c]

[a] Bank of England,[b] SERC   [c]London School of Economics, CfM, CEPR

April 2016

Abstract

Using the universe of housing market transactions in England and Wales in the last twenty years, we document a robust pattern of history dependence. Sale prices and selling probabilities today are affected by the housing market conditions in the period in which properties were bought. We systematically investigate the causes of history dependence, assessing the role of three main hypotheses: reference dependence (including, but not limited to, loss aversion); down-payment effects; and match quality. We complement our analysis with data on house listings, which we match to actual sales.

Key words: housing market, fluctuations, down-payment effects, anchoring, loss aversion, thick-market effects, match quality

*For helpful comments, we would like to thank Francesco Caselli, Per Krusell, and seminar participants at the Bank of England and LSE.

1 Introduction

Movements in housing markets have a profound effect on the business cycle and households' finances. Indeed, housing markets have played a critical role in the recent financial crisis and arguably in the very slow recovery that followed. Understanding the sources and impact of housing market movements is critically important to design appropriate macroeconomic and housing policies.

Our paper starts by documenting a robust pattern of history dependence in house prices and transactions. Specifically, we find that the strength of the housing market in the year a house was last transacted influences the price at which the house sells next. Moreover, we document that houses previously bought during booms are less likely to be sold. The results are based on 20 million housing transactions from England and Wales and are robust to the use of repeat sales; in other words, the results are not driven by changes in the composition of the houses transacted.

History dependence in housing markets is clearly at odds with a frictionless model in which the value of a house (and its transactability) depend exclusively on the future stream of dividends (rental value) the property delivers. The results echo the literature on labour markets, which has documented cohort (or vintage) effects, whereby workers first hired during booms have permanently higher wages.

After documenting these housing market regularities, we explore three main channels that can potentially account for the history dependence in the data: 1) Anchoring—and within anchoring, the potential for loss aversion or asymmetries in reference dependence; 2) Down-payment requirements; and 3) Thick-market effects.

The idea behind the first channel we study (anchoring) goes back to Tversky and Kahneman (1982) and builds on a well-established result from laboratory experiments: agents tend to show a bias in their final estimates of a given value that overweighs possibly irrelevant initial cues. In the context of the housing market, sellers may give excessive weight to the price they paid (vis-à-vis the market evolution of prices) when posting new prices; if they bought at high prices, this will lead to higher advertised prices and low time in the market. Past prices may influence the reservation price for a further reason if the seller exhibits loss aversion as in Kahneman and Tversky (1979). If the seller regards the previous price as a reference point with respect to which gains and losses are measured, then a high past price leads to a high reservation price. Unlike anchoring, loss aversion generates an asymmetry in the effect of past prices that can be contrasted with the effect of simple anchoring.

The idea behind the second channel (down payment) goes back to Stein (1995). The purchase of a house normally requires a significant down payment. For repeat buyers, a large percentage of their down payment comes from the proceeds of the sale of their old homes, and, importantly, a majority of home sales are to repeat buyers. Hence, owners who bought in a house boom will have, all else equal, limited home equity; they will then have higher reservation prices and be less likely to sell than owners of comparable houses bought at lower prices, as the money they would be left with after their sale would be insufficient for a down payment on a new property. Accordingly, potential sellers who previously bought at high prices will tend to post higher prices, spend more time on the market and, conditional on a transaction, sell at higher prices than sellers of comparable houses bought at lower prices.

The third channel (market thickness) starts from the premise that in house booms or thick markets, in which there are many houses for sale, potential buyers can form better matches (i.e., they can find their "ideal houses"); the higher quality of the matches increases buyers' willingness to pay and leads to higher transaction prices (Ngai and Tenreyro, 2014). Home owners in high-quality matches have higher reservation prices and hence are less likely to sell than owners who bought in thin markets and formed poor matches with their houses.

These three channels lead to different models of behaviour and to potentially very different policy implications. Understanding and quantifying their relevance is a first step to inform the design of policies aimed at preventing or reacting to future crises. In the context of the current recovery, the analysis can shed new light on the causes for the remarkably low speed, illustrated in Figure 1. The figure shows the monthly quality-adjusted average price and the monthly total number of transactions in England and Wales over 1995-2014. Transactions reached their peak in 2007 and then declined sharply. Prices reached their peak slightly afterwards (between the end of 2007 and the beginning of 2008), subsequently fell, and only after 2009 experienced a resurgence. This national picture hides a heterogeneous regional situation where locations such as London have now nominal prices higher than the 2007 peak, whereas other regions have still to regain the lost ground.

Three main results stand out from our investigation. Anchoring is an important determinant of both house prices and quantities. This reference dependence is symmetric with respect to gains and losses and appears more clearly in actual prices and selling probabilities than in listed prices. Loss aversion, while statistically significant, can account for a very small share of the variation in house prices or quantities. Moreover, the analysis shows that cash-financed transactions display all the patterns observed in the general data, suggesting that credit frictions are not the main determinant of history dependence. Consistent with this, the degree of leverage of mortgages also does not seem to explain away history dependence. Finally, there is virtually no evidence that

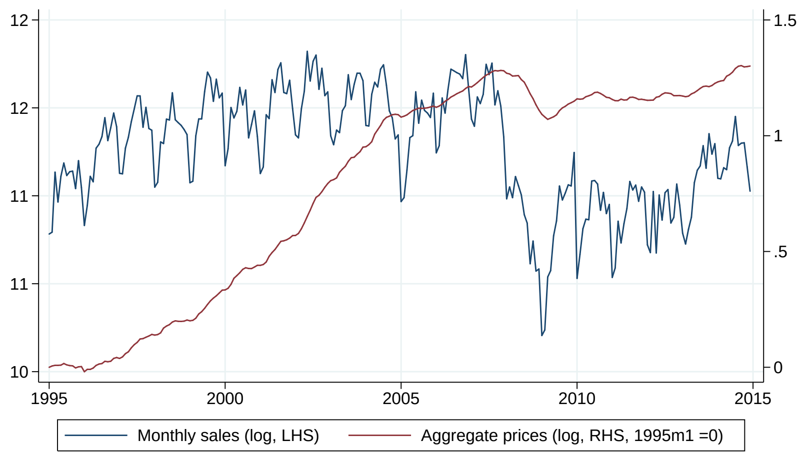

Figure 1: Monthly house prices and sales, England and Wales

Notes: Data are taken from the England and Wales Land Registry, described in Section 4. Average house prices are quality adjusted as described in the same section.

thick-market effects play a major role in low-frequency housing cycles, in contrast with their importance in determining prices and quantities at higher frequencies (for instance at seasonal frequencies).

Our paper differs from previous contributions in the literature in a number of ways. First, it uses a database that contains the universe of all housing transactions in England and Wales for the last twenty years. The size of the sample allows us to highlight precise patterns in the data. Second, our analysis allows us to identify anchoring effects, which are distinct from the more often studied loss-aversion effects. Third, by separating cash- from mortgage-financed transactions, our paper can isolate the role of behavioral and matching explanations from leverage-based mechanisms such as down-payment requirements. Fourth, our study is the first to empirically analyze the presence of thick-market effects at lower frequencies for the whole universe of housing transactions in a country. (Previous studies on thick markets focused on high-frequency movements or were confined to smaller geographical areas and time periods.) Finally, the paper provides an analysis of prices as well as selling probabilities. This allows us to study the housing market more comprehensively.

The paper is organized as follows. Section 2 discusses the relation with the existing literature. Section 3 describes the methodology. Section 4 presents the analysis of actual market outcomes (prices and quantities) in the England and Wales Land Registry. Section 5 runs a similar analysis on house listings on a major UK property portal. Section 6 presents concluding remarks.

2 Literature

Our paper is related to three strands of the literature: (1) studies of loss aversion and reference dependence in asset markets, especially housing; (2) studies of the role of financial frictions in determining people's mobility; and (3) labor economics papers that analyze history dependence in wages.

Asset market literature The seminal paper on loss aversion in the housing market is Genesove and Mayer (2001) who analyse the effect of loss aversion on listing prices, actual prices, and time on the market. Using data on individual property listings in the Boston condominium market at weekly intervals between 1990 and 1997, they show that sellers who previously bought at a higher price set higher advertised prices as they are averse to nominal losses. They also report no significant effects of loss aversion on transacted prices but significant effects on time on the market.

Anenberg (2011) analyze similar issues in a 18-year dataset of sales in the San Francisco Bay Area. While Genesove and Mayer (2001) find that effects on listed prices do not translate significantly onto actual transacted prices, Anenberg (2011) finds stronger and more significant effect of loss aversion on transacted prices.

Beggs and Graddy (2009) study art auctions of Modern, Impressionists, and Contemporary paintings in London and New York; and find evidence of reference dependence in prices. Differently from Genesove and Mayer (2001), they report evidence of anchoring (which implies a symmetric effect of gains and losses) but no evidence of loss aversion.

Despite the differences in scope and markets studied, our paper finds strong evidence of reference dependence in line with Beggs and Graddy (2009). We identify a statistically significant effect of loss aversion both on actual and listed prices, as well as selling probabilities, although the effects account for very little variation in all three variables in England and Wales. Our findings are broadly consistent with those in Genesove and Mayer (2001) and Anenberg (2011), but they also highlight that loss aversion has played a very limited role in the market we analyze and certainly cannot explain away the history dependence observed in the data.

Mobility The gyrations in the housing market of the recent years have stimulated a number of studies on the relation between house prices and mobility, in which the role of mortgage financing and loss aversion is often critical.

Ferreira et al. (2012), for instance, use data from the American Housing Survey to analyze the effect of negative equity and mortgage interest rate lock-in on household mobility. Engelhardt (2003) use the panel of the National Longitudinal Survey of Youth to assess the impact of loss aversion and down-payment requirements on household mobility.

These studies focus on household mobility with an eye on its labour market consequences. In this paper, we focus specifically on housing sales, abstracting from considerations on the mobility of households.

