Peak-Bust Rental Spreads

Marco Giacoletti[*] Christopher A. Parsons[†]

University of Southern California University of Southern California

November, 2020

Landlords appear to use stale information when setting rents. Among over 43,000 California rental houses in 2018-2019, those last purchased during 2005-2007 (the peak) rent for 2-3% more than those purchased during 2008-2010 (bust). Neither house nor landlord characteristics explain this "peak-bust rental spread." To clarify the mechanism, we test cross-sectional predictions from a simple theory of rent-setting. We find empirical support for both *distorted beliefs* and *prospect theory*. In the first, past sales prices distort landlords' current estimates of house values/rents. In the second, monthly payments establish (recurring) reference points, against which gains or losses are measured.

Keywords: Distorted beliefs, prospect theory, liquidity constraints, residential rents, behavioral biases in real estate

JEL Classification: D40, G00, G40, R31

[*]Marshall School of Business, University of Southern California. **Email:** mgiacole@marshall.usc.edu

[†]Marshall School of Business, University of Southern California. **Email:** parsonsc@marshall.usc.edu

1 Introduction

Consider a landlord who has just received a notice to vacate from the existing resident. How does she determine what rent to ask of a new potential tenant? Like any retailer, she faces a trade-off: a higher rent offers more revenue in the event of a transaction, but due to uncertain demand, also increases the probability of the house not renting. The optimal rent balances these marginal effects.

However, a number of unique features of real estate imply potential departures from the fully rational benchmark described above. First, about half of U.S. rentals are owned by "mom and pop" investors who may lack the sophistication, information, or cognitive resources to set rents efficiently.[1] Second, since information on the cross-section of prices is not readily accessible, and houses are heterogeneous, landlords prone to valuation errors, or who use non-informative heuristics, may have a hard time realizing any errors they might make.[2] Finally, unlike financial markets, in which rational and informed traders can bet against others' mistakes, real estate markets are comparatively inefficient. Segmentation, high transactions costs, and the inability to short-sell limit arbitrage opportunities.

In this paper, we explore whether a landlord's exposure to stale information — aggregate house prices over a decade ago — influences the rent she sets today.[3] Among online rental listings for over 43,000 houses in California from December 2018 to March 2019, we find that landlords having last purchased their houses at the peak of the real estate boom (2005-2007) charge rents 2-3% higher than landlords who acquired their houses in the ensuing bust (2008-2010). We refer to this as the *peak-bust rental spread*. An alternative way of estimating

[1] https://www.huduser.gov/portal/pdredge/pdr-edge-frm-asst-sec-061118.html

[2] Previous work has shown that real estate investors may be prone to valuation errors. For example, they may form expectations on future prices based on the over-extrapolation of past trends (see Cheng, Raina, and Xiong, 2014, Glaeser and Nathanson, 2017, and Gao, Sockin, and Xiong, 2019), they may overweight irrelevant information provided by individuals in their social network (see Bailey, Cao, Kuchler, Stroebel, and Wong, 2018), or they may just not be fully informed on the quality of individual houses or neighborhoods (see Stroebel and Kurlat, 2015).

[3] Personal experiences have been linked to a wide variety of financial/economic decisions and outcomes. See Nagel and Malmenadier (2011, 2016), and Bernile, Bhagwat, and Raghavendra (2017).

the effect is to regress rent on the logarithm of the Case-Shiller house price index at the time of last purchase. The estimated sensitivity is 0.054 (with a t-stat of approximately 10.5), suggesting that for run ups like 2001-2006 (100%), cross-sectional rents would differ by about 5%, even after more than a decade.

Two immediate questions arise. The first is whether historical house prices really are stale, or whether they tell us about current house quality, services provided by landlords, or other attributes that matter to tenants. The second regards the mechanism. If the timing of historical purchases is orthogonal to rental services, why do landlords appear to take it into account at all?

Starting with the first question, a key aspect of our empirical design is that the measure of staleness does not use the actual historical sales price of individual houses. In a seminal paper studying loss aversion among condominium sellers in Boston, Genesove and Mayer (2001) describe the econometric challenges associated with using individual purchase prices as reference points against which gains and losses are measured. In particular, sales prices partly reflect durable, but unobservable aspects of housing quality. Thus, were we to estimate a cross-sectional regression of current (t) house (i) rents $R_{i,t}$ on observable characteristics $X_{i,t}$ and historical ($\tau < t$) sales prices $p_{i,\tau}$, the concern would be that past sales prices may proxy for street noise, views, micro-environment, or other attributes not captured by hedonic characteristics X.[4]

In hopes of minimizing this issue, our analysis features cross-sectional regressions of rents on a sequence of "acquisition vintage" dummy variables, or on house price index values at the time of acquisition (in addition to standard hedonic attributes), and thus exploits only fluctuations in house price *index values* as our source of variation in purchase prices. While still possible that residual differences in unobserved house or landlord attributes exist across

[4]Genesove and Mayer (2001) characterize both the size and direction of the bias introduced by unobserved house characteristics in estimates of loss aversion among sellers. Though simulations suggest that the absolute size of such bias is small, their analysis also accounts for unobserved quality by including past sales residuals (in logarithms) in regressions of current asking prices. For robustness, we also employ this approach (see Table 4).

vintages, the size of any bias is limited to that caused by unobserved heterogeneity between large, diversified groups containing thousands of houses.

The most direct evidence against vintage-level differences in landlord or house quality is that the peak-bust rental spread is not sensitive to controls for either, despite a large increase in explanatory power. For example, with only zip code fixed effects ($R^2 = 63\%$), regressing rents on the (logarithm of) the house price index value at time of purchase gives a coefficient of 0.042. However, controlling for various property and landlord characteristics not only sharply increases the R^2 to 87%, but also increases the coefficient on the lagged index (0.054). Oster (2019), building upon Altonji, Elder, and Taber (2005), describes how the comparison of both coefficients and R^2 can be used to place bounds on the size of remaining omitted variable bias.[5] Applying the test to our setting, to explain the relation between rents and lagged property index values, a somewhat pathological correlation structure is needed: rents would have to exhibit a sensitivity to unobservables over eight times higher to that involving observables, and with the opposite sign.

Given this, it is perhaps unsurprising that when we compare observable house and/or landlord characteristics across acquisition vintages, either minimal differences are observed (houses), or when more significant differences are observed (landlords), they go in the wrong direction. Indeed, for the latter case, we do find that the composition of landlords is strongly related to acquisition vintage, with bust-acquired houses being more likely to be purchased by corporate entities and/or professional investors. However, this variation works against the result, because all else equal, investors tend to ask higher, rather than lower, rents. In any case, the magnitude of any effect is negligible; controlling for landlord type has almost no impact on our main estimate. Whether comparing between the peak- and bust-vintages, or across all years using the Case-Shiller index, the estimated magnitude is nearly identical

[5]The procedure compares the coefficient(s) of interest — here, the difference in average rents across acquisition vintages — when increasingly informative sets of controls are added to the regression. This cross-regression sensitivity to observables (i.e., how much the estimated coefficient declines as the R^2 increases) is then used to place bounds on the true, bias-free effect. Intuitively, the more stable the estimated coefficient, particularly with a substantial increase in R^2, the less likely it is that unobservables represent a significant source of bias.

whether we control for landlord type or not.

To further address heterogeneity in landlords' quality or preferences (e.g., discount rates or risk tolerance), we conduct a placebo test using Texas. Although data on landlords is sparser, price-to-rent ratios were and remain much lower in Texas rather than in California, and thus, would likely have been equally (or more) attractive to investors "reaching for yield" in the wake of the crisis. Critically however, Texas did not experience the same house price volatility as did California through the early 2000s (see Figure 4). Finding a complete absence of a comparable peak-bust rental spread thus suggests that realized historical price volatility, rather than the changing composition of investors, is the key determinant of the observed patterns.

A final test of unobserved quality compares time on market between houses purchased during different times. To the extent that acquisition years distort asking rents relative to the quality of rental services, we would also expect that landlords influenced by high (low) historical purchase prices would have a lower (higher) probability of renting their property. Indeed, when we test this empirically, we find that houses acquired at the peak sit in inventory about 6% longer than those acquired during the bust.

The remainder of the paper addresses possible mechanisms. We explore two possibilities. The first is via *distorted beliefs*. Here, the idea is that landlords may not be perfectly informed about the quality of their houses, and thus, may use historical prices to improve their estimates. To see how distorted beliefs may arise, note that prices can vary due to either: 1) discount and capitalization rates, which primarily describe time-series changes, and 2) fundamentals impacting cash flows, which primarily describe the cross-section. One way that landlords may hold incorrect beliefs is by underweighting the importance of discount rate variation, which will cause them to overweight the importance of prices as signals of quality. To see this, suppose that discount rates in the housing sector are 5% at date t_1, and 10% at date t_2. A landlord that underestimates this variation – say, using 6% and 9% respectively – will mistakenly believe that a house purchased at t_1 (t_2) has better (worse) fundamentals than it actually does.

Another reason that stale signals may enter into landlords' rent-setting decisions is through *reference-dependent preferences*, whereby landlords face monetary or psychological costs when rent falls below a reference level. The importance of reference dependence in real estate has been demonstrated in the sales market. Genesove and Mayer (2001) find that sellers are reluctant to set asking prices below their original purchase prices. As the authors discuss, this can be reconciled with loss aversion (i.e., psychological costs from incurring nominal losses), as well as liquidity constraints, since cash-poor sellers may not have sufficient reserves to pay off an existing loan, and finance a down payment on a new house (Stein, 1995, Genesove and Mayer, 1997).

Exploring these issues in the rental market offers two additional contributions. First, it provides a complementary way to appreciate the findings in Birru (2015), which shows that when the consecutive transaction is of the same type (e.g., buying and then selling a stock) transforming the reference point – as with a stock split in equity markets – eliminates the disposition effect entirely. Our results suggest that when the transaction type *differs* (e.g., buying and then renting a house), the agent may regard the transformed reference point as relevant for calculating gains and losses in the new context. For example, the landlord's monthly payment is a natural, recurring benchmark,[6] raising the possibility that past transactions may exert a nearly continuous influence on current rental markets, in addition to their punctuated impact on sales.

Second, the nature of landlords' payoffs in the rental versus selling market allows for a sharper distinction between loss aversion and liquidity motives. Although both mechanisms make the same prediction for sales (i.e., buying for a high price means trying to sell for a high price), this is not necessarily the case when renting. A key reason for this distinction is

[6]The sum of a landlord's expenses will depend on the purchase price via mortgage installments and, due to peculiarities in the California tax code, property taxes. In California, property taxes are not marked to market annually, but per Proposition 13 of 1978, are instead limited to appreciate no more than 2% annually. Consequently, current taxes depend on historical prices, and given volatility in aggregate prices, the time since purchase.

that while rental houses are purely financial assets, owner-occupied homes provide additional sources of consumption utility. These considerations, as described by Stein (1995) and Andersen, Badarinza, Liu, Marx, and Ramadorai (2019), can alter sellers' utility *when a transaction occurs*, such as having to downsize if a sale fails to generate sufficient cash for a down payment. For a landlord however, the reverse pattern is likely more realistic. *When a transaction does not occur*, the full costs of mortgage service, taxes, maintenance, and other expenses must be funded using the landlord's cash reserves. If the costs of "funding" the property are convex, – e.g., if paying \$1,000 to maintain a vacant house is more than twice as painful as paying \$500 – then the landlord's expected utility is particularly sensitive to the house not renting, which causes her to ask less, and increase the probability of a transaction.

To formalize these effects, we develop a highly stylized model of rent-setting, adapted from Lazear (1986). In a benchmark case with no reference dependence, the model captures a key trade-off faced by landlords: asking a higher rent increases revenue if the house rents, but also decreases the chance that it does. This trade-off is altered when we allow the landlord's utility to also depend on a reference point, C. When rent drops below this threshold, the landlord experiences a loss, which increases as the gap between C and rent widens. Critically, the shape of the loss function determines whether a high reference point causes the landlord to set a rent that is higher, or lower, than in the benchmark case.

Under classic prospect theory, the function is steepest just below C, and the landlord is most concerned about maximizing rent conditional on the house renting. However, if the loss function gets steeper as the gap between C and rent increases, the landlord is most concerned about maximizing the chance that the house rents. Consequently, loss-averse landlords set rents above the benchmark case — a specific manifestation of prospect theory's general prediction of risk-seeking in the domain of losses — whereas liquidity-constrained landlords choose lower rents, attempting to hedge against an even larger loss incurred in the no-rent state.

In the final part of the paper, we design empirical tests intended to identify which among the different economic mechanisms are most responsible for the observed patterns: distorted beliefs, preferences based on prospect theory, or liquidity constraints.

Starting with distorted beliefs, the reliance on historical prices should be weaker when landlords are more informed about their houses' values. We test this prediction in the data by comparing listings that differ in terms of the availability of similar properties, such as those matched on size or number of bedrooms. Intuitively, a landlord whose house is relatively unique can extract only minimal information from nearby active rental listings, or from recent rentals or sales, thus leading her to rely more heavily on past prices. In the cross-section, we indeed find that the sensitivity between rents and historical price indexes is stronger for houses with fewer active nearby (within one mile) comparable listings. Moreover, the sensitivity is stronger also for properties that have unique characteristics with respect to other houses in their zip code, currently listed or not.

Isolating the effects of reference dependence is complicated by the fact that historical purchase prices may simultaneously influence reference points (e.g., through higher monthly payments) and landlords' beliefs. In hopes of more cleanly making this distinction, we compare landlords that choose different leverage at time-of-purchase, while controlling for purchase price. Our main comparison is between two groups of landlords: 1) those employing little to no leverage, with loan-to-value (LTV) ratios below 50%, and 2) those using leverage aggressively (over 95% LTV). For the first group, the sum of all monthly obligations would typically be lower than market rents, and thus, the loss function would play little, if any, role on rent-setting. For the second, back-of-the-envelope calculations indicate that most landlords would face paper losses on a monthly basis.[7] As described above, prospect theory predicts higher rents for the high-LTV group, while the liquidity constraints explanation predicts lower rents. No difference would suggest that reference points are irrelevant.

We find that, after controlling for house and landlord quality, landlords with little to no

[7]We lack further data on landlords' monthly expenses, and therefore cannot precisely measure gains and losses for individual landlords. To the extent that the high-LTV group contains landlords not experiencing losses, the estimated magnitudes will be biased downward.

leverage ask almost 5% less than landlords in the high-LTV group, which is consistent with the predictions of prospect theory. As further evidence against liquidity constraints, we show that the peak-bust spread is present also for properties that are more likely to be held by deep pocketed investors, for whom liquidity concerns are likely less important. For this test, we split our sample based on zip code-level income and/or the level of monthly rent, finding a peak-bust rental spread at least as large in wealthy areas and for expensive houses, which is inconsistent with liquidity constraints playing a primary role.

Combining these observations with the prior tests, we interpret the overall evidence as suggesting that both distorted beliefs and prospect theory contribute to the observed patterns. Because the tests regarding each mechanism exploit different, and in some cases orthogonal,[8] sources of variation, both could be relevant for a given landlord. That is, whereas a low-LTV landlord having purchased for a high price might ask a higher rent due only to optimistic beliefs, a high-LTV landlord also purchasing for a high price will charge even more. Empirically, we find that these effects are nearly independent, as reflected in each marginal effect being almost unchanged in the presence of the other.

The rest of the paper is summarized in Figure 1. Section 2 describes the data, and characterizes the main results on the peak-bust rental spread. With the basic pattern established, we then address whether fundamental (Section 3) or behavioral factors (Section 4) explain our findings. In the fundamental (left) branch, we consider heterogeneity in both houses, landlords, and unobservable characteristics. We reference the specific tests we perform in colored blocks. In the behavioral (right) branch, we first present a simple model (Section 4.1) that allows for acquisition vintages to influence rents through their impact on beliefs or preferences. We then return to the data in Sections 4.2 and 4.3, and design tests intended to disentangle the role played, respectively, by biased beliefs and preferences. Section 5 concludes.

[8]For example, some of our tests for information quality exploit differences across zip codes, whereas the LTV tests for reference points are within zip codes.

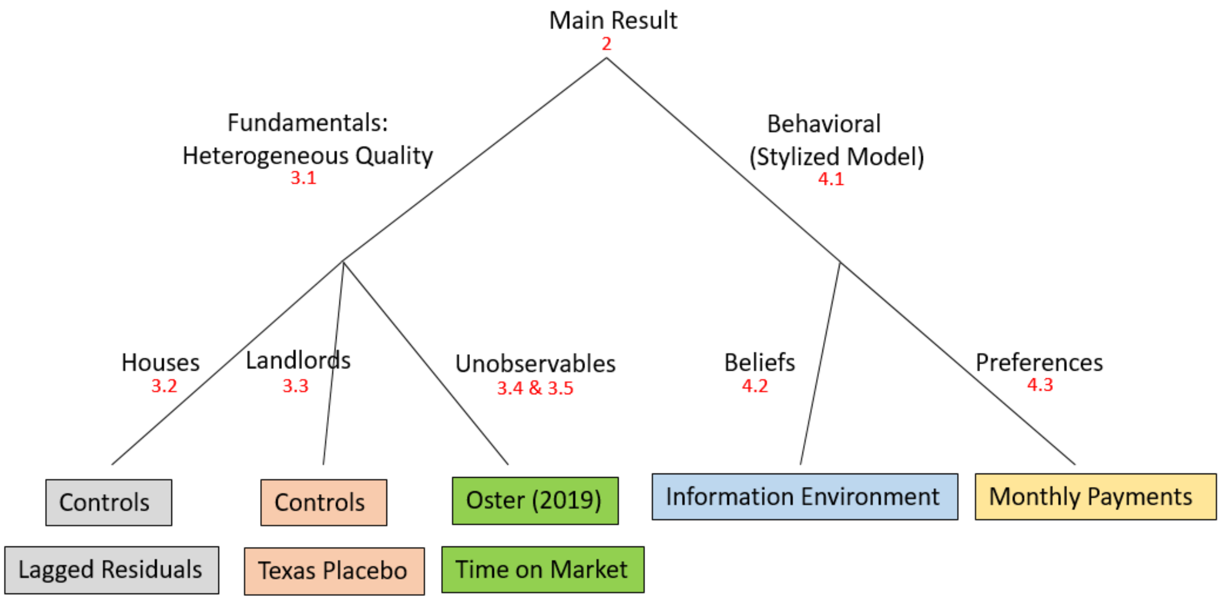

Figure 1: Roadmap of the paper.

2 Historical Purchase Timing and Current Rents

2.1 Data

Our dataset is constructed by collecting rental listings from the website of a major online rental listings service. The final dataset of advertised rents is based on listings for houses located in the state of California, collected from December 2018 through March 2019. We exclude from the dataset houses and apartments for which the number of bedrooms, the number of bathrooms, size or information on the last sale date are not available. We also exclude houses last purchased before 1980, houses that have a last purchase date more recent than their construction date (presumably a data error), and listings that have been online for more than 300 days at the time of collection.[9]

Our main dataset consists of 44,237 rental listings which, in our analysis, we pool and treat

[9]We collect information on listings at two weeks intervals. The filters reduce the total number of observations (all appearances of listings across all collection waves) from 217,166 to 88,611. Then, in our main analysis, when the same listing appears multiple times across different collection waves, we store in the dataset a single observation based on the most recent information collected. This reduces the sample to 44,237 unique listings.

as a single cross-section. Figure 1 shows the geographic distribution of these listings across California zip codes, while Table 1 reports summary statistics of the data. The majority of the properties are single family residencies (SFRs). Roughly 17.7% of properties are condo or apartments, 7.7% are townhouses, and 5.5% are multi-family residencies. Less than 1% are studios. Slightly less than 40% of listings are posted by real estate agents, suggesting that roughly 60% correspond to smaller, non-institutional investors, consistent with the overall fraction of mom-and-pop investors observed in national surveys. At the time of data collection, on average, a typical listing has been online for 39 days (median 23 days). The average monthly rent in the dataset is slightly above $3,500 (median $2,750), with a standard deviation of $3,400.

