Testing for Information Asymmetries in Real Estate Markets

Pablo Kurlat
Stanford University and NBER

Johannes Stroebel
NYU Stern, NBER, and CEPR

In housing markets, neighborhood characteristics are a key source of information heterogeneity: sellers are usually better informed about neighborhood values than buyers are, but some sellers and buyers are better informed than their peers are. Consistent with predictions from a new framework for analyzing such markets with heterogeneous assets and differentially informed agents, we find that changes in the composition of sellers toward more informed sellers and sellers with a larger supply elasticity predict subsequent house price declines. This effect is larger for houses with more price exposure to neighborhood characteristics, and smaller for houses bought by buyers that are more informed. (*JEL* G14, D53, D82, R21, R31)

In many settings, market participants have differential information about important characteristics of heterogeneous assets. Akerlof (1970), for example, analyzes a situation in which sellers of used cars have superior information relative to potential buyers. In other markets, sellers are better informed than buyers are on average, but not all buyers and sellers are equally well informed. Consider the residential real estate market: transaction prices include payments for both the land and the structure, each of which is hard to value and can be a source of asymmetric information between market participants.[1] On average, home sellers are likely to have better information than potential buyers have about both neighborhood and house characteristics. In addition, however, some

This version: April 2015. This paper was previously circulated as "Knowing your Neighborhood: Asymmetric Information and Real Estate Markets." We thank Viral Acharya, Sumit Agarwal, Yakov Amihud, Aurel Hizmo, Andreas Fuster, Mark Garmaise, Stefano Giglio, Bob Hall, Theresa Kuchler, Stijn van Nieuwerburgh, Matteo Maggiori, Holger Mueller, Chris Parsons, Monika Piazzesi, Laura Starks, Florian Scheuer, Amit Seru, Wei Xiong, Jaime Zender, one anonymous referee, as well as seminar and conference participants at Stanford University, Chicago Booth, Berkeley Haas, NYU Stern, Northwestern Kellogg, the 2013 Summer Real Estate Symposium, Universidad Torcuato di Tella, and the 2014 Texas Finance Festival. We thank Trulia for providing data. Supplementary data can be found on *The Review of Financial Studies* web site. Send correspondence to Johannes Stroebel, Finance Department, New York University Stern School of Business, New York, NY 10012; telephone: (650) 888-3441. E-mail: johannes.stroebel@stern.nyu.edu.

[1] For example, the value of a house's structure includes hard-to-observe aspects of construction quality (Stroebel 2014). Similarly, local amenities such as crime rates or school quality change constantly (Guerrieri, Hartley, and Hurst 2013), which makes it hard for market participants to value a neighborhood correctly.

doi:10.1093/rfs/hhv028 Advance Access publication May 4, 2015

of the possible buyers or sellers might have an information advantage relative to their peers. For example, real estate agents might be particularly well informed about neighborhood gentrification patterns and demographic trends, and buyers who have previously lived in the same neighborhood face less of an information disadvantage compared with buyers who are moving from farther away.

We argue that such asymmetric information is substantial and has key implications for equilibrium housing market outcomes. In particular, we document the importance of two aspects of information asymmetry: differences in information about neighborhood characteristics and information differentials within buyers and within sellers.

Our empirical analysis is guided by the predictions from a new theoretical framework for analyzing markets with many heterogeneous assets and differentially informed agents, based on Kurlat (2014). In this framework, an agent's valuation of a property depends on characteristics of both the neighborhood and the structure. Homeowners' valuation of their current unit also includes an idiosyncratic shock that captures, for example, the need to move for job-related reasons. All potential buyers value a property identically based on characteristics of the neighborhood and the structure, both of which they do not observe perfectly. We model information quality as the ability to differentiate between properties of different overall value, and assume that some agents can do this better than others can.

In this framework, the presence of asymmetric information about neighborhood characteristics has several implications for equilibrium outcomes. First, the composition of sellers in a neighborhood should predict subsequent price changes for houses in that neighborhood. This is because some owners are more likely to sell in response to changes in hard-to-observe (and thus partially unpriced) neighborhood characteristics. For example, owners that are more responsive might include those with better information about neighborhood characteristics. Imagine a neighborhood that has experienced a recent adverse demographic trend. Even though this change is not yet common knowledge, and therefore not yet fully reflected in neighborhood house prices, real estate agents living in the neighborhood might be particularly aware of it. They choose to move out of the neighborhood at higher rates, hoping to sell their own house before the demographic change becomes known to all buyers and reduces neighborhood home values. In equilibrium, therefore, the proportion of owners that are more informed among sellers in a neighborhood is indicative of partially unpriced neighborhood characteristics. As some of these characteristics are revealed to all market participants, home prices will adjust toward the true property value. Hence, the composition of sellers at the time a house was purchased should be correlated with the subsequent price appreciation of that house.

A second equilibrium prediction from this framework is that the effect of seller composition on subsequent house price appreciation will be stronger for houses with a value that is more dependent on neighborhood characteristics.

We refer to these houses as having a high neighborhood-β. Third, buyers that are more informed should obtain higher appreciation on average, because they are able to select better houses among the heterogeneous pool of houses on sale. Fourth, the appreciation obtained by buyers that are more informed should be less sensitive to hard-to-observe neighborhood characteristics (and hence seller composition) than that of buyers that are less informed. Informed buyers select which house to buy based on their combined information about both the structure and the neighborhood. Therefore they trade them off in a way that less-informed buyers do not: conditional of buying from a worse neighborhood, more-informed buyers are more selective on the structure, which reduces the effect of neighborhood characteristics on the value of the houses they buy.

We test these predictions empirically, using nearly 20 years of transaction-level house price data from Los Angeles County, covering about 1.5 million property sales. We first document that average neighborhood price appreciation correlates with changes in the composition of sellers as suggested by our theoretical framework. We focus on three measures of seller composition. First, we argue that real estate professionals should be particularly well informed about changes in the quality of their neighborhood. Using data on the universe of real estate licenses issued by the California Department of Real Estate, we find that a one standard deviation increase in the share of real estate professionals among sellers in a neighborhood predicts a decline in future annual neighborhood price appreciation of 13 basis points. Second, we argue that owners of houses with a value that is more affected by neighborhood characteristics (higher neighborhood-β houses) should respond more elastically to changes in neighborhood characteristics, as the value of their house is more affected when the neighborhood changes. Because neighborhood characteristics are primarily capitalized in the land component of a property's value, we propose the share of land in the total property value assigned by the tax assessor as a proxy for a property's neighborhood-β. We verify this by showing that the prices of properties with a larger land share do in fact respond more to changes in average neighborhood prices. Consistent with our empirical predictions, we find that a one standard deviation increase in the average land share in value of houses sold in a neighborhood is predictive of future neighborhood-level price declines of 79 basis points annually. Finally, we argue that longer-tenure residents are less elastic in their decision to move, and show that a one standard deviation increase in the share of sellers who have only recently moved to a neighborhood predicts neighborhood price declines of 47 basis points annually.

The estimated relationship between seller composition and house price growth is economically large. For a home buyer with a loan-to-value ratio of 80%, a 50-basis-points-higher annual capital gain translates into 2.5-percentage-points-higher annualized holding period return.

In addition to measuring the relationship between seller composition and neighborhood-level house prices, we also test directly whether seller composition is correlated with observable changes in neighborhood characteristics. Using data from the California Department of Education and the Home Mortgage Disclosure Act, we show that the share of socioeconomically disadvantaged students in local schools, as well as the income of new home buyers move with the composition of sellers in a neighborhood. We also show that the effect of changes in seller composition on subsequent price changes is indeed significantly larger for houses with a higher neighborhood-β.

Although this evidence is highly consistent with the presence of asymmetric information about neighborhood characteristics in housing markets, the significant autocorrelation of house price changes means that a relationship between seller characteristics and subsequent price changes is by itself insufficient evidence for the presence of asymmetric information. For example, it could be that all market participants are equally aware of future neighborhood-level price declines, but more-responsive owners react more strongly to them. To rule out such alternative explanations, we control for past neighborhood-level price changes in our regressions, which removes the commonly predictable component of house price changes. The correlation between seller composition and subsequent house price changes is unchanged.

In addition, we also consider how house price appreciation varies with the characteristics of the buyer, and argue that these findings are uniquely explained by information asymmetries. We find that real estate agents purchase houses that experience almost a full percentage point higher subsequent annualized capital gains. We also identify a second set of buyers who are likely to be better informed about neighborhood characteristics. Specifically, using the geo-coded address of all transacted properties combined with the identity of the transactors, we identify buyers who previously owned a house relatively close to the property they are purchasing. We argue that these buyers are likely to be better informed about neighborhood characteristics, and find that they indeed purchase houses that experience above-average subsequent capital gains. This is hard to reconcile with an explanation in which all agents are equally informed. Crucially, we also show that the effect of seller composition on price appreciation is smaller for houses bought by real estate agents and for houses bought by individuals who have previously lived closer to the house they are purchasing. This is consistent with the prediction that the capital gains of more-informed buyers should be less sensitive to hard-to-observe neighborhood characteristics than those of less informed ones. Models of price predictability for reasons other than asymmetric information do not generate these predictions.

Our paper contributes to the literature considering the role of asymmetric information between various agents in housing and mortgage markets (Garmaise and Moskowitz 2004; Levitt and Syverson 2008; Stroebel 2014); a related body of research analyzes the ability of better-informed agents to exploit superior information to time changes in house prices (Bayer, Geissler, and Roberts 2011; Cheng, Raina, and Xiong 2013;

Chinco and Mayer 2014).[2] Relative to this literature, the current paper is the first to document that neighborhood characteristics provide a significant source of information asymmetry in housing markets, allowing homeowners to time market movements. It also is the first with an explicit focus on understanding market outcomes when some sellers and some buyers are better informed than their peers are. Such differential information across both buyers and sellers is not only a realistic feature of housing markets, but it also generates unique predictions that allow us to identify cleanly the presence of asymmetric information in housing markets with some degree of price predictability.