Labour market literature Beaudry and DiNardo (1991) is one of the first papers to emphasize history dependence in the labor market. The authors take a standard wage equation and show that the unemployment rate *when the contract started* is a significant determinant of today's wages. They interpret their findings as a result of wage stickiness and insurance contracts (firms insure workers against fluctuations in income over the business cycle). Their results have been replicated in a number of studies and for different countries: for instance, Grant (2003) shows that the results hold for a different period; McDonald and Worswick (1999) show they hold for Canada; and Devereux and Hart (2007) for the United Kingdom. Hagedorn and Manovskii (2013) offer a different interpretation for this result. In their model, wages depend only on contemporaneous match quality (productivity). However, the distribution of current match productivities depend on the past evolution of market tightness.

A closely related set of studies in this literature focuses on the effect of market conditions *at the time of labor market entry*. Kahn (2010) uses the National Longitudinal Survey of Youth, whose respondents graduated from college between 1979 and 1989. She estimates the effects of both national and state economic conditions at time of college graduation on labor market outcomes for the first two decades of a career. Oreopoulos et al. (2012) also shows that initial labour market conditions have long-term effects on the earnings of college graduates and (less) on the earnings of noncollege workers.

3 Identifying history dependence

The (log) house price is usually modeled as:

$$p_t = X\beta + \delta_t + v_t + w_t, \tag{1}$$

where X is a vector of housing characteristics, δ_t is the aggregate house price level at time t, v_t denotes unobservable characteristics of the property, and w_t is an idiosyncratic component uncorrelated with aggregate conditions and property characteristics.

To empirically investigate history dependence in prices we add to the equation the aggregate price level at the time when the property was bought:

$$p_t = X\beta + \delta_t + \gamma\hat{\delta}_0 + e_t, \tag{2}$$

where $\hat{\delta}_0$ is the price index estimated for time 0 (when the property was bought) using an equation similar to (1) but without the unobserved term, and γ is the estimated history dependence coefficient. The error is now denoted as e_t (instead of w_t) as it incorporates the unobservable factor v_t.[1] Note that we could have introduced p_0, the price of the individual property when it was previously bought, instead of $\hat{\delta}_0$. However, since p_0 itself can be decomposed into specific characteristics of the property and aggregate market conditions, as in equation (2), appropriate substitution will lead to equation (2).[2] By focusing on the aggregate component of the past price, we sidestep the problem that p_0 also contains unobservable characteristics that could on its own create history dependence.

Figure 1 reveals that, for most of the sample period, England and Wales house prices have been trending upwards. Keeping current sale year constant, such a trend leads to a correlation between property tenure and past aggregate prices ($\hat{\delta}_0$). For instance, a property that has been only two years with an owner will often have a higher $\hat{\delta}_0$ than a property that has been eight years with the same owner. We therefore also control for the duration of the tenure (DUR_t), measured as the number of years between two sales. Such variable has the added advantage of controlling for some time-variant unobserved property characteristics such as depreciation. It is likely that depreciation follows a nonlinear pattern; hence we allow for DUR_t to enter non-parametrically through a third-degree polynomial.[3]

To measure the effect of history dependence on transaction probabilities, we start from an equation similar to (1) but with a 0/1 indicator as dependent variable. This indicator takes the value one when the property was sold in a given year, and zero otherwise. Using this approach, a property appears in the dataset each year after its first registered sale (before the first sale we do not observe DUR_t).

Mechanisms To capture loss aversion we split past aggregate prices ($\hat{\delta}_0$) to distinguish between potential gains and losses in the estimating equation. Our measures are simply the aggregate price growth between two sales, truncated at zero:

$$GAIN_t = \max\left(\hat{\delta}_t - \hat{\delta}_0, 0\right). \tag{3}$$

[1]Notice that including $\hat{\delta}_0$ is equivalent to including a linear term in losses (and gains) with respect to the aggregate index, $\hat{\delta}_0 - \delta_t$, because δ_t is already part of the regression. Strictly speaking, if losses were computed using a measure of aggregate prices calculated in advance, $\hat{\delta}_t$ (in a similar way to which $\hat{\delta}_0$ is computed), the equivalence would not be perfect, because δ_t is estimated in an equation that includes the loss term. However, in practice, results with $\hat{\delta}_0 - \hat{\delta}_t$ are almost identical; they are available from the authors.

[2]The only difference is that the error will contain also past unobservables (v_0).

[3]In the empirical analysis we also interact this polynomial with a dummy for properties that were bought new (rather than in the secondary market), because depreciation could be quicker for new properties and, for those houses, tenure corresponds to the age of the property.

$$LOSS_t = \max\left(\hat{\delta}_0 - \hat{\delta}_t, 0\right). \tag{4}$$

From the perspective of a property owner, a loss will be computed as $p_0 - \hat{p}_t$, where $\hat{p}_t$ is the (log) expected selling price of the property today and p_0, as before, is the price at which the property was bought. A formula such as (4) above assumes that the (log) expected selling price is $\hat{p}_t = p_0 + (\delta_t - \delta_0)$. In other words, the expected price equals the original price indexed by the (quality-adjusted) house price inflation during the period, $\delta_t - \delta_0$. This assumption is used by, among others, Shan (2011), who evaluates the impact of capital gain taxes on housing transactions. However, Genesove and Mayer (2001) use a different measure, $LOSS_t = \max(\delta_0 - \delta_t + \epsilon_0, 0)$. This assumes that the loss from the perspective of the seller is $\hat{p}_0 - \hat{p}_t$, where $\hat{p}_0$ was the expected selling price at the time of purchase, conditional on observables. To us, it seems more natural for prospective sellers to take the initial purchase price p_0 as given and then apply the average price inflation to this actual purchase price.[4]

To capture market thickness, we use the aggregate number of transactions occurring in the market when the property was bought, and indicate it with ϕ_0. This is a proxy for market thickness; other possible measures, such as the average time on the market for properties or the average searching time for house buyers, are not available over this long time frame.

Econometric identification The main challenge in estimating these effects come from the presence of unobserved housing characteristics (v), which could be correlated with our variables of interest ($\hat{\delta}_0$, ϕ_0, $GAIN_t$, $LOSS_t$). These unobserved characteristics can be time-invariant (such as a fixed feature of the property) or time-variant (such as depreciation or different likelihood of properties being renovated in any given period). We control for depreciation in our baseline specification through $f(DUR_t)$ and its interaction with a dummy indicating whether a property is new. To control for the remaining threats (fixed property features and different renovation likelihoods) we use three approaches. The first approach includes unit-level house-fixed effects to partial out time-invariant unobservables. The second approach includes unit-level fixed effects and restricts the sample to flats, as flats are less likely to change their value by a lot after a renovation (their size, a critical determinant of price, usually cannot be altered). The third approach restricts the sample to properties that were bought new, because this eliminates the problem of renovations (albeit it does not allow us to use unit-level fixed effects). When appropriate, we include regressions with unit-level fixed effects in the main tables of the paper. We show regressions with flats and new properties in the Appendix.

[4] Moreover, with this measure we insulate our estimation from the econometric problems associated to the presence of the residual e_0 in the nonlinear loss measure.

4 Market prices and transactions

The first part of this section describes our main data source, the England and Wales Land Registry (LR). It contains twenty years of residential transactions from January 1995 to December 2014. We then describe how we compute our measure of aggregate house prices, and estimate history dependence for prices and quantities. Last but not least, we explore a number of possible mechanisms for the patterns we uncover.

4.1 Data and summary statistics

The LR records all residential property transactions, with few exceptions.[5]

For each sale, the LR contains the precise postcode, the street name, the street number, and the apartment number if the property belongs to a multi-unit building. The LR records three attributes of the property: its type (flat, terraced, semi-detached, detached); whether the property is new; and the tenure type of the property (freehold or leasehold). The variable Date of Transfer in LR is the day written on the transfer deed, that is, the date of completion, when keys and funds change hands.

The analysis presented in this paper relies on the identification of repeat sales on the same property—we need information on the previous purchase of a property to make inference about history dependence. We consider two sales as happening on the same property when they share the same postcode, street name, street number, apartment number (if any), and property type (flat, terraced, semi, detached). Table 1 shows descriptive statistics for the LR and distinguishes between 'sales' and 'properties' to highlight the role of repeat sales in the analysis. Figure A3 in the Appendix shows the number of years elapsed between sales of the same property.

The dataset contains close to 20 million sales for twenty years of data, that is, approximately one million sales per year. The number of properties to appear in the dataset is only 12 million, meaning that there are 7 million properties with more than one sale. Moving to the right columns of Table 1 means restricting attention to sales that happened in later years. We use these more restricted samples (as opposed to the main sample, which we denote as *Sample 1*) because more information is available in later years. Since 2002, the LR dataset includes a variable ('charge') which indicates the use of a mortgage to purchase the property[6]—hence we label as *Sample 2* the subset of transactions whose

[5] The exceptions are listed at http://www.landregistry.gov.uk/market-trend-data/public-data/price-paid-data. Most of these excluded transactions refer to sales through company structures or investment purchases. In practice, many investment purchases are nevertheless included in the data—see Bracke (2015) who studies buy-to-rent investments. A public version of the dataset is available online (http://www.landregistry.gov.uk/market-trend-data/public-data/price-paid-data).

[6] This variable is not available in the public dataset but can be purchased from the Land Registry.

Table 1: Summary statistics: price analysis

Notes: The table shows descriptive statistics of the sales included in the England and Wales Land Registry for the years 1995-2014. The first column contains information on all these sales. The second column describes the dataset used to analyse history dependence in prices: it is made of all properties which have at least two sales in the dataset, and excludes for each property the first of such sales. (The first sale is used to compute the market conditions—aggregate prices and number of transactions—at the time of the purchase.) The third column is similar to the second but only refers to properties whose first sale took place after 2001, as for this sample we can tell whether the property was purchased with a mortgage. Finally, the fourth column describes properties whose first sale took place after March 2005 and can potentially be matched to the Product Sales Database (PSD), a dataset of residential mortgages where we can identify the initial LTV with which a house was bought. (All residential mortgages in the United Kingdom have been recorded starting in April 2005. The PSD is described in more detail in subsection 4.3.)
The first two rows of the table report the number of transactions in the dataset and the number of individual properties that appear at least once in the dataset, respectively. The variables from Flat to New are dummy indicators included in the empirical analysis; the table reports their average.