In our dataset rent is measured as the latest available listed rent for each property. Unlike property sale agreements (or sale deeds), rental agreements are not subject to official record keeping by county or city registries. Thus, there are no centralized repositories of contractual prices that can be accessed by practitioners and researchers. Previous academic papers studying the rental market have relied either on market-level estimates for large geographical areas (that is, country-level or regional-level indices, as in Jorda, Knoll, Kuvshinov, Schularick, and Taylor (2019)), or on a combination of listed rents and surveys (see Genesove (1999), Gilbukh, Haughwout, and Tracy (2017), Demers and Eisfeldt (2018) and Begley, Loewenstein, and Willen (2019)), or on estimates reported by owners or other market participants (Eichholtz, Korevaar, Lindenthal, and Tallec (2020)). The same is true for practitioners. For instance, the zip code-level rental indices published by Zillow (Zillow Rental Indexes) use listed rents as their main data input.[10]

If there are off-market negotiations between tenants and landlords, the latest listed rent and the actual contractual rent paid by the tenant might differ. However, we believe this is unlikely to substantially affect our results. As we explain in detail in the following sections, our analysis relies on the comparison of average rents across different purchase vintages, so

[10]See the *Methodology* description published at https://www.zillow.com/research/zillow-rent-index-methodology-2393/

any difference between listed rents and contractual rents would affect our results only to the extent that it systematically varies across purchase vintages.

Section 2.2 provides direct evidence suggesting that this is not the case. While the likelihood and impact of off-market negotiations should be increasing in the bargaining power of tenants, we find that the magnitude of our estimates of the peak-bust rental spread is not affected by the degree of bargaining power of tenants in the local zip code. Moreover, as we show in section 3.5, purchase vintages do not only affect listed rents, but also time on the market, since peak-purchased houses remain listed online longer than houses purchased during the bust. Thus, to match the evidence we present in the data, not only landlords who purchased during the peak would have to be more willing to lower rents in off-market negotiations, but they would also have to be unwilling to lower listed rents ahead of off-market negotiations, thus letting their houses sit on the market for a longer period of time.

2.2 Empirical Estimation

Our key empirical exercise consists of comparing otherwise similar houses purchased at different points during the peak-bust cycle that occurred from the mid 2000s through the early 2010s. Figure 2 shows some preliminary patterns. For 2-bedroom homes (single family residences, townhouses or apartments) in the city of Los Angeles, we plot the average advertised rent/square foot for a sequence of six, non-overlapping acquisition vintages: 2002-2004, 2005-2007, and so on, until 2017-2019. The blue bars report average rent per square foot for each purchase vintage, and for comparison, the red line is the average historical value of the S&P Case-Shiller index value for Los Angeles in the years of each purchase vintage. The correlation is easily apparent, with rents being higher for houses purchased at the peak (2005-2007) and/or during the recovery (after 2013), and lower for houses purchased during the bust (2008-2013).[11]

[11]This effect is across rental properties purchased at different points in time within the same metropolitan area, and is different than the "memory" effects that have been observed among renters by Bordalo, Gennaioli, and Shleifer (2019), who show that tenants moving to a less (more) expensive metropolitan area tend to rent

An immediate concern is that houses acquired at different times may differ in terms of quality, location, services, or other relevant attributes. We thus generalize the patterns above in a regression framework that includes controls for various dimensions of house quality:

$$\log(R_i) = \Gamma I_{p,i} + \mathcal{B}_{ctrl} X_{ctrl,i} + a_z + e_i, \tag{1}$$

where R_i is the monthly rent for listing i, $I_{p,i}$ is a vector of acquisition vintage dummies, $X_{ctrl,i}$ is a vector of house i's characteristics, and a_z is a family of zip code fixed effects. To keep the number of categories manageable (as we will report the coefficients for each vintage in our tables), and to focus more precisely on the specific years of interest, all properties acquired prior to 1990 are lumped together, as are properties purchased from 1990-1994 and 1995-1998. Starting in 1999, we group vintages into 3-year increments: 1999-2001, 2002-2004, and continuing through 2017-2019. The latter group is the reference category in our estimates.

The set of house characteristics, $X_{ctrl,i}$, is extensive, and includes dummies for the number of bedrooms and bathrooms in the property, as well as indicators for properties that offer only street parking (no garage or parking slot), for properties that have shared laundry, for townhouses, condos or apartments and studios. Other characteristics include the log square feet size of the property, age and age squared, and dummies for rental properties that forbid pets, that do not have air conditioning, and that provide a refrigerator, a dishwasher, hardwood flooring, and/or forced air heating and central air conditioning. For condos and apartments, we also include the floor on which the apartment is located. In some specifications, we also control for the log number of days the property has been listed online (log number of days-in-inventory), as well as whether the listing was posted by a real estate agent.

Columns 1, 2, 3 and 4 of Table 2 report parameter estimates for different specifications of Equation (1), depending on which control variables are included. Starting with the first column, the difference between peak (2005-2007) and bust (2008-2010) vintage is $5.5 - 2.8 = 2.7\%$, and

more (less) expensive apartments in their new locations.

$5 - 2.8 = 2.2\%$ when compared to the subsequent three years (2011-2013). When including the full set of controls for house characteristics, the peak bust difference is $5.2 - 2.8 = 2.4\%$, while the difference with respect to 2011-2013 is $4.4 - 2.8 = 1.6\%$. Across all specifications, the differences are statistically significant at the 1% level.[12]

Relative to the 2017-2019 acquisition vintage, individual estimates of Γ are negative and statistically significant, indicating that houses most recently acquired rent for the highest amounts. One potential explanation for this is that purchase dates may correspond with renovations, so that houses acquired more recently may have more, and/or more contemporary, updates not captured by $X_{ctrl,i}$. Consequently, if the relationship between current rents and time since last purchase were monotonic, it would be difficult to disentangle the "stale price" effect from that related to recent renovations.

The non-monotonicity of real estate prices in California is what makes it a near-ideal setting for making this distinction. Because prices dropped so precipitously from 2008-2010 compared to the three years prior (2005-2007), comparing rents between these particular vintages represents a lower bound on the stale-price effect, since the older vintage is also the most (historically) expensive. Accordingly, it is perhaps not surprising that if we control explicitly for the time elapsed since last purchase, the estimate of the peak-bust rental spread – the difference in the point estimates between the 2005-07 and 2008-10 vintages – increases slightly to 3.5%.

Figure 3 provides a graphical representation of the vintage-to-vintage patterns, plotting the individual elements of Γ based on the fully controlled specification in column 4 of Table 2 (top panel), against changes in the historical price index over the same years (bottom panel).[13] When compared, the two panels of Figure 3 show a clear positive co-movement.

[12]The reader might be concerned that part of the difference between the 2005 to 2007 and the 2008 to 2010 dummy is driven by houses that last transacted in 2008. At that point, the housing market had not yet reached its bottom, and the purchase prices for some rental properties might still have been (relatively) high. In Table A.1 in Appendix A we repeat the analysis excluding all properties last sold in 2008. The results are virtually identical to the ones reported in Table 2.

[13]Whereas Figure 2 uses the Case-Shiller Index for Los Angeles, Figure 3 calculates a similar index, replacing the dependent variable (log rent) with the log purchase price of each listed property i in Equation

Recalling that the data reflect asked, not necessarily contracted, rents, one concern is that off-market negotiations – particularly peak-acquiring landlords being forced into accepting lower contractual rents – may cause the actual effects to be smaller than what implied by our estimates. Although, as explained in section 2.1, we generally believe that such concessions were unusual in California during our sample period, in column 5 of Table 2 we repeat the full specification from column 4, but restricting the sample to houses that were listed in zip codes where listed properties cleared the market quickly. In particular, we restrict the sample to zip codes where the average number of days listed properties spend online is below the median across California zip codes in our dataset. While in these zip codes landlords should be marginally less willing to negotiate for reductions in rents, estimates of the peak-bust rental spread in column 5 are nearly identical compared to the ones from the full sample reported in column 4.

The last two columns of Table 2 show the results when, instead of using vintage dummies to capture extreme highs and lows of the historical price index at time-of-purchase, we use these index values directly. In column 6, the acquisition dummy indicators are replaced with a single variable, the logarithm of $p_{last,zip}$, which corresponds to zip code-specific price indices (ZHVI) published and maintained by Zillow.[14] Column 7 uses the logarithm of $p_{last,CA}$, the average of the Case-Shiller values for San Diego, Los Angeles and San Francisco from January 1987 through June 2019. In both cases, with zip code fixed effects, the coefficients can be interpreted as within-zip elasticities, i.e., the percent rent by which two houses in the same zip code would differ on average, for every percent change in the price index prevailing when they were respectively acquired.

Interestingly, the coefficients in columns 6 and 7 are almost identical, with an implied

(1). This allows us to use the entire sample, including houses not in major metropolitan areas, for which Case-Shiller estimates are not generated. Parameter estimates are reported in Table A.2 in nominal (columns 1 and 3) and real (columns 2 and 4) terms, the latter using December 2018 as the reference date. The figure looks nearly identical if we use an equal- or weighted-average of Case-Shiller index values for all major cities in California.

[14] ZVHI values are based on hedonic, proprietary machine-learning algorithms. The reported frequency is monthly, and because ZHVI is not uniformly populated, coverage can vary by zip code. The longest available series start in April 1996.

elasticity of 0.054, and t-statistics exceeding ten. Observing that average prices in California dropped by 25%-40% (depending on location) from the peak in Summer 2006 through the bottom in Fall 2011, this is consistent with the estimated peak-bust rental spread of 2.4%, which implies that 5-10% of historical price fluctuations remain imprinted in current market rents, even a decade later. Although the rental market is our primary interest, Table A.3 in the Appendix repeats the analysis among a sample of house *sales*, using actual close prices.[15] Here too, we observe a nearly identical sensitivity to past index prices, suggesting again similar dynamics between asked and contractual prices.

3 Heterogeneous Quality

An important concern is that the cross-sectional patterns we document might be driven, at least in part, by unobserved heterogeneity across housing units. If either house quality or services vary in the same way as do fluctuations in historical prices, beyond what is captured by controls, then the relation between stale prices and rents will be spurious. We begin in Section 3.1, describing the key identification problem, as well as how our main source of price variation (acquisition years) offers important advantages for minimizing the concern. The following sections then present additional tests intended to address remaining observable and unobservable heterogeneity in house quality and landlords/services quality.

3.1 Why Acquisition Vintages Rather than Historical Prices?

Recall that our key empirical strategy, shown in Equation (1), relates asked rents to the year(s) in which a property was purchased, not to the original purchase price. By focusing on fluctuations in past index values, our hope is to isolate price variation largely orthogonal to unmeasured quality attributes, and thus identify an effect on rent-setting that is not due

[15] We use data from Corelogic on resales of single family residences in California over the period from January 2017 to December 2018.

to omitted heterogeneity in quality.

To see the advantage of using index values rather than historical prices, note that rental houses are financial assets, which can be valued using standard techniques. At any time t, the price of financial asset i, $p_{i,t}$, is the sum of its expected future cash flows $C_{i,t+1}$, $C_{i,t+2}$,..., each deflated by a discount rate r_{t+1}, r_{t+2},... that compensates investors for time preferences and risk. For simplicity, assume that, at time t, prices are formed with the expectation of a constant growth rate in cash flows, $\frac{C_{t+1}}{C_t} = \bar{g}_t$, and a constant discount rate $\bar{r}_t$. With these assumptions, the price of housing unit i at time t can be written as

$$p_{i,t} = \frac{E[C_{i,t+1}]}{\bar{r}_t - \bar{g}_t} = \frac{\rho E[R_{i,t+1}]}{\bar{r}_t - \bar{g}_t}, \tag{2}$$

where $\bar{r}_t$ and $\bar{g}_t$ are the discount and growth rates prevailing at time t, which are assumed to be identical for all houses in the cross-section. While possible that financial risk and growth rates may differ across the state of California, most of this variation should be captured by zip code fixed effects, leaving only risk differences between houses within a few miles of each other, which we presume to be small.[16]

$R_{i,t+1}$ is the expected rent, of which fraction $1 - \rho$ is lost to taxes and other costs. Let $log\,(R_{i,t+1}) = \phi_{i,t} + \beta_t X_{i,t}$, where $\beta_t X_{i,t}$ represents a mapping of observable characteristics onto the flow of rental services, and $\phi_{i,t}$ a multiplier representing quality attributes not spanned by $X_{i,t}$. Rearranging terms, we have:

$$\epsilon_{i,t} = log(p_{i,t}) - \beta_t X_{i,t} = log(\rho) + \phi_{i,t} - log(\bar{r}_t - \bar{g}_t). \tag{3}$$

The left hand side represents the pricing error from a hedonic regression of sales prices on

[16]For example, it is plausible that house values are more sensitive to the macroeconomy in large cities. Standard asset pricing models would indicate that investors demand a risk premium for holding assets with such higher exposure, and should be reflected in the discount rate. For waht concerns growth rates in rents, numerous studies find that they are small and stable, on the order of 0.5-1% above inflation. See Shiller (2006), Campbell, Morris, Gallin, and Martin (2009), Giglio, Maggiori, and Stroebel (2015) and Giglio, Maggiori, and Stroebel (2016).

observable house characteristics. These residuals can be attributed to differences in taxes and/or operating efficiency (ρ), unobservable quality (ϕ), and capitalization ("cap") rates ($\bar{r} - \bar{g}$).[17]

Equation (3) clarifies the reason why including the historical price of house i, $p_{i,\tau<t}$, in a time-t cross-sectional regression of rents is potentially problematic. Because $p_{i,\tau}$ is a function of $\phi_{i,\tau}$, a positive correlation with $\phi_{i,t}$ means that variation in $p_{i,\tau}$ will also capture variation in current quality, invalidating the interpretation of $p_{i,\tau}$ as a measure of staleness. Fortunately, the same decomposition also indicates a natural solution.

Since the nature of the identification problem is *cross-sectional* — house j has a higher historical price residual than house i because it is better in some way — achieving identification from the *time series* offers an appealing alternative. As indicated by Equation (3), fluctuations in capitalization rates $(r - g)$ induce fluctuations in asset prices, even when the house's cash flow characteristics, particularly ϕ, do not change. Indeed, an important goal of commercially published real estate indices, such as Case-Shiller or Zillow's "Home Value Index" (ZHVI), is to account for both observable and unobservable dimensions of house quality, and thus capture time-series variation in capitalization rates, rather than compositional changes.[18]

Accordingly, in hopes of breaking the linkage between historical prices and unobservable quality, our analysis focuses on historical fluctuations in index prices. Since such fluctuations exhibit such a distinct temporal signature in California, they can also be inferred from examining year-of-purchase directly. Indeed, recall from the last two columns of Table 2 that the estimated rent-index sensitivities are nearly identical to those implied by the cross-vintage comparisons.

Although vintage-level unobservable quality may vary, a number of factors suggests that

[17] In real estate, it is common to quote prices as multiples of *net operating income (NOI)*, which approximate the cash flows available to equity investors (debt costs are not part of NOI). The ratio $\frac{NOI}{p_t}$ is equal to the capitalization rate, or "cap rate." Whereas NOI does not perfectly correspond to free cash flow (the numerator in Equation (2)) — for example ignoring investment costs — this discrepancy makes little difference for the discussion here.

[18] The Case-Shiller index is constructed from repeated sales. Zillow's proprietary algorithm is inherently cross-sectional, and may include measures of quality not incorporated in standard hedonic analysis.

such heterogeneity is likely too small to meaningfully affect our results. Most importantly, the main source of variation is limited to that differing, on average, between acquisition vintages containing thousands of houses each. Moreover, for our estimates to be biased upward, any such vintage-level average differences must be negatively correlated with the capitalization rate prevailing at time t, in order to generate a positive correlation with historical prices. The following sections show that accounting for cross-vintage heterogeneity usually goes in the opposite direction than what would be needed to account for the peak-bust rental spread in the data. Consequently, omitted house (or landlord) characteristics do not appear capable of providing a credible alternative to the interpretation of stale prices impacting current rents.

3.2 House Heterogeneity

3.2.1 Distribution of house characteristics across vintages

To the extent that unobservable heterogeneity between vintages is correlated with differences in observable characteristics, comparing house characteristics may help alleviate the concern that peak-acquired houses are better than bust-acquired ones. Table 3 reports various means of house characteristics for the four acquisition vintages 2005-2007, 2008-2010, 2011-2013, and 2014-2016. In general, differences are minimal, and when they are significant, often go the opposite way of the main finding. For example, bust-acquired houses tend to be more newly constructed (about two years) than either of the immediately following vintages in 2011-2013 and 2014-2016, despite having lower rents. The main comparison of interest, however, is between the peak- and bust-acquired houses, between which we fail to find any statistically significant difference in size, number of bedrooms, number of bathrooms, likelihood of being a condominium/apartment (versus a single family home), age, laundry services (for apartments) and parking policy.

3.2.2 Lagged hedonic residuals

As an additional test of the importance of unobservables on the acquisition-vintage fixed effects, we implement the empirical strategy introduced by Genesove and Mayer (2001) and Beggs and Graddy (2009), which uses past sales residuals from hedonic regressions to measure quality attributes not spanned by observable characteristics. We first estimate a regression similar to Equation (1):

$$\log(p_{i,last}) = \mathcal{B}_{ctrl} X_{ctrl,i} + a_z + a_y + v_i, \tag{4}$$

where $p_{i,last}$ is the last purchase price of the house, $X_{ctrl,i}$ is the same vector of controls used in column 4 of Table 2, a_z is a zip code fixed effect, and a_y a year-of-last-sale fixed effect.[19] Denoting $\pi_{i,last}$ as the hedonic estimate from Equation (4) for house $i's$ most recent sale, we then estimate several variants of Equation (1) that incorporate the residual $\log(p_{i,last}) - \pi_{i,last} = v_i$.

Table 4 presents the results. In the first column, note that although the coefficient on the residuals from past sales $\log(p_{i,last}) - \pi_{i,last}$ is strongly and positively associated with current rents, its inclusion has almost no impact on the sequence of acquisition vintage coefficients. In particular, the estimate of the peak-bust rental spread remains stable at 2.5%, with a p-value less than 1%. Column 2 replaces the acquisition vintage indicators with the logarithm of the last sales price, along with the lagged residual.

Our preferred specification eschews actual historical sales prices due to their potential correlation with unobserved quality. However, the specification in column 2 includes lagged residuals, along with contemporaneous house characteristics, so that the estimate of the coefficient on $\log(p_{i,last})$ is driven more by changes in index prices, rather than cross-sectional heterogeneity.

Recall from column 4 of Table 2 that a peak-bust rental spread of 2.4% implies a rent

[19]Virtually identical results obtain including separate fixed effects for each zip code $\times$ year-of-last-sale pair.

elasticity to historical prices of about 0.054 — i.e., a 10% drop in index prices corresponds to a 0.54% drop in rents. This is similar to, although somewhat lower than, the estimated coefficient on $\log(p_{i,last})$, from column 2, which is 0.083 ($t = 16.7$).[20] Controlling for house age and time on market leaves the results unchanged (column 3), as does interacting lagged residuals with purchase vintages, which allows the value of unobservable characteristics to appreciate with the overall price index (column 4).

3.3 Landlord Heterogeneity

Rather than peak- and bust-acquired houses differing in terms of their property characteristics, systematic differences in owners and/or landlords may be responsible for the peak-bust rental spread. For example, some landlords may provide superior service for their rental properties, and consequently, charge higher rent. A second possibility is that holding costs for empty properties may differ across landlords, and thus, alter the trade-off between higher asking rents and a lower probability of matching with a tenant. In this section, we consider the joint hypothesis that: 1) the types of owners acquiring rental property varies through the housing cycle, and 2) the predicted direction on asking rents parallels that of the peak-bust rental spread.

3.3.1 Controlling for landlord type

We begin by supplementing our set of rental listings with data from Corelogic deed files, which identify the name of the most recent buyer. Linking each individual listing to the Corelogic data requires an exact property match, resulting in 34,473 listings (78% of the data) where the property owner can be identified. For 25,000 of these properties we can observe the full set of house and rental characteristics. Two designations in the deed records are

[20] One possible reason, as discussed in Genesove and Mayer (2001) is that in addition to capturing changes in the housing index, and quality attributes not spanned by controls, transactions prices may reflect over- or underpayment relative to fundamentals. See the discussion on pages 1238-1241 of Genesove and Mayer (2001) for a detailed discussion of this issue, and the potential biases introduced.

significant: those for absentee owners and for owners who are corporations or legal entities. The former are directly flagged in the deed record files. For the latter, Corelogic identifies a number of these directly, which we supplement by searching for key words in buyer names, such as "LLC," "corporate," "investment," "INC," and other words frequently associated with corporate entities.