We also contribute to a large body of literature that has tested the predictions from trading models with asymmetrically informed agents in markets other than real estate. One important set of papers analyzes correlations between the trading behavior of firm insiders and subsequent stock returns. For example, Lorie and Niederhoffer (1968) measure the predictive properties of insiders transactions, and find that they forecast large movements in stock prices. See Jaffe (1974), Finnerty (1976), Seyhun (1986, 1992), Lin and Howe (1990), and Coval and Moskowitz (2001) for related studies. Easley, Hvidkjaer, and Ohara (2002), Kelly and Ljungqvist (2012), and Choi, Jin, and Yan (2013) show the empirical importance of information asymmetries in asset pricing models in the style of Grossman and Stiglitz (1980), applied to the stock market. In our setting, we show that the share of informed and elastic sellers predicts future neighborhood-level capital gains, suggesting that they, too, have insider information about characteristics of the neighborhood. However, in real estate markets, the autocorrelation in house prices means that a relationship between sellers' behavior and subsequent price changes by itself is not sufficient evidence for asymmetric information. To rule out alternative explanations, we analyze unique predictions from a framework with differentially informed buyers and sellers.

On the theoretical side, our framework builds on the analysis of Kurlat (2014), extended to account for the indivisibility of houses and for heterogeneity among sellers. We model an environment where some buyers and sellers have different quality of information, and, even though there is no aggregate noise, the difficulty in telling apart different property types prevents information aggregation. This model is related to the literature that, following Akerlof (1970), has analyzed competitive equilibria in settings with asymmetric information. There are two main strands of this literature. The first is concerned with cases where assets are hard to distinguish from each other and traders on one side of the market are equally uninformed (Wilson 1980; Hellwig 1987; Gale 1992, 1996; Dubey and Geanakoplos 2002; Guerrieri, Shimer, and Wright 2010). The second analyzes cases where assets are clearly distinct and some traders are more informed than others, but there is noise that prevents

[2] In other settings, more informed investors have been shown to time market movements to their advantage (Brunnermeier and Nagel 2004; Temin and Voth 2004; Cohen, Frazzini, and Malloy 2008).

information aggregation (Grossman and Stiglitz 1980; Kyle 1985; Admati 1985; Caballe and Krishnan 1994; Kyle and Xiong 2001; Kodres and Pritsker 2002).[3]

1. Empirical Predictions

In this section we describe the residential real estate market as an environment with information asymmetries about home values between different market participants. We propose several predictions that allow us to analyze empirically the importance of these information asymmetries. These predictions follow from a theoretical framework derived from Kurlat (2014). A formal statement of our model, a definition of the appropriate equilibrium concept, and the derivation of the predictions are presented in the Online Appendix.

Suppose the value of house h in neighborhood n is comprised of the sum of the value of the structure and the value of the land. The parameter θ_n measures the attractiveness of the neighborhood; η_h measures the quality of the structure. Total property value is $v_{hn} = \beta_h \theta_n + \eta_h$. The parameter β_h captures how sensitive the value of a particular house is to neighborhood characteristics. It can be interpreted as a factor loading, and we will refer to it as the property's neighborhood-β.

Potential buyers have identical preferences, and their willingness to pay is v_{hn}. Current owners receive idiosyncratic shocks to how well they are matched to their house. These shocks capture, for example, job-related relocation needs. An owner would be willing to sell his house for $v\varepsilon$, where ε is idiosyncratic and may take values lower than 1; this creates gains from trade. Different subsets of owners have different distributions of ε shocks: this distribution is more dispersed for some owner types than it is for others.

Every owner knows the value of the structure of his own house, η_h. Further, each owner has some information about the quality of his own neighborhood, θ_n. Some owners have information about neighborhood quality that is more precise compared with the information of others. Buyers can be either informed or uninformed. If they are uninformed, then all houses look alike to them.[4] If they are informed, they have some information about the overall value of the house, v_{hn}.

We assume that the marginal buyer is uninformed. This assumption is consistent with our empirical measures on informedness, which indicate that there are relatively few informed buyers. This implies that all houses that look identical to uninformed buyers will trade at the same price.[5] Some owners

[3] Sockin and Xiong (2014) build a model to analyze information aggregation and learning by symmetrically but imperfectly informed agents in housing markets.

[4] This is a normalization. There could be some public information, such as each house's number of bedrooms or widely known features of the neighborhood, which is already built into the distributions of θ_n and η_h.

[5] Because we normalize the information of the least-informed buyer to zero, this translates to all houses trading at the same price. If even the least informed buyers were aware of, say, the number of bedrooms of a property, the

choose to sell their houses and some choose to keep them, depending on the quality of their house, their information about the quality of the neighborhood, and their idiosyncratic shock. Some houses will be bought by informed buyers and some by uninformed ones. Informed buyers use their superior information to pick better houses among the ones on offer, while paying the same price as uninformed buyers.

Over time, some buyers resell their house. By then, at least part of the information about v has become publicly available. This information is therefore reflected in the resale price, along with any subsequent shocks. The appreciation obtained during a buyer's tenure is thus a (noisy) measure of v at the time of purchase.[6] Consequently, the average appreciation obtained by buyers in a particular neighborhood is a measure of θ. Put differently, high-θ neighborhoods are underrated: they are better than is commonly known, and better than is currently reflected in local house prices. As θ becomes public information over time, houses in these neighborhoods should experience higher-than-average appreciation. Conversely, houses that are bought in overrated, low-θ neighborhoods should experience lower-than-average subsequent appreciation.

The selection of different subsets of owners into selling will be different in underrated, high-θ neighborhoods than in overrated, low-θ neighborhoods. Therefore the composition of sellers in a neighborhood is a good indicator of whether the neighborhood is over- or underrated and should be predictive of future appreciation.

Consider first classifying owners by their level of information. More-informed owners are more likely to sell in overrated neighborhoods; less-informed owners are equally likely to sell in all neighborhoods, simply because they are less aware of the neighborhood's over- or underratedness. This leads to the following prediction:

Prediction 1.a. The fraction of informed sellers in a neighborhood should be negatively associated with the subsequent appreciation of houses in that neighborhood.

Consider next categorizing owners by the β_h (neighborhood-β) of their homes. The overall value of a high-β_h house will be very sensitive to neighborhood quality. This means that in an overrated neighborhood, high-β_h houses will be more overrated than low-β_h houses. Therefore, if current owners

statement would imply that all houses with the same number of bedrooms trade at the same price. The assumption that the marginal trader is uninformed contrasts with Kyle (1985) and related models, where informed traders are marginal and their presence can move prices closer to fundamentals. Given that our sample contains relatively few informed buyers, this is not the case in our setup. See the Online Appendix for details.

[6] In our empirical section, we focus on annualized appreciation during an owner's tenure, no matter how long the ownership spell is. Implicitly, this assumes that information is revealed at a roughly constant rate during an ownership spell. We experimented with specifications that assume something other than uniform annualized appreciation but found no evidence in favor of these alternative specifications.

have superior information about neighborhood characteristics, owners of high-β_h houses will more strongly select into selling in overrated neighborhoods, compared with owners of low-β_h houses. This leads to the following prediction:

Prediction 1.b. The average neighborhood-β of houses sold in a neighborhood should be negatively associated with the subsequent appreciation of houses in that neighborhood.

Finally, consider categorizing owners by how long they have lived in their current house. Longer-tenure owners are likely to have more extreme idiosyncratic shocks. For instance, they are more likely to have experienced large changes in family structure that make them mismatched with their house (very low ε). However, they are also more likely to have gotten so attached to their house that they are reluctant to move (very high ε). If this is true, few longer-tenure owners will be close to the point of indifference between selling and not selling, compared with shorter-tenure owners. Therefore their collective selling decisions will react less strongly to over- or undervaluation of a neighborhood as compared with shorter-tenure owners. This leads to the following prediction:

Prediction 1.c. The proportion of longer-tenure sellers in a neighborhood should be positively associated with the subsequent appreciation of houses in that neighborhood.

The next prediction is about the relative appreciation of different houses within a neighborhood. The effects of neighborhood over- or undervaluation will be greater for higher-β_h houses; those houses, by definition, have a higher factor loading on the value of the neighborhood. This implies that when seller composition indicates that θ is high, one should expect to see not just higher subsequent appreciation overall, but a disproportionate effect on high-β_h houses. This leads to the following prediction:

Prediction 2. Seller composition in a neighborhood should predict more subsequent appreciation (in absolute value) for high-β houses.

If, as we assume, more-informed buyers are able to select better houses at the same price, then this immediately implies the following prediction:

Prediction 3. More-informed buyers should obtain higher appreciation.

Now compare the differential appreciation obtained by informed buyers over uninformed buyers conditional on buying in a high-θ or low-θ neighborhood. Because informed buyers choose houses based on their overall value, they trade off neighborhood and house quality. This means that in an overrated, low-θ neighborhood they are only willing to buy high-η houses. In an underrated, high-θ neighborhood, they are more willing to buy houses with lower η. Uninformed buyers, by contrast, are equally likely to buy high-η or low-η houses in high-θ or low-θ neighborhoods. This implies the following prediction:

Prediction 4. The differential appreciation obtained by informed buyers should be negatively associated with seller compositions that predict high neighborhood appreciation.

In other words, both being bought by an informed buyer and being located in a neighborhood where seller composition indicates high-θ should predict a high appreciation for a given house, but the interaction of these two variables should be negative.

In the following section we test these empirical predictions, providing evidence that asymmetric information about neighborhood characteristics is an important feature of residential real estate markets.

2. Data Description

To conduct the empirical analysis, we combine a number of datasets. The first dataset contains information on the universe of ownership-changing housing deeds in Los Angeles County between June 1994 and the end of 2011. We observe approximately 7.15 million deeds covering such transactions. Properties are uniquely identified via their Assessor Parcel Number (APN). Variables in this dataset include property address (including the latitude and longitude of each property), contract date, transaction price, type of deed (e.g., Intrafamily Transfer Deed, Warranty Deed, or Foreclosure Deed), and the identity of the buyer and seller. It also reports the amount and duration of the mortgage and the identity of the mortgage lender. Figure 1 shows the location of each of the properties with transactions in our dataset. From this dataset, we extract all arms-length transactions for which transaction prices reflect the true market value of the property. This procedure, which excludes, among others, intrafamily transfer deeds and foreclosure deeds, is described in the Online Appendix. There are about 1.45 million arms-length transactions.