	All Land Registry	**Sample 1** Complete Previous sale > 1994	**Sample 2** Mortgage vs cash Previous sale > 2001	**Sample 3** Mortgage info Previous sale > 2004
Sales	19,628,516	7,527,731	3,199,389	1,385,653
Properties	12,089,086	5,038,658	2,570,092	1,234,381
Median price	124,500	145,000	165,000	176,500
Flat	0.18	0.19	0.22	0.24
Terraced	0.31	0.34	0.34	0.32
Semi	0.28	0.27	0.26	0.25
Detached	0.23	0.20	0.19	0.19
Lease	0.23	0.24	0.27	0.28
New	0.10	0.00	0.00	0.00

Table 2: Summary statistics: transactions analysis

Notes: The table shows descriptive statistics of the dataset used to analyse the transaction probability of properties in any given year. The dataset is created by taking the LR samples (whose descriptive statistics are shown in Table 1) and expanding them so that each house has an observation in each year since its first appearance in the LR. (For the empirical analysis we create a variable which equals one if property i sells in year t, and zero otherwise.) This table and the regressions reported in the paper refer to a 10% random sample of the data, where the sampling is done on postcodes. The three columns in the table correspond to the same three samples of the data as described in Table 1.

	Sample 1 10% sample, extended Previous sale > 1995	**Sample 2** Mortgage vs cash Previous sale > 2001	**Sample 3** Mortgage info Previous sale > 2004
Observations	13,790,771	6,782,041	3,644,508
Sales	721,364	300,536	127,680
Properties	1,169,610	863,576	637,284
Median price	90,000	144,950	169,950
Flat	0.19	0.22	0.24
Terraced	0.30	0.31	0.31
Semi	0.29	0.28	0.28
Detached	0.25	0.23	0.22
Lease	0.21	0.24	0.25
New	0.10	0.10	0.10

previous purchase happened after 2001. Since 2005, the UK Financial Conduct Authority (FCA) has been recording information on all owner-occupier mortgages into the Product Sales Database (PSD)—hence we label as *Sample 3* the subset of transactions that can be matched into the PSD. These more restricted samples contain more flats and, therefore, more leasehold properties.[7] There are no new properties in column 2, 3, and 4 of the Table, as it is expected, since these are part of repeat-sale pairs and the first purchase (which could refer to a new build) is not part of the sample.

To estimate the impact of history dependence on a property's selling probability (and, in aggregate, on the number of transactions) we reshape and expand the dataset so that each house has an observation in each year since its first appearance in the LR (its first sale after 1995). With 12 million properties and 20 years, the final extended datasets has over 120 million rows (the average property appears for the first time in the middle of the sample, meaning that we can follow it for 10 years). To keep the empirical analysis manageable, we extract a 10% random sample of the data.[8] We create a variable, q_{it}, which equals one if property i sells in year t, and zero otherwise. We treat the first sale as missing because we do not observe DUR_t before that observation. Figure A5 in the Appendix shows the trend of the average q_{it} over the years. This trend mirrors the number of aggregate transactions showed in Figure 1. Table 2 replicates the descriptive statistics of Table 1 for this extended dataset.

We compute the aggregate level of house prices needed to measure history dependence by running a standard regression of the form:

$$p_{ijt} = \alpha_j + X_i\beta + \delta_t + e_{it} \tag{5}$$

where δ_t is a set of time dummies at monthly frequencies from 1995 to 2014. The coefficients on these dummies are shown in Figure 1. Running a repeat sales regression, rather than using postcode fixed effects, produces nearly identical results. Not surprisingly the index we estimate tracks very closely published series. Figure A4 in the Appendix compares our aggregate measure with the Nationwide house price index (not seasonally adjusted).

Figure 2 shows gains and losses for the sales used in the price analysis. Given the aggregate movement in house prices shown in Figure 1, for most households in England and Wales homeownership has produced gains rather than losses. The sample contains 496,574 sales with a realized nominal loss (out of 7.5 million transactions). Conditional

[7] A leasehold is a tenancy arrangement by which someone buys a property for a limited number of years, usually 99, 125 or 999. It is usually associated with flats. See Bracke et al. (2014) and Giglio et al. (2015).

[8] Since most of the regressions reported use postcode fixed effects, the random sampling is based on postcodes.

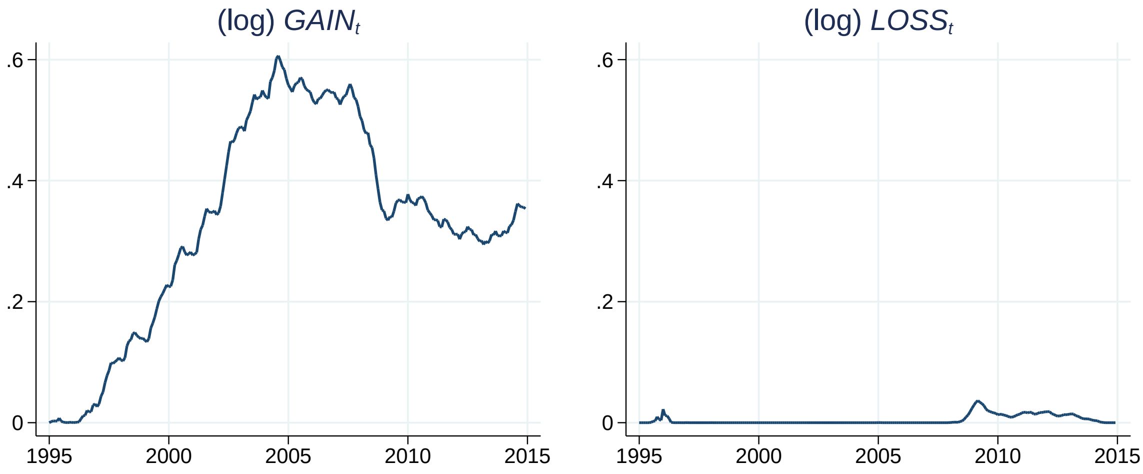

Figure 2: Gains and losses when selling

Notes: The charts show average (log) gains and losses for properties sold in any given month of the sample used in the price analysis. The average gain is 41.4 percent, with a standard deviation of 34.2 percent. The average loss is 0.4 percent, with a standard deviation of 1.8 percent.

on realizing a loss, the average (log) loss is 5.3 percent; conditional on realizing a gain, the average (log) gain is 44.3 percent. It is suggestive to notice that the drop in the average number of transactions per year shown in Figure 1 overlaps nicely with the drop in average gains experienced by sellers.

4.2 History dependence

Prices Table 3 contains regressions with the current sale price of a house as the dependent variable. All regressions control for property type as measured by the LR (flat, terraced, semi-detached or detached property; new or second-hand property; property sold as leasehold or freehold) as well as the number of years elapsed since the current sellers have bought the property (DUR_t).

The first two columns of the Table show the results of regressing today's prices on the prices of previous purchases of the same properties. This is for descriptive purposes only, since any coefficient on previous prices may be capturing the effect of unobserved property characteristics rather than pure history dependence. As expected, the two regressions—the second of which also interacts the previous price with DUR_t—yield a large and significant coefficient for the main variable of interest.

The remaining four columns explore the effect of past aggregate prices (δ_0). The first two columns use postcode fixed effects ('PC6' indicates a full UK postcode, which is

Table 3: History dependence in house prices

Notes: The table contains regressions where the dependent variable is the current sale price of a house. All regressions control for property type as measured by the Land Registry (flat, terrached, semi-detached or detached property; new or second-hand property; property sold as leasehold or freehold) for a nonparametric function of the number of years between sales interacted with whether the property was bought new. 'PC6, M' indicates full postcode and month fixed effects and 'Unit, M' indicates unit-level fixed effect regressions and monthly fixed effects.

Dependent variable:	Transacted price (p_{ijt}) **Sample 1** (1995-2014)					
	(1)	(2)	(3)	(4)	(5)	(6)
Previous price (p_0)	0.344 (0.000)	0.348 (0.000)				
$p_0 \times DUR_{it}$		-0.001 (0.000)				
Previous aggr. prices (δ_0)			0.042 (0.001)	0.068 (0.001)	0.042 (0.001)	0.077 (0.001)
$\delta_0 \times DUR_{it}$				-0.004 (0.000)		-0.006 (0.000)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	PC6, M		PC6, M		Unit, M	
N	7,527,731	7,527,731	7,527,731	7,527,731	7,527,731	7,527,731
Fixed effects	957,732	957,732	957,732	957,732	5,038,657	5,038,657
R-squared	0.924	0.924	0.908	0.908	0.981	0.981

associated with 10-15 housing units)—because our data do not contain much information on housing characteristics, it is important to use detailed geographical fixed effects to control both for physical property figure and for the value of location. The last two column refer to the unit-level fixed effect regressions (and also includes monthly fixed effects to control for the average price in England and Wales).

Effects are between 0.04 and 0.08, indicating a 4-8 percent increase of the sale price compared to another equivalent house which was bought in a different period at half the price. The interaction with DUR_t is negative and significant in all specifications, implying that history dependence fades over time.

Columns (1) and (2) of the two panels of Table A.1 in the Appendix contain the two robustness checks for the price regression mentioned in the Methodology section: unit-level fixed effects restricting the sample to flats and a sample only made of properties which were bought new. The restricted samples replicate the results of the main analysis. If anything, the history dependence coefficients are larger than in Table 3.