Regarding the first part of the joint hypothesis above, owner/landlord composition varies strongly with aggregate house prices. The fraction of houses with absentee owners at the time of purchase increased sharply during the bust. Whereas 25.5% of listings last purchased in the years from 2005-2007 were originally absentee owned, this increased to 38.3% for properties last purchased from 2008-2010, and again to 53.6% for acquisitions from 2011-2013. Similar patterns are observed for corporate owners and legal entities, which increased from 2.5% in 2005-2007, to 11.2% in 2008-2010, and finally to 14% in 2011-2013.[21]

We then return to our benchmark specification in Equation (1), but now include controls for landlord type. Table 5 presents our results. To ensure that the sample containing landlord information is comparable to the full dataset, the first two columns show the results of re-estimating our baseline specifications from prior analysis, without new variables. The estimated peak-bust rental spread remains very similar (2.2%, $p < 0.01$), both with (column 2) and without (column 1) the lagged hedonic residual included.

We then augment the specification with indicators for absentee and/or corporate owners. Column 3 shows that accounting for absentee owners makes virtually no difference in the rent spread between peak- and bust-acquired houses (2.2%, $p < 0.01$). Moreover, absentee owners — which increased in frequency during the bust — are associated with significantly *higher* rather than lower rents, the opposite pattern implied by the peak-bust rental spread. The final column adds an indicator for corporate and legal entities, which we assume are absentee owners, and thus, constitute a subset of the absentee owner group. This distinction,

[21]Similar patters in the bust and post-crisis period for several residential real estate markets in the United States are discussed also in Xiao and Xiao (2019).

shown in column 4, indicates that of the absentee owners, it is corporations and legal entities who are most responsible for the higher estimated rent. Yet, despite the inclusion of controls for landlord type, the peak-bust rental spread remains stable at 2.3%, significant at the 1% level.[22]

3.3.2 A placebo

We replicate our analysis by linking acquisition vintages to current rents in Texas. Geographically large, populous, growing, and demographically diverse, the U.S. state with the second largest economy represents arguably the best single-state comparison to California (the largest economy), though admittedly important differences remain. One is that relative to California, Texas real estate is considerably less expensive. As of 2019, a market report by *Cushman and Wakefield* indicates that every property class in both commercial and residential (multi-family) real estate has a higher price-to-rent ratio in California's major cities compared to those in Texas, despite prices in the latter having more than doubled over the last decade.[23] Unfortunately, since information in Corelogic's deed records is less detailed for Texas compared to California, we cannot explicitly control for landlord characteristics as before. However, given that current income (rather than appreciation) is the prime consideration for corporate real estate investors, it seems reasonable that plummeting bond yields during the crisis — the rate on 10-year U.S. Treasury bonds was over 5% in June 2007 but less than 2.5% in December of 2008 — would have also tilted the composition of Texas's landlords toward having more investors, as observed in California.[24]

Another important difference is that the extreme price volatility observed in California did not occur in Texas. Whereas Los Angeles and San Francisco saw prices rise by 121% and

[22] We have also interacted both landlord types with the acquisition vintage indicators. Though the peak-bust rental spread is present across both landlord groups, the estimates are smaller, noisier and less significant for absentee landlords and corporate landlords.

[23] See http://www.cushmanwakefield.com/en/research-and-insight/2019/us-q1-2019-marketbeat.

[24] Evidence of real estate investors reaching for yield in the U.S. is provided by Glaeser (2013) and Duca and Ling (2018). Korevaar (2019) provides historical evidence based on transactions from the 17th and 18th century in the Netherlands.

91% respectively from January 2000 to January 2005, only to fall by 21% and 27% over the following five years, Dallas's Case-Shiller index indicates a rise of only 16% from 2000-2005. More importantly, for all major cities in Texas, prices *rose* from 2005-2010, including Dallas, where real estate values increased by a modest 0.5%. Comparable price changes, based on the "All Transactions Price Data" from the St. Louis Federal Reserve Bank, for Houston were 26% (2000-2005) and 14% (2005-2010), for Austin 21% and 23%, and for San Antonio 25% and 20%.

Together, these observations suggest that if stale prices influence rents, we would not expect to find a comparable peak-bust rental spread in Texas, whereas if compositional changes in landlord/owners were responsible for the peak-bust spread observed in California, we likely would. Figure 4 reports the results. As with the analogous figure for California, we plot the estimated coefficients for a sequence of purchase/acquisition vintage dummies. The top panel corresponds to a cross-sectional regression of current rents, and the bottom panel to a panel regression, where the dependent variable is the logarithm of the acquisition price for each house. Together, the analysis in Texas shows a fairly smooth trajectory in historical prices, and a complete absence of a peak-bust rental spread between 2005-2007 compared to 2008-2010. In fact, the estimated coefficient for the latter vintage is higher than that for the former, although jointly, no coefficient on any of the acquisition vintage dummies is significant.

3.4 Statistical Diagnosis of Unobservables

Evidence in the previous sections shows that heterogeneous quality across houses and landlords is unlikely to drive our findings. Still, unobservable dimensions of house or landlord quality may vary across vintages in a way that explains the peak-bust spread in California. Thus, in this section we introduce empirical tests directly aimed at identifying the potential confounding effects of unobserved heterogeneity.

A formal treatment of unobservable variable bias is developed by Altonji, Elder, and Taber

(2005), and subsequently extended in Oster (2019). The goal is to place plausible bounds on the size of omitted variable bias, given structural assumptions about the effects of both unobservable and observable factors on the outcome variable of interest (here rents). The typical procedure is to estimate regressions with progressively more controls, measuring: 1) how much the R^2 increases, and 2) how much the coefficient of interest (here the coefficients on purchase/acquisition vintage dummies) is reduced. The final step is to extrapolate to a hypothetical regression that controls for all relevant factors, i.e., those both observed by the econometrician, as well as those not observed. In this hypothetical estimation, which involves the maximum possible explanatory power, R^2_{max}, the question is whether we would still estimate a significant effect for the coefficient(s) of interest, and if so, its magnitude.[25]

The answer depends on the extent to which unobservables behave similarly to observables, in terms of their marginal impact on the estimated treatment effect. If this sensitivity, δ, is equal to one, then the effect of unobservables is identical to that of the observables, and the extrapolation to R^2_{max} respects the "slope" between the long and short regressions. For example, if control variables increase the R^2 from 30% to 50%, with β declining from 1 to 0.7, one would hypothesize that in a theoretical regression with $R^2_{max} = 90\%$ –an increase twice that observed from the short to long regressions ($2 \times 20\% = 40\%$)– the coefficient of interest would suffer an additional decline of $2 \times (1 - 0.7) = 0.6$, for an estimated bias-free treatment effect of 0.1. Likewise, the cases $\delta < 1$ and $\delta > 1$ correspond, respectively, to unobservables having a smaller and larger impact on the treatment effect than observables.

Because δ is itself unobservable, a common implementation of this procedure is to ask, for a given long-short regression pair, and a hypothetical R^2_{max}, how large would δ need to be in order to fully explain the observed effect. Denoting this threshold $\bar{\delta}$, high values imply a required sensitivity that is larger (in absolute value) than the one with respect to changes in observable characteristics. For more details on the methodology developed by Oster (2019),

[25] R^2_{max} might still be less than 1 due to, for example, measurement error, but is still expected to be very high.

see Appendix B.[26]

Table 6 presents three variants of this calculation. In the tests we use only the sample of properties for which the full set of house controls is available, so that point estimates from the short regressions, while very similar, are not identical to the ones reported in column 1 of Table 2. In column 1 of Table 6, the treatment variable is the "bust" indicator, taking a value of one if the house was purchased at the depth of the crisis (2008-2010). The short regression includes only zip code fixed effects and the other (non-bust) acquisition dummy variables ($R^2 = 62.7\%$), while the inclusion of house characteristics in the long regression increases the R^2 to 85.9%. The negative value for $\bar{\delta} = -8.56$ at $R^2_{max} = 100\%$ reflects the observation that the estimated treatment effect becomes larger (in absolute value), relative to the other acquisition vintages, in the long versus short regression. That is, for the bias-free treatment effect to be zero, not only would unobservables have to display more than eight times the sensitivity to the treatment effect as do observables, but with opposite sign. Absent such unusual correlation structure, the estimates here suggest that the rent reduction observed among bust-acquired houses is unlikely to be explained by unobserved heterogeneity in housing quality.

The second column adds the peak-vintage coefficient in relief. In addition to a dummy for the bust (2008-10), we also include an indicator for the period that spans both these years and the immediately proceeding peak years (2005-2010). With both variables in the regression together (along with vintages outside the 2005-2010 window), a significant coefficient on the bust-vintage dummy indicates that, relative to 2005-2007, the years 2008-2010 are associated with lower rent. Column 2 shows this is the case, both in the short (3.3%) and long (2.4%) regressions. Unlike in column 1, where we observed a strengthening of the effect, column 2 indicates a decline, but the effect is very modest. To drive the coefficient to zero, even if R^2_{max} were 100%, unobservables would have to exert an influence on the treatment effect over

[26]This methodology is used in many recent empirical studies in urban economics (see for example Clarke and Freedman (2019) and Albert and Viladecans-Marsal (2019)), and finance (see for example Heimer, Myrseth, and Schoenle (2019), Dougal, Gao, Mayew, and Parsons (2019) and Gao, Huang, and Goldstein (2019)).

4.5 times larger than the one of observable factors.

Finally, in the third column we consider the specification of Equation 1 in which we replace the year of purchase vintages with a single variable: the level of the zip code house price index at the time of last purchase (see column 6 of Table 2). The short regression includes only the zip code index and zip code fixed effects, while the long regression extends the set of controls to include all available information on house and rental characteristics. The treatment coefficient of interest is now the one for the zip code index. Similar to what observed for column 1, the estimated treatment effect becomes larger as we include more controls. Thus, we find a negative value for $\bar{\delta} = -9.24$ at the maximum R^2 (100%), which implies that, to explain the treatment coefficient, unobservables should have an effect on rents more than nine times larger than the one of observable characteristics, and this effect would have to be of the opposite sign with respect to the one of observables.

3.5 Time on Market

To the extent that acquisition years distort asking rents relative to fundamental values, we would expect this to be at least partially reflected in the time a house sits on the market before renting. More specifically, controlling for quality, houses purchased during the peak (at high prices) are expected to rent more slowly than houses acquired during the bust (at low prices).

To test this prediction, we estimate the number of days each listing was posted online before renting. We make three assumptions. The first is that when a house listing disappears online, this implies that the house has been rented. Although a house's failure to rent may result in it being put up for sale, or in cases of proposed rentals that are currently owner-occupied, simply taken off the rental market, we believe that such situations are rare. The second assumption stems from the way the data set was assembled. We crawled the online portal over consecutive two-week intervals, and thus we only observe whether a listing disappeared between collection dates, rather than the exact date it rented. When a listing disappears, we assume that it rented halfway between the collection dates. This means that each listing will be mis-measured by seven days on average, with a minimum (maximum) of one (thirteen) day(s). The final assumption is that the online portal occasionally features "zombie" listings, where the house has rented, but the advertisement has not been removed. As described in section 2.1, we dropped from the data the roughly 4% of observations with listing times exceeding 300 days. This filter has a minimal impact on the results, and alternative cutoffs give similar estimates.

Overall, the California rental market during the data collection period appears very tight, with over one-third of houses leaving the market within two weeks. Our main interest is whether, conditioned on observables, peak-acquired houses take longer to rent than bust acquired houses. On average, the difference is relatively modest, with the former taking 45 days to rent versus 41 for the latter. Although a relatively small difference, a Kolmogorov-Smirnov statistic rejects the hypothesis that the empirical distributions of time on the market for peak and bust acquired groups, reported in Figure 5, are drawn from the same distribution ($p < 0.05$). While the difference is small for much of the range, houses acquired during the peak (blue dots) become overrepresented starting at about 50 days, and continuing through the right tail of the distribution.

Regression analysis confirms a small, but statistically significant disparity in time on the market. We estimate:

$$\log(days_i) = \beta_r \log(R_i) + \mathcal{B}_{ctrl} X_{ctrl,i} + a_z + e_i, \tag{5}$$

where $days_i$ is the number of days on market. Although most of our observations correspond to completed spells, approximately 25% remain outstanding as of our final collection date (mid-March 2019). Consequently, in Table 7, we present maximum-likelihood estimates from an accelerated failure time (AFT) model that accounts for right-censoring.[27] $X_{ctrl,i}$ represents

[27] The residual e_i is assumed to have an extreme value distribution, and the baseline hazard function is Weibull with shape parameter κ. The shape parameter determines whether the hazard rate in the model is

the same vector of listing characteristics used in Equation (1), excluding the log of inventory days, and a_z is a zip code fixed effect.

The first two columns of Table 7 indicate that across the entire sample, after controlling for characteristics, more expensive houses take longer to rent. With the logarithm of rent by itself (column 1), the estimate suggests that in the cross-section, a house with 5% higher rent will take 3% longer to clear the market, or about 1.3 days on average. Column 2 shows that this effect is concave, a 5% higher rent would lead to a 2.8 days delay on average.

Column 3 focuses specifically on the difference in time on market between the peak-and bust-acquired vintages, with only houses acquired from 2005-2010 in the estimation. After controls for house characteristics and zip code, the point estimate is 6%, although with marginal statistical significance ($p < 0.10$). Note that while of the same order of magnitude, the estimate in column 3 (6.3%) is two to four times larger than that implied by either columns 1 or 2. One possible explanation is that despite the attributes spanned by observable characteristics, and those captured by lagged hedonic residuals, rents in the cross-section nevertheless contain information about house quality. However, this is *not* true for rent disparities caused by stale information — i.e., acquisition vintages —, for which we indeed find a larger effect on inventory times.

The results in the last column of Table 7, along with the average size of the peak-bust rental spread, can be used to produce a back-of-the-envelope estimate of the loss in expected revenue for a landlord acquiring her house at the bust, compared to one acquiring her house at the peak. At an average rent of $3,524 (the sample average), a 2.4% reduction implies a revenue loss of $1,015 a year. Although on average the house will rent 6% faster (3 days), this expediency is offset by the lower rent only up to 3 months into the lease term. Then, the total difference in revenue between peak- and bust-acquired rental houses over a year is about $700 on average. While we believe this rough calculation is useful for benchmarking the all-in distortion implied by stale information the rental market, we reiterate that whereas

increasing ($\kappa > 1$), decreasing ($\kappa < 1$) or constant ($\kappa = 1$) in the number of days on market.

the effects of stale prices on rents appear to be estimated relatively precisely across a variety of specifications, due to the measurement error discussed above, we are less confident in the exact estimate of the analogous spread for time on market.

4 Why do landlords consider stale information?

In this section, we take the relation between historical index prices and current rents as given, and attempt to better understand the reasons why stale information enters landlords' optimization problems. To fix ideas, we first present a simple model (Section 4.1) that formalizes a key trade-off landlords face when setting rents – higher revenue conditionally on renting versus a lower probability of renting – and then allow for two possible deviations from this fully rational benchmark. In the first, landlords hold *incorrect beliefs*, in that they mistakenly overweight past (stale) prices as signals of house quality. In the second, landlords have *preferences* influenced by reference points, which in turn depend on historical purchase prices. After characterizing the empirical predictions, we return to the data, and test the implications of the different mechanisms.

4.1 A Simple Model of Rent Setting

4.1.1 Rational Benchmark

We begin with a simple, one-period model, based on Lazear (1986), involving a landlord attempting to rent out a house she owns. When setting rent, the key trade-off she faces is between maximizing revenue if the house rents, and lowering the probability of a transaction.[28]

[28]Our treatment of the trade-off is highly stylized (see also Stull, 1978 and Allen, Rutherford, and Thomson, 2009), and does not attempt to formally model the two-sided search problem involving both landlords and tenants. One drawback of our approach is that the demand curve takes a particularly simple form – in particular the probability a house rents is linear in rent – whereas more sophisticated frameworks (e.g., Genesove and Han, 2012, Head, Huw, and Sun, 2014 and, most importantly, Guren, 2018) indicate that this may not be realistic. Since we use the model only to establish a rational benchmark, and then derive qualitative implications when distorted beliefs and/or reference points are added, we believe that the fact of relying on a simplified framework does not change the key predictions that we explore in the following

There is a single landlord, whose house has value $v \sim U[0,1]$ to any potential tenant. The landlord, knowing only the distribution of v but not its realization, posts a take-it-or-leave-it rent R, for which the tenant can rent the house ($v \geq R$), or if not ($v < R$), search for another one. By ignoring any costs the landlord may face, the model equates expected revenue with expected profit.[29] There is a single period, all parties are risk-neutral, and the reservation utility for landlords and tenants are both zero. The landlord's problem is:

$$\max_R (1 - F[R])R = \max_R (1 - R)R, \tag{6}$$

which gives $R^* = \frac{1}{2}$, and expected revenue of $\frac{1}{4}$. The landlord's ex-ante lack of information about v means that she cannot price discriminate, and thus trades off the probability of renting $(1 - R)$ against the rent conditional on a successful match (R). The expected surplus from trade is $\frac{1}{4} + \int_{R^*}^{1}(v - R^*)dv = \frac{3}{8} < E[v] = \frac{1}{2}$. The efficiency loss obtains from situations where $R > v$, but nevertheless, the house fails to rent.

4.1.2 Distorted Beliefs

Now consider that the landlord of house i believes (mistakenly) that $v \sim U[\Delta, 1 + \Delta]$, where $\Delta \in \{-X, X\}; 0 \leq X \leq \frac{1}{2}$. The first case, $\Delta = -X$, corresponds to undervaluation, whereas the second case, $\Delta = X$, implies a landlord who is positively biased about the distribution of v. The landlord will now set

$$R^* = \frac{1 + 2\Delta}{2}, \tag{7}$$

which is either too high, or too low, by X.

If the house is undervalued by the landlord, the expected revenue is $\frac{1}{4} - X^2$, which is strictly worse than in the benchmark case. Although the house now rents more often than

sections.

[29] This assumption is innocuous for costs that do not depend on rent, such as periodic maintenance, taxes, or mortgage service. On the other hand, the solution will generally differ if the cost of providing rental services is not constant across potential tenants, or if rent acts as a screening device.

before, this is more than offset by lower rent in the event of a successful match.[30] In the limit of $X = \frac{1}{2}$, the house always rents ($R^* = 0$), and the tenant captures the entire surplus. When the house is overvalued, the landlord's expected revenue remains $\frac{1}{4} - X^2$. However, as X approaches $\frac{1}{2}$, R^* approaches 1, and the probability of renting the house diminishes to zero, so that the combined surplus goes to zero.

4.1.3 Preferences

As an alternative, return to the benchmark case, where the landlord holds correct beliefs about v. However, in addition to maximizing expected profit, the landlord is influenced by a reference point $C > 0$, against which rent R is directly compared. If $R \geq C$, the landlord's objective function is as above, i.e., she maximizes expected revenue as in Equation (6). However, for $R < C$, the landlord experiences an additional utility loss $g(C - R) > 0$, where $g'(\cdot) > 0$.

There are two primary ways to motivate the existence of reference point C, and the landlord's loss function around it, $g(\cdot)$. One is through *liquidity constraints*, whereby C constitutes the landlord's periodic (e.g., monthly) expenses related to the rental property. To the extent that the rent R fails to exceed C, the landlord must fund the deficit from other sources. Here, g captures the frictions associated with such cash-flow management, such as liquidating other assets (i.e., transactions costs), delaying other consumption, or the opportunity cost of the landlord's time. Under the liquidity-management hypothesis, we specify convexity in the loss function, or $g'' > 0$. With this formulation, we wish to capture the intuition that landlords find it relatively easy to accommodate modest rent-cost deficits, but as the size of the shortfall increases, face additional costs that raise more than proportionally.

Reference-dependence can also arise through *prospect theory preferences* (Kahneman and

[30]The tenant is, of course, better off, with an expected utility of $\int_{R^*}^{1}(v - R^*)dv = \frac{1}{2}(X + \frac{1}{2})^2 > \frac{1}{8}$. The total surplus is $\frac{3}{8} + \frac{X}{2} - \frac{X^2}{2}$, increases as the landlord becomes more pessimistic about her house's value: as X increases, R^*, and consequently the probability of the house not renting, both approach zero.