The second dataset contains the universe of residential property tax assessment records for the year 2010. This dataset includes information on property characteristics such as construction year, owner-occupancy status, lot size, building size, and the number of bedrooms and bathrooms. The tax assessment records also include an estimate of the market value of the property in January 2009, split into a separate assessment for the value of the land and that of the structure. This will be important, because the price of properties with a larger share of total value constituted by land should change more in response to neighborhood characteristics. In other words, we propose that the land share

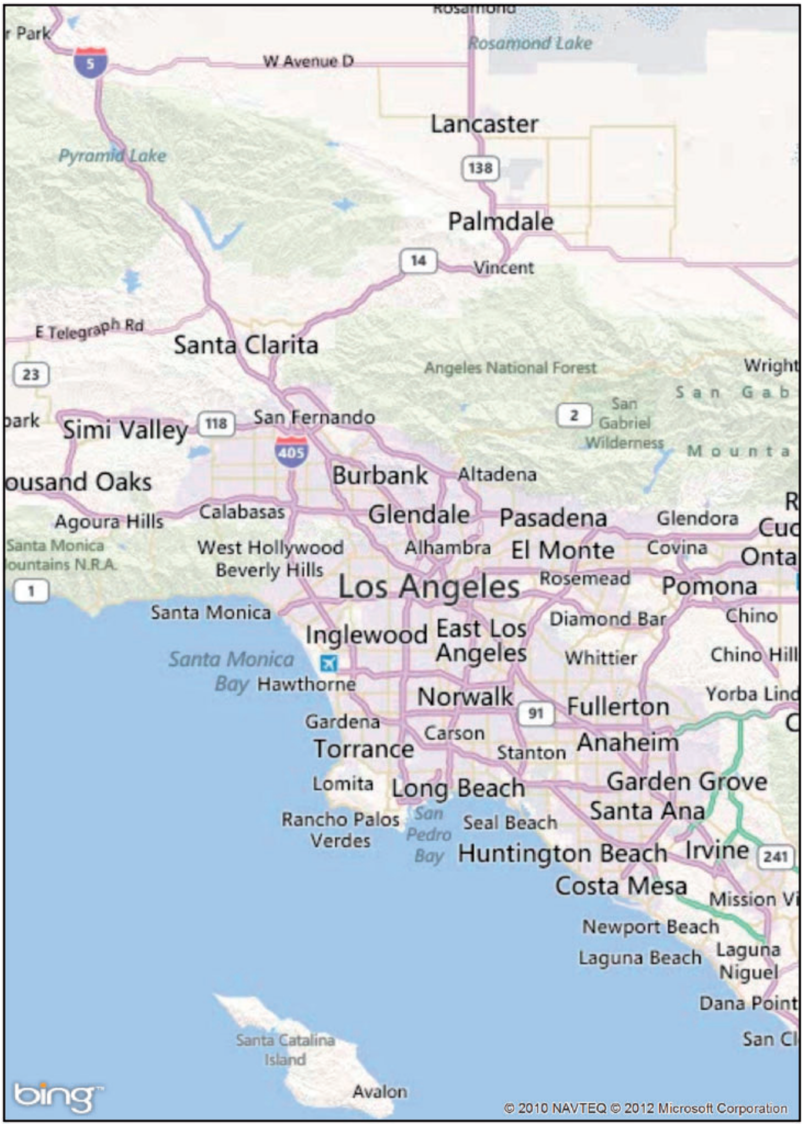

Figure 1
Transaction sample
Maps show the location of all houses for which we observe a transaction between June 1994 and December 2011 for the Los Angeles area.

in total value might be a good proxy for neighborhood-β, the factor loading of property values on neighborhood characteristics. Section 3.1 shows empirically that this is indeed the case. To check whether the relative values assigned to land and property by the tax assessor appear realistic, Figure 2 shows how the

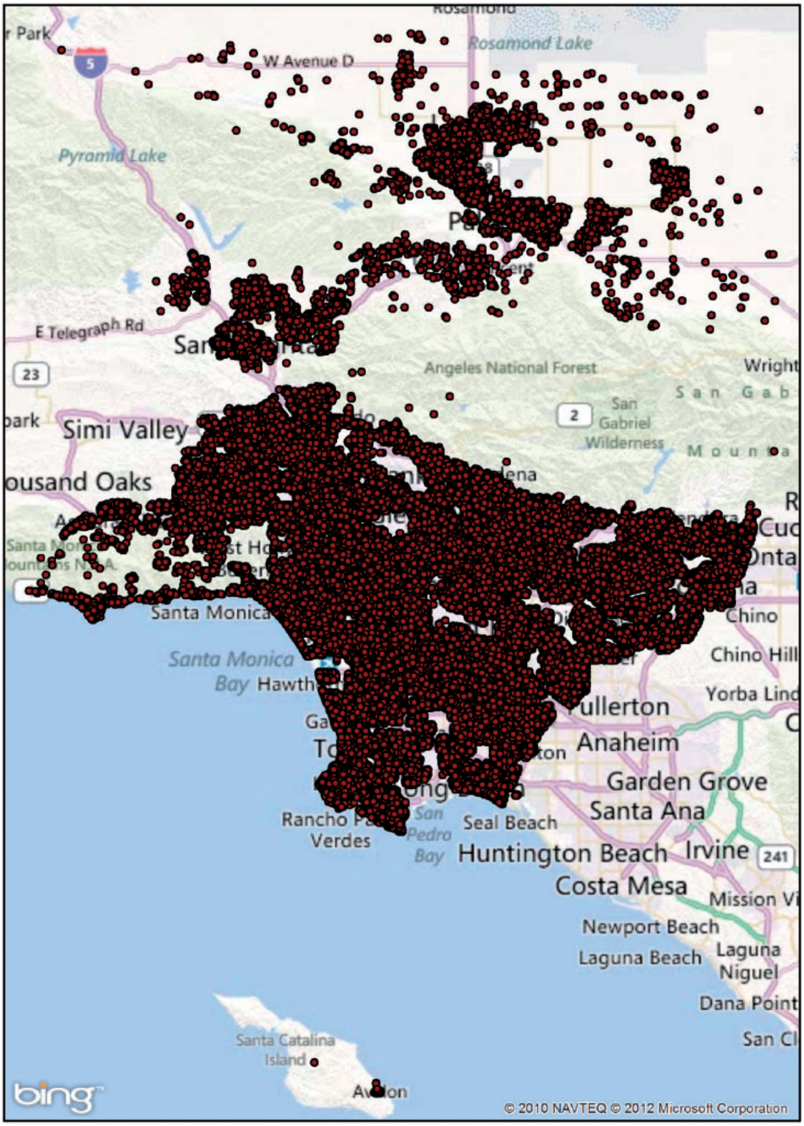

Figure 1
Continued

fraction of total value that is constituted by land varies across Los Angeles County. As one might expect, land is more valuable relative to the structure in the downtown area and near the coast. Importantly for our purposes, there is also significant variation in the land share measure for houses that are relatively

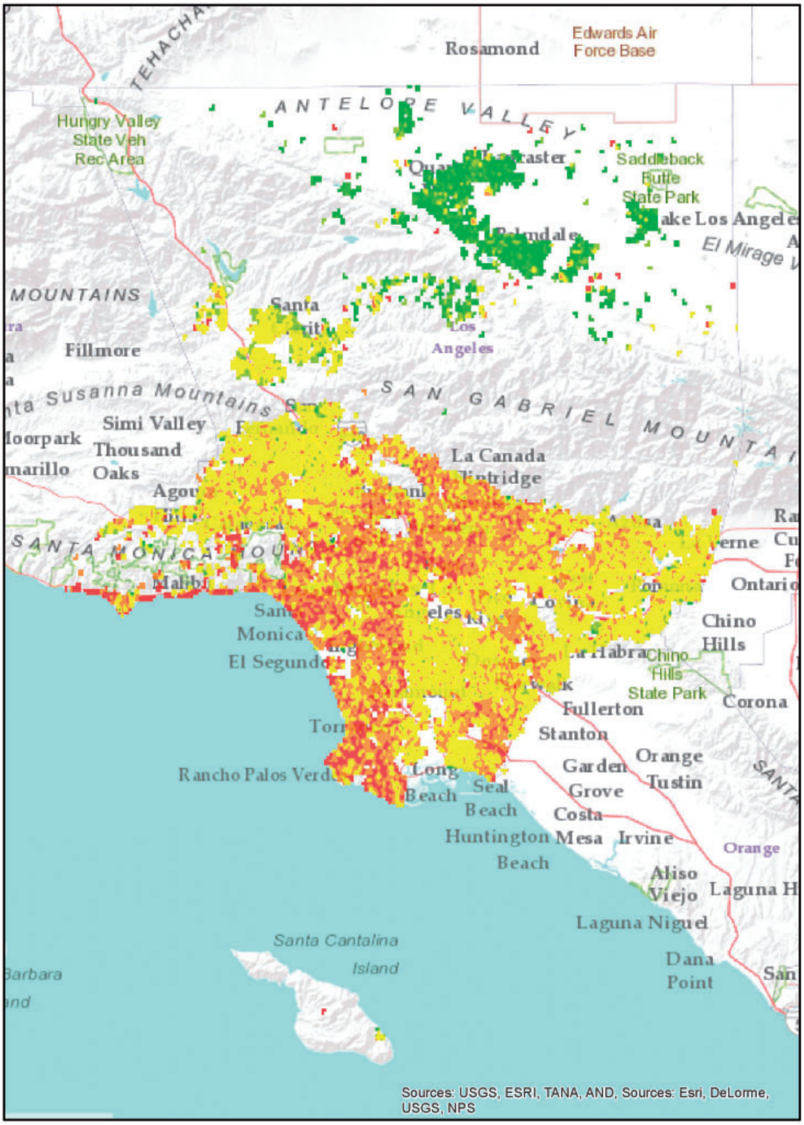

Figure 2
Heat map of land share in property value
Map shows the distribution of the fraction of total property values that is made up from land, as reported in the assessment records. Land share in total value is increasing from green to red.

close to each other (i.e., in the same "neighborhood"). The Online Appendix provides an example of two neighboring properties with different land share.[7]

[7] One might be concerned that we only measure the land share in 2009, in particular if the measure was not stable over time. Unfortunately, the data provider for the assessor data generally overwrites past assessments with new

We also use data from the California Department of Real Estate on the universe of real estate agent and broker licenses issued in California since 1969. We propose that real estate professionals may be particularly well informed about changes in neighborhood characteristics relative to other buyers and sellers. We merge this license data to the housing transaction data using the name of the transactors reported in the property deeds. This allows us to identify properties that have been bought or sold by a real estate professional. In particular, we classify a property as having been bought or sold by a real estate professional if at the time of sale there was an active real estate agent or broker license issued in Los Angeles county to somebody of that name.[8]

3. Results

3.1 Measuring neighborhood-β

One of the characteristics that differentiate houses in the environment described in Section 1 was the neighborhood-β (neighborhood factor loading) of the individual homes. That is, houses differ in how much their value varies as neighborhood characteristics change. To test the resulting empirical predictions, we must first determine a measure of each house's neighborhood-β. As suggested above, neighborhood characteristics have a larger effect on the land-value component of a property than on the structure-value component. This is because in the long run, it is the land rather than the structure that capitalizes neighborhood amenities (see Arnott and Stiglitz 1979; Davis and Heathcote 2007; Albouy 2009).