Transactions The goal of the transactions analysis is to investigate whether the purchase price of a property affects the probability that a house sells in any subsequent period. We assume the following linear model analogous to equation (2):

$$\Pr(q_{jt} = 1) = \alpha_j + X\beta + \delta_t + \hat{\delta}_0 + f(DUR_t) + e_t. \tag{6}$$

Table 4 shows the results for history dependence in housing quantities in a similar manner as Table 3 does for prices. The dataset is in a property-year format and therefore we control for aggregate conditions with year fixed effects (rather than month fixed effects as in Table 3).

The coefficient on the previous price (p_0) is -0.013 in the first column and -0.026 when the previous price is interacted with DUR_t. These are substantial effects since the average selling probability in the sample is 0.052 (see Figure A5). The interaction has a positive coefficient, indicating a declining importance of history dependence as years go by.

Columns (3) and (4) contain the main specification, with postcode fixed effects. The coefficient without interaction indicates that doubling aggregate house prices at the time of purchase decreases the selling probability by 0.5 percent (or, equivalently, causes a 10 percent decline relative to the average annual selling probability of 5 percent). The coefficient gets larger when the interaction with DUR_t is included.

Effects are much larger in the regressions with unit fixed effects. By construction, only properties that eventually sell affect the coefficients of these regressions (for properties that never sell—after the first purchase—the dependent variable is always zero and gets

Table 4: History dependence in transactions

Notes: The structure of the table and the variables used in the regressions are analogous to Table 3. The dependent variable is an indicator of whether a house was sold in any given year. The dataset used for the regressions is an expanded version of the LR where each property appears in every year after its first recorded sale.
To make the number of observations manageable, the time dimension in the fixed effects is the year (Y) rather than the month (M) as in Table 3. For the same reason, results are reported for a 10% random sample of properties in the data.

Dependent variable:	Selling probability (q_{ijt}) **Sample 1** (1995-2014)					
	(1)	(2)	(3)	(4)	(5)	(6)
Previous price (p_0)	-0.013 (0.000)	-0.026 (0.000)				
$p_0 \times DUR_{it}$		0.002 (0.000)				
Previous aggr. prices (δ_0)			-0.005 (0.001)	-0.020 (0.001)	-0.136 (0.001)	-0.119 (0.002)
$\delta_0 \times DUR_{it}$				0.002 (0.000)		-0.002 (0.000)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	PC6, Y		PC6, Y		Unit, M	
N	13,790,771	13,790,771	13,790,771	13,790,771	13,790,771	13,790,771
Fixed effects	114,986	114,986	114,986	114,986	1,169,609	1,169,609
R-squared	0.025	0.025	0.025	0.025	0.138	0.138

absorbed by the fixed effect). Conditional on selling at least once after the property enters the dataset, the average annual selling probability is 10.4 percent. The effect of history dependence in these regressions is still large, predicting an fall of 12-14 percent in the conditional selling probability when past aggregate prices double.

4.3 Loss aversion, down payment, and match quality

Prices To distinguish loss aversion from symmetric anchoring to the previous purchase price, we separate gains from losses. To investigate match quality, we include the log of the total number of transactions in England and Wales at the moment the house was purchased by the current seller as a measure of market thickness.

The first column in Table 5 shows the results for *Sample 1.* The coefficient on gains remains similar to the coefficient on past aggregate prices in Table 3, but with opposite sign. Since most sellers in the sample realised a gain, the movements in the past aggregate price index map quite well into movements in gains. The coefficient on losses is substantially higher than the one on gains, indicating the presence of loss aversion. This regularity holds in all the remaining regressions of the paper, in line with the previous literature. However, quantitatively, $LOSS_t$ can account for only a very small fraction of the variation in prices. Specifically, if we use as a measure of quantitative significance the ratio $\frac{\hat{\gamma_{LOSS}}\sigma_{LOSS}}{\sigma_{p_0}}$, which indicates the share of variation in p_0 that can be accounted by the variation in LOSS, the result is a modest 0.5 percent (the standard deviation of log prices is 0.7). In contrast, the effect of simple anchoring is quantitatively bigger. A similar statistical measure as above indicates that δ_0 can account for 2 percent of the variation in transacted prices.

The coefficient on market thickness (represented by ϕ_0, the number of transactions in the market at the time the house was bought) is low at 0.2 percent. The standard deviation of the log of past transactions is 0.25; hence the share of current price variation predicted by variation in ϕ_0 is less than 0.1 percent.

According to Stein (1995), mortgage leverage is important in determining housing market activity. Because homeowners need a downpayment to move to their next property, they are reluctant to sell at a loss, which would translate in no equity to be put in the purchase of their next house. We check this mechanism by switching our attention to *Sample 2*, which has information on whether a property was bought with a mortgage. To highlight down-payment effects, we interact our main variables of interest ($GAIN_t$, $LOSS_t$) with an indicator variable for properties that were bought with a mortgage. Before switching to these new regressions, we make sure in the second column of Table 5 that the regression without the interaction terms has similar coefficients than the regression in *Sample 1.*

Table 5: Mechanisms, price regressions

Notes: The Table shows regressions to investigate the mechanisms behind history dependence. The regression in the first column of the table shares the same sample as Table 3. The other columns refer to *Sample 2*, where information is available on whether a property was purchased with a mortgage. (The statistics for this sample are shown in column (3) of Table 1.) We created an indicator variable associated with mortgage-funded transactions which we include both in levels and interacted with the main variables of interest.

Dependent variable:	Transacted price (p_{ijt})			
	Sample 1 (1995-2014)	**Sample 2** (2002-2014)		
	(1)	(2)	(3)	(4)
$GAIN_t$	-0.040	-0.054	-0.029	-0.036
	(0.001)	(0.001)	(0.002)	(0.002)
Mortgage x $GAIN_t$			-0.027	-0.013
			(0.002)	(0.002)
$LOSS_t$	0.196	0.161	0.186	0.153
	(0.006)	(0.007)	(0.011)	(0.011)
Mortgage x $LOSS_t$			-0.045	-0.005
			(0.012)	(0.012)
Market size when bought (ϕ_0)	-0.002	0.003	0.001	0.012
	(0.000)	(0.001)	(0.001)	(0.001)
Mortgage x ϕ_0				-0.015
				(0.001)
Bought with mortgage			0.032	0.195
			(0.001)	(0.012)
Controls	Yes	Yes	Yes	Yes
Controls, Month FE × Mortgage				Yes
Fixed effects	PC6, M	PC6, M		
N	7,527,731	3,199,389	3,199,389	3,199,389
Fixed effects	957,732	783,108	783,108	783,108
R-squared	0.908	0.911	0.911	0.911

Columns (3) and (4) contain the interactions of the variables of interest with the mortgage dummy. The baseline coefficients are now to be interpreted as the effect on properties bought without a mortgage. These effects show the importance of behavioral forces unrelated to financing frictions. On top of these effect, we would expect a larger effect of $LOSS_t$ on properties financed with a mortgage, and a lower effect of $GAIN_t$. In apparent contrast with Stein (1995), we find instead a larger effect of $GAIN_t$ and a smaller effect of $LOSS_t$, as compared with properties financed by cash. However, the effects highlighted by Stein (1995) are likely to be stronger for properties sold with a high LTV—we check that with *Sample 3*, in the next subsection.

While the regression in column (3) only includes the interaction of the mortgage indicator with $GAIN_t$ and $LOSS_t$, the regression in column (4) also includes the interaction with ϕ_0, past market thickness, as well as with all other controls and time dummies to capture any potential difference in those coefficients between cash- and mortgage-financed properties.

In the second and third column of the table, the market size at the time of purchase (ϕ_0) has still a small effect, between 0.001 and .003, from which it does not appear that thick market effects, at least when measured through the overall number of housing transactions at the time of purchase, are a major force driving house prices.[9] However, when all controls and the monthly dummies are interacted with the mortgage indicator in column (4), there does seem to be a thick market effect for cash-financed properties, whereas the effect is not there for mortgage-financed ones—a result that requires further investigation.

The robustness checks in the two panels of Table A.1 in the Appendix are consistent with the results presented here, except for the loss term in the sample of properties bought new, which has a negative effect on prices.

Transactions Results for selling probabilities are presented in Table 6. In this case both the effect of $GAIN_t$ and $LOSS_t$ are stronger for mortgage-financed properties, consistent with Stein (1995). In fact, different from the price regression, most of the effect of gains on transactions comes from properties bought with a mortgage. The coefficient on $LOSS_t$ is significantly higher than the coefficient on $GAIN_t$, although the difference between the two is reduced moving from *Sample 1* to *Sample 2*. As with the price analysis, because losses are such a rare occurrence, the large coefficient on $LOSS_t$ does not translate into a large quantitative effect. Using a similar statistics as before, we find that loss aversion can account for 0.8 percent of the variation in selling probabilities.

[9]We got similar results when we computed the overall number of housing transactions at a more local level.

Table 6: Mechanisms, transactions regressions

Notes: The table is similar to Table 5 except that the time dimension is year (Y) and the dependent variable is an indicator for whether the property was sold (q_{jt}).