Tversky, 1979, 1991, 1992). A key attribute of prospect theory is "loss aversion," whereby the loss function g is particularly steep for small, positive values of $C - R$. In contrast to the loss function motivated by liquidity constraints considerations, the loss function is concave ($g'' < 0$) under prospect theory.[31] Although the landlord always prefers smaller deficits (recall that $g' > 0$ in all cases), under prospect theory, she is particularly concerned with minimizing the chance of incurring even small losses.[32]

Under both liquidity constraints and prospect theory, the reference point, C, may depend on the price a landlord paid for her house. One reason, which applies to every property in our sample, is due to peculiarities in the California tax code. Proposition 13, passed by the state legislature in 1978, mandated that property taxes on newly sold property be no more than 1% of the purchase price; furthermore, subsequent percentage increases in taxes were limited to 2%. The combination of these factors means that property tax bills remain closely tied to original acquisition prices, even decades later. A second reason, is due to mortgage installments. In general, higher purchase prices translate to larger mortgages, and therefore higher monthly payments.[33]

With reference-dependent utility, the landlord's optimization problem becomes:

$$\max_R (1 - R)R - (1 - R)g(C - R) - Rg(C), \tag{8}$$

which we present in expanded form to highlight the difference with the benchmark case, stemming from the second and third terms. The last term captures the additional loss in expected utility if the house does not rent. Given that $g'(.) > 0$, and that $g(0) = 0$, it follows that $g(C)$ is the maximum utility loss the landlord can experience, which occurs with probability R. The second (middle) term captures the landlord's loss in the event that the

[31] Since R enters with a negative as the argument, g exhibits convexity with respect to R, consistent with traditional formulations of prospect theory utility functions.

[32] Consistent with this prediction, Anenberg (2011) finds strong evidence in the sales market that homeowners facing nominal losses on their housing investments and owners with high LTV ratios sell for higher prices.

[33] Argyle, Nadauld, and Palmer (2020) present evidence that monthly-payment considerations impact the demand for automobiles, and attribute these effects to liquidity constraints.

house rents, $g(C-R)$, multiplied by $1-R$, the chance the house rents at posted rent R.

The optimal rent is now:

$$R^* = \begin{cases} \frac{1-g(C)}{2} & \text{if } C \leq \hat{C} \\ 1 - \frac{1+g(C)-g(C-R^*)}{2+g'(C-R^*)} & \text{otherwise,} \end{cases}$$

where $\hat{C}$ is defined by $g'(0)[\hat{C}-1]+2\hat{C}+g(\hat{C})-1=0$. The cutoff $\hat{C}$ characterizes the value of the reference point below which, in equilibrium, $R^* > C$. In this region, C still plays a role in the landlord's decision, but only insofar as she reduces the rent, in order to avoid the maximum loss, $-g(C)$, which is experienced if the house fails to rent. This leads to our first proposition.

Proposition 1 *For $\hat{C} >> C \approx 0$, $R^* \approx \frac{1}{2}$.*

When the landlord's reference point is small, the optimal rent is close to the no-reference point benchmark, which follows from $g(0)=0$. In our later empirical analysis, we will develop empirical proxies for C, which generate cross-sectional dispersion in the size of landlords' reference points. For sufficiently small values of these proxies, the rents we observe constitute our approximation to the benchmark case with no reference dependence.

Our second proposition pertains to higher values for C. For $C > \hat{C}$, the landlord will experience a utility loss not only if the house doesn't rent, $-g(C)$, but also a (smaller) loss if it does, $-g(C-R)$. This additional consideration leads to an upward, discontinuous rent shift at $\hat{C}$. Crucially, the size of this shift, as well as how R^* varies with further increases in C, depend on the shape of $g(\cdot)$.

Proposition 2 *For $C > \hat{C}$, under prospect theory, $R^* > \frac{1}{2}$, the benchmark with no reference dependence. Likewise, in the liquidity constraints case, $R^* < \frac{1}{2}$.*

To build some intuition for Proposition 2, consider first the linear case, $g''(\cdot)=0$, where $g(C-R)=K(C-R)$, with $K>0$. With a linear specification, changes in R have equal, but opposite effects on the utility losses experienced by the landlord. On the one hand, a lower value for R decreases her chance of experiencing the maximum possible loss ($-KC$), but at the same rate, increases her chance of experiencing a smaller loss, $-K(C-R)$, with the marginal effects with respect to C exactly offsetting. Thus, although the presence of the reference point unambiguously decreases the landlord's expected utility, the specific value of C plays no role, and R^* equals the benchmark value of $\frac{1}{2}$.

This is not true when the loss function $g(\cdot)$ exhibits curvature. Unlike the case in which $g''(\cdot) = 0$, either a concave (corresponding to prospect theory) or convex loss function (liquidity constraints) leads to an unambiguous relation with the benchmark case with no reference dependence. Provided that $C > \hat{C}$, then $R^* > \frac{1}{2} \Leftrightarrow g'(C - R^*) > 2[g(C) - g(C - R^*)]$, which can be written as

$$\frac{g'(C - R^*)}{\frac{g(C)-g(C-R^*)}{R^*}} > 2R^*.$$

Observing that the left hand side is the ratio of the local slope of g evaluated at $C - R^*$, compared to the average slope from C to $C - R^*$, it follows directly that $R^* > \frac{1}{2} \Leftrightarrow g''(\cdot) < 0$, and $R^* < \frac{1}{2} \Leftrightarrow g''(\cdot) > 0$.

Intuitively, Proposition 2 illustrates a key trade-off between the two effects of reference dependence on the optimal rent-setting behavior of landlords. When the loss function is fairly flat around the reference point, but steepening as the $C - R$ gap widens ($g''(\cdot) > 0$), landlords are mostly concerned about renting the house and avoiding large losses, even if this requires dropping the rent, and accepting a modest loss. Such modest losses are, of course, exactly what landlords with prospect theory preferences ($g''(\cdot) < 0$) find least tolerable on the margin. For them, the diminishing sensitivity of the loss function creates an incentive for risk taking in the loss region, a key implication of prospect theory. In our context, such risk-taking is operationalized via higher rent: even though this reduces the probability of renting (and reduces expected revenue), it is offset by the benefit of narrowing $C - R$, in the event that the house successfully rents.

4.2 Empirical Evidence: Distorted beliefs

In this section, we explore whether distorted beliefs might be the mechanism driving the peak-bust rental spread in the data. In particular, we consider the possibility that landlords may regard stale signals as informative about house quality, leading them to set high (low) rents after having purchased at the peak (bust) of the prior real estate cycle. The nature of the inference problem, as described in Section 3.1, is one in which a landlord faces uncertainty about her house's quality. Because sales prices capitalize both observable and unobservable dimensions of house quality, past transactions are informative about the persistent component of each. Unless the landlord is already perfectly informed, perhaps from recent comparable transactions, her current estimate of quality will depend in part on the price she paid.[34]

At the same time, sales prices are also a function of discount factors, irrespective of house quality. One way that landlords may come to hold distorted beliefs is by conflating this source of price variation (which occurs mostly in the time-series), with that caused by heterogeneity in quality (which is mostly about the cross-section). More specifically, if a landlord underestimates the time-series volatility of capitalization rates, she will systematically overestimate house quality ($\Delta = X$) when rates are low such as in 2006, and underestimate house quality ($\Delta = -X$) when rates are high, such as in 2010. In this way, the peak-bust rental spread can be understood as a combination of the landlord's acquisition timing, and her incomplete understanding of the information structure.[35]

Our empirical tests exploit cross-sectional variation in the number, and quality, of signals a landlord is likely to have received from nearby, comparable transactions. The key prediction

[34]Information on the cross-section of prices for similar properties may not be readily available, both because prices are observed rarely, due to the general illiquidity of the housing market, and because houses are heterogeneous, and so the observed prices might not be informative. Segmentation in housing demand is directly documented by Piazzesi, Schneider, and Stroebel (2020), and also explored by studies that focus on local spillovers, such as Campbell, Giglio, and Pathak (2011) and Anenberg and Kung (2014).

[35]While sharing some of the key prediction of classical anchoring — irrelevant information being treated as relevant — the behavior described here may be more accurately characterized as "anchoring via misattribution," in the sense that landlords' reliance on irrelevant signals stems from a misunderstanding of the information structure, rather than from alternative psychological foundations for anchoring. See, for example, Mussweiler and Strack (2001), Epley and Gilovich (2006), Bergman, Ellingsen, Johannesson, and Svensson (2010) and other contributions mentioned in the literature review by Furham and Boo (2011).

is that the relation between rents and past index values is weaker when these signals are more numerous or more informative.

We develop five proxies for the information environment faced by each landlord. First, for each listing we count the number of comparable rental properties advertised at the same time, which we define quite strictly, as properties with the same number of bedrooms, within 75 square feet of living area, and located within one mile.[36] The mean house has 0.31 comparables within one mile, and the standard deviation of the number of comparables is 0.48. Our second and third measures estimate a house's "atypicality" with respect to the remainder of the zip code. For a given listing, each measure of atypicality is the absolute value of the difference between the relevant attribute and the zip code average. In absolute value, the median rental property in our sample differs from the mean in the same zip code by 0.7 bedrooms and by 480 square feet in terms of size. In our regressions, we will scale this measures by the standard deviation of each characteristics within the postcode in which a property is located. The fourth and fifth measures capture heterogeneity at the zip code, rather than listing level. By calculating the standard deviation in the number of bedrooms and size for each zip code, we hope to expand the notion of comparability beyond that contained only in active listings. Our intuition is that in homogenous areas, landlords can generally extract more information from recent rental listings or even sales, including those that predate the beginning of our sample period. For both heterogeneity measures, we compared zip codes above and below the median.

The results are presented in Table 8. To keep the presentation manageable, we measure historical price fluctuations using zip code-level price indices, allowing us to interact a single variable (rather than nine acquisition vintage dummies) with the above proxies intended to capture variation in the information environment. Consistent with our predictions, we find that the effect of historical price indices on rents is stronger when the information environment is weaker. The coefficients on all five interaction terms are statistically significant. For the

[36] We obtained geo-location coordinates (latitude and longitude) for the 44,237 listings included in our final dataset.

first, a one standard deviation increase in the number of comparable listings decreases the sensitivity to past index prices by 35%. Likewise, a one standard deviation increase in atypicality for size (bedrooms) increases the rent-historical index sensitivity by 130% (20%). At the zip code level, those with higher-than-median variation in property size have a rent-historical index sensitivity about 80% larger than those below the median. When ranking zip codes based on heterogeneity in the number of bedrooms, the comparable effect is a 60% increase in sensitivity.

4.3 Empirical Evidence: Preferences

A second reason why stale prices may matter for rent setting today is via reference-dependent utility functions, whereby landlords experience periodic (e.g., monthly) disutility when the cash flow received from a rental property falls below a predetermined threshold. As Barberis (2013) discusses in his survey of prospect theory, testing for reference dependence is challenging because outside the laboratory, there are often many candidates for possible reference points, and the researcher lacks guidance on which, if any, is relevant in a given context.

Fortunately, our setting offers what would seem to be a natural candidate: a landlord's monthly obligations related to the rental property. Unlike owning a stock, where the present value of all expected cash flows is paid up front, owning real estate "backloads" at least some, and depending on leverage, most of these costs. Consequently, landlords face regular outflows for maintenance, taxes, and mortgage service, which constitute together a highly salient benchmark against which rents might be compared. In this section, we exploit variation in landlords' use of debt financing to generate dispersion in monthly payments, and then relate this dispersion to the cross-section of rents.

Recall from section 4.1.3 that simply incorporating reference-dependence into the landlord's objective function is not enough to deliver a clear empirical prediction for rent-setting. As Proposition 2 indicates, reference-dependence may either increase or decrease asking rents, depending on the curvature of the loss function. Under prospect theory, the loss function is steepest just below C, causing the landlord to increase rent, in hopes of closing the gap. Although this simultaneously increases the chance that the house doesn't rent, on the margin, the convexity of the loss function ($g''(C - R) > 0$) renders the first consideration more important than the second. On the other land, when landlords are more concerned about insuring against large losses ($g''(C - R) < 0$), lower rents are optimal, since the increase in the probability of renting has the largest impact on expected utility. Thus, by examining whether cross-sectional variation in C leads to more, or less, aggressive rent-setting by landlords, we can infer whether prospect theory, or liquidity constraints, appears more important for generated the peak-bust rental spread.

For a subset of the properties in our sample, Corelogic reports the original amount borrowed by buyers who used a mortgage to finance their house purchase. For example, for the same San Francisco house above, a buyer that uses 80% mortgage financing in 2019 would borrow $80\% \times \$1.335M = \$1.07M$. At 4% for 30 years, the corresponding monthly payment would be \$5,100. On the other hand, the monthly payment using only 50% debt for the purchase would be only \$3,200. The analysis here controls for the purchase price, and instead focuses on variation in the loan-to-value (LTV) ratio.[37]

We first compare rents, controlling for historical purchase prices, between landlords with low-LTV (50% or less) and high-LTV (95% or higher) ratios. The low-LTV group represents the control group, for which we expect C to play little if any role in the landlord's rent-setting decisions. Recall from Proposition 1 that if C is small, the optimal rent is close to the benchmark without reference dependence.[38] The high-LTV group represents the intent-to-treat group. Among these landlords, we assume that the high interest costs — in addition to taxes and other expenses — tip the balance such that the typical highly-levered

[37]We exclude from the data observations for which the LTV ratio at the time of the last purchase is equal to zero. We make this choice because $LTV = 0$ may imply that the property was purchased with cash, or that the buyer obtained financing through a channel alternative to the mortgage market. Such alternative forms of debt financing are not recorded by Corelogic.

[38]Specifically, recall that if $C < \hat{C}$, then in equilibrium, the house rents for more than the reference point (i.e., $R > C$), and $g(R - C) = 0$. Although the model predicts that rents for $0 < C < \hat{C}$ will be lower than rent in the no-reference-dependence case, we assume that for very low LTV values, C is small enough that this is negligible.

landlord cannot avoid a monthly "loss" ($R < C$) at current rents.

Although we lack detailed data on maintenance expenses for each property, back-of-the-envelope calculations suggests that the LTV cutoffs used above are reasonable for identifying landlords likely to experience monthly gains or losses. The average rent in our sample is \$3,500 per month. Allocating 20% to property maintenance and other miscellaneous expenses (e.g., exterior maintenance), the average rental house yields perhaps \$2,800 monthly before taxes and mortgage service. As described above, property taxes in California depend on purchase, not market, prices. Using the average historical purchase price of \$590,000, the monthly property tax bill will be about \$550 per month, leaving \$2,250 prior to mortgage service. Using a 30-year term, and an interest rate of 4%, the LTV which generates a monthly payment (principal and interest only) equal to the rent is 80%.[39] We thus select LTV cutoffs, in hopes of grouping landlords into those likely to face a monthly deficit versus surplus.[40]

Among the approximately 4,000 landlords with LTV ratios either below 50% or above 95%, we estimate:

$$\begin{aligned} \log(R_i) = \phi_{LTV,high} I_{LTV_{i,last}>0.95} + \phi_p \log(p_{i,last}) + \rho(\log(p_{i,last}) - \pi_{i,last}) + \\ + \mathcal{B}_{ctrl} X_{ctrl,i} + a_z + u_i, \end{aligned} \tag{9}$$

where $LTV_{i,last}$ is the loan-to-value (LTV) ratio for the house at the time of purchase. The vector $X_{ctrl,i}$ contains controls for both house and landlord characteristics (the latter being the dummies for absentee landlords and landlords who are legal or corporate entities, as discussed in detail in section 3.3). The key covariate is the dummy transformation of LTV, $I_{LTV_{i,last}>0.95}$, which takes a value of one when the LTV is greater than 0.95, leaving the low-LTV group as the reference category. Importantly, note that in contrast to our benchmark

[39] For this calculation, we use 4%, which corresponds to (approximately) the minimum rate available to investors over the last 15 years. By doing so, we assume that landlords having initially borrowed when rates were higher would have subsequently refinanced, and thus, lowered their monthly payments.

[40] A landlord borrowing 95% of the purchase price would face a monthly payment of \$2,676, requiring the landlord to contribute \$426 to cover the shortfall. In contrast, a landlord with 50% LTV would have a monthly surplus of \$842.

estimates, we now control explicitly for lagged purchase prices, $\log(p_{i,last})$, and instead focus on variation in monthly payments through variation in LTV. As a control for unobserved heterogeneity in house quality, we also include the lagged hedonic residual from Equation (4), $\log(p_{i,last}) - \pi_{i,last}$.

The first column of Table 9 indicates a positive and significant coefficient estimate for $\phi_{LTV,high}$. Holding purchase price constant, landlords having borrowed 95% or more set rent 4.5% ($p < 0.01$) higher than those with LTV ratios less than 50%. Column 2 adds interactions between lagged hedonic residuals ($\log(p_{i,last}) - \pi_{i,last}$) and acquisition vintages, allowing the rental value of unobserved heterogeneity to appreciate with the housing index since purchase. However, consistent with our previous tests that include the lagged residual, this inclusion makes virtually no difference for rent-setting. In addition, the coefficient on $I_{LTV_{i,last}>0.95}$ is nearly identical whether we omit, or control, for landlord type (e.g., absentee or corporate landlords). Although not a perfect control, this suggest that the positive sensitivity of rents to LTV is unlikely to be explained by differences in risk tolerance across landlords.

The next three columns show the results when we include all observations for which we have LTV data, not just those at the extremes of the distribution. Columns 3 and 4 maintain the dummy variable specification, with the reference category now being houses for which $0.50 < LTV \leq 0.95$. The estimates here, which parallel the controls employed in columns 1 and 2, indicate that the rent-LTV relation is fairly uniform across the distribution. Moving from the high-LTV group to the middle-LTV group is associated with a reduction in asking rents of 2.25%, with another drop of 2.34% when transitioning to the low$-LTV$ group. Column 5 shows the result when the raw value of LTV is included as a covariate. With a t-statistic of nine, the coefficient on LTV suggests that an increase in the LTV ratio of 0.25 — similar to moving from the middle category to either extreme in columns 1 and 2— leads to a rent increase of about 2%.

In addition to the analysis involving monthly payments in Table 9, we present two additional pieces of evidence in Tables A.4 and A.5 in the Appendix. While we lack wealth and/or income data on landlords, particularly for non-corporations for which we expect loss averse preferences to be most likely, we use average house rents (Table A.4) and zip code-level income (Table A.5) as a crude proxy for the financial position of landlords.[41] Here, the idea is that the maximal per-period loss possible – the house not renting – would likely have a smaller utility impact on deep-pocketed versus poorer and/or financially constrained landlords. If so, then according to Proposition 2, reference-dependence based on liquidity constraints would predict an *inverse* relation between rent and index values at time of purchase – the opposite of what we observe – and in addition, attenuating as landlord wealth increases. Instead, what both tables indicate is that the peak-bust rental spread is not only present among zip codes with higher average income and for more expensive houses, but is stronger than in the complementary samples. Together with the evidence on monthly payments, these findings are consistent with loss averse landlords minimizing cost-rent deficits provided that the house rents, and contrary to liquidity-constrained landlords attempting to hedge large losses by dropping rents.

To this point, our empirical analysis is designed to isolate the effects of either distorted beliefs or reference dependence. However, because these tests involve different sources of variation, they do not rule out that, for a given landlord, *both* mechanisms may be simultaneously at work. Evidence of this is shown in the last three columns (5, 6 and 7) of Table 9. In columns 5 and 6 respectively, we explain rents as a function of lagged acquisition prices, $log(p_{last})$, and LTV, separately, and then include both together in column 7. That we observe minimal change in the coefficients – and if anything a slight strengthening – suggests that in the cross-section, variation in loan-to-value ratios are largely orthogonal to purchase prices, and that the effects on rent-setting are approximately additive.