In this section we show that the land share in total value of each house as identified by the tax assessor is indeed a good proxy for the neighborhood-β of that house. We consider a zip code as the neighborhood of interest. For each pair of arms-length transactions of house i, located in zip code n, with first sale in quarter q_1, and second sale in quarter q_2, we calculate the annualized capital gain of the house between the two transactions, $CapGain_{i,n,q_1,q_2}$. In addition, we measure average price movements in that zip code over the same period, $ZipCapGain_{n,q_1,q_2}$. We do this by determining the annualized change in the median transaction price. In addition, we construct a measure of the land share in total value for each house, $LandShare_i$, by exploiting the separate valuations of the land and the structure component in the assessor records. We then run

assessments, which makes it hard to construct a contemporaneous measure for all years in our sample. However, we are primarily interested in a ranking of properties by land share as a measure of the exposure to neighborhood shocks, rather than the exact value share. To see whether our land share measure is stable over time, we have obtained additional assessor estimates for each house for January 2012; these allow us to construct a second land share measure for each house. The Spearman's rank correlation coefficient of the land share measure between the two dates is $\rho = 0.9723$. This suggests that the measure is a stable relative ranking of houses in terms of their exposure to neighborhood level shocks.

[8] This will introduce measurement error, because we misclassify people with common names to be real estate agents. However, although this might lead to attenuation bias, it should not introduce a systematic bias into our analysis. Consistent with this, in the Online Appendix we show that our results are strengthened if we drop the 100 and 1,000 most common names from the matches with real estate agents.

Table 1
Land share as neighborhood-β

	(1) Capital gain	(2) Capital gain	(3) Capital gain
Zip code capital gain	0.997***	0.955***	0.966***
	(0.004)	(0.013)	(0.015)
Zip code capital gain × land share		0.068***	0.061***
		(0.017)	(0.021)
R-squared	0.793	0.793	0.808
N	391,536	391,534	286,134

Table 1 shows results from Regression 1. We include all sales pairs between June 1994 and December 2011. In Column (3) we restrict the sales pairs to be from zip codes with at least 5,000 transactions in this period. Standard errors are clustered at the zip code level. Significance levels: * ($p < 0.10$), ** ($p < 0.05$), and *** ($p < 0.01$).

Regression 1 for all repeat sale pairs between June 1994 and December 2011. The results are presented in Table 1.

$$CapGain_{i,n,q_1,q_2} = \alpha_1 + \alpha_2 ZipCapGain_{n,q_1,q_2} + \alpha_3 ZipCapGain_{n,q_1,q_2} \times LandShare_i + \epsilon_i \qquad (1)$$

In Column (1) we drop the interaction between $ZipCapGain$ and $LandShare$. The coefficient on $ZipCapGain$ shows that, reassuringly, on average, house price movements closely track movements of the zip code median price. In Column (2) we include the interaction. The positive coefficient α_3 shows that houses with a larger land share in total value move more in the direction of the market, both when prices increase and when prices decrease. This suggests that the land share of a house is indeed a proxy for the neighborhood-β of that house. In Column (3) we only include transaction pairs from zip codes with at least 5,000 transactions between June 1994 and and December 2011. For those zip codes, the measurement of average neighborhood-level price changes is more precise. The results are unchanged when looking at this subsample.

3.2 Changes in seller composition predict price changes

In this section we test Prediction 1, which says that if sellers have superior information about neighborhood characteristics, then changes in the composition of sellers in a neighborhood should be predictive of future price changes of homes in that neighborhood. We regress the annualized capital gain of houses between two arms-length transactions, $CapGain$, on control variables and the composition of sellers in neighborhood n and quarter q_1. We focus on three measures of seller composition, suggested by Predictions 1.a., 1.b., and 1.c., respectively: (a) the fraction of sellers that are real estate professionals, and are thus particularly well informed about neighborhood characteristics, (b) the average land share of transacted houses, and (c) the time sellers have lived in their home.[9] Table 2 shows summary statistics of

[9] Our measurement of home tenure is censored, because for sellers who initially bought a property before the beginning of our sample period (June 1994), we cannot observe the actual length of tenure, but only know that it

Table 2
Summary statistics for seller composition

Variable	Neighborhood	Mean	Standard deviation	
			Unconditional	Conditional
Share informed sellers	Zip code	0.043	0.031	0.027
	Census tract	0.043	0.052	0.050
Average seller land share	Zip code	0.594	0.114	0.045
	Census tract	0.594	0.122	0.056
Seller share tenure > 3	Zip code	0.789	0.079	0.068
	Census tract	0.789	0.120	0.111

Table 2 shows summary statistics for the seller composition by quarter and neighborhood for two definitions of neighborhood: zip code and four-digit census tract. Standard deviations are shown both unconditionally and conditional on the neighborhood (i.e., showing the within-neighborhood standard deviation). The sample period for "share informed sellers" and "average seller land share" is June 1994 to December 2011; for the share of sellers with tenure exceeding 3 years, the sample period is July 1997 to December 2011.

the seller composition variables for two definitions of a neighborhood: a zip code and a four-digit census tract. We show both the sample-wide standard deviation, as well as the within-neighborhood standard deviation.

We then run Regression 2 using different geographies as our definition of a neighborhood. The regression includes neighborhood fixed effects, as well sales quarter pair fixed effects, to remove aggregate (Los Angeles-wide) market movements in house prices over time.

$$CapGain_{i,n,q_1,q_2} = \alpha + \beta_1 SellerComposition_{n,q_1} + X_i' \beta_2 + \xi_n + \phi_{q_1,q_2} + \epsilon_i \quad (2)$$

Table 3 shows the results from Regression 2 when we consider a neighborhood to be a zip code. Standard errors are clustered at the initial quarter by zip code level. Column (1) analyzes the effect of the share of real estate agents among home sellers on the subsequent return of homes without controlling for property or financing characteristics. A one conditional standard deviation increase in the share of sellers that are real estate professionals is associated with a 12-basis-points decline in the annualized return of houses.

In Column (2) we add a large set of control variables X_i, including information on the property (age, building size, number of bedrooms and bathrooms, information on pool and air conditioning, and property type) and the mortgage financing (the loan-to-value ratio, the mortgage duration, and whether it is a VA, FHA, or jumbo mortgage). The Online Appendix describes these control variables in more detail, and provides summary statistics. The estimated correlation between changes in the seller composition and subsequent returns is unchanged by the addition of these control variables. This suggests that the correlation is not driven by observable differences in the composition of houses that might confound our estimates of the effect of seller composition. This is

must have been longer than the time since the beginning of the sample. To deal with this, we define a long-tenure seller to be someone who moved into the neighborhood more than 3 years ago. We then consider the effect of the share of long-tenure sellers among the total population of sellers, and only look at the return between transaction pairs where $q_1 >$ Q2 1997. Results are not sensitive to the choice of 3 years as the cut-off value.

Table 3
Effect of seller composition in zip code on capital gains

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share informed sellers	−4.454***	−4.737***					−4.571***	−3.4578*
	(1.715)	(1.663)					(1.436)	(1.882)
Average seller land share			−16.80***	−17.65***			−13.87***	−13.86***
			(2.021)	(2.004)			(1.876)	(2.462)
Share in zip of tenure > 3					6.813***	6.863***	6.018***	5.196***
					(0.656)	(0.622)	(0.578)	(0.691)
Fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Controls: Property + financing	·	✓	·	✓	·	✓	✓	✓
Sample restrictions	·	·	·	·	·	·	·	Drop 2006+
R-squared	0.626	0.636	0.628	0.638	0.647	0.658	0.659	0.418
$\bar{y}$	12.58	12.56	12.58	12.56	13.72	13.70	13.70	18.37
N	394,803	391,837	394,803	391,837	302,570	300,108	300,108	162,777

Table 3 shows results from Regression 2. The dependent variable is the annualized capital gain of a property between two sequential arms-length sales. The seller composition variables are measured at the quarter × zip code level. All specifications include sales quarter pair fixed effects and zip code fixed effects. Columns (2), (4), (6), (7), and (8) control for characteristics of the property (property size, property age, property type, number of bedrooms and bathrooms, and whether the property has a pool or air conditioning) and characteristics of the financing (mortgage type, mortgage duration, and loan-to-value ratio). Standard errors are clustered at the initial quarter × zip code level. Columns (1) through (4) include sales pairs where the first sale was after June 1994. Columns (5) through (8) include sales pairs where the first sale was after June 1997. Column (8) drops all sales pairs with at least one transaction in or after 2006. Significance levels: * ($p<0.10$), ** ($p<0.05$), and *** ($p<0.01$).

comforting, because we argue that the correlation is driven by hard-to-observe information that current inhabitants have about neighborhood characteristics. In the Online Appendix we show that these results are not driven by mismatches of common names when identifying realtors among the sellers. In addition, the Online Appendix provides further evidence that the relationship is driven by the superior information of informed sellers, and not by a possible role of realtors as liquidity-providing dealers that hold inventory.[10]

In Columns (3) and (4) we consider the effect of changes in the composition of transacted houses toward those with a higher land share in total value (higher neighborhood-β). We argued that an increase in the average land share of transacted homes should predict future declines in neighborhood prices, because the owners of homes with a higher neighborhood-β should sell faster in response to hard-to-observe, partially unpriced negative neighborhood shocks. The results in Column (4) suggest that a one conditional standard deviation increase in the average land share of sold houses is indeed associated with a 79-basis-points decline in subsequent annualized capital gains of houses in that neighborhood.

In Columns (5) and (6) we analyze the effect of a change in the proportion of long-tenured sellers. The results in Column (6) suggest that a one conditional standard deviation decrease in the share of sellers who lived in their house for more than 3 years is associated with a 47-basis-points decline in annualized capital gains of houses in that neighborhood. This is consistent with a setting in which owners that have only recently moved into a neighborhood respond more elastically when neighborhood characteristics change. In Column (7) we jointly include all three measures of seller composition in a neighborhood. The magnitude of the estimated contribution of each of the three measures falls somewhat, as one would expect if each is a noisy measure of the same underlying neighborhood characteristics.