Dependent variable:	Selling probability (q_{ijt})			
	Sample 1 (1995-2014)	**Sample 2** (2002-2014)		
	(1)	(2)	(3)	(4)
$GAIN_t$	0.005	0.019	-0.007	-0.000
	(0.001)	(0.001)	(0.001)	(0.002)
Mortgage x $GAIN_t$			0.037	0.026
			(0.001)	(0.002)
$LOSS_t$	-0.086	-0.040	-0.009	-0.020
	(0.004)	(0.004)	(0.007)	(0.008)
Mortgage x $LOSS_t$			-0.037	-0.021
			(0.007)	(0.009)
Market size when bought (ϕ_0)	0.003	-0.007	-0.006	-0.007
	(0.000)	(0.001)	(0.001)	(0.001)
Mortgage x ϕ_0				0.000
				(0.001)
Bought with mortgage			-0.005	-0.073
			(0.000)	(0.019)
Controls	Yes	Yes	Yes	Yes
Controls, Year FE × Mortgage				Yes
Fixed effects	PC6, Y	PC6, Y		
N	13,790,771	6,782,041	6,782,041	6,782,041
Fixed effects	114,986	110,321	110,321	110,321
R-squared	0.025	0.029	0.029	0.029

Moving from *Sample 1* to *Sample 2* changes the sign of the thick market effect, which goes from 0.003 to -0.007. A negative effect of ϕ_0 is more consistent with intuition; a hot market at the time of purchase raises the average quality of matches, which in turn leads homeowners to wait longer before moving. Since properties in *Sample 2* have been bought more recently on average, the change in sign and magnitude of the coefficient could mean that these effects are stronger at shorter tenure durations (i.e., when DUR_t is lower).

4.4 Further analysis with mortgage information

To improve on the previous regressions, we interact our variables of interest with an indicator of properties sold with a high Loan-To-Value ratio (LTV). The information on LTV comes from the PSD, a confidential dataset containing information on all individual homebuyer mortgages in the UK since April 2005. The PSD only collects information on new sales (mortgages or re-mortgages), excluding alterations or top-ups of the loan.

Available variables include mortgage characteristics such as loan size, the LTV at the time of origination, length in years, interest rate, and whether the borrower is a first time buyer; the postcode of the property (rather than the complete address), and buyers' age and income. The data are transmitted from banks and other mortgage providers to the Financial Conduct Authority on a quarterly basis.

To create our high-LTV indicator, we first match the PSD loans with LR transactions. The match is based on postcode and property price, which we require to coincide exactly. When there are multiple matches for one LR sale, the mortgage whose origination date is closest to the LR completion date is chosen. According to the LR cash-vs-mortgage data, out of the 1,385,653 sales in *Sample 3*, 980,220 are of properties that were bought with a mortgage. At the moment we match 432,319 of these with a loan in the PSD. The match is therefore not perfect, and the control group of mortgaged properties with a low LTV must be interpreted as a contaminated control group. To define a high-LTV mortgage, we choose a threshold of 80 percent, which corresponds to the median LTV at the time of purchase.

Prices Before analyzing the effect of interactions, it is useful to check that the coefficients on gains, losses, and market thickness for *Sample 3* are similar to the ones for *Sample 1* and *Sample 2*. The first column of Table 7 shows that they are.

It is also useful to check whether the interactions with the mortgage indicator yield a similar effect to the one in *Sample 2*. The second column shows that this is true for the loss coefficient, which is reduced in the interaction with mortgages, but not for the gain coefficient. It now appears that sellers who bought with a mortgage are less willing to reduce the price when they are experiencing a gain.

The third column introduces the interaction with the high-LTV indicator. Consistent with Stein (1995), the interaction coefficient indicates a lower willingness to reduce the price when gaining and a higher willingness to increase the price when facing losses.

Transactions Similar to the price regressions, Table 8 shows that the effect of gains on selling probability gets bigger in more recent samples. Most of the effect comes from mortgage-funded properties and, within these, properties funded with a high LTV.

The effect of losses in *Sample 3* is similar to the one in *Sample 2*. There is an effect of losses on cash-financed properties in all specifications, but the effect on mortgage-financed properties is slightly larger. The largest effect is for properties bought with high leverage.

The coefficient on market thickness is the same as in *Sample 2*; the last column shows a relatively large effect of market thickness on high-LTV properties.

Table 7: High-LTV mortgages, price regressions

Notes: The regressions in this table are similar to the ones reported in columns (2)-(4) in Table 5, except that (a) the subset of the data used is Sample 3 and we include in the regressions the interactions of the variables of interest with an indicator variable flagging properties that were bought with a LTV greater than 80 percent.

Dependent variable:	Transacted price (p_{ijt})				
	Sample 3 (2005-2014)				
	(1)	(2)	(3)	(4)	(5)
$GAIN_t$	-0.095	-0.120	-0.122	-0.088	-0.086
	(0.009)	(0.012)	(0.012)	(0.016)	(0.016)
Mortgage x $GAIN_t$		0.031	0.030	0.016	0.019
		(0.011)	(0.011)	(0.019)	(0.020)
($LTV > 80\%$) x $GAIN_t$			0.019	0.020	0.009
			(0.013)	(0.013)	(0.024)
$LOSS_t$	0.158	0.181	0.177	0.115	0.114
	(0.008)	(0.014)	(0.014)	(0.016)	(0.016)
Mortgage x $LOSS_t$		-0.039	-0.046	0.030	0.025
		(0.015)	(0.016)	(0.019)	(0.019)
($LTV > 80\%$) x $LOSS_t$			0.055	0.055	0.083
			(0.021)	(0.021)	(0.024)
Market size when	0.001	-0.000	0.000	0.005	0.005
bought (ϕ_0)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Mortgage x ϕ_0				-0.004	-0.003
				(0.002)	(0.002)
($LTV > 80\%$) x ϕ_0					-0.006
					(0.003)
Bought with		0.020	0.021	0.090	0.075
mortgage		(0.001)	(0.001)	(0.037)	(0.038)
$LTV > 80\%$			-0.011	-0.011	0.098
			(0.001)	(0.001)	(0.056)
Controls	Yes	Yes	Yes	Yes	Yes
Controls, Month FE × Mortgage				Yes	Yes
Controls, Month FE × ($LTV > 80\%$)					Yes
Fixed effects			PC6, M		
N	1,385,653	1,385,653	1,385,653	1,385,653	1,385,653
Fixed effects	566,868	566,868	566,868	566,868	566,868
R-squared	0.931	0.932	0.932	0.932	0.932

Table 8: High-LTV mortgages, transactions regressions

Notes: The table is similar to Table 7 except that the time dimension is year (Y) and the dependent variable is an indicator for whether the property was sold (ϕ_0).

Dependent variable:			Selling probability (q_{ijt}) **Sample 3** (2005-2014)		
	(1)	(2)	(3)	(4)	(5)
$GAIN_t$	0.040	-0.003	-0.004	0.015	0.016
	(0.004)	(0.006)	(0.006)	(0.009)	(0.009)
Mortgage x $GAIN_t$		0.058	0.050	0.023	0.029
		(0.006)	(0.006)	(0.010)	(0.010)
$(LTV > 80\%)$ x $GAIN_t$			0.049	0.051	0.027
			(0.008)	(0.008)	(0.013)
$LOSS_t$	-0.044	-0.021	-0.019	-0.024	-0.024
	(0.004)	(0.007)	(0.007)	(0.008)	(0.008)
Mortgage x $LOSS_t$		-0.028	-0.012	-0.002	-0.009
		(0.008)	(0.008)	(0.010)	(0.010)
$(LTV > 80\%)$ x $LOSS_t$			-0.087	-0.087	-0.039
			(0.009)	(0.009)	(0.011)
Market size when	-0.006	-0.006	-0.006	-0.006	-0.006
bought (ϕ_0)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Mortgage x ϕ_0				-0.001	0.001
				(0.001)	(0.001)
$(LTV > 80\%)$ x ϕ_0					-0.008
					(0.002)
Bought with		-0.005	-0.005	-0.017	-0.035
mortgage		(0.000)	(0.000)	(0.020)	(0.020)
$LTV > 80\%$			0.004	0.004	0.097
			(0.001)	(0.001)	(0.025)
Controls	Yes	Yes	Yes	Yes	Yes
Controls, Year FE × Mortgage				Yes	Yes
Controls, Year FE × $(LTV > 80\%)$					Yes
Fixed effects			PC6, Y		
N	3,644,508	3,644,508	3,644,508	3,644,508	3,644,508
Fixed effects	104,845	104,845	104,845	104,845	104,845
R-squared	0.035	0.035	0.035	0.036	0.036

Table 9: Quantitative effects of changes in gains and losses

Notes: TO BE COMPLETED

	Levels			Changes (Δ)	
	2007	2008	2014	2007 – 2008	2007 – 2014
Average selling probability (%)	7.1	3.1	4.1	4.0	3.0
Average (log) $GAIN$ (%)	59.0	48.6	37.6	10.4	21.4
Average (log) $LOSS$ (%)	0.0	0.4	0.0	-0.4	0.0
% of Δ selling prob explained by $\Delta GAIN/LOSS$: $\frac{\gamma_{GAIN}\Delta GAIN+\gamma_{LOSS}\Delta LOSS}{\Delta \text{selling probability}}$					
Estimates using aggregate price indices					
Sample 1 (1995-2014): $\gamma_{GAIN} = 0.005$, $\gamma_{LOSS} = -0.086$				2.1%	3.6%
Sample 2 (2002-2014): $\gamma_{GAIN} = 0.019$, $\gamma_{LOSS} = -0.040$				5.5%	14.3%
Sample 3 (2005-2014): $\gamma_{GAIN} = 0.040$, $\gamma_{LOSS} = -0.044$				10.8%	28.5%
Estimates using regional price indices					
Sample 1 (1995-2014): $\gamma_{GAIN} = 0.005$, $\gamma_{LOSS} = -0.090$				2.2%	3.6%
Sample 2 (2002-2014): $\gamma_{GAIN} = 0.024$, $\gamma_{LOSS} = -0.078$				7.0%	17.1%
Sample 3 (2005-2014): $\gamma_{GAIN} = 0.054$, $\gamma_{LOSS} = -0.091$				14.9%	38.5%

4.5 How much can we explain of the post-2007 fall in transactions?

The fall in housing market activity that occurred in the UK after 2007 is one of the motivations for our paper. As shown in Figure 1, the aggregate number of transactions did not return to its pre-crisis level even after six years, in 2014. Figure 2 shows there were significant changes in the average gains and losses over this period. The goal of this subsection is to understand whether these movements can explain part of the low activity in the UK housing market given the effects estimated by our regressions. We examine both the fall in transactions between 2007 and 2008, and for the one between 2007 and 2014—the most recent year for which we have data.