To further corroborate the idea that distorted beliefs play a role in the data, independently of preferences, in Table 10, rather than comparing high-LTV and low-LTV borrowers, we

[41] We provide separate estimates of the peak-bust rental spread for listings: with monthly rent above and below median, and: from zip codes with income above/below median, based on estimates from the 2016 SOI Individual Income Tax Statistics published by the Internal Revenue Services.

focus only on the latter group, re-analyzing the peak-bust rental spread for only this subset. Column 1 shows the results for the $LTV \leq 50\%$ cohort,[42] whereas column 2 uses a cutoff of 75%. In both cases, for which we expect loss aversion to play a minimal role in rent-setting, the peak-bust rental spread is at least as large as our benchmark estimates. Columns 3 and 4 repeat the analysis using the zip code-level index at the time of purchase. In these cases as well, historical index prices appear strongly related to current rents, suggesting that even when landlords expect positive cash flows from renting, stale information over a decade ago continues to impact rent-setting decisions, likely through distorted beliefs.

5 Conclusion

Using data on approximately 43,000 rental houses in California, we ask whether stale signals — historical house prices over a decade prior — influence the rent-setting behavior of landlords today. We find that rents for houses acquired when aggregate prices were high (2005-2007) are 2-3% higher than houses acquired during the ensuing correction (2008-2010). More generally, the cross-section of rents is related to the timing of past purchase: current rents are related to the value of the Case-Shiller index for major cities in California at the date of acquisition.

We then explore three families of possible explanations: unobserved heterogeneity in house or service quality, distorted beliefs, and reference-dependence. We fail to find support for the first alternative, and in some cases (e.g., landlords), the variation goes the wrong way toward explaining the patterns. On the other hand, we do find support for the latter two mechanisms. Distorted beliefs appear relevant because, when landlords are better informed about the values of their houses, stale signals matter less. Preferences appear also to matter, as landlords with high loan-to-value ratios set higher rents than those with less debt usage, even controlling for purchase price. Critically, the direction of this effect – high LTV leading to high rents – allows us to be more specific about the manner in which reference points

[42] Estimates of the purchase/acquisition vintage dummies for this cohort are noisier than in the previous tables, due to the limited sample size.

matter to landlords. If the source of reference-dependent utility stemmed from liquidity constraints or other costs related to cash management, we would expect (as we show in a simple model) landlords to set lower rents, whereas psychological costs stemming from loss aversion would rather lead them to set higher rents. Our findings on this point thus suggest that, at least in part, prospect theory considerations appear relevant for landlords.

We regard our empirical evidence on distorted beliefs as important, insofar as it documents how market participants in the housing market use past prices to set rent expectations. This link between past prices and expected dividends (rents) has been suggested in previous work on extrapolative expectations (see for example Glaeser and Nathanson, 2017 and Kaplan, Mitman, and Violante, 2017). However, existing contributions have mainly focused on investors extrapolating from recent price growth, and on how this behavior may amplify market trends. Our study finds that, even after a decade, past price fluctuations may continue to influence current estimates of values, similar to macroeconomic experiences that also appear to influence beliefs over very long horizons (Nagel and Malmenadier, 2011, 2016). In this sense, our results suggest that backward looking behavior may also contribute to sluggishness in the time series of rents, even in aggregate.

Studying rents rather than prices represents also, in our view, an important contribution to the existing behavioral literature. Most studies of reference dependence in financial markets study consecutive transactions in the same market — e.g., buying and selling a house (Genesove and Mayer, 2001), buying and selling a stock (Odean, 1998), issuing debt at different times (Dougal, Engelberg, Parsons, and Van Wesep, 2015). The evidence presented in this paper suggests that in addition to these "intramarket" effects, reference points appear capable of *spanning* markets: those established in one context (sales) can influence outcomes in another (rents).

To the extent that this generalizes, the impact of reference points is potentially much larger than previously recognized. For example, do landlords that buy when aggregate prices are high spend more on maintenance expenses? Or, even beyond real estate, when short-selling, do institutions having bought their shares for high prices (relative to the current market) charge higher fees when lending their shares? In such settings, reference points may exert a near-continuous "flow" influence, in addition to their saltatory impact at time-of-sale. For assessing magnitudes, this is particularly important for real estate: whereas only about 6% of U.S. homes sell every year (see Piazzesi and Schneider, 2009), almost half of housing units in California are rented.

References

Albert, Sole-Olle, and Elisabet Viladecans-Marsal, 2019, Housing Booms and Local Spending, *Journal of Urban Economics*.

Allen, Marcus T., Ronald C. Rutherford, and Thomas A. Thomson, 2009, Residential Asking Rents and Time on the Market, *Journal of Real Estate Finance and Economics*.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber, 2005, Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools, *Journal of Political Economy*.

Andersen, Steffen, Christian Badarinza, Lu Liu, Julie Marx, and Tarun Ramadorai, 2019, Reference Dependence in the Housing Market, *Working Paper*.

Anenberg, Elliot, 2011, Loss Aversion, Equity Constraints and Seller Behavior in Real Estate Markets, *Regional Science & Urban Economics*.

Anenberg, Elliot, and Edward Kung, 2014, Estimates of the Size and Source of the Price Declines Due to Nearby Foreclosures, *American Economic Review*.

Argyle, Bronson, Taylor Nadauld, and Christopher Palmer, 2020, Monthly Payment Targeting and the Demand for Maturity, *Review of Financial Studies*.

Bailey, Michael, Ruiqing Cao, Theresa Kuchler, Johannes Stroebel, and Arlene Wong, 2018, The Economic Effects of Social Networks: Evidence from the Housing Market, *Journal of Political Economy*.

Barberis, Nicholas C., 2013, Thirty Years of Prospect Theory in Economics: A Review and Assessment, *Journal of Economic Perspectives*.

Beggs, Alan, and Kathrin Graddy, 2009, Anchoring Effects: Evidence from Art Auctions, *American Economic Review*.

Begley, Jaclene, Laura Loewenstein, and Paul S Willen, 2019, The Rent-Price Ratio During the Boom and Bust: Measurement and Implications, *Working Paper*.

Bergman, Oscar, Tore Ellingsen, Magnus Johannesson, and Cicek Svensson, 2010, Anchoring and Cognitive Ability, *Economics Letters*.

Bernile, Gennaro, Vineet Bhagwat, and Rau P. Raghavendra, 2017, What Doesn't Kill You Will Only Make You More Risk-Loving: Early-Life Disasters and CEO Behavior, *Journal of Finance*.

Birru, Justin, 2015, Confusion of Confusions: A Test of the Disposition Effect and Momentum, *Review of Financial Studies*.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, 2019, Memory and Reference Prices: an Application to Rental Choice, *American Economic Review Papers and Proceedings*.

Campbell, John Y., Stefano Giglio, and Parag Pathak, 2011, Forced Sales and House Prices, *American Economic Review*.

Campbell, Sean D., Davis A. Morris, Joshua Gallin, and Robert F. Martin, 2009, What Moves Housing Markets: A Variance Decomposition of the Rent-Price Ratio, *Journal of Urban Economics*.

Cheng, Ing-Haw, Sahil Raina, and Wei Xiong, 2014, Wall Street and the Housing Bubble, *American Economic Review*.

Clarke, Wyatt, and Matthew Freedman, 2019, The Rise and Effect of Homeowners Associations, *Journal of Urban Economics*.

Demers, Andrew, and Andrea Eisfeldt, 2018, Total Returns to Single Family Rentals, *Working Paper*.

Dougal, Casey, Joseph Engelberg, Christopher A. Parsons, and Edward D. Van Wesep, 2015, Anchoring on Credit Spreads, *Journal of Finance*.

Dougal, Casey, Pengjie Gao, William J. Mayew, and Christopher A. Parsons, 2019, What's in a (School) Name? Racial Discrimination in Higher Education Bond Markets, *Journal of Financial Economics*.

Duca, John V., and David C. Ling, 2018, The Other (Commercial) Real Estate Boom and Bust: The Effects of Risk Premia and Regulatory Capital Arbitrage, *Journal of Banking and Finance*.

Eichholtz, Piet, Matthijs Korevaar, Thies Lindenthal, and Ronan Tallec, 2020, The Total Return and Risk of Residential Real Estate, *Working Paper*.

Epley, Nicholas, and Thomas Gilovich, 2006, The Anchoring-and-Adjustment Heuristic, *Psychological Science*.

Furham, Adrian, and Hua Chu Boo, 2011, A Literature Review of the Anchoring Effect, *The Journal of Socio-Economics*.

Gao, Meng, Jiekun Huang, and Itay Goldstein, 2019, Informing the Market: The Effect of Modern Information Technologies on Information Production, *Review of Financial Studies*.

Gao, Zhenyu, Michael Sockin, and Wei Xiong, 2019, The Economic Consequences of Housing Speculation, *Review of Financial Studies*.

Genesove, David, 1999, The Nominal Rigidity of Apartment Rents, *Review of Economics and Statistics*.

Genesove, David, and Lu Han, 2012, Search and matching in the housing market, *Journal of Urban Economics*.

Genesove, David, and Christopher Mayer, 1997, Equity and Time to Sale in the Real Estate Market, *American Economic Review*.

Genesove, David, and Christopher Mayer, 2001, Loss Aversion and Seller Behavior: Evidence from the Housing Market, *Quarterly Journal of Economics*.

Giglio, Stefano, Matteo Maggiori, and Johannes Stroebel, 2015, Very Long-Run Discount Rates, *Quarterly Journal of Economics*.

Giglio, Stefano, Matteo Maggiori, and Johannes Stroebel, 2016, No-Bubble Condition: Model-Free Tests of in Housing Markets, *Econometrica*.

Gilbukh, Sonia, Andrew Haughwout, and Joseph Tracy, 2017, The Price to Rent Ratio: A Macroprudential Application, *Working Paper*.

Glaeser, Edward L., 2013, A Nation of Gamblers: Real Estate Speculation and American History, *American Economic Review*.

Glaeser, Edward L., and Charles G. Nathanson, 2017, An extrapolative model of house price dynamics, *Journal of Financial Economics*.

Guren, Adam, 2018, House Price Momentum and Strategic Complementarity, *Journal of Political Economy*.

Head, Allend, Lloyd-Ellis Huw, and Hongfei Sun, 2014, Search, Liquidity, and the Dynamics of House Prices and Construction, *The American Economic Review*.

Heimer, Rawley Z, Kristian Ove R. Myrseth, and Raphael S. Schoenle, 2019, YOLO: Mortality Beliefs and Household Finance Puzzles, *Journal of Finance*.

Jorda, Oscar, Katharina Knoll, Dmitry Kuvshinov, Moritz Schularick, and Alan M Taylor, 2019, The Rate of Return on Everything, *Quarterly Journal of Economics*.

Kahneman, Daniel, and Amos Tversky, 1979, Prospect Theory: An Analysis of Decision under Risk, *Econometrica*.

Kahneman, Daniel, and Amos Tversky, 1991, Loss Aversion in Riskless Choice: A Reference-Dependent Model, *Quarterly Journal of Economics*.

Kahneman, Daniel, and Amos Tversky, 1992, Advances in Prospect Theory: Cumulative Representation of Uncertainty, *Journal of Risk and Uncertainty*.

Kaplan, Greg, Kurt Mitman, and Giovanni L. Violante, 2017, The Housing Boom and Bust: Model Meets Evidence, Working paper, NBER Working Paper.

Korevaar, Matthijs, 2019, Reach for Yield and Real Estate: The Impact of Investor Demand on House Prices, *Working Paper*.

Lazear, Edward, 1986, Retail Pricing and Clearance Sales, *American Economic Review*.

Mussweiler, Thomas, and Fritz Strack, 2001, The Semantics of Anchoring, *Organizational Behavior and Human Decision Processes*.

Nagel, Stefan, and Ulrike Malmenadier, 2011, Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?, *Quarterly Journal of Economics*.

Nagel, Stefan, and Ulrike Malmenadier, 2016, Learning from Inflation Experiences, *Quarterly Journal of Economics*.

Odean, Terrance, 1998, Are Investors Reluctant to Realize Losses?, *Journal of Finance*.

Oster, Emily, 2019, Unobservable Selection and Coefficient Stability: Theory and Evidence, *Journal of Business Economics and Statistics*.

Piazzesi, Monika, and Martin Schneider, 2009, Momentum Traders in the Housing Market: Survey Evidence and a Search Model, *American Economic Review*.

Piazzesi, Monika, Martin Schneider, and Johannes Stroebel, 2020, Segmented Housing Search, *American Economic Review*.

Shiller, Robert, 2006, Long-Term Perspectives on the Current Boom in Home Prices, *The Economist's Voice*.

Silverman, B., 1984, *Density Estimation for Statistics and Data Analysis*. (Chapman and Hall).

Stein, Jeremy, 1995, Prices and Trading Volume in the Housing Market: A Model with Downpayment Constraints, *Quarterly Journal of Economics*.

Stroebel, Johannes, and Pablo Kurlat, 2015, Testing for Information Asymmetries in Real Estate Markets, *Review of Financial Studies*.

Stull, William J., 1978, The Landlord's Dilemma: Asking Rent Strategies in a Heterogeneous Housing Market, *Journal of Urban Economics*.

Xiao, Steven, and Serena Xiao, 2019, Market Power and Consumer Welfare: Evidence from Home Rental Markets, *Working Paper*.

Tables

	N. Obs.	Mean	Median	Std. Dev.
size (sqft)	44,237	1,753	1,548	1,052
1 bed	44,237	0.0564	0	0.2308
2 bed	44,237	0.2410	0	0.4277
3 bed	44,237	0.3915	0	0.4881
4-plus bed	44,237	0.3111	0	0.4629
1 bath	44,237	0.1636	0	0.3699
2 bath	44,237	0.3776	0	0.4848
3 bath	44,237	0.3323	0	0.4710
4-plus bath	44,237	0.1266	0	0.3326
condo/apt.	44,237	0.1765	0	0.3813
townhouse	44,237	0.0770	0	0.2667
multi-family	44,237	0.0549	0	0.2277
studio	44,237	0.0068	0	0.0823
age	38,648	41.53	38	25.85
listing by agent	44,237	0.3944	0	0.4887
inventory days	42,811	39.44	23	48.05
monthly rent ($)	44,237	3,524	2,750	3,409
last purchase price ($)	42,938	590,092	408,000	702,820
last purchase year	44,237	2010.6	2013	7.12

Table 1: Summary statistics for the rental listings dataset. The data consist of listings posted from December 2018 through March 2019, across the state of California.

	(1) Log Rent	(2) Log Rent	(3) Log Rent	(4) Log Rent	(5) Log Rent zip Inv. Days < Median	(6) Log Rent	(7) Log Rent
log($p_{last,ZIP}$)						0.0536***	
						(0.00511)	
log($p_{last,CA}$)							0.0539***
							(0.00461)
dummy 1980s	-0.129***	-0.104***	-0.105***	-0.0945***	-0.0698***		
	(0.0151)	(0.00913)	(0.0122)	(0.0123)	(0.0124)		
dummy 1990:1994	-0.0927***	-0.0802***	-0.0742***	-0.0657***	-0.0500***		
	(0.0130)	(0.00777)	(0.00994)	(0.00999)	(0.00955)		
dummy 1995:1998	-0.0510***	-0.0630***	-0.0538***	-0.0487***	-0.0352***		
	(0.0100)	(0.00622)	(0.00758)	(0.00770)	(0.00633)		
dummy 1999:2001	-0.0510***	-0.0649***	-0.0575***	-0.0543***	-0.0371***		
	(0.00875)	(0.00574)	(0.00699)	(0.00698)	(0.00706)		
dummy 2002:2004	-0.0381***	-0.0506***	-0.0425***	-0.0400***	-0.0366***		
	(0.00715)	(0.00471)	(0.00523)	(0.00522)	(0.00532)		
dummy 2005:2007	-0.0280***	-0.0341***	-0.0288***	-0.0275***	-0.0275***		
	(0.00786)	(0.00456)	(0.00518)	(0.00517)	(0.00523)		
dummy 2008:2010	-0.0554***	-0.0580***	-0.0536***	-0.0518***	-0.0504***		
	(0.00678)	(0.00407)	(0.00450)	(0.00453)	(0.00458)		
dummy 2011:2013	-0.0503***	-0.0486***	-0.0452***	-0.0442***	-0.0465***		
	(0.00615)	(0.00399)	(0.00432)	(0.00431)	(0.00469)		
dummy 2014:2016	-0.0174***	-0.0168***	-0.0154***	-0.0161***	-0.0231***		
	(0.00557)	(0.00366)	(0.00389)	(0.00389)	(0.00403)		
age			-0.00154***	-0.00149***	-0.00140***	-0.00146***	-0.00147***
			(0.000416)	(0.000398)	(0.000406)	(0.000414)	(0.000399)
age-sq			9.77e-06**	9.66e-06***	7.19e-06*	9.72e-06**	9.68e-06***
			(3.88e-06)	(3.67e-06)	(3.78e-06)	(3.86e-06)	(3.68e-06)
log(inventory days)			0.0167***	0.0166***	0.00887***	0.0171***	0.0167***
			(0.00165)	(0.00158)	(0.00163)	(0.00164)	(0.00158)
log(size)		0.380***	0.378***	0.372***	0.267***	0.376***	0.372***
		(0.0173)	(0.0187)	(0.0185)	(0.0230)	(0.0191)	(0.0185)
agent listing		-0.0101***	-0.0123***	-0.0143***	-0.0108***	-0.0155***	-0.0144***
		(0.00240)	(0.00266)	(0.00266)	(0.00268)	(0.00273)	(0.00269)
shared laundry		-0.0553***	-0.0530***	-0.0553***	-0.0429***	-0.0574***	-0.0553***
		(0.00779)	(0.00833)	(0.00791)	(0.00891)	(0.00820)	(0.00789)
townhouse		-0.0966***	-0.107***	-0.106***	-0.0831***	-0.105***	-0.106***
		(0.00574)	(0.00648)	(0.00634)	(0.00516)	(0.00649)	(0.00635)
condo		-0.110***	-0.120***	-0.123***	-0.102***	-0.122***	-0.122***
		(0.00748)	(0.00884)	(0.00877)	(0.0108)	(0.00900)	(0.00877)
multi-family		-0.136***	-0.146***	-0.140***	-0.110***	-0.143***	-0.141***
		(0.00757)	(0.00877)	(0.00863)	(0.00886)	(0.00897)	(0.00866)
street parking		-0.0354***	-0.0400***	-0.0521***	-0.0361***	-0.0552***	-0.0539***
		(0.0109)	(0.0124)	(0.0123)	(0.0136)	(0.0128)	(0.0123)
studio		-0.380***	-0.407***	-0.404***	-0.363***	-0.404***	-0.401***
		(0.0304)	(0.0352)	(0.0346)	(0.0413)	(0.0354)	(0.0346)
refrigerator				0.0235***	0.0151**	0.0236***	0.0232***
				(0.00390)	(0.00609)	(0.00400)	(0.00390)
dishwasher				-0.00122	-0.00636**	-0.00142	-0.00145
				(0.00288)	(0.00280)	(0.00293)	(0.00288)
no pets				-0.0188***	-0.0216***	-0.0186***	-0.0186***
				(0.00273)	(0.00315)	(0.00281)	(0.00273)
hardwood floor				0.0317***	0.0314***	0.0319***	0.0317***
				(0.00276)	(0.00284)	(0.00282)	(0.00275)
forced heat				0.0123***	0.0169***	0.0126***	0.0126***
				(0.00292)	(0.00307)	(0.00293)	(0.00293)
no AC				-0.0267***	-0.00235	-0.0259***	-0.0261***
				(0.00804)	(0.00523)	(0.00777)	(0.00804)
central AC				0.0301***	0.0109**	0.0300***	0.0303***
				(0.00383)	(0.00444)	(0.00381)	(0.00383)
floor (condo)				0.00297**	0.00228**	0.00272**	0.00287**
				(0.00126)	(0.00116)	(0.00125)	(0.00124)
bedrm dummies	NO	YES	YES	YES	YES	YES	YES
bathrm dummies	NO	YES	YES	YES	YES	YES	YES
zip code FE	YES	YES	YES	YES	YES	YES	YES
F-stat 05:07 - 08:10	10.49	24.65	20.29	20.08	14.51	-	-
p-value 05:07 - 08:10	0.0012	0.0000	0.0000	0.0000	0.0002	-	-
F-stat 05:07 - 11:13	7.87	9.92	9.76	10.51	9.00	-	-
p-value 05:07 - 11:13	0.0051	0.0017	0.0018	0.0012	0.0028	-	-
Obs	43,136	43,136	36,556	36,556	17,906	34,749	36,448
R-sq	0.629	0.858	0.857	0.859	0.859	0.858	0.859

Table 2: Impact of acquisition vintage on current asked rents (see Equation (1)). In columns from 1 to 5, the Table also reports the F-statistics and relative p-values for two tests. The first one is a test of the null that the dummy coefficient for properties last purchased from 2005 to 2007 and the dummy coefficient for properties last purchased from 2008 to 2010 are equal. The second is a test of the null that the coefficients for the 2005 to 2007 and the 2011 to 2013 dummies are equal. In column 5, the sample is restricted to zip codes where the average number of days listings remain on the website is below the median across California zip codes. In column 6, we replace the acquisition vintage dummies with the log of the ZHVI price index in the month when the house was last purchased and in the zip code where the house is located (log($p_{last,zip}$)). In column 7, we replace the acquisition vintage dummies with the log of the level, in the month when the house was last purchased, of a house price index for California constructed by the authors (log($p_{last,CA}$)). Standard errors are reported in parenthesis and are clustered by zip code.