One concern with these results is that they might be driven by effects that were particular to the financial crisis. Los Angeles house prices peaked in September 2006, and fell precipitously in the years after. To ensure that our results are not unique to that part of the sample, in Column (8), we drop all transaction pairs where at least one of the two transactions occurred in or after the year 2006. Unsurprisingly, the average annualized house price appreciation over this period is higher, at 18%. The estimated coefficients on seller composition are almost identical, suggesting that the patterns in Table 3 are not unique to the housing-bust period. The coefficients are also very similar when we exclude all data after 2003, dropping the years of the most significant house price increases.

In the Online Appendix we provide various additional robustness checks to this analysis. In particular, we show that the results are not driven by a selection

[10] Excluding short ownership spells by realtors, which could reflect inventory-holding, does not change the results.

Table 4
Effect of seller composition in census tract on capital gains

	(1)	(2)	(3)	(4)	(5)	(6)
Share informed sellers	−1.496***			−0.332		
	(0.501)			(0.601)		
Average seller land share		−8.130***			−3.893***	
		(0.872)			(0.770)	
Share in census tract of tenure > 3			2.962***			1.913***
			(0.202)			(0.313)
Fixed effects	$q_1 \times q_2$, Census tr.	$q_1 \times q_2$, Census tr.	$q_1 \times q_2$, Census tr.	$q_1 \times q_2 \times$ Zip code, Census tr.	$q_1 \times q_2 \times$ Zip code, Census tr.	$q_1 \times q_2 \times$ Zip code, Census tr.
Controls: Property + financing	✓	✓	✓	✓	✓	✓
R-squared	0.638	0.639	0.660	0.687	0.687	0.705
$\bar{y}$	12.56	12.56	13.70	12.56	12.56	13.70
N	391,802	391,802	300,084	391,802	391,802	300,084

Table 4 shows results from Regression 2. The dependent variable is the annualized capital gain of a property between two sequential arms-length sales. The seller composition variables are measured at the quarter × four-digit census tract level. Columns (1) through (3) include sales quarter pair fixed effects and census tract fixed effects; Columns (4) through (6) include sales quarter pair × zip code fixed effects in addition to census tract fixed effects. All specifications control for characteristics of the property (property size, property age, property type, number of bedrooms and bathrooms, and whether the property has a pool or air conditioning) and characteristics of the financing (mortgage type, mortgage duration, and loan-to-value ratio). Standard errors are clustered at the initial quarter × census tract level. Columns (1), (2), (4), and (5) include sales pairs where the first sale was after June 1994, Columns (3) and (6) include sales pairs where the first sale was after June 1997. Significance levels: * ($p<0.10$), ** ($p<0.05$), and *** ($p<0.01$).

into the sample of repeat sales, or the presence of "flippers" among the short-tenured sellers. We also show that the results extend to considering subsequent ownership periods of the house.

The results presented in Table 3 show that the characteristics of sellers within a zip code are correlated with subsequent neighborhood price changes. However, there might be additional relevant information about the immediate neighborhood of a particular property that is not reflected in the composition of all sellers in a zip code, but only in the composition of sellers in the more-immediate vicinity of the property. Table 4 reports results from Regression 2 with a neighborhood being defined as a four-digit census tract. Although there are 293 unique zip codes in our sample, there are 1,255 unique four-digit census tracts.

The results in Columns (1) through (3) include census tract fixed effects in addition to the sales quarter pair fixed effects. As before, increases in the share of informed sellers and the average land share of transacted homes predict subsequent declines in neighborhood-level capital gains, whereas an increase in the average tenure of sellers predicts higher neighborhood-level capital gains. The magnitudes of the estimated effects are smaller than the ones estimated at the zip code level, probably because there is more noise in the measures of seller composition and this results in attenuation bias. Columns (4) through (6) include an interaction of zip code fixed effects with the sales quarter pair fixed effects, in addition to census tract fixed effects. This allows the time movement of house prices to differ by zip code. Here, the identification comes from differential variation of seller composition across census tracts within the same zip code. Because this removes the effect of neighborhood characteristics that are common for different census tracts within the same zip code, the estimated coefficients are unsurprisingly smaller.

One question is why less-informed home buyers do not condition their choice of house or neighborhood on the composition of sellers if it is truly informative about neighborhood characteristics. One reason is that, in practice, this information is unavailable or extremely hard to obtain in real time. For example, it usually takes months before deed records are updated and accessible to the public. Further, the bulk transaction-level deeds information that would be required to analyze changes in the seller composition is not directly provided by Los Angeles County, but rather is only accessible through commercial data vendors at costs that are prohibitive to individual home buyers. In addition, the significant transaction costs in the housing market make this market unattractive to arbitrageurs who might have the resources to purchase real-time data access.[11]

3.3 Changes in seller composition among forced moves

Our interpretation of the correlation between seller composition and subsequent appreciation is that more-informed sellers (or sellers that are more affected by neighborhood trends) will choose to sell their house when they have superior information about neighborhood characteristics.

A related prediction is that the composition of sellers that sell for reasons other than their information about future neighborhood price changes should be less correlated with subsequent neighborhood appreciation. Consider the set of current homeowners that have received an attractive job offer in another city. Many of them will sell their house and move to take up this offer, independently of whether their neighborhood is currently overvalued or undervalued. This is true for both informed and uninformed owners. In other words, after receiving a large-enough moving shock ε, even informed sellers in an undervalued, high-θ neighborhood are likely to sell their house. Therefore, among people with such a moving shock, the proportion of informed sellers should be less predictive of future prices than it is among the population of households without the exogenous moving shock.

To test this, we identify a group of homes that are sold by individuals who have plausibly received such an exogenous shock that prompts them to sell the house. In particular, we identify transactions where we observe either a death or divorce of the original owner in the 12 months prior to the transaction.[12] Approximately 5% of all sales are identified as being "forced." We then compute

[11] Standard realtor fees are about 6% of the purchase price (Piazzesi, Schneider, and Stroebel 2015).

[12] A divorce of owners is recorded when an intrafamily transfer deed is filed that transfers ownership from common ownership of a married couple to individual ownership of one of the two partners. See the Online Appendix on the procedure to identify transactions following the death of an owner.

Table 5
Effect of seller composition for forced moves on capital gains

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share informed sellers	−0.347*	−0.368**					−0.505***
	(0.176)	(0.173)					(0.177)
Average seller land share			−0.966***	−0.968***			−1.226***
			(0.266)	(0.261)			(0.303)
Share in zip of tenure > 3					0.439***	0.422***	0.452***
					(0.098)	(0.097)	(0.105)
Fixed effects	✓	✓	✓	✓	✓	✓	✓
Controls: Property + financing	·	✓	·	✓	·	✓	✓
R-squared	0.628	0.638	0.627	0.638	0.647	0.658	0.658
$\bar{y}$	12.71	12.69	12.68	12.66	13.77	13.75	13.79
N	372,185	369,509	378,714	375,973	293,142	290,858	286,236

Table 5 shows results from Regression 2. The dependent variable is the annualized capital gain of a property between two sequential arms-length sales. The seller composition variables are measured at the quarter × zip code level among households that experienced a death or divorce within the 12 month prior to the sale. All specifications include sales quarter pair fixed effects and zip code fixed effects. Columns (2), (4), (6), and (7) also control for characteristics of the property (property size, property age, property type, number of bedrooms and bathrooms, and whether the property has a pool or air conditioning) and characteristics of the financing (mortgage type, mortgage duration, and loan-to-value ratio). Standard errors are clustered at the initial quarter × zip code level. Columns (1) through (4) include sales pairs where the first sale was after June 1995, Columns (5) through (7) include sales pairs where the first sale was after June 1998. Significance levels: * (p <0.10), ** (p <0.05), and *** (p <0.01).

our measure of seller composition in a zip code only within that subset of sales. The average share of real estate agents among forced sellers is of similar magnitude, at around 3.4% of transactions, relative to 4.3% in the full sample. The average land share of sold houses, and the share of long-tenured sellers, are also of equal magnitude in the sample of forced moves as they are in the sample of all sales. We then repeat Regression 2 using the seller composition measured in the sample of forced sellers. The results are presented in Table 5.

While seller composition is still related to the subsequent price appreciation, the effect is an order of magnitude smaller than for the entire sample. This suggests that even following a divorce, whether or not the owners sell depends to some degree on their information. However, many of the sales are driven by other factors, such as the need to move after a divorce or the desire to split an inheritance among multiple descendants. Because these shocks are independent of whether a neighborhood is overvalued or undervalued, seller shares are less predictive of future capital gains. This evidence provides comfort that the correlation of seller composition and subsequent appreciation in the full sample is driven by an endogenous selection of better-informed households selling their property.

3.4 Predictability in house prices

In markets such as the stock market, for which we have strong theoretical and empirical reasons to believe that they are relatively efficient and frictionless, an uninformed marginal investor should not be able to predict price changes. In such markets, when some traders' behavior predicts price changes, this constitutes strong evidence that these traders are better informed than the marginal investor. In housing markets, however, it is a well-established fact that aggregate price changes are at least somewhat predictable (see the study by Ghysels et al., 2013, and the references therein). This predictability complicates the interpretation of our results up to now as purely tests for asymmetric information: finding a correlation between seller composition and subsequent returns is a necessary, but not a sufficient condition to detect asymmetric information. One alternative explanation could be that seller composition predicts appreciation because more elastic groups of owners simply respond more to commonly anticipated changes in neighborhood-level house prices, rather than to private information.[13]

We argue in two ways that alternative explanations that do not account for asymmetric information are unable to explain our findings. First, in this section we explicitly control for the main source of house price predictability in the literature, by showing that the correlation between seller composition and subsequent returns remains unchanged after conditioning on past price changes. Second, in Section 3.7, we provide strong evidence for Predictions 3 and 4, which are unique to a model with asymmetric information. These tests consider the interaction of buyer and seller informedness, and show that the effect of seller composition on subsequent returns is particularly big for houses purchased by uninformed buyers. This is inconsistent with a story in which seller composition is driven by price movements that are predictable by all market participants.

Case and Shiller (1989, 1990) find that house price appreciation in the short-run is positively serially correlated. This suggests a possible alternative interpretation for our findings. Maybe the composition of sellers predicts appreciation because elastic groups of owners react more strongly to changes that are commonly predictable based on past appreciation. If this were the case, then controlling for past capital gains in Regression 2 should significantly reduce the correlation between seller composition and returns. To show that this is not the case, we control for past capital gains of houses in the zip code over the past 12 and 24 months in Regression 2.[14]

The results in Table 6 show that, indeed, past capital gains in a zip code have strong predictive power for future capital gains. However, importantly, the inclusion of past returns as control variables does not affect the magnitude or statistical significance of the estimated relationship between seller composition and future returns. This suggests that the predictive power of seller composition for future returns is not driven by the autocorrelation of returns. In other words, sellers are reacting to information beyond what is contained in past returns.