Table 9 shows our calculations. In the upper half of the table we collect the factual data: the differences in selling probabilities and the differences in average gains and losses for the relevant years. The lower half of the table contains the calculations based on the coefficients from the empirical analysis. Because the coefficients change depending on whether *Sample 1*, *Sample 2*, or *Sample 3* is used, we compute the contribution of gains and losses with all the coefficients.

We see from the upper half of Table 9 that most of the action comes from changes in average gains. Average losses were null in 2007, reached an unremarkable 0.4 percent in 2008 and fell back to null again in 2014. By contrast, average gains went down by 10 percentage points between 2007 and 2008, and by further 10 percentage points between 2008 and 2014. In general the explanatory power for the 2007-2014 difference is greater than the one for 2007-2008, because as time went by average gains kept on dimishing—some homeowners with large gains sold in the meantime, and more homeowners bought their houses when prices were already high. Since the coefficient on gains changes substantially depending on the sample, the explanatory power of gains and losses ranges from 2 to almost 30 percent. On the one hand, later samples have properties with a shorter holding period (DUR_t), which could push history dependence up. On the other hand, if history dependence has increased over time (perhaps because of the increased leverage of homebuyers) it is more appropriate to use more recent estimates. On balance, we believe the large drop in average gains experienced by UK homeowners has had a significant role in the post-2007 decline in housing activity.

Regional analysis In the analysis so far we have used prices aggregated at the national level to compute the effect of gains and losses on selling probability. A more granular approach could increase the explanatory power of the model because people are likely to be influenced by local, rather than national, prices. We re-estimate the transactions regressions using regional aggregate house prices instead of national ones, and obtain very similar, albeit slightly larger coefficients. The results from these specfication are at the bottom of the lower panel.

5 Listing prices and time on the market

This section is based on data from Zoopla, a major UK property portal. Using this source allows us to study listing prices and time on the market for properties that were advertised for sale in England and Wales after 2008. Some of these properties were later sold and became part of the LR records.

5.1 Data and summary statistics

Zoopla is the second UK property portal in terms of traffic. Its dataset starts in November 2008.[10] In this paper we restrict our attention to sale adverts where an address can be precisely identified.[11] The dataset contains information on the address of properties, listing prices, and property attributes (such as property type and number of bedrooms).

Zoopla collects data only from estate agents, not individual sellers. In the UK, most transactions occur via estate agents (in 2010, only 11 percent of homes were sold privately

[10]Notice that the availability of Zoopla only since end 2008—i.e., after the crisis—is not a major constraint for our analysis since we can still know from the Land Registry the previous, pre-crisis purchase price for these properties, and study how it relates to advertised prices.

[11]Since information is supplied to Zoopla by estate agents logging in the portal, adverts can sometimes contain missing or wrong information. We only use advertisements with a complete address.

Table 10: Zoopla Summary statistics

Notes: The table contains descriptive statistics for sale listings that appeared on the Zoopla property portal between January 2009 and December 2014. A listing is labelled 'sold' when it matches a LR sale within 12 months of being posted online.

	Zoopla Sample 1		**Zoopla Sample 2**		**Zoopla Sample 3**	
	All listings	Sold listings	All listings	Sold listings	All listings	Sold listings
	Previous sale > 1994		Previous sale >2001		Previous sale > 2004	
Sales	2,485,382	1,021,092	1,897,169	757,813	1,317,877	512,485
Properties	1,957,504	979,891	1,497,020	728,295	1,052,301	494,287
Median price (listed)	185,000	189,995	180,000	189,950	180,000	189,950
Median price (sold)		183,000		181,500		183,000
Flat	0.16	0.15	0.18	0.16	0.19	0.18
Terraced	0.32	0.34	0.33	0.35	0.34	0.35
Semi	0.29	0.29	0.28	0.29	0.27	0.28
Detached	0.23	0.22	0.21	0.20	0.20	0.19
Bedrooms	2.85	2.82	2.80	2.78	2.76	2.74
Lease	0.21	0.19	0.23	0.20	0.24	0.21

– see Office of Fair Trading (2010)).

The Zoopla data reveal the intention to put a property on the market. To verify whether a specific property was actually sold, we check whether we can match the listing with a LR transaction on the same property address within one year of the appearance of the listing. Table A2 in the Appendix describes the steps to carry out the match.

Similar to Table 1, Table 10 shows the descriptive statistics for the Zoopla dataset. The table contains two more variables with respect to Table 1—the listed price and the number of bedrooms—and distinguishes between properties that appeared on Zoopla (All listings) and properties that effectively sold (Sold listings). In the empirical analysis, we concentrate on the latter group to compare the effect on listed price against the effect on the transacted price for the same sample of properties.

5.2 History dependence in listed prices and time on the market

Prices We run the regression in equation (2) both on the listed and transacted price of properties in the Zoopla sample. Results are reported in Table 11; similar to the analysis in the previous section, we include in the regressions the interaction with the mortgage indicator for properties in *Zoopla Sample 2*, and the interaction with the high-LTV indicator for properties in *Zoopla Sample 3*.

As expected, loss aversion generally has a positive and bigger impact on listed prices than on actual prices, implying that the price effect of loss aversion arises from the seller and gets partially reduced in the interaction with the buyer of the property. By contrast, the effect of gains (anchoring) is ambiguous. In the regressions on *Zoopla Sample 1* and *Zoopla Sample 2*, the effect of gains is *lower* on listed prices, suggesting that the effect of gains arise from the interaction between seller and buyer. According to this view, sellers with large gains should accept larger discounts from their asking price. However, results on *Zoopla Sample 3* seem to suggest the opposite.

In terms of leverage effects, properties bought with a mortgage display a lower effect of gains and losses on both the listed and the final price. By contrast, properties bought with a high LTV have a lower effect of gains and a higher effect of losses, consistent with Table 7.

Transactions With listing data we can check whether historical conditions influence the hazard rate at which a house sells once it has been advertised on the property portal. Similar to the analysis of the unconditional selling probability, we expand the Zoopla samples so that each property-month combination has its row, and construct a dummy variable which equals one only when the property is matched with a LR transaction because it was sold.

Figure A6 shows the average selling probability grouped by date month (on the left-hand side graph) or months after listing (on the right-hand side graph). The average monthly selling probability is 0.077.

Table 12 shows that gains have a positive effect on selling probability; the effect is small in *Zoopla Sample 1* and *Zoopla Sample 2*, and large in *Zoopla Sample 3*. The interaction between gains and mortgage indicators has an ambiguous sign. Losses have a negative and large effect, and the effect is larger for houses bought with a high-LTV mortgage. The thick market effect (ϕ_0) is negative and significant, albeit not huge.

Table 11: History dependence in listed and actual prices

Notes: The table has a similar structure to Table 7. 'Listed' refers to the selling price advertised on the Zoopla portal. 'Transacted' refers to the actual transaction price recorded in the LR.

Dependent variable:	Price (p_{ijt})					
	Listed	Transacted	Listed	Transacted	Listed	Transacted
	Zoopla Sample 1		**Z. Sample 2**		**Z. Sample 3**	
	1995-2014		2002-2014		2005-2014	
$GAIN_t$	-0.001	-0.010	-0.028	-0.037	-0.158	-0.052
	(0.002)	(0.002)	(0.004)	(0.004)	(0.019)	(0.019)
Mortgage x $GAIN_t$			0.026	0.027	0.120	0.103
			(0.004)	(0.004)	(0.019)	(0.019)
$(LTV > 80\%)$ x $GAIN_t$					0.031	0.027
					(0.021)	(0.021)
$LOSS_t$	0.124	0.084	0.183	0.141	0.155	0.135
	(0.007)	(0.007)	(0.015)	(0.016)	(0.018)	(0.019)
Mortgage x $LOSS_t$			-0.078	-0.073	-0.072	-0.070
			(0.017)	(0.017)	(0.021)	(0.021)
$(LTV > 80\%)$ x $LOSS_t$					0.068	0.059
					(0.022)	(0.023)
Market size when	-0.001	-0.001	0.005	0.003	0.002	0.005
bought (ϕ_0)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Mortgage x ϕ_0			-0.006	-0.005	-0.002	-0.002
			(0.002)	(0.002)	(0.002)	(0.002)
$(LTV > 80\%)$ x ϕ_0					-0.002	-0.003
					(0.002)	(0.002)
Bought with			0.069	0.063	0.022	0.028
mortgage			(0.018)	(0.018)	(0.023)	(0.024)
$LTV > 80\%$					0.021	0.038
					(0.026)	(0.027)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	PC6, M		PC6, M		PC6, M	
N	1,021,092	1,021,092	757,813	757,813	512,485	512,485
Fixed effects	470,042	470,042	392,843	392,843	304,857	304,857
R-squared	0.969	0.968	0.973	0.972	0.977	0.977

Table 12: History dependence on selling probability given listing

Notes: The table shows the results from a linear probability model for the monthly selling probability of a property advertised on the Zoopla portal.