	2005:2007	2008:2010	2011:2013	2014:2016
size (sqft)	1755.5	1723.8	1736.7	1754.2
	[-1.5674]	-	[-0.7595]	[-1.7650]
fraction 1 bed	0.0600	0.0500	0.0566	0.0644
	[-1.9191]	-	[-1.5135]	[-3.2931]
fraction 2 beds	0.2507	0.2307	0.2277	0.2503
	[-2.0405]	-	[0.3770]	[-2.4754]
fraction 3 beds	0.3900	0.4066	0.3954	0.3741
	[1.4833]	-	[1.1972]	[3.6199]
fraction 4+ beds	0.2993	0.3126	0.3203	0.3111
	[1.2553]	-	[-0.8647]	[0.1752]
fraction 1 bath	0.1618	0.1576	0.1599	0.1625
	[-0.4965]	-	[-0.3266]	[-0.7136]
fraction 2 baths	0.3951	0.3771	0.3887	0.3568
	[-1.6127]	-	[-1.2534]	[2.2821]
fraction 3 baths	0.3239	0.3448	0.3314	0.3422
	[1.9310]	-	[1.4853]	[0.3052]
fraction 4+ plus baths	0.1192	0.1204	0.1199	0.1385
	[0.1677]	-	[0.0841]	[-2.8932]
fraction condo/apt.	0.1702	0.1606	0.1709	0.2032
	[-1.1336]	-	[-1.4548]	[-5.9103]
age	38.61	38.79	40.68	40.26
	[0.3010]	-	[-3.7169]	[-2.8338]
fraction street parking	0.0138	0.0097	0.0107	0.0118
	[-1.6848]	-	[-0.5373]	[-1.0934]
fraction shared laundry	0.0588	0.0496	0.0547	0.0548
	[-1.7868]	-	[-1.1805]	[-1.2562]
Obs	3,331	4,434	7,172	8,721

Table 3: Mean of house characteristics for rental houses last purchased in four different vintages: years from 2005 to 2007, from 2008 to 2010, from 2011 to 2013 and from 2014 to 2016. We report in square brackets t-statistics for a test of the difference in means between the 2008 to 2010 vintage and each one of the other three vintages.

	(1) Log Rent	(2) Log Rent	(3) Log Rent	(4) Log Rent
dummy 1980s	-0.0884*			
	(0.0500)			
dummy 1990:1994	-0.0752***			
	(0.0253)			
dummy 1995:1998	-0.0616***			
	(0.0124)			
dummy 1999:2001	-0.0652***			
	(0.00873)			
dummy 2002:2004	-0.0394***			
	(0.00644)			
dummy 2005:2007	-0.0252***			
	(0.00595)			
dummy 2008:2010	-0.0505***			
	(0.00489)			
dummy 2011:2013	-0.0437***			
	(0.00457)			
dummy 2014:2016	-0.0153***			
	(0.00408)			
$\log(p_{last})$		0.0829***	0.0788***	0.0788***
		(0.00495)	(0.00501)	(0.00501)
$\log(p_{last}) - \pi_{last}$	0.148***	0.0655***	0.0696***	0.0427**
	(0.0128)	(0.0118)	(0.0117)	(0.0173)
$(\log(p_{last}) - \pi_{last})$ ×dummy 1980s				0.143**
				(0.0558)
$(\log(p_{last}) - \pi_{last})$ × dummy 1990:1994				-0.0517
				(0.0814)
$(\log(p_{last}) - \pi_{last})$ × dummy 1995:1998				0.109**
				(0.0536)
$(\log(p_{last}) - \pi_{last})$ × dummy 1999:2001				0.0872**
				(0.0388)
$(\log(p_{last}) - \pi_{last})$ × dummy 2002:2004				0.0930*
				(0.0485)
$(\log(p_{last}) - \pi_{last})$ × dummy 2005:2007				0.123***
				(0.0370)
$(\log(p_{last}) - \pi_{last})$ × dummy 2008:2010				-0.0106
				(0.0261)
$(\log(p_{last}) - \pi_{last})$ × dummy 2011:2013				0.0260
				(0.0225)
$(\log(p_{last}) - \pi_{last})$ × dummy 2014:2016				0.0380*
				(0.0211)
age	-0.00124***		-0.000877***	-0.000886***
	(0.000307)		(0.000301)	(0.000306)
age-sq	7.58e-06***		4.90e-06*	4.82e-06*
	(2.63e-06)		(2.53e-06)	(2.61e-06)
log(inventory days)	0.0164***		0.0160***	0.0160***
	(0.00172)		(0.00171)	(0.00171)
additional controls	YES	YES	YES	YES
bedrm dummies	YES	YES	YES	YES
bathrm dummies	YES	YES	YES	YES
zip code FE	YES	YES	YES	YES
F-stat 05:07 - 08:10	14.21	-	-	-
p-value 05:07 - 08:10	0.0002	-	-	-
F-stat 05:07 - 11:13	8.60	-	-	-
p-value 05:07 - 11:13	0.0034	-	-	-
Obs	29,640	29,640	29,640	29,640
R-sq	0.859	0.858	0.860	0.860

Table 4: Impact of acquisition vintage on current asked rent, after controlling for unobservable characteristics, using the approach in Genesove and Mayer (2001). p_{last} is the last purchase price of the rental property, while π_{last} is the hedonic estimate of the log price based on Equation (4). The *additional controls* include all the controls used in column 4 (excluding age, age squared and log inventory days) of Table 2. Standard errors are reported in parenthesis and are clustered by zip code.

	(1) Log Rent	(2) Log Rent	(3) Log Rent	(4) Log Rent
dummy 1980s	-0.0373	-0.0576	-0.0554	-0.0584
	(0.0285)	(0.0445)	(0.0447)	(0.0448)
dummy 1990:1994	-0.0866***	-0.0788***	-0.0765***	-0.0777***
	(0.0268)	(0.0260)	(0.0260)	(0.0261)
dummy 1995:1998	-0.0588***	-0.0558***	-0.0538***	-0.0552***
	(0.00818)	(0.00801)	(0.00782)	(0.00782)
dummy 1999:2001	-0.0570***	-0.0553***	-0.0533***	-0.0544***
	(0.00691)	(0.00680)	(0.00658)	(0.00655)
dummy 2002:2004	-0.0407***	-0.0377***	-0.0360***	-0.0365***
	(0.00636)	(0.00636)	(0.00623)	(0.00621)
dummy 2005:2007	-0.0336***	-0.0306***	-0.0292***	-0.0296***
	(0.00600)	(0.00615)	(0.00609)	(0.00608)
dummy 2008:2010	-0.0565***	-0.0523***	-0.0515***	-0.0530***
	(0.00464)	(0.00479)	(0.00474)	(0.00473)
dummy 2011:2013	-0.0460***	-0.0435***	-0.0438***	-0.0453***
	(0.00473)	(0.00475)	(0.00479)	(0.00475)
dummy 2014:2016	-0.0145***	-0.0139***	-0.0131***	-0.0134***
	(0.00490)	(0.00478)	(0.00478)	(0.00476)
absentee			0.00691**	0.00391
			(0.00347)	(0.00363)
corporation/legal entity				0.0178***
				(0.00683)
$(\log(p_{last}) - \pi_{last})$		0.178***	0.179***	0.178***
		(0.0113)	(0.0115)	(0.0115)
age	-0.00119***	-0.00101***	-0.00102***	-0.00102***
	(0.000387)	(0.000379)	(0.000378)	(0.000378)
age-sq	1.23e-05***	1.09e-05***	1.09e-05***	1.09e-05***
	(3.22e-06)	(3.22e-06)	(3.22e-06)	(3.22e-06)
log(inventory days)	0.0157***	0.0146***	0.0146***	0.0145***
	(0.00171)	(0.00157)	(0.00156)	(0.00157)
bedrm dummies	YES	YES	YES	YES
bathrm dummies	YES	YES	YES	YES
zip code FE	YES	YES	YES	YES
F-stat $_{05:07\text{ - }08:10}$	16.75	14.56	15.52	16.93
p-value $_{05:07\text{ - }08:10}$	0.0000	0.0001	0.0001	0.0000
Obs	24,992	24,672	24,672	24,672
R-sq	0.869	0.876	0.876	0.876

Table 5: Impact of acquisition vintage on current rents, controlling for both house characteristics and landlord type. We control for landlord type by adding to the specification of Equation (1) dummies equal to one if the landlord was absentee or a corporate/legal entity at the time of purchasing the property. Corporate/legal entities are a subset of absentee landlords. Estimates are based on the sample of rental listings matched with Corelogic deed records. Standard errors are reported in parenthesis and are clustered by zip code.

	(1) Equation (1)	(2) Equation (1) with 2005:2010 dummy	(3) Equation (1) no vintage dummies
Treatment Coefficient	dummy 2008:2010	dummy 2008:2010	zip code price index
$\bar{\delta}$	-8.5641	4.6083	-9.2355
Coeff. Short	-0.0412	-0.0330	0.0435
R-square Short	0.631	0.631	0.631
Coeff. Long	-0.0518	-0.0244	0.0536
R-square Long	0.859	0.859	0.858
Vintage Dummies in Short Reg.	YES	YES	NO
zip code FE in Short Reg.	YES	YES	YES

Table 6: Statistical diagnosis of unobservables: The Table reports the value of $\bar{\delta}$ that would make the unbiased estimate of the treatment coefficient equal to zero (see Section 3.4 and Appendix B). The Table also reports estimates of the treatment coefficients and R-squares for the "short" and "long" regressions. The short regressions in columns 1 and 2 include in the conditioning information only acquisition vintage dummies and zip code fixed effects. The long regressions include all controls from column 4 of Table 2. In column 1, year of purchase dummies have the same the specification as in Equation (1), and the treatment of interest is the dummy coefficient for properties last purchased in the years from 2008 to 2010. In column 2 the specification is changed. We replace the dummy for the 2005 to 2007 vintage with a dummy equal to one for all houses purchased from 2005 to 2010. In this new specification, the treatment of interest is the dummy for houses purchased from 2008 to 2010, which directly measures the peak-bust rental spread. Finally, in column 3 we consider a specification in which, instead of a vintage dummy, the treatment variable is the zip code house price index at the time of purchase (same specification as in column 6 of Table 2). The short regression includes only the zip code price index and zip code fixed effects, while the long regression includes zip code fixed effects along with the full set of controls for house and rental characteristics.

	(1) Log Inv. Days AFT	(2) Log Inv. Days AFT	(3) Log Inv. Days AFT 2005:2010
log(R)	0.597***	1.324***	
	(0.0409)	(0.404)	
log(R)-sq		-0.0430*	
		(0.0235)	
dummy 2005:2007			0.0633*
			(0.0334)
additional controls	YES	YES	YES
bedrm dummies	YES	YES	YES
bathrm dummies	YES	YES	YES
zip code FE	YES	YES	YES
Weibull Par. 95% C.I.	[1.39,1.42]	[1.39,1.42]	[1.55,1.64]
Obs	36,331	36,331	6,061

Table 7: Effect of asked rents on time on market, based on the accelerated failure time (AFT) model in Equation (5). The dependent variable is the log of the number of days the house remains listed. In column 3 the sample is restricted to homes last purchased between January 2005 and December 2010. Standard errors are reported in parenthesis and are clustered by zip code.

	(1) Log Rent	(2) Log Rent	(3) Log Rent	(4) Log Rent	(5) Log Rent
$log(p_{last,zip})$	0.0527***	0.0198**	0.0349***	0.0287***	0.0299***
	(0.00569)	(0.00841)	(0.00678)	(0.00563)	(0.00518)
$log(p_{last,zip}) \times log(1 + NumMatch)$	-0.0388***				
	(0.00705)				
$log(1 + NumMatch)$	0.502***				
	(0.0943)				
$log(p_{last,zip}) \times AtypicalSize$		0.0275***			
		(0.00817)			
$AtypicalSize$		-0.363***			
		(0.107)			
$log(p_{last,zip}) \times AtypicalBeds$			0.0104*		
			(0.00599)		
$AtypicalBeds$			-0.104		
			(0.0783)		
$log(p_{last,zip}) \times I(HighStdSizeZip)$				0.0202**	
				(0.00859)	
$log(p_{last,zip}) \times I(HighStdBedsZip)$					0.0178**
					(0.00837)
$\log(p_{last}) - \pi_{last}$	0.129***	0.115***	0.117***	0.130***	0.130***
	(0.0105)	(0.0107)	(0.0109)	(0.0106)	(0.0106)
$\sigma(Information\ Environment\ Variable)$	0.48	1.15	0.73	-	-
additional controls	YES	YES	YES	YES	YES
bedrm dummies	YES	YES	YES	YES	YES
bathrm dummies	YES	YES	YES	YES	YES
zip code FE	YES	YES	YES	YES	YES
Obs	28,844	21,780	21,763	28,844	28,844
R-sq	0.861	0.868	0.868	0.861	0.861

Table 8: Information environment and peak-bust rental spread. The table reports estimates of the interaction terms between the (log of the) zip code price index at the time of purchase ($\log(p_{last,zip})$) and a measure of the quality of information available at the time of listing. We consider five different measures. Our first measure ($NumMatch$) is equal, for each rental listing, to the number of other listings located within a one-mile radius of the considered listing, and that have the same number of bedrooms and size within 75 square feet. We then construct two meausure of "atypicality" of the listed house ($AtypicalSize$ and $AtypicalBeds$), measured as the absolute value of the difference between the average size and number of bedrooms in the zip code, and the size and number of bedrooms of the listing. The differences are standardized by the standard deviation of each characteristic within the zip code. Finally, we use two zip code-level measures, that assess the degree of heterogeneity of local properties. These are dummies ($I(HighStdSizeZip)$ and $I(HighStdBedsZip)$) equal to one if the standard deviation of size, or the standard deviation of the number of bedrooms, is above the median across zip codes. Standard errors are reported in parenthesis and are clustered by zip code.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Log Rent	Log Rent	Log Rent	Log Rent	Log Rent	Log Rent	Log Rent
	Only *High* and *Low* LTV_{last}						
$I(LTV_{last} > 95\%)$	0.0491***	0.0498***	0.0235***	0.0237***			
	(0.00950)	(0.00934)	(0.00407)	(0.00407)			
$I(LTV_{last} \leq 50\%)$			-0.0243***	-0.0247***			
			(0.00680)	(0.00672)			
LTV_{last}						0.0702***	0.0933***
						(0.00991)	(0.00988)
$log(p_{last})$	0.0844***	0.0868***	0.0678***	0.0678***	0.0652***		0.0710***
	(0.0111)	(0.0108)	(0.00508)	(0.00507)	(0.00507)		(0.00510)
$\log(p_{last}) - \pi_{last}$	0.0894***	0.0744	0.147***	0.156***	0.154***	0.220***	0.155***
	(0.0285)	(0.0841)	(0.0122)	(0.0349)	(0.0356)	(0.0350)	(0.0352)
$\log(p_{last}) - \pi_{last} \times$ vintage	NO	YES	NO	YES	YES	YES	YES
additional controls	YES	YES	YES	YES	YES	YES	YES
bedrm dummies	YES	YES	YES	YES	YES	YES	YES
bathrm dummies	YES	YES	YES	YES	YES	YES	YES
zip code FE	YES	YES	YES	YES	YES	YES	YES
Obs	3,947	3,947	19,730	19,730	19,730	19,730	19,730
R-sq	0.894	0.895	0.872	0.872	0.872	0.871	0.872

Table 9: Effect of initial leverage at the time of purchase on current rents. LTV_{last} is the loan-to-value ratio of the mortgage on the rental property at the time of purchase, while p_{last} is the purchase price, and π_{last} is the hedonic estimate of the log purchase price based on regression Equation (4). Estimates are based on the sample of rental listings matched with Corelogic deed records. In columns 1 and 2 the sample is restricted to properties with original LTV_{last} greater than 0.95 or smaller or equal than 0.50 (*High* and *Low* LTV_{last}). The *additional controls* contain controls for house characteristics as well as for landlord type (dummies equal to one for absentee landlords and landlords who are corporate/legal entities). Standard errors are reported in parenthesis and are clustered by zip code.

	(1)	(2)	(3)	(4)
	Log Rent	Log Rent	Log Rent	Log Rent
	$LTV_{last} \leq 0.50$	$LTV_{last} \leq 0.75$	$LTV_{last} \leq 0.50$	$LTV_{last} \leq 0.75$
$\log(p_{last,zip})$			0.0926***	0.0407***
			(0.0245)	(0.00922)
dummy 1980s		0.439***		
		(0.0325)		
dummy 1990:1994	0.295	0.0725		
	(0.455)	(0.115)		
dummy 1995:1998	-0.144***	-0.0533***		
	(0.0534)	(0.0206)		
dummy 1999:2001	-0.112***	-0.0612***		
	(0.0328)	(0.0130)		
dummy 2002:2004	-0.0209	-0.0214*		
	(0.0278)	(0.0121)		
dummy 2005:2007	-0.00926	-0.0232*		
	(0.0396)	(0.0120)		
dummy 2008:2010	-0.0840***	-0.0601***		
	(0.0286)	(0.00840)		
dummy 2011:2013	-0.0435	-0.0525***		
	(0.0334)	(0.00746)		
dummy 2014:2016	-0.0302	-0.0152*		
	(0.0248)	(0.00792)		
$\log(p_{last}) - \pi_{last}$	0.231***	0.222***	0.217***	0.214***
	(0.0537)	(0.0202)	(0.0493)	(0.0197)
age	-0.00118	-0.00144**	-0.00132	-0.00156**
	(0.00192)	(0.000670)	(0.00194)	(0.000685)
age-sq	5.86e-06	1.20e-05**	7.17e-06	1.33e-05**
	(1.78e-05)	(5.71e-06)	(1.82e-05)	(5.86e-06)
log(inventory days)	0.0141*	0.0156***	0.0128*	0.0155***
	(0.00725)	(0.00247)	(0.00710)	(0.00250)
additional controls	YES	YES	YES	YES
bedrm dummies	YES	YES	YES	YES
bathrm dummies	YES	YES	YES	YES
zip code FE	YES	YES	YES	YES
Obs	1,193	8,424	1,181	8,330
R-sq	0.885	0.871	0.888	0.871

Table 10: Further evidence on distorted beliefs: Effect of acquisition vintage on current rents for homes that were purchased with a loan-to-value ratio (LTV_{last}) smaller or equal than 0.50 (columns 1 and 3), or smaller or equal than 0.75 (columns 2 and 4). In columns 3 and 4, we replace the acquisition vintage dummies with the log of the ZHVI price index in the month when the house was last purchased and in the zip code where the house is located ($\log(p_{last,zip})$). The *additional controls* contain controls for house characteristics as well as for landlord type (dummies equal to one for absentee landlords and landlords who are corporate/legal entities). Standard errors are reported in parenthesis and are clustered by zip code.