[13] Any Los Angeles-wide price predictability is already controlled for through the ϕ_{q_1,q_2} fixed effects.

[14] Similar results are achieved when controlling for price changes over the past 3, 6, and 36 months.

Table 6
Effect of seller composition on capital gains: Control for past capital gains

	(1)	(2)	(3)	(4)	(5)	(6)
Share informed sellers	−6.190***			−4.658***	−4.703***	−4.705***
	(1.754)			(1.442)	(1.459)	(1.455)
Average seller land share		−17.99***		−13.94***	−13.98***	−14.02***
		(2.054)		(1.863)	(1.865)	(1.861)
Share in zip of tenure > 3			6.851***	6.000***	5.981***	5.976***
			(0.616)	(0.574)	(0.576)	(0.575)
Capital gain past year	0.752***	0.659***	0.641**	0.634**		0.492**
	(0.237)	(0.223)	(0.281)	(0.273)		(0.240)
Capital gain past two years					0.413*	0.213
					(0.224)	(0.208)
Fixed effects, property and financing controls	✓	✓	✓	✓	✓	✓
R-squared	0.635	0.636	0.658	0.659	0.659	0.659
$\bar{y}$	12.98	12.98	13.70	13.70	13.70	13.70
N	367,633	367,633	300,067	300,067	300,011	299,995

Table 6 shows results from Regression 2. The dependent variable is the annualized capital gain of a property between two sequential arms-length sales. The seller composition and past capital gains variables are measured at the quarter × zip code level. All specifications include sales quarter pair fixed effects, zip code fixed effects, and control for characteristics of the property (property size, property age, property type, number of bedrooms and bathrooms, and whether the property has a pool or air conditioning) and characteristics of the financing (mortgage type, mortgage duration, and loan-to-value ratio). Standard errors are clustered at the initial quarter × zip code level. Columns (1) and (2) include sales pairs where the first sale was after June 1995. Columns (3) through (6) include sales pairs where the first sale was after June 1997. Significance levels: * ($p<0.10$), ** ($p<0.05$), and *** ($p<0.01$).

3.5 Seller composition and neighborhood demographics

So far our evidence has shown that changes in seller composition can predict the future appreciation of houses, and can do so over and above what would be predictable from past house price changes. This is a prediction from a model in which sellers react to ongoing changes in neighborhood characteristics that are difficult for potential buyers to observe. In this section we test whether changes in seller composition actually predict changes in neighborhood characteristics that are observable at the zip code level.

We use two datasets that contain annual zip code-level demographic information. None of these data were available to home buyers at the time of purchasing the house, which means that demographic shifts measured in these data were not easily observable in real time. The first dataset contains information from the Home Mortgage Disclosure Act's (HMDA) Loan Application Registry, which provides details on all mortgage application in major Metropolitan Statistical Areas. It includes details on the year of mortgage application, the census tract of the house, and the applicant's income. We use these data to construct an annual zip code-level measure of the average income of all mortgage applicants. Importantly, because mortgage applications measure the average income of the flow of households into a neighborhood, they reflect changes faster than measures of the average income within a neighborhood. We then run Regression 3, where we regress this income measure on the seller composition in that year. As in Section 3.4, we control for the capital gain of houses in the zip code over the past year to make sure we are not just capturing a differential elasticity to persistent but observable demographic shifts. Results are very similar when excluding this variable. We also include fixed effects for the zip code, ξ_n, and the calendar year, ϕ_y. We cluster standard errors at the zip code level, to allow for an arbitrary time-series correlation of the residuals.

$$ZipCode\_Demographics_{n,y} = \alpha + \beta_1 SellerComposition_{n,y} + \beta_2 PastCapGain_{n,y} + \xi_n + \phi_y + \epsilon_{n,y} \quad (3)$$

The results are shown in Columns (1) through (4) in Panel A of Table 7. A one conditional standard deviation increase in the share of informed sellers is associated with a $\$1,400$ decline in the average household income of mortgage applicants in that zip code (relative to an average $\$126,000$). A one conditional standard deviation increase in the average land share of sellers corresponds to a \$2,600 decline in the average income reported in mortgage applications. A one conditional standard deviation increase in the share of short-tenured sellers corresponds to a \$2,700 decrease in mortgage applicants' income. This evidence suggests that sellers do indeed react to changes in neighborhood demographics, which are hard to observe in real time.

We also use a second dataset to provide us with annual zip code-level demographic information. In particular, we obtain data from the California Department of Education on the demographics of the student population between 2000 and 2011 at the school level. From these data we construct, for each zip code and year, a student-population weighted measure of demographics of all schools in that zip code, and measure the share of students that are classified as socioeconomically disadvantaged.[15] Columns (5) through (8) of Table 7, Panel A, show results from Regression 3, replacing $ZipCode\_Demographics$ with the share of socioeconomically disadvantaged students. The results show that a one conditional standard deviation increase in the share of informed sellers coincides with a 0.67-percentage-point shift in the demographics of the student population towards socioeconomically disadvantaged students (relative to an average of 61%). Similarly, a one conditional standard deviation increase in the average land share of transacted homes corresponds to an increase in the share of socioeconomically disadvantaged students by 0.5 percentage points. Finally, a one conditional standard deviation increase in the share of short-tenured sellers is associated with a 0.5-percentage-point increase in the share of children that are economically disadvantaged.

Our basic results showed that sellers are reacting to information that is not yet reflected in prices. The evidence presented in Table 7 gives a sense of what that information could be about: changes in neighborhood-level demographic variables. These demographic shifts are hard for buyers to observe in real time

15 A "socioeconomically disadvantaged" student is defined as (i) a student neither of whose parents have received a high school diploma, or (ii) a student who is eligible for the free school lunch program.

Table 7
Seller characteristics and zip code demographics

Panel A: Contemporaneous demographics

	Average mortgage applicant income				Share of disadvantaged students			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Appreciation past year	−16.84***	−16.84***	−17.45***	−14.99***	−1.459	−0.992	−1.066	−1.834
	(5.394)	(5.347)	(5.340)	(5.364)	(1.623)	(1.617)	(1.617)	(1.619)
Share informed sellers	−50.06**			−35.84*	25.15***			23.72***
	(20.35)			(20.32)	(6.134)			(6.121)
Average seller land share		−58.10***		−51.90***		10.76***		8.747***
		(10.71)		(10.74)		(3.186)		(3.202)
Share in zip of tenure > 3			39.74***	36.18***			−7.566***	−6.659***
			(7.060)	(7.069)			(1.922)	(1.932)
Fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
R-squared	0.930	0.930	0.930	0.931	0.968	0.968	0.968	0.969
$\bar{y}$	126.0	126.0	126.0	126.0	60.53	60.53	60.53	60.53
N	4,157	4,157	4,157	4,157	3,106	3,106	3,106	3,106

Panel B: Demographics in one year

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Appreciation past year	−30.08***	−30.68***	−31.10***	−29.64***	−1.480	−1.345	−1.455	−1.482
	(4.323)	(4.299)	(4.297)	(4.325)	(1.389)	(1.379)	(1.378)	(1.389)
Current mortgage applicant income	0.530***	0.528***	0.530***	0.527***				
	(0.013)	(0.013)	(0.013)	(0.013)				
Current share disadv. student					0.583***	0.584***	0.581***	0.582***
					(0.017)	(0.017)	(0.017)	(0.018)
Share informed sellers	−36.24**			−32.57**	2.778			2.576
	(16.28)			(16.35)	(5.341)			(5.352)
Average seller land share		−20.64**		−18.59**		−2.715		−3.344
		(8.614)		(8.667)		(2.956)		(2.971)
Share in zip of tenure > 3			6.651	5.053			−3.429**	−3.557**
			(5.773)	(5.792)			(1.697)	(1.705)
Fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
R-squared	0.956	0.956	0.956	0.956	0.979	0.979	0.979	0.979
$\bar{y}$	129.4	129.4	129.4	129.4	60.62	60.62	60.62	60.62
N	4,147	4,147	4,147	4,147	2,846	2,846	2,846	2,846

Table 7 shows results from Regression 3 for the years 1996 through 2011 in Columns (1) through (4), and for the years 2000 through 2011 in Columns (5) through (8). The unit of observation is zip code by year. The dependent variables are the average income of all mortgage applicants (Columns 1–4) and the share of all students that are classified as socioeconomically disadvantaged (Columns 5–8); in Panel A, the dependent variable is measured in the current year, and in Panel B it is measured in the subsequent year. Each specification includes zip code fixed effects and year fixed effects, and controls for past price appreciation. Standard errors are clustered at the zip code level. Significance levels: * ($p<0.10$), ** ($p<0.05$), and *** ($p<0.01$).

(e.g., because the relevant data are usually only released with significant delay), but sellers who live in the neighborhood are likely to have better information. As these demographic shifts become common knowledge, they will be reflected in prices. This finding explains why current seller composition predicts future capital gains.

In addition, we test whether current seller composition also has some predictive power for future demographics, over and above what is predictable using current demographics. Panel B of Table 7 presents results from Regression 4. The dependent variable is the demographic measure in the next year; control variables are past capital gains, contemporaneous seller composition, and contemporaneous demographics, in addition to zip code and year fixed effects.

$$ZipCode\_Demo_{n,y+1} = \alpha + \beta_1 SellerComposition_{n,y} + \beta_2 PastCapGain_{n,y} + \beta_3 ZipCode\_Demo_{n,y} + \xi_n + \phi_y + \epsilon_{n,y} \quad (4)$$

Unsurprisingly, current demographics are strongly correlated with future demographics. In addition, we can see that current seller composition is related to future demographics in the zip code, even if not all specifications are statistically significant. This suggests that current owners do not only have an information advantage in detecting contemporaneous demographic shifts, but might also have an insight into predicting future demographic changes.