Dependent variable:	selling probability (q_{ijt}) *once listed*					
	Listed	Transacted	Listed	Transacted	Listed	Transacted
	Zoopla Sample 1		**Z. Sample 2**		**Z. Sample 3**	
	1995-2014		2002-2014		2005-2014	
$GAIN_t$	0.009	0.003	0.006	0.010	0.036	0.104
	(0.001)	(0.003)	(0.002)	(0.006)	(0.009)	(0.028)
Mortgage x $GAIN_t$			0.008	-0.003	0.060	-0.090
			(0.002)	(0.006)	(0.008)	(0.029)
($LTV > 80\%$) x $GAIN_t$					0.039	-0.004
					(0.010)	(0.032)
$LOSS_t$	-0.092	-0.113	-0.050	-0.090	-0.062	-0.083
	(0.004)	(0.010)	(0.007)	(0.022)	(0.008)	(0.027)
Mortgage x $LOSS_t$			-0.044	-0.026	-0.003	-0.022
			(0.008)	(0.024)	(0.009)	(0.030)
($LTV > 80\%$) x $LOSS_t$					-0.067	-0.069
					(0.010)	(0.032)
Market size when	-0.004	-0.005	-0.012	-0.008	-0.013	-0.003
bought (ϕ_0)	(0.000)	(0.001)	(0.001)	(0.002)	(0.001)	(0.003)
Mortgage x ϕ_0			0.009	0.004	0.011	-0.002
			(0.001)	(0.002)	(0.001)	(0.003)
($LTV > 80\%$) x ϕ_0					-0.001	-0.002
					(0.001)	(0.003)
Bought with			-0.110	-0.047	-0.131	0.014
mortgage			(0.009)	(0.026)	(0.011)	(0.034)
$LTV > 80\%$					0.012	0.025
					(0.013)	(0.038)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	PC6, M		PC6, M		PC6, M	
N	13,269,760	4,759,719	10,077,509	3,527,745	6,870,940	2,374,872
Fixed effects	667,789	470,042	592,673	392,843	495,001	304,857
R-squared	0.098	0.206	0.112	0.225	0.132	0.251

6 Conclusions

This paper investigates history dependence in the housing market using the universe of housing transactions in England and Wales in the last twenty years.

We show evidence consistent with anchoring and loss aversion affecting both house prices and selling probabilities. Our data allow us to separate properties which were bought with a mortgage and properties which were bought with cash. For a subsample of the data, we can also separate out properties which were bought with a high-LTV mortgage. While a large part of the effects we identify remain strong and significant for properties bought without a mortgage, we do find larger effects of the loss term for houses financed through a mortgage and in particular high-LTV ones, consistent with downpayment effects as in Stein (1995).

Thick market effects, whereby properties bought in a hot market are less likely to be sold because they consitute a better match with the tastes of the buyer, are not a significant determinant of house prices, on average, but appear to have a noticeable and significant effect on properties bought with cash. They also seem to have a small but noticeable effect on selling probability, which warrants further investigation.

In the last part of the paper, we merge our data with listings from a major UK property portal. We find that listing prices display more loss aversion than achieved prices. Similar, the probability of selling a property in each month subsequent to its appearance on the portal is much affected by losses but not so much by gains.

References

Anenberg, E. (2011). Loss aversion, equity constraints and seller behavior in the real estate market. *Regional Science and Urban Economics*, 41(1):67–76.

Beaudry, P. and DiNardo, J. (1991). The effect of implicit contracts on the movement of wages over the business cycle: Evidence from micro data. *Journal of Political Economy*, pages 665–688.

Beggs, A. and Graddy, K. (2009). Anchoring effects: Evidence from art auctions. *American Economic Review*, 99(3):1027–39.

Bracke, P. (2015). How much do investors pay for houses? Bank of England working papers 549, Bank of England.

Bracke, P., Pinchbeck, T., and Wyatt, J. (2014). The Time Value of Housing: Historical Evidence from London Residential Leases. SERC Discussion Papers 0168, Spatial Economics Research Centre, LSE.

Devereux, P. J. and Hart, R. A. (2007). The spot market matters: evidence on implicit contracts from britain. *Scottish Journal of Political Economy*, 54(5):661–683.

Engelhardt, G. V. (2003). Nominal loss aversion, housing equity constraints, and household mobility: evidence from the united states. *Journal of Urban Economics*, 53(1):171–195.

Ferreira, F., Gyourko, J., and Tracy, J. (2012). Housing busts and household mobility: an update. *Economic Policy Review*, (Nov):1–15.

Genesove, D. and Mayer, C. (2001). Loss aversion and seller behavior: Evidence from the housing market. *The Quarterly Journal of Economics*, 116(4):1233–1260.

Giglio, S., Maggiori, M., and Stroebel, J. (2015). Very Long-Run Discount Rates. *The Quarterly Journal of Economics*, forthcoming.

Grant, D. (2003). The effect of implicit contracts on the movement of wages over the business cycle: Evidence from the national longitudinal surveys. *Industrial & Labor Relations Review*, 56(3):393–408.

Hagedorn, M. and Manovskii, I. (2013). Job selection and wages over the business cycle. *The American Economic Review*, 103(2):771–803.

Kahn, L. B. (2010). The long-term labor market consequences of graduating from college in a bad economy. *Labour Economics*, 17(2):303–316.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, pages 263–291.

McDonald, J. T. and Worswick, C. (1999). Wages, implicit contracts, and the business cyle: Evidence from canadian micro data. *Journal of Political Economy*, 107(4):884–892.

Ngai, L. R. and Tenreyro, S. (2014). Hot and cold seasons in the housing market. *American Economic Review*, 104(12):3991–4026.

Office of Fair Trading (2010). Home buying and selling: A market study. Technical report.

Oreopoulos, P., von Wachter, T., and Heisz, A. (2012). The short-and long-term career effects of graduating in a recession. *American Economic Journal: Applied Economics*, 4(1):1–29.

Shan, H. (2011). The effect of capital gains taxation on home sales: Evidence from the Taxpayer Relief Act of 1997. *Journal of Public Economics*, 95(1-2):177–188.

Stein, J. C. (1995). Prices and Trading Volume in the Housing Market: A Model with Down-Payment Effects. *The Quarterly Journal of Economics*, 110(2):379–406.

Tversky, A. and Kahneman, D. (1982). Judgements of and by representativeness. In Kahneman, D., Slovic, P., and Tversky, A., editors, *Judgement under Uncertainty: Heuristics and Biases*. Cambridge University Press, Cambridge.

A Appendix

A.1 Robustness check for price regressions

Table A1: Price robustness regressions

Notes: This table replicates the analysis of Table 3 for two subsets of the data: flats for which we can use unit-level fixed effects and properties which were bought new.

Dependent variable:	Sale price (p_{ijt})				
	1995-2014			**2002-2014**	
	(1)	(2)	(3)	(4)	(5)
Previous aggr. prices (δ_0)	0.054	0.094			
	(0.002)	(0.003)			
$\delta_0 \times DUR_{it}$		-0.006			
		(0.000)			
$GAIN_t$			-0.056	-0.057	0.047
			(0.002)	(0.005)	(0.006)
Mortgage x $GAIN_t$					-0.139
					(0.006)
$LOSS_t$			0.145	0.104	0.182
			(0.015)	(0.020)	(0.028)
Mortgage x $LOSS_t$					-0.133
					(0.031)
Market size when bought (ϕ_0)			-0.005	0.002	0.001
			(0.001)	(0.002)	(0.002)
Bought with mortgage					0.039
					(0.001)
Controls	Yes	Yes	Yes	Yes	Yes
Fixed effects		Unit, M		Unit, M	
N	1,445,912	1,445,912	1,445,912	697,269	697,269
Fixed effects	908,911	908,911	908,911	542,321	542,321
R-squared	0.978	0.978	0.978	0.985	0.985

Dependent variable:	Sale price (p_{ijt})				
	1995-2014			**2002-2014**	
	(1)	(2)	(3)	(4)	(5)
Previous aggr. prices (δ_0)	0.104	0.111			
	(0.005)	(0.006)			
$\delta_0 \times DUR_{it}$		-0.002			
		(0.001)			
$GAIN_t$			-0.105	-0.133	-0.131
			(0.005)	(0.008)	(0.009)
Mortgage x $GAIN_t$					0.002
					(0.006)
$LOSS_t$			-0.071	-0.026	-0.137
			(0.025)	(0.027)	(0.042)
Mortgage x $LOSS_t$					0.154
					(0.045)
Market size when bought (ϕ_0)			0.006	0.004	0.004
			(0.002)	(0.003)	(0.003)
Bought with mortgage					0.013
					(0.002)
Controls	Yes	Yes	Yes	Yes	Yes
Fixed effects		PC6, M		PC6, M	
N	838,622	838,622	838,622	323,861	323,861
Fixed effects	154,024	154,024	154,024	77,185	77,185
R-squared	0.915	0.915	0.915	0.886	0.886

A.2 One-step regressions

An alternative specification is to ignore previous estimates of aggregate prices and use dummies for previous purchase years. The estimating equations become:

$$p_{ijt}(\text{or } q_{ijt}) = \alpha_j + X_i\beta + \delta_t + \sum_s \delta_{0s} + e_{it}.$$

The coefficients for the δ_{0s}, the previous purchase years, are shown in Figure A1 and A2.There are fewer observations in the last years so standard errors are wider. The cohort of 2007 achieved the highest prices.