Figures

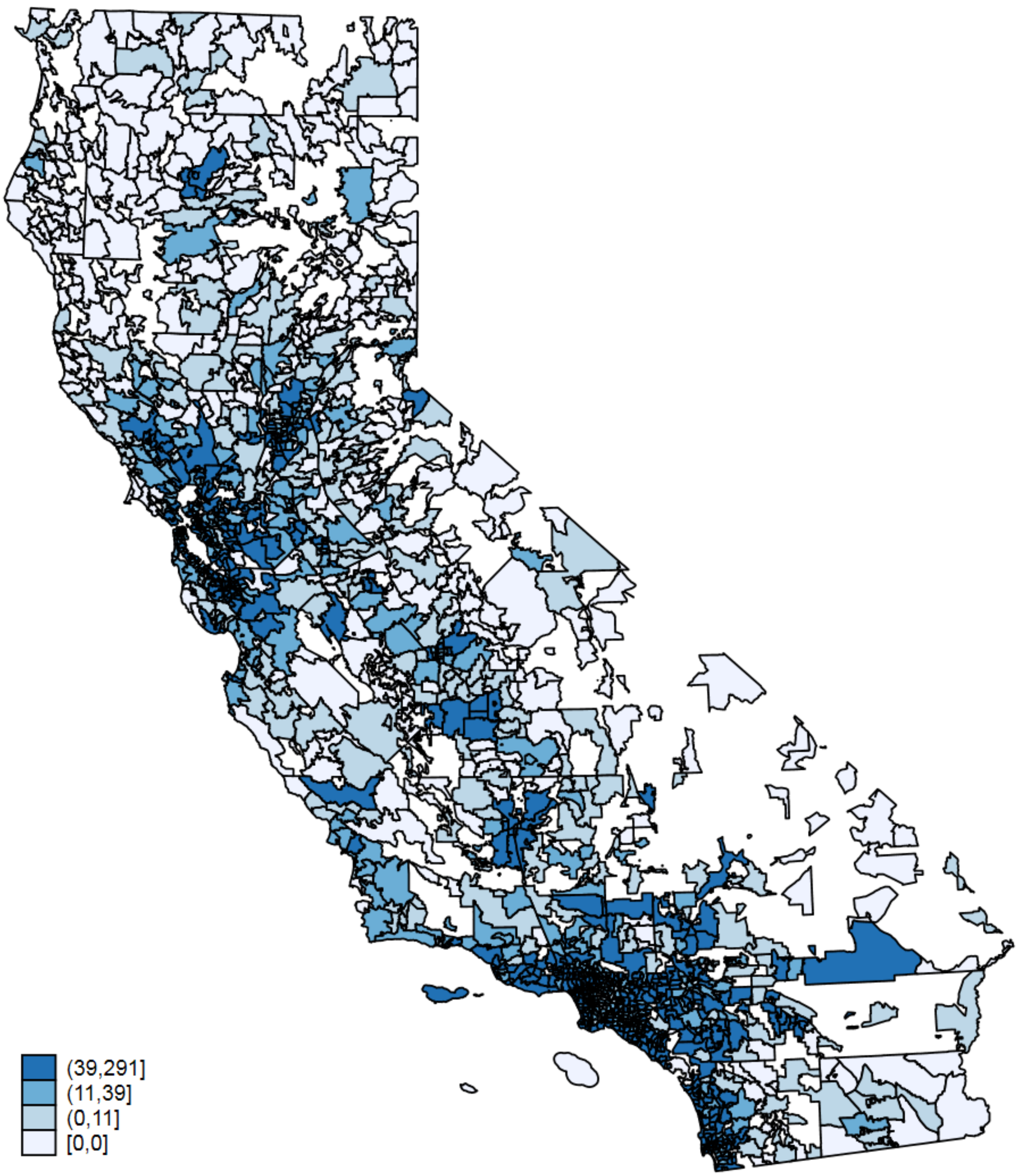

Figure 1: Heat map showing the number of rental listings per zip code. The data consist of listings posted from December 2018 through March 2019, across the state of California.

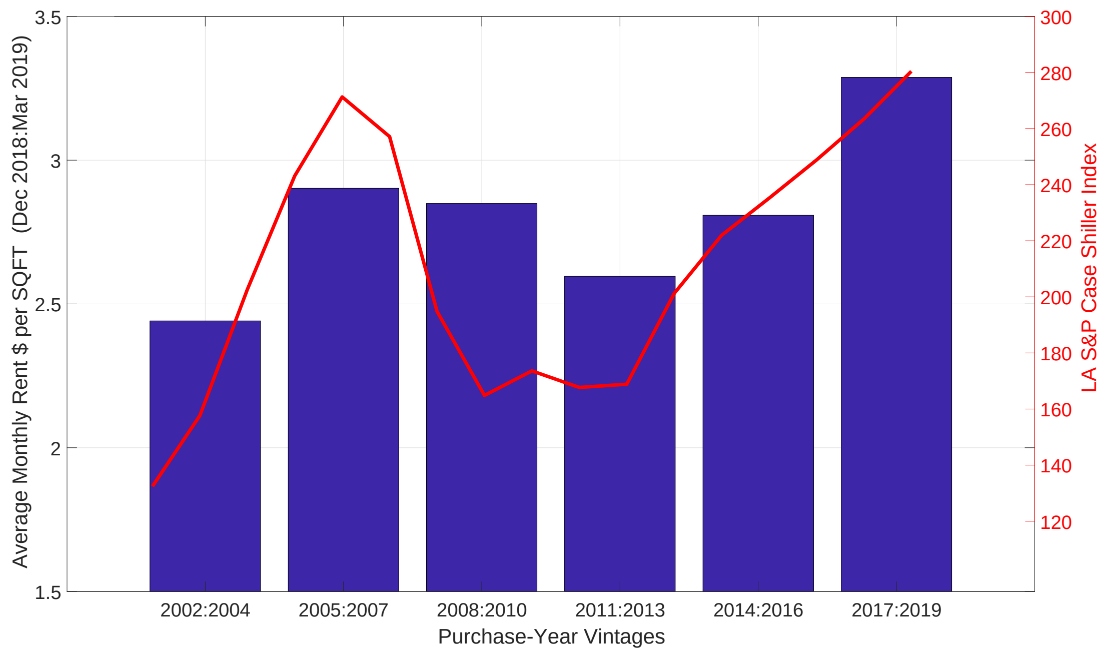

Figure 2: The blue bars show the average monthly rent per square foot for listings of 2 bedroom rental properties in the city of Los Angeles, by purchase/acquisition vintage. The sample of rental listings has been collected over the period from December 2018 to March 2019. We compare the pattern in the cross-section of rents (blue bars) against the time series evolution of aggregate house prices across acquisition vintages, which is measured using the S&P Case-Shiller repeated sale index for Los Angeles (the red line). The index is set equal to 100 in January 2000.

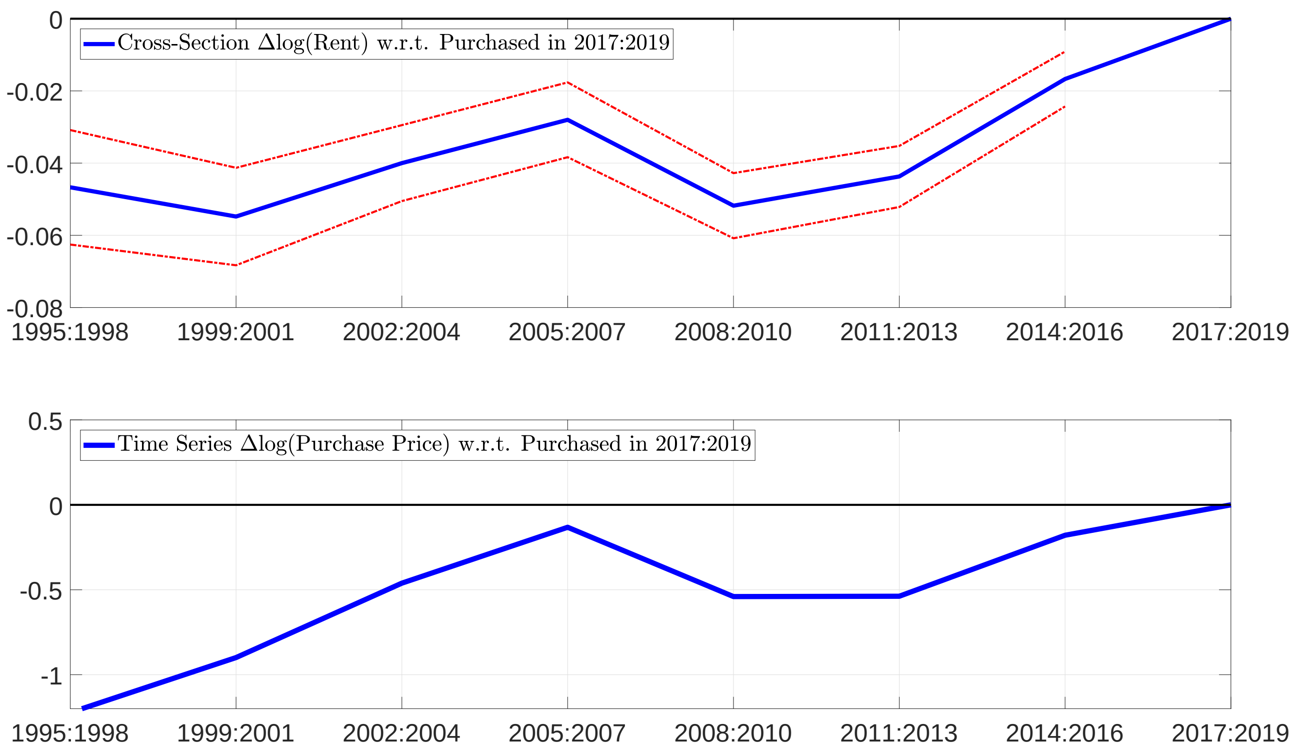

Figure 3: The top panel shows estimates of purchase/acquisition vintage dummies from the specification of Equation (1) in column 4 of Table 2. The dependent variable is log monthly rent, and the dummies measure the log differences in the cross-section of current rents between the rent asked by houses purchased in each vintage and the rent asked by houses purchased in 2017-2019. The bottom panel shows purchase/acquisition vintage year dummies from the regression specification in column 2 of Table A.2. The dependent variable is the log of the last purchase price of the house, and the estimates of the dummies in this second panel can be interpreted as an historical house price index, showing the log difference between purchase prices in each vintage and purchase prices in the 2017-2019 vintage. All estimates are based on rental listings from the state of California, collected over the period from December 2018 through March 2019.

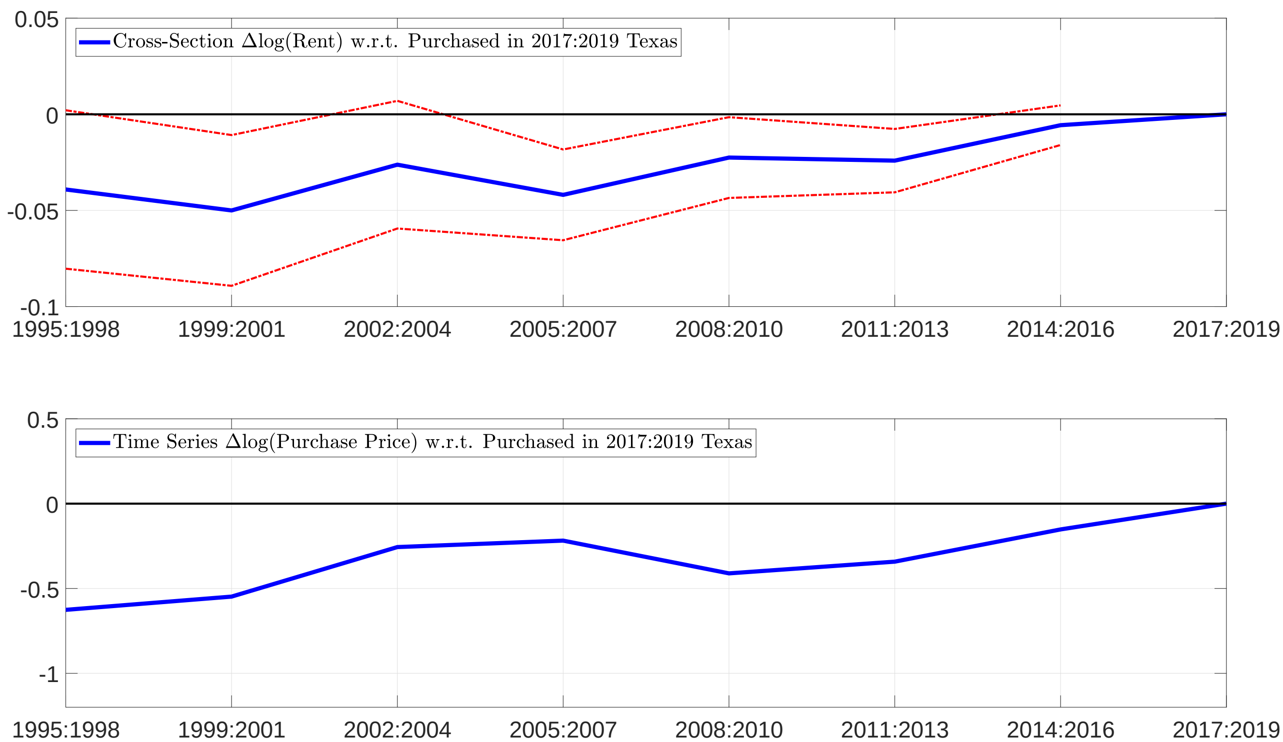

Figure 4: While Figure 3 reports evidence based on the main rental listings data for California, the Figure above reports evidence based on rental listings from the state of Texas, collected over the period from December 2018 through March 2019. Similar to Figure 3, the top panel shows estimates of purchase/acquisition vintage dummies from the specification of Equation (1) in column 4 of Table 2. The dependent variable is log monthly rent, and the dummies measure the log differences in the cross-section of current rents between the rent asked by houses purchased in each vintage and the rent asked by houses purchased in 2017-2019. The bottom panel shows purchase/acquisition vintage year dummies from the regression specification in column 2 of Table A.2, estimated using purchase prices for the Texas rental properties. The dependent variable is the log of the last purchase price of the house, and the estimates of the dummies in this second panel can be interpreted as an historical house price index, showing the log difference between purchase prices in each vintage and purchase prices in the 2017-2019 vintage.

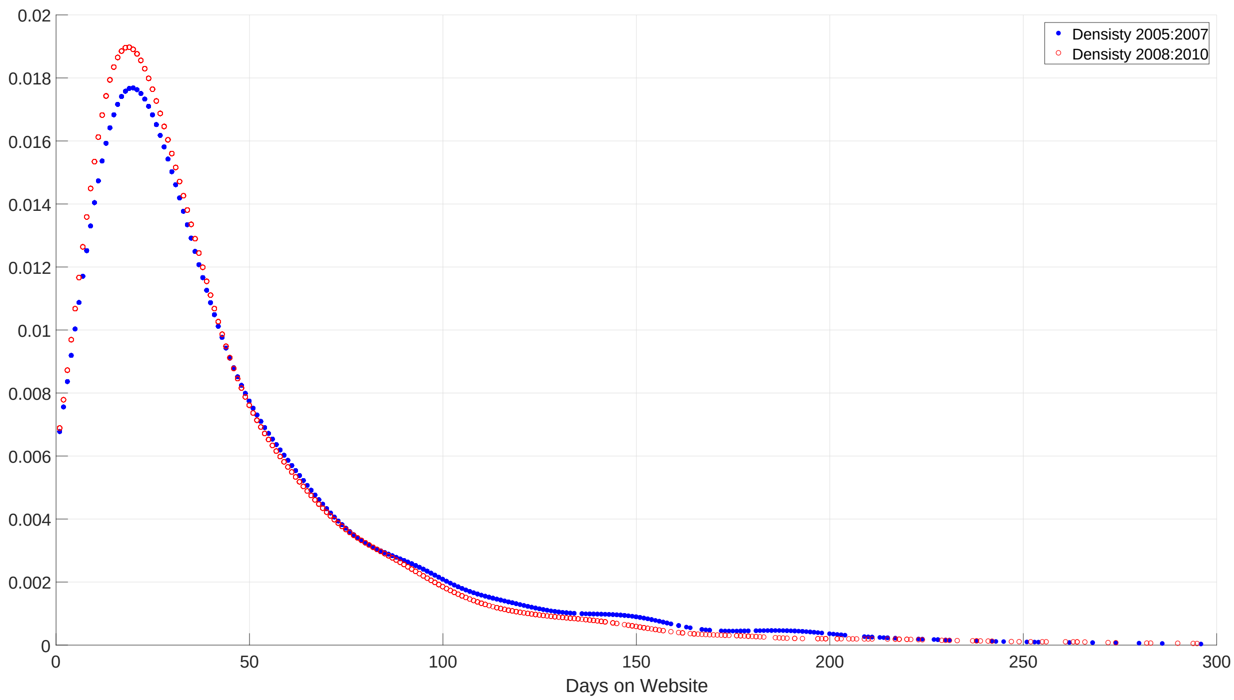

Figure 5: Distribution of time on market (inventory days) for houses purchased at the peak (2005-2007) and the bust (2008-2010) of housing markets in California. The Figure shows kernel density estimates for the number of days until the listing was removed, or the number of days the listing had been available till our last collection date, for houses that were still listed at that point. Densities are estimated using a Normal kernel with optimal bandwith chosen according to the methodology in Silverman (1984).

Appendix

A Additional Tables

	(1)	(2)	(3)	(4)
	Log Rent	Log Rent	Log Rent	Log Rent
	Excluding Houses Last Purchased in 2008			
dummy 1980s	-0.128***	-0.104***	-0.103***	-0.0926***
	(0.0149)	(0.00912)	(0.0122)	(0.0123)
dummy 1990:1994	-0.0916***	-0.0798***	-0.0738***	-0.0656***
	(0.0129)	(0.00782)	(0.00996)	(0.00996)
dummy 1995:1998	-0.0508***	-0.0620***	-0.0520***	-0.0470***
	(0.0101)	(0.00655)	(0.00784)	(0.00795)
dummy 1999:2001	-0.0495***	-0.0650***	-0.0582***	-0.0551***
	(0.00861)	(0.00562)	(0.00678)	(0.00677)
dummy 2002:2004	-0.0380***	-0.0502***	-0.0423***	-0.0400***
	(0.00711)	(0.00474)	(0.00528)	(0.00528)
dummy 2005:2007	-0.0302***	-0.0344***	-0.0297***	-0.0283***
	(0.00784)	(0.00457)	(0.00520)	(0.00517)
dummy 2008:2010	-0.0649***	-0.0587***	-0.0545***	-0.0526***
	(0.00771)	(0.00475)	(0.00514)	(0.00515)
dummy 2011:2013	-0.0506***	-0.0481***	-0.0449***	-0.0439***
	(0.00608)	(0.00395)	(0.00426)	(0.00425)
dummy 2014:2016	-0.0197***	-0.0176***	-0.0161***	-0.0167***
	(0.00549)	(0.00358)	(0.00381)	(0.00380)
age			-0.00152***	-0.00145***
			(0.000418)	(0.000398)
age-sq			9.79e-06**	9.60e-06***
			(3.90e-06)	(3.67e-06)
log(inventory days)			0.0163***	0.0162***
			(0.00164)	(0.00157)
log(size)		0.389***	0.389***	0.382***
		(0.0167)	(0.0179)	(0.0177)
agent listing		-0.00881***	-0.0109***	-0.0128***
		(0.00243)	(0.00269)	(0.00270)
shared laundry		-0.0587***	-0.0567***	-0.0582***
		(0.00809)	(0.00861)	(0.00816)
townhouse		-0.0946***	-0.105***	-0.104***
		(0.00578)	(0.00646)	(0.00633)
condo		-0.106***	-0.116***	-0.119***
		(0.00754)	(0.00892)	(0.00886)
multi		-0.129***	-0.139***	-0.133***
		(0.00752)	(0.00862)	(0.00848)
street parking		-0.0351***	-0.0395***	-0.0514***
		(0.0111)	(0.0125)	(0.0125)
studio		-0.370***	-0.392***	-0.390***
		(0.0304)	(0.0351)	(0.0345)
refrigerator				0.0232***
				(0.00395)
dishwasher				-0.000921
				(0.00287)
no pets				-0.0185***
				(0.00282)
hardwood floor				0.0316***
				(0.00282)
forced heat				0.0106***
				(0.00301)
no AC				-0.0268***
				(0.00807)
central AC				0.0316***
				(0.00400)
floor (condo)				0.00323**
				(0.00126)
bedrm dummies	NO	YES	YES	YES
bathrm dummies	NO	YES	YES	YES
zip code FE	YES	YES	YES	YES
F-stat 05:07 - 08:10	12.85	19.38	16.25	16.28
p-value 05:07 - 08:10	0.0004	0.0000	0.0001	0.0001
Obs	42,930	42,930	36,486	36,486
R-sq	0.625	0.857	0.856	0.858

Table A.1: Robustness of the effect of acquisition vintage on current rents. We replicate the analysis in the first 4 columns of Table 2, after removing from the data rental listings for houses last purchased in 2008. The Table also shows the F-statistic and the relative p-value of a test of the null that the dummy coefficient for properties last purchased from 2005 to 2007 and the dummy coefficient for properties last purchased from 2008 to 2010 are equal. Standard errors are reported in parenthesis and are clustered at the zip code level.