3.6 Importance of neighborhood-β

In this section we consider to what extent the effect of neighborhood seller composition varies across different houses within the same neighborhood. This tests Prediction 2. Because neighborhood amenities are capitalized in the land value of a property, we would expect the effect of seller composition on price changes to be larger for houses with a larger land-share component in total value. To test whether this is indeed the case, we run Regression 5, where $LandShare_i$ is the house-specific share of total value made up by land, as reported in the tax assessor data. The coefficient of interest is β_3, which measures the increase in the responsiveness of capital gains to seller composition when the house has a larger land share.

$$CapGain_{i,n,q_1,q_2} = \alpha + \beta_1 SellerComposition_{n,q_1} + \beta_2 LandShare_i + \beta_3 SellerComposition_{n,q_1} \times LandShare_i + X_i' \beta_2 + \xi_n + \phi_{q_1,q_2} + \epsilon_i \quad (5)$$

The results are presented in Table 8, for neighborhoods defined as both zip codes and four-digit census tracts. The effect of all three measures of seller composition is bigger for houses with a larger land share. Column (1) shows that a move from the 25th to the 75th percentile of the land-share distribution (i.e., 47% land share to 75% land share) increases the response of annualized capital gains to a one conditional standard deviation increase in the share of

Table 8
Effect of seller composition by land share

	(1)	(2)	(3)	(4)	(5)	(6)
Land share	−0.605***	7.974***	−4.344**	−0.681***	7.672***	−3.202***
	(0.107)	(0.961)	(2.026)	(0.098)	(0.435)	(1.068)
Share informed sellers	−1.604			1.385		
	(1.153)			(0.435)		
Land share × share informed sellers	−10.02***			−4.470***		
	(2.332)			(0.839)		
Average seller land share		−7.633***			1.651***	
		(1.953)			(0.637)	
Land share × average seller land share		−15.45***			−14.68***	
		(1.592)			(0.783)	
Share in NH of tenure > 3			4.273***			1.197*
			(1.024)			(0.634)
Land share × share in NH of tenure > 3			4.123***			2.746***
			(1.476)			(0.918)
Neighborhood	Zip code	Zip code	Zip code	Census tr.	Census tr.	Census tr.
Fixed effects, property and financing controls	✓	✓	✓	✓	✓	✓
R-squared	0.637	0.638	0.659	0.638	0.640	0.660
$\bar{y}$	12.56	12.56	13.70	12.56	12.56	13.70
N	391,837	391,837	300,108	391,802	391,802	300,084

Table 8 shows results from Regression 5. The dependent variable is the annualized capital gain of a property between the two repeat sales. The seller composition variables are measured at the quarter × zip code level in Columns (1) through (3), and at the quarter × four-digit census tract level in Columns (4) through (6). Columns (1) through (3) include sales quarter pair fixed effects and zip code fixed effects, whereas Columns (3) through (6) include sales quarter pair and census tract fixed effect. All specifications control for characteristics of the property (property size, property age, property type, number of bedrooms and bathrooms, and whether the property has a pool or air conditioning) and characteristics of the financing (mortgage type, mortgage duration, and loan-to-value ratio). Columns (3) and (6) include sales pairs where the first sale was after June 1997, and all other columns include sales pairs where the first sale was after June 1994. Standard errors are clustered at the initial quarter × zip code level. Significance levels: * (p <0.10), ** (p <0.05), and *** (p <0.01).

informed sellers in a zip code by 8 basis points. A similar move in the land-share distribution increases the response of capital gains to a one conditional standard deviation increase in the average land share of sellers by 19 basis points (Column 2). Moving from the 25th to the 75th percentile of the land-share distribution increases the response of capital gains to a one conditional standard deviation change in the share of short-tenured sellers by 8 basis points (Column 3). Columns (4) through (6) show similar effects when we define a neighborhood as a census tract.

3.7 Relative informedness of buyers

In an environment with important information asymmetries, we would also expect that more informed buyers obtain higher average appreciation (Prediction 3), and that this effect is especially strong conditional on buying houses from overrated neighborhoods (Prediction 4). To test these predictions, we construct three measures of better-informed buyers. Our first measure presumes that real estate professionals are more informed about the true value of houses on sale, and tests the predictions by replacing $InformedBuyer_i$ in Regression 6 with a dummy variable for whether or not the buyer was a real

Table 9
Effect of buyer characteristics: Real estate professionals

	(1)	(2)	(3)	(4)
Real estate professional	0.737***	0.530***	0.393	1.778***
	(0.046)	(0.073)	(0.270)	(0.630)
Share informed sellers		−5.142***		
		(0.873)		
Real estate professional × share informed sellers		4.703***		
		(1.497)		
Average land share			−17.73***	
			(0.767)	
Real estate professional × average land share			0.586	
			(0.461)	
Share in zip of tenure > 3				6.956***
				(0.379)
Real estate professional × share in zip of tenure > 3				−1.329*
				(0.791)
Fixed effects, property and financing controls	✓	✓	✓	✓
R-squared	0.637	0.637	0.638	0.659
$\bar{y}$	12.56	12.56	12.56	13.70
N	391,837	391,837	391,837	300,107

Table 9 shows results from Regression 6. The dependent variable is the annualized capital gain of a property between two sequential arms-length sales. In Columns (1) through (3), sales pairs are included when the first sale was after June 1994, and in Column (4) when the first sale was after June 1997. All specifications include sales quarter pair fixed effects, zip code fixed effects, and control for characteristics of the property (property size, property age, property type, number of bedrooms and bathrooms, and whether the property has a pool or air conditioning) and characteristics of the financing (mortgage type, mortgage duration, and loan-to-value ratio). Standard errors are clustered at the initial quarter × zip code level. Significance levels: * (p <0.10), ** (p <0.05), and *** (p <0.01).

estate agent.

$$\begin{aligned} CapGain_{i,n,q_1,q_2} = & \alpha + \beta_1 SellerComposition_{n,q_1} + \beta_2 InformedBuyer_i \\ & + \beta_3 SellerComposition_{n,q_1} \times InformedBuyer_i \\ & + X_i' \beta_2 + \xi_n + \phi_{q_1,q_2} + \epsilon_i \end{aligned} \quad (6)$$

Table 9 shows the results from this regression. In Column (1), which tests Prediction 3, we do not include the measure of seller composition or its interaction with the informed buyer measure. Real estate agents purchase houses that outperform by about 74 basis points annually, relative to otherwise observationally similar houses purchased by individuals that are not real estate agents. This is consistent with real estate agents being better at picking good deals from the set of homes on offer.[16] In Columns (2) through (4), we show that, consistent with Prediction 4, the difference in the capital gain of houses purchased by real estate agents and those bought by other individuals is particularly big in neighborhoods that are predicted to underperform. As discussed above, the reason for this is that in underrated neighborhoods,

16 One might conjecture that some of this outperformance accrues to realtors that are not buying with better information, but rather act as inventory-holding, liquidity-providing dealers in this market. In the Online Appendix we show that this cannot explain our findings.

Table 10
Effect of buyer characteristics: Same zip code

	(1)	(2)	(3)	(4)
Same zip	1.105***	0.905***	0.172	3.869***
	(0.072)	(0.119)	(0.452)	(0.822)
Share informed sellers		−4.980***		
		(1.722)		
Same zip × share informed sellers		4.827**		
		(2.248)		
Average land share			−17.72***	
			(1.998)	
Same zip × average land share			1.564**	
			(0.753)	
Share in zip of tenure > 3				7.071***
				(0.625)
Same zip × share in zip of tenure > 3				−3.462***
				(1.065)
Fixed effects, property and financing controls	✓	✓	✓	✓
R-squared	0.637	0.637	0.638	0.659
$\bar{y}$	12.56	12.56	12.56	13.70
N	391,837	391,837	391,837	300,107

Table 10 shows results from Regression 6. The dependent variable is the annualized capital gain of a property between two sequential arms-length sales. In Columns (1) through (3), sales pairs are included when the first sale was after June 1994, and in Column (4) when the first sale was after June 1997. All specifications include sales quarter pair fixed effects, zip code fixed effects, and control for characteristics of the property (property size, property age, property type, number of bedrooms and bathrooms, and whether the property has a pool or air conditioning) and characteristics of the financing (mortgage type, mortgage duration, and loan-to-value ratio). Standard errors are clustered at the initial quarter × zip code level. Significance levels: * (p <0.10), ** (p <0.05), and *** (p <0.01).

informed buyers find most houses to be a good deal, and thus behave similarly to uninformed buyers, who cannot tell good and bad houses apart. In overrated neighborhoods, however, informed buyers use their information to select only homes that are a particularly good bargain, whereas uninformed buyers continue to be unable to tell good and bad houses apart.

We construct two additional measures of the relative informedness of buyers, which exploit that we can observe the names of all buyers and sellers in the deeds records. First, for every buyer of a house, we check whether we observe someone with the same name to have bought or sold a different house in the same zip code in the past year. Because having lived in the same zip code should provide buyers with better information, as compared with that of buyers who have not done so, we classify such buyers as informed.[17] About 5% of all houses are bought by individuals who have previously lived in the same zip code. The results are presented in Table 10.

Column (1) shows that those buyers who previously owned a house in the same zip code purchase homes that have a 1.1-percentage-point higher annualized capital gain. This is consistent with Prediction 3. Columns (2)

[17] We only observe the previous location for individuals who were previous owners in a neighborhood. This means that we will assign a value of “0” to current buyers who have previously rented, even if they lived in the same neighborhood. This will downward-bias our estimate of the coefficients on $InformedBuyer$ and its interaction with the seller composition.

Table 11
Effect of buyer characteristics: Distance to previous home

	(1)	(2)	(3)	(4)
Log(distance to previous home)	−0.281***	−0.272***	0.022	−1.096***
	(0.018)	(0.030)	(0.085)	(0.235)
Share informed sellers		−6.132***		
		(1.840)		
Log(distance to previous home) × share informed sellers		−0.249		
		(0.626)		
Average land share			−19.56***	
			(1.240)	
Log(distance to previous home) × average land share			−0.512***	
			(0.144)	
Share in zip of tenure > 3				4.973***
				(0.930)
Log(distance to previous home) × share in zip of tenure > 3				1.067***
				(0.305)
Fixed effects, property and financing controls	✓	✓	✓	✓
R-squared	0.630	0.630	0.631	0.679
$\bar{y}$	12.82	12.82	12.82	13.73
N	99,472	99,472	99,472	68,972

Table 11 shows results from Regression 6. The dependent variable is the annualized capital gain of a property between two sequential arms-length sales. In Columns (1) through (3) sales pairs are included when the first sale was after June 1994, and in Column (4) when the first sale was after June 1997. All specifications include sales quarter pair fixed effects, zip code fixed effects, and control for characteristics of the property (property size, property age, property type, number of bedrooms and bathrooms, and whether the property has a pool or air conditioning) and characteristics of the financing (mortgage type, mortgage duration, and loan-to-value ratio). Standard errors are clustered at the initial quarter × zip code level. Significance levels: * (p <0.10), ** (p <0.05), and *** (p <0.01).

through (4) test Prediction 4. Again, the coefficients on the interaction terms suggest that the effect of seller composition on the capital gains of houses bought by neighborhood insiders is significantly lower.