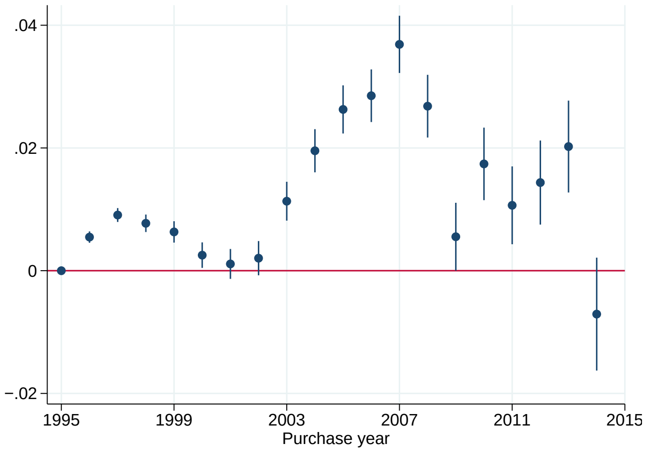

Figure A1: History dependence in prices: effect of year of purchase

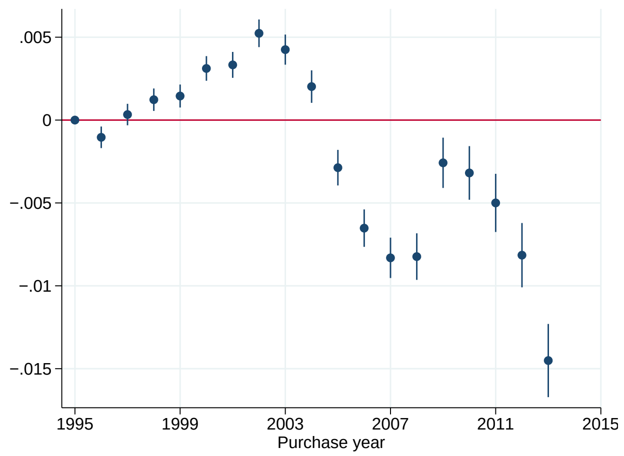

Figure A2: History dependence in quantities: effect of year of purchase

A.3 Additional figures

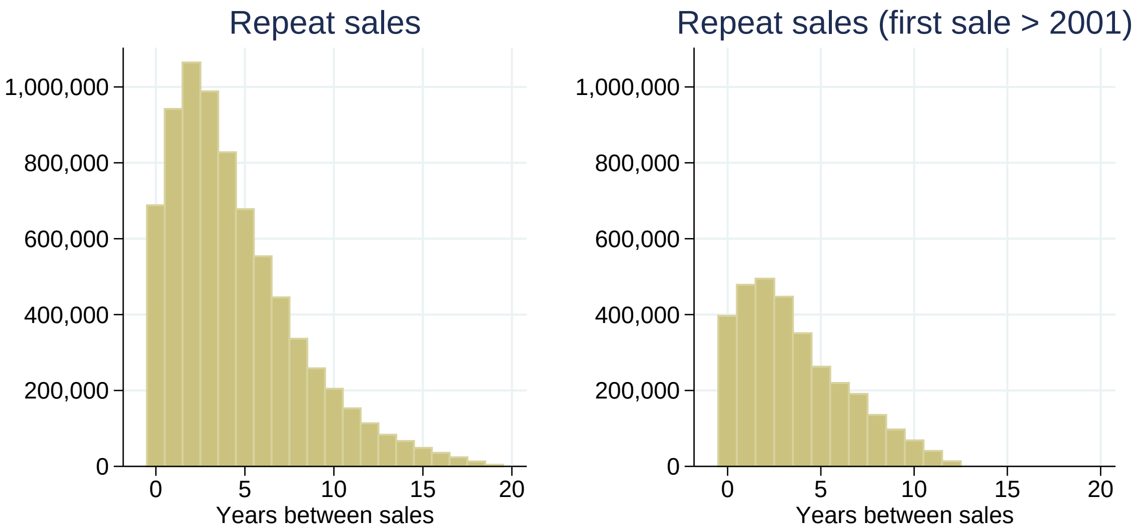

Figure A3: Years between sales

Notes: Since being part of the sample depends on whether there are at least two sales on that same property, the durations are skewed towards short intervals—a common feature of studies based on repeat sales.

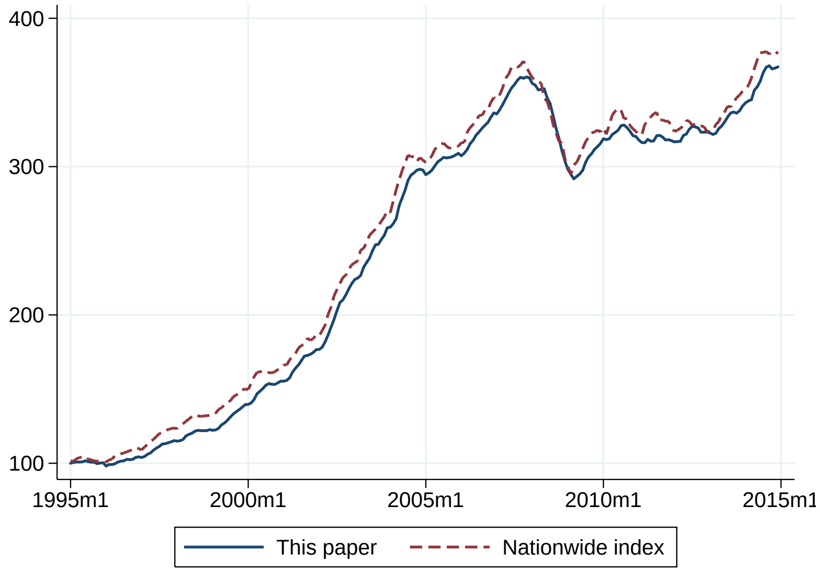

Figure A4: Aggregate house prices

Notes: Quality-adjusted aggregate house prices for this paper are estimated using 1995-2014 Land Registry data from England and Wales. The line in chart plots the exponentiated δ_t coefficients from the regression $p_{ijt} = \alpha_j + X_i\beta + \delta_t + e_{it}$, where α_j represents postcode fixed effects and X_i is a vector of housing characteristics included in the Land Registry: type of property, whether the property is new, and whether the property is sold under a leasehold arrangement.
The Nationwide quality-adjusted aggregate price are estimated using data on transactions funded with a mortgage from Nationwide. The series and the methodology are available at `http://www.nationwide.co.uk/about/house-price-index/download-data`.

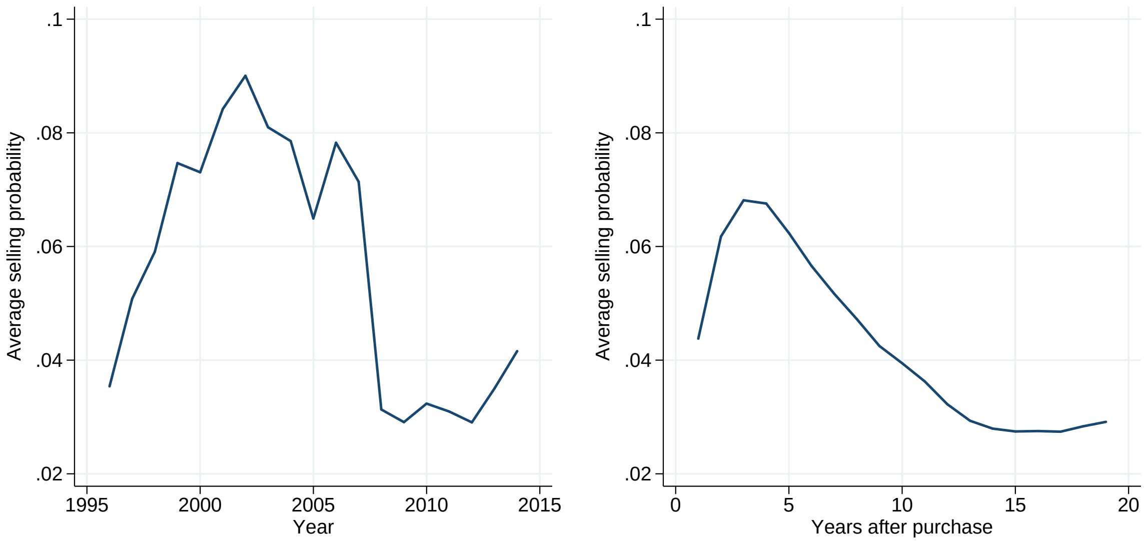

Figure A5: Aggregate selling probability

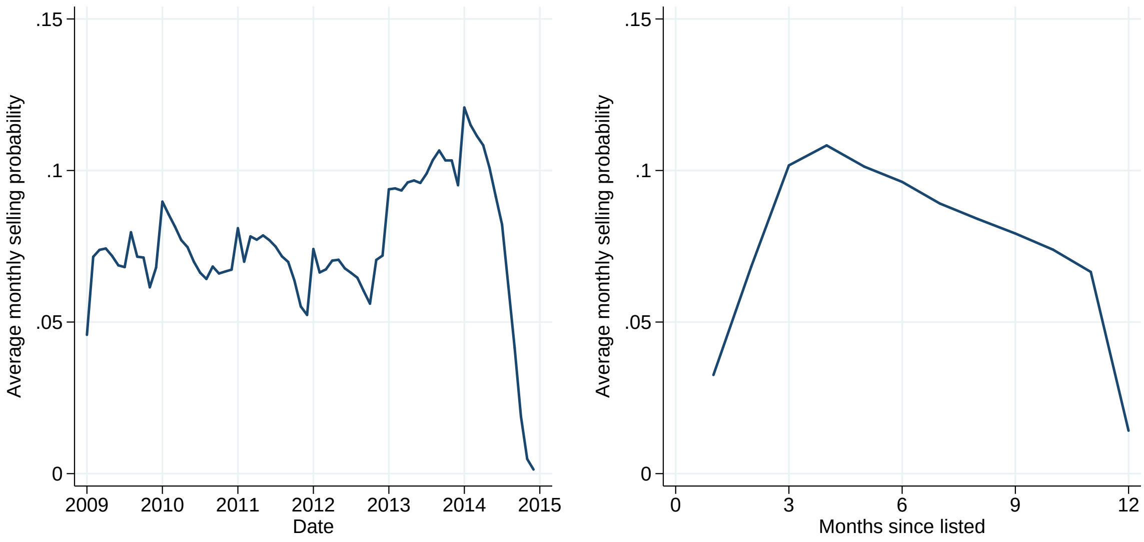

Figure A6: Selling probability after listing

A.4 Additional tables

Table A2: Preparing the data

Notes: This table shows how the sample size changes when moving from the original data source to the final sample used in the analysis.

	Sample size
WhenFresh/Zoopla dataset	12,942,865
Exclude rent listings	8,013,886
Exclude listings after 31/12/2014	7,307,582
Exclude listings in Scotland or Northern Ireland	7,001,252
Drop if listing creating date > listing deleting date	6,994,288
Drop if initial price missing	6,944,503
Duplicates in address if <180 days between listing dates (keep earliest)	4,538,231
Exclude first and last percentile of initial price	4,449,444
Can retrieve previous Land Reg sale	2,485,382