	(1) Log Purchase Price	(2) Log Purchase Price	(3) Log Purchase Price Dec 2018 $	(4) Log Purchase Price Dec 2018 $
dummy 1980s	-1.475***	-1.447***	-0.684***	-0.656***
	(0.0361)	(0.0382)	(0.0341)	(0.0351)
dummy 1990:1994	-1.236***	-1.244***	-0.713***	-0.719***
	(0.0215)	(0.0227)	(0.0219)	(0.0233)
dummy 1995:1998	-1.204***	-1.222***	-0.785***	-0.803***
	(0.0156)	(0.0146)	(0.0157)	(0.0146)
dummy 1999:2001	-0.877***	-0.900***	-0.533***	-0.556***
	(0.0141)	(0.0129)	(0.0140)	(0.0128)
dummy 2002:2004	-0.444***	-0.462***	-0.163***	-0.182***
	(0.0126)	(0.0118)	(0.0126)	(0.0118)
dummy 2005:2007	-0.124***	-0.134***	0.0987***	0.0882***
	(0.0153)	(0.0135)	(0.0155)	(0.0136)
dummy 2008:2010	-0.550***	-0.542***	-0.394***	-0.386***
	(0.0126)	(0.0112)	(0.0127)	(0.0112)
dummy 2011:2013	-0.545***	-0.538***	-0.437***	-0.431***
	(0.0109)	(0.00976)	(0.0109)	(0.00977)
dummy 2014:2016	-0.174***	-0.180***	-0.120***	-0.127***
	(0.00837)	(0.00719)	(0.00844)	(0.00721)
age		-0.00437***		-0.00432***
		(0.000824)		(0.000805)
age-sq		3.31e-05***		3.27e-05***
		(8.09e-06)		(7.90e-06)
log(size)		0.533***		0.534***
		(0.0209)		(0.0211)
shared laundry		0.0648***		0.0667***
		(0.0184)		(0.0184)
townhouse		-0.180***		-0.179***
		(0.0107)		(0.0108)
condo		-0.177***		-0.178***
		(0.0112)		(0.0112)
multi-family		-0.174***		-0.175***
		(0.0152)		(0.0152)
street parking		0.00958		0.00754
		(0.0269)		(0.0269)
studio		0.309***		0.310***
		(0.0573)		(0.0574)
bedrm dummies	NO	YES	NO	YES
bathrm dummies	NO	YES	NO	YES
zip code FE	YES	YES	YES	YES
Obs	41,882	36,574	41,561	36,264
R-sq	0.642	0.764	0.612	0.747

Table A.2: Average purchase prices by acquisition vintage for the rental properties in our sample. In columns 1 and 2, the dependent variable is the nominal purchase price of the property, while in columns 3 and 4 the dependent variable is the real purchase price (expressed in terms of December 2018 dollars). Standard errors are reported in parenthesis and are clustered by zip code.

	(1) Log Price	(2) Log Price	(3) Log Price	(4) Log Price
$\log(p_{last,zip})$	0.0550***	0.0830***	0.0482**	0.0407*
	(0.0129)	(0.0147)	(0.0214)	(0.0212)
$\log(p_{last,zip})\times$ absentee				0.0381***
				(0.00716)
$\log(p_{last}) - \pi_{last}$			0.305***	0.305***
			(0.0147)	(0.0147)
age		-0.00242***	-0.00246***	-0.00239***
		(0.000443)	(0.000395)	(0.000395)
age-sq		2.12e-05***	2.15e-05***	2.15e-05***
		(3.72e-06)	(3.27e-06)	(3.28e-06)
log(size)		0.580***	0.580***	0.579***
		(0.00982)	(0.00941)	(0.00941)
has fireplace		0.0208***	0.0245***	0.0234***
		(0.00397)	(0.00397)	(0.00390)
has pool		0.0908***	0.0944***	0.0921***
		(0.00569)	(0.00523)	(0.00541)
has parking space		0.0542***	0.0479***	0.0468***
		(0.00838)	(0.00873)	(0.00855)
has garage		-0.00165	0.00696	0.00648
		(0.00584)	(0.00612)	(0.00604)
absentee				-0.0862***
				(0.00831)
bedrm dummies	NO	YES	YES	YES
bathrm dummies	NO	YES	YES	YES
zip code year-month FE	YES	YES	YES	YES
Obs	216,823	189,488	136,360	136,360
R-sq	0.789	0.874	0.880	0.881

Table A.3: Impact of acquisition vintage on current house sale prices of single family residences in California over the period from January 2017 to December 2018. Acquisition vintages are captured by the log of the ZHVI price index in the month when the house was last purchased and for the zip code where the house is located ($\log(p_{last,zip})$). Standard errors are reported in parenthesis and are clustered by zip code. *absentee* is a dummy equal to one for absentee owners, who are more likely to treat the house as an investment asset, as is the case for landlords of rental properties.

	(1) Log Rent	(2) Log Rent	(3) Log Rent	(4) Log Rent	(5) Log Rent	(6) Log Rent
		≥ Median Rent			< Median Rent	
dummy 1980s	-0.109***	-0.101***	-0.0923***	-0.0958***	-0.0978***	-0.0863***
	(0.0148)	(0.0189)	(0.0191)	(0.00972)	(0.0142)	(0.0143)
dummy 1990:1994	-0.0925***	-0.0870***	-0.0789***	-0.0677***	-0.0564***	-0.0468***
	(0.0111)	(0.0141)	(0.0140)	(0.0104)	(0.0125)	(0.0125)
dummy 1995:1998	-0.0680***	-0.0596***	-0.0549***	-0.0503***	-0.0373***	-0.0313***
	(0.00964)	(0.0112)	(0.0113)	(0.00696)	(0.00825)	(0.00828)
dummy 1999:2001	-0.0690***	-0.0636***	-0.0611***	-0.0544***	-0.0426***	-0.0383***
	(0.00791)	(0.00957)	(0.00961)	(0.00712)	(0.00819)	(0.00808)
dummy 2002:2004	-0.0464***	-0.0368***	-0.0357***	-0.0530***	-0.0454***	-0.0414***
	(0.00745)	(0.00792)	(0.00794)	(0.00542)	(0.00621)	(0.00612)
dummy 2005:2007	-0.0269***	-0.0193**	-0.0183**	-0.0428***	-0.0408***	-0.0376***
	(0.00692)	(0.00775)	(0.00768)	(0.00547)	(0.00617)	(0.00615)
dummy 2008:2010	-0.0507***	-0.0474***	-0.0470***	-0.0630***	-0.0583***	-0.0552***
	(0.00645)	(0.00695)	(0.00701)	(0.00498)	(0.00543)	(0.00545)
dummy 2011:2013	-0.0442***	-0.0420***	-0.0416***	-0.0514***	-0.0476***	-0.0452***
	(0.00660)	(0.00703)	(0.00700)	(0.00447)	(0.00466)	(0.00466)
dummy 2014:2016	-0.00787	-0.00523	-0.00621	-0.0253***	-0.0251***	-0.0251***
	(0.00576)	(0.00605)	(0.00606)	(0.00414)	(0.00435)	(0.00440)
age		-0.00239***	-0.00221***		-0.00124***	-0.00128***
		(0.000448)	(0.000448)		(0.000412)	(0.000400)
age-sq		2.09e-05***	1.96e-05***		5.81e-06*	6.24e-06*
		(3.88e-06)	(3.80e-06)		(3.51e-06)	(3.39e-06)
log(inventory days)		0.0164***	0.0161***		0.0145***	0.0149***
		(0.00235)	(0.00223)		(0.00217)	(0.00211)
log(size)	0.513***	0.513***	0.503***	0.239***	0.230***	0.227***
	(0.0225)	(0.0249)	(0.0248)	(0.0169)	(0.0180)	(0.0179)
agent listing	-0.0105***	-0.0130***	-0.0151***	-0.00840***	-0.0102***	-0.0127***
	(0.00352)	(0.00385)	(0.00392)	(0.00289)	(0.00323)	(0.00327)
shared laundry	-0.0714***	-0.0623***	-0.0628***	-0.0442***	-0.0465***	-0.0498***
	(0.0111)	(0.0116)	(0.0110)	(0.00956)	(0.0102)	(0.00958)
townhouse	-0.0938***	-0.100***	-0.0982***	-0.0834***	-0.0954***	-0.0954***
	(0.00810)	(0.00914)	(0.00891)	(0.00725)	(0.00769)	(0.00765)
condo	-0.101***	-0.105***	-0.107***	-0.101***	-0.113***	-0.119***
	(0.0114)	(0.0135)	(0.0134)	(0.00800)	(0.00956)	(0.00967)
multi	-0.124***	-0.126***	-0.121***	-0.132***	-0.148***	-0.141***
	(0.0108)	(0.0129)	(0.0126)	(0.0105)	(0.0112)	(0.0110)
street parking	-0.0300*	-0.0429**	-0.0569***	-0.0418***	-0.0441***	-0.0512***
	(0.0167)	(0.0180)	(0.0179)	(0.0131)	(0.0151)	(0.0150)
studio	-0.288***	-0.306***	-0.304***	-0.416***	-0.438***	-0.437***
	(0.0450)	(0.0506)	(0.0496)	(0.0355)	(0.0413)	(0.0408)
refrigerator			0.0186***			0.0219***
			(0.00555)			(0.00451)
dishwasher			-0.00146			0.000119
			(0.00467)			(0.00327)
no pets			-0.0155***			-0.0212***
			(0.00410)			(0.00365)
hardwood floor			0.0282***			0.0332***
			(0.00401)			(0.00337)
forced heat			0.00152			0.0191***
			(0.00438)			(0.00355)
no AC			-0.0346***			0.00196
			(0.0107)			(0.00788)
central AC			0.0423***			0.0158***
			(0.00619)			(0.00404)
floor (condo)			0.00229*			0.00425**
			(0.00134)			(0.00177)
bedrm dummies	YES	YES	YES	YES	YES	YES
bathrm dummies	YES	YES	YES	YES	YES	YES
zip code FE	YES	YES	YES	YES	YES	YES
F-stat 05:07 - 08:10	6.65	8.01	8.62	14.95	10.96	10.66
p-value 05:07 - 08:10	0.0101	0.0048	0.0034	0.0001	0.0010	0.0011
Obs	22,188	19,181	19,181	21,956	18,259	18,259
R-sq	0.824	0.824	0.827	0.822	0.818	0.821

Table A.4: Effect of acquisition vintage on current rents, for rental properties demanding rents, respectively, above and below the median in our sample. We replicate the analysis in columns 2, 3 and 4 of Table 2 within two subsamples of the data. In columns 1, 2 and 3 the sample is restricted to listings that ask monthly rent greater or equal than the median. In columns 4, 5 and 6 the sample is restricted to listings that ask monthly rent below the median. The Table also shows the F-statistic and p-value of a test of the null that the dummy coefficient for properties last purchased from 2005 to 2007 and the dummy coefficient for properties last purchased from 2008 to 2010 are equal. Standard errors are reported in parenthesis and are clustered by zip code.

	(1)	(2)	(3)	(4)	(5)	(6)
	Log Rent	Log Rent	Log Rent	Log Rent	Log Rent	Log Rent
	$\geq$ Median zip code Income			< Median zip code Income		
dummy 1980s	-0.117***	-0.118***	-0.112***	-0.0526***	-0.0418***	-0.0327***
	(0.0149)	(0.0185)	(0.0190)	(0.00817)	(0.00978)	(0.00978)
dummy 1990:1994	-0.0847***	-0.0816***	-0.0757***	-0.0594***	-0.0497***	-0.0415***
	(0.0123)	(0.0154)	(0.0153)	(0.00815)	(0.00951)	(0.00962)
dummy 1995:1998	-0.0702***	-0.0648***	-0.0625***	-0.0334***	-0.0241***	-0.0185***
	(0.00986)	(0.0111)	(0.0113)	(0.00510)	(0.00616)	(0.00606)
dummy 1999:2001	-0.0720***	-0.0693***	-0.0679***	-0.0371***	-0.0261***	-0.0223***
	(0.00799)	(0.00943)	(0.00944)	(0.00501)	(0.00613)	(0.00598)
dummy 2002:2004	-0.0532***	-0.0470***	-0.0468***	-0.0368***	-0.0299***	-0.0257***
	(0.00716)	(0.00777)	(0.00777)	(0.00379)	(0.00436)	(0.00427)
dummy 2005:2007	-0.0375***	-0.0309***	-0.0315***	-0.0259***	-0.0224***	-0.0194***
	(0.00712)	(0.00804)	(0.00800)	(0.00398)	(0.00442)	(0.00439)
dummy 2008:2010	-0.0568***	-0.0545***	-0.0556***	-0.0414***	-0.0371***	-0.0338***
	(0.00668)	(0.00741)	(0.00750)	(0.00354)	(0.00373)	(0.00371)
dummy 2011:2013	-0.0415***	-0.0396***	-0.0409***	-0.0367***	-0.0329***	-0.0305***
	(0.00573)	(0.00609)	(0.00608)	(0.00294)	(0.00313)	(0.00311)
dummy 2014:2016	-0.00769	-0.00723	-0.00900	-0.0226***	-0.0215***	-0.0209***
	(0.00561)	(0.00597)	(0.00599)	(0.00304)	(0.00321)	(0.00322)
age		-0.00106**	-0.000945**		-0.00195***	-0.00198***
		(0.000417)	(0.000419)		(0.000360)	(0.000345)
age-sq		1.06e-05***	9.98e-06***		9.72e-06***	9.94e-06***
		(3.73e-06)	(3.72e-06)		(3.43e-06)	(3.28e-06)
log(inventory days)		0.0175***	0.0169***		0.00494***	0.00543***
		(0.00198)	(0.00191)		(0.000954)	(0.000937)
log(size)	0.532***	0.530***	0.522***	0.149***	0.140***	0.139***
	(0.0223)	(0.0244)	(0.0245)	(0.0119)	(0.0128)	(0.0127)
agent listing	-0.0152***	-0.0173***	-0.0187***	-0.00523***	-0.00526***	-0.00691***
	(0.00361)	(0.00395)	(0.00401)	(0.00195)	(0.00204)	(0.00202)
shared laundry	-0.0502***	-0.0471***	-0.0499***	-0.0316***	-0.0320***	-0.0347***
	(0.00963)	(0.0103)	(0.0101)	(0.00482)	(0.00515)	(0.00503)
townhouse	-0.0854***	-0.0885***	-0.0868***	-0.0502***	-0.0585***	-0.0593***
	(0.00808)	(0.00920)	(0.00904)	(0.00443)	(0.00470)	(0.00470)
condo	-0.0954***	-0.0950***	-0.1000***	-0.0692***	-0.0807***	-0.0805***
	(0.0113)	(0.0131)	(0.0131)	(0.00494)	(0.00555)	(0.00575)
multi	-0.0957***	-0.0956***	-0.0911***	-0.0998***	-0.114***	-0.110***
	(0.0137)	(0.0154)	(0.0154)	(0.00568)	(0.00656)	(0.00647)
street parking	-0.0293*	-0.0436**	-0.0560***	-0.0239**	-0.0245**	-0.0280**
	(0.0163)	(0.0172)	(0.0171)	(0.0108)	(0.0122)	(0.0119)
studio	-0.109**	-0.128**	-0.136**	-0.367***	-0.375***	-0.371***
	(0.0529)	(0.0598)	(0.0590)	(0.0221)	(0.0251)	(0.0250)
refrigerator			0.0217***			0.00369
			(0.00473)			(0.00308)
dishwasher			-0.00286			0.00246
			(0.00459)			(0.00226)
no pets			-0.00900**			-0.0240***
			(0.00401)			(0.00225)
hardwood floor			0.0234***			0.0239***
			(0.00373)			(0.00229)
forced heat			-0.00550			0.0166***
			(0.00414)			(0.00238)
no AC			-0.0333***			0.00507
			(0.00999)			(0.00635)
central AC			0.0324***			0.00860***
			(0.00519)			(0.00249)
floor (condo)			0.00333*			0.000326
			(0.00190)			(0.000775)
bedrm dummies	YES	YES	YES	YES	YES	YES
bathrm dummies	YES	YES	YES	YES	YES	YES
zip code FE	YES	YES	YES	YES	YES	YES
F-stat 05:07 - 08:10	9.58	11.03	12.01	14.02	7.49	7.67
p-value 05:07 - 08:10	0.0021	0.0010	0.0006	0.0002	0.0064	0.0058
Obs	22,262	19,623	19,623	21,740	17,662	17,662
R-sq	0.769	0.773	0.776	0.798	0.798	0.802

Table A.5: Effect of acquisition vintage on current rents, for rental properties in zip codes with average household income, respectively, above and below the median in our sample. We replicate the analysis in columns 2, 3 and 4 of Table 2 within two subsamples of the data. In columns 1, 2 and 3 the sample is restricted to listings for houses located in zip codes with average income greater or equal than the median (based on 2016 zip code-level average income calculated by the IRS). In columns 4, 5 and 6 the sample is restricted to listings for houses located in zip codes with average income below the median. The Table also shows the F-statistic and p-value of a test of the null that the dummy coefficient for properties last purchased from 2005 to 2007 and the dummy coefficient for properties last purchased from 2008 to 2010 are equal. Standard errors are reported in parenthesis and are clustered by zip code.

B Unobservable Selection and Coefficient Stability

An abbreviated version of the analysis in Oster (2019) is sketched below. Consider the following regression equation:

$$y_i = \beta x_i + w_{1,i} + w_{2,i} + e_i$$

where x_i is the treatment of interest for observation i, $w_{1,i}$ is a scalar capturing the observed characteristics of i and $w_{2,i}$ is a scalar capturing unobserved quality of i. Assume that $w_{1,i} = \psi \tilde{W}_{1,i}$, where $\tilde{W}_{1,i}$ is a vector of observable characteristics. Also assume that w_1 and w_2 are orthogonal, i.e. w_2 captures variation unobserved quality that is not spanned by the observable characteristics. There will always be a scalar δ such that:

$$\delta \frac{\sigma_{1,x}}{\sigma_1^2} = \frac{\sigma_{2,x}}{\sigma_2^2}$$

where $\sigma_{1,x} = Cov(w_1, x)$, $\sigma_{2,x} = Cov(w_2, x)$, $\sigma_1^2 = Var(w_1)$ and $\sigma_2^2 = Var(w_2)$. Now, consider a "short" regression, including only the treatment x_i as a control; β° and R° are the regression coefficient and the R^2 from the short regression. Consider then the "long" regression including both x_i and $w_{1,i}$, with output $\hat{\beta}$ is $\hat{R}$. Oster (2019) shows that, under some restrictive assumptions:

$$\beta^* - \hat{\beta} \approx \delta \left(\hat{\beta} - \beta^\circ\right) \frac{1 - \hat{R}}{\hat{R} - R^\circ},$$

where β^* is the unbiased estimator of the population value of β. The bias is increasing in the difference in the slope estimates between the "short" to the "long" regression, and is decreasing in the difference in R^2. The exact representation in the equation above holds only under the assumption that the relative contributions of each observable control (each element of $\tilde{W}_{1,i}$) to x is the same as its contribution to y. This assumption does not hold in general in the data. However, Oster (2019) shows that even after removing restrictive assumptions, a consistent estimator of β^* can still be found. This estimator will retain the key properties and intuition of the basic case. Moreover, Oster (2019) shows that her framework can be further extended to a case where the short regression includes in the conditioning information not only x, but also some of the observable controls. This last case is the one we implement in our study.