We also generate a second, more continuous measure of buyer informedness based on the previous residence of buyers. For those buyers that we observe selling a house anywhere in Los Angeles County within 12 months of the purchase, we construct a measure of the log-distance in kilometers between the house they sold and the house they bought to proxy for $InformedBuyer_i$ in Regression 6.[18] This variable has a mean of 2.01, and a standard deviation of 1.48. We conjecture that the further these buyers previously lived from the house they are now purchasing, the less likely they are to have information about neighborhood trends. The results are presented in Table 11. The sample size is smaller than for our other regressions, because we do not always find a previous seller with the same name. The balance of homes is bought either by people who were previously renters or by people moving from outside of Los Angeles County.

[18] For houses bought by an individual with a name that shows up more than once as a seller in the previous 12 months we take the distance to the geographically closest sale. The results are very similar when we pick the average distance across all observed sales.

Column (1) shows that buyers who previously lived further away buy houses that have a lower appreciation than otherwise similar houses bought by persons who lived closer by. This suggests that individuals who lived closer had superior information about characteristics of the neighborhood that allowed them to pick better deals. Again, Columns (2) through (4) show that the capital-gains difference between those houses bought by neighborhood insiders and those bought by outsiders is particularly big in overrated neighborhoods (i.e., those where the share of informed sellers and the average land share of sold homes is high, and the share of long-tenure sellers is low).

In this section we provided evidence that buyers who have had experience in the same zip code, who lived closer by, and who are real estate professionals purchase houses that subsequently outperform otherwise similar homes bought by less-informed agents. Their superior information allows them to pick better properties. This is consistent with Prediction 3. In addition, and consistent with Prediction 4, this advantage is lower when the predicted neighborhood quality is higher. This provides further evidence for our interpretation that the correlation between seller composition and subsequent price changes is driven by superior information of the sellers. If seller composition were just a proxy for commonly known information that is not observed by the econometrician, this would not explain why better-informed buyers outperform, or why seller composition is less predictive of returns for houses bought by more-informed buyers.

4. Conclusion

In many markets, sellers are better informed than buyers are about the true value of the traded asset. In addition, there might also be information heterogeneity among both buyers and sellers. We argue that residential real estate is an example of this type of market. Sellers are better informed than buyers are about both neighborhood characteristics and structural attributes of a house, but among both buyers and sellers, some are better informed than others are. We propose a new framework for empirically studying such markets with many heterogeneous assets and differentially informed agents. We then analyze the universe of housing transactions in Los Angeles County between 1994 and 2011, to quantify the effect of this type of asymmetric information on equilibrium market outcomes. We find that changes in the seller composition toward more informed sellers and sellers with a larger supply-elasticity predict subsequent price declines of houses in that neighborhood. This finding is unaffected by the inclusion of past price changes as a control variable, and is larger for houses with a value that depends more on neighborhood characteristics.

Importantly, our framework allows us to consider the role of differentially informed buyers. This generates a set of additional predictions that are unique to a model with asymmetric information, and allows us to reject alternative explanations that rely on a differential elasticity of reacting to commonly known information, for example, because of differential transaction costs. We find that more informed buyers buy houses that experience higher *ex post* appreciation. Importantly, we also find that the correlation between seller composition and subsequent returns is smaller for houses bought by more informed buyers. Overall, our findings suggest that homeowners have superior information about important neighborhood characteristics, and exploit this information to time local market movements.

It is well known that asymmetric information can severely undermine the liquidity of markets. Many markets deal with this problem through some combination of regulations, such as laws against insider trading, and contractual practices, such as seller warranties. In real estate markets, legal disclosure requirements and the involvement of real estate agents are intended in part to mitigate the natural information advantage of sellers over buyers. Our results suggest that there remains substantial information asymmetry, involving hard-to-observe features of both neighborhoods and houses, that is immune to these remedies. Further, the differential information is not limited to a difference between buyers and sellers, but rather it exists within each of these groups, which creates an advantage for those who are more informed relative to their peers.

Even though our empirical analysis focuses on the residential real estate market, the information structure we consider is similar in other important financial markets. For example, in the venture-capital market, the success of start-up firms is a combination of both industry-level factors and the skills of the individual entrepreneurs. Some venture capitalists are better at identifying promising companies (either in promising industries, or with skilled entrepreneurs) than others are (Hochberg, Ljungqvist, and Lu 2007). Because investment term sheets are usually not publicly disclosed and venture capital investments are indivisible (i.e., it is not possible for other firms to automatically co-invest with more informed venture capitalists at the same terms), less-informed venture capitalists cannot learn about the value of individual companies by observing prices paid by more-informed investors. Similar empirical tests could therefore be conducted to test for the magnitude of asymmetric information both between venture-capital investors as well as between entrepreneurs.

References

Admati, A. 1985. A noisy rational expectations equilibrium for multi-asset securities markets. *Econometrica* 53:629–57.

Akerlof, G. 1970. The market for "lemons": Quality uncertainty and the market mechanism. *Quarterly Journal of Economics* 84:488–500.

Albouy, D. 2009. What are cities worth? Land rents, local productivity, and the value of amenities. Working Paper, 14981, NBER.

Arnott, R., and J. Stiglitz. 1979. Aggregate land rents, expenditure on public goods, and optimal city size. *Quarterly Journal of Economics* 93:471–500.

Bayer, P., C. Geissler, and J. Roberts. 2011. Speculators and middlemen: The role of flippers in the housing market. Working Paper, 16784, NBER.

Brunnermeier, M., and S. Nagel. 2004. Hedge funds and the technology bubble. *Journal of Finance* 59:2013–40.

Caballe, J., and M. Krishnan. 1994. Imperfect competition in a multi-security market with risk neutrality. *Econometrica* 62:695–704.

Case, K., and R. Shiller. 1989. The efficiency of the market for single-family homes. *American Economic Review* 79:125–37.

———. 1990. Forecasting prices and excess returns in the housing market. *Real Estate Economics* 18:253–73.

Cheng, I.-H., S. Raina, and W. Xiong. 2013. Wall street and the housing bubble. Working Paper, 18904, NBER.

Chinco, A., and C. Mayer. 2014. Distant speculators and asset bubbles in the housing market. Working Paper, 19817, NBER.

Choi, J., L. Jin, and H. Yan. 2013. Informed trading and expected returns. Working Paper, 19420, NBER.

Cohen, L., A. Frazzini, and C. Malloy. 2008. The small world of investing: Board connections and mutual fund returns. *Journal of Political Economy* 116:951–79.

Coval, J., and T. Moskowitz. 2001. The geography of investment: Informed trading and asset prices. *Journal of Political Economy* 109:811–41.

Davis, M., and J. Heathcote. 2007. The price and quantity of residential land in the united states. *Journal of Monetary Economics* 54:2595–620.

Dubey, P., and J. Geanakoplos. 2002. Competitive pooling: Rothschild-stiglitz reconsidered. *Quarterly Journal of Economics* 117:1529–70.

Easley, D., S. Hvidkjaer, and M. Ohara. 2002. Is information risk a determinant of asset returns? *Journal of Finance* 57:2185–221.

Finnerty, J. 1976. Insiders activity and inside information: A multivariate analysis. *Journal of Financial and Quantitative Analysis* 11:205–15.

Gale, D. 1992. A walrasian theory of markets with adverse selection. *Review of Economic Studies* 59:229–55.

———. 1996. Equilibria and pareto optima of markets with adverse selection. *Economic Theory* 7:207–35.

Garmaise, M., and T. Moskowitz. 2004. Confronting information asymmetries: Evidence from real estate markets. *Review of Financial Studies* 17:405–37.

Ghysels, E., A. Plazzi, R. Valkanov, and W. Torous. 2013. Forecasting real estate prices. In Handbook of economic forecasting, vol. 2, Part A, ed. G. Elliott and A. Timmermann, 509-80, Elsevier.

Grossman, S., and J. Stiglitz. 1980. On the impossibility of informationally efficient markets. *American Economic Review* 70:393–408.

Guerrieri, V., D. Hartley, and E. Hurst. 2013. Endogenous gentrification and housing price dynamics. *Journal of Public Economics* 100:45–60.

Guerrieri, V., R. Shimer, and R. Wright. 2010. Adverse selection in competitive search equilibrium. *Econometrica* 78:1823–62.

Hellwig, M. 1987. Some recent developments in the theory of competition in markets with adverse selection. *European Economic Review* 31:319–25.

Hochberg, Y., A. Ljungqvist, and Y. Lu. 2007. Whom you know matters: Venture capital networks and investment performance. *Journal of Finance* 62:251–301.

Jaffe, J. 1974. Special information and insider trading. *Journal of Business* 47:410–28.

Kelly, B., and A. Ljungqvist. 2012. Testing asymmetric-information asset pricing models. *Review of Financial Studies* 25:1366–413.

Kodres, L., and M. Pritsker. 2002. A rational expectations model of financial contagion. *Journal of Finance* 57:769–99.

Kurlat, P. 2014. Asset markets with heterogeneous information. Working Paper, Stanford University.

Kyle, A. 1985. Continuous auctions and insider trading. *Econometrica* 53:1315–35.

Kyle, A., and W. Xiong. 2001. Contagion as a wealth effect. *Journal of Finance* 56:1401–40.

Levitt, S., and C. Syverson. 2008. Market distortions when agents are better informed: The value of information in real estate transactions. *Review of Economics and Statistics* 90:599–611.

Lin, J.-C., and J. Howe. 1990. Insider trading in the otc market. *Journal of Finance* 45:1273–84.

Lorie, J., and V. Niederhoffer. 1968. Predictive and statistical properties of insider trading. *Journal of Law & Economics* 11:35–53.

Piazzesi, M., M. Schneider, and J. Stroebel. 2015. Segmented housing search Working Paper, 20823, NBER.

Seyhun, N. 1986. Insiders' profits, costs of trading, and market efficiency. *Journal of Financial Economics* 16:189–212.

———. 1992. Why does aggregate insider trading predict future stock returns? *Quarterly Journal of Economics* 107:1303–31.

Sockin, M., and W. Xiong. 2014. Learning about the neighborbood: A model of housing cycles.

Stroebel, J. 2014. Asymmetric information about collateral values. *Journal of Finance* Forthcoming.

Temin, P., and H.-J. Voth. 2004. Riding the south sea bubble. *American Economic Review* 94:1654–68.

Wilson, C. 1980. The nature of equilibrium in markets with adverse selection. *The Bell Journal of Economics* 11:108–30.