Information Frictions and Housing Market Dynamics

Elliot Anenberg

2012-48

Information Frictions and Housing Market Dynamics*

Elliot Anenberg[†]

July 18, 2012

Abstract

This paper examines the effects of seller uncertainty over their home value on the housing market. Using evidence from a new dataset on home listings and transactions, I first show that sellers do not have full information about current period demand conditions for their homes. I incorporate this type of uncertainty into a dynamic search model of the home selling problem with Bayesian learning. Simulations of the estimated model show that information frictions help explain short-run persistence in price appreciation rates and a positive (negative) correlation between price changes and sales volume (time on market).

1 Introduction

Since the seminal work of Stigler [1961], economists have long recognized the importance of imperfect information in explaining the workings of a variety of markets. Surprisingly, given its importance to the macro economy, little work has focused on the effects of imperfect information in the housing market.[1] The housing market is a classic example of a market affected by imperfect information. Each house is a unique, differentiated asset; trading volume of comparable homes tends to be thin due to high transaction costs; and market conditions are highly volatile over time. These features of the housing market make it difficult for sellers to determine their home values at any point in time.

In this paper I model the effect of this type of seller uncertainty on the housing market. The model adds a framework for seller uncertainty and Bayesian learning in the spirit of Lazear [1986] to the typical features of the dynamic micro search models in the housing literature (Carrillo [2010], Horowitz [1992], Salant [1991]). I estimate the model and use it to test whether uncertainty is important for explaining several key stylized facts about housing market dynamics that have attracted much attention in the literature, in part because they are inconsistent with the predictions of standard asset pricing models.

The first fact is that price appreciation rates display predictability in the short-run. In their seminal papers, Case and Shiller [1989] and Cutler et al. [1991] find that a 1 percent increase in real annual house prices is associated with a .2 percent increase the next year, adjusting for changes in the nominal interest rate. Numerous other studies have also documented this persistence, and it has led some to question the efficiency of the housing market because it cannot be explained by fundamentals.[2] Thus, an important question is whether the amount of short-run momentum found in the data is consistent

---

[1]Levitt and Syverson [2008] and Taylor [1999] are examples of studies that focus on the effect of information asymmetries on micro features of the data, but less is understood about the broader effects of information frictions on housing market dynamics.

[2]For example, Case and Shiller [1989] and Glaeser and Gyourko [2006] find that persistence in rents and wages cannot explain the amount of serial correlation in prices, and thus cite inefficiency as a likely explanation. See Cho [1996] for a survey of the literature on house price dynamics.

with a rational model of the housing market.

Another key feature of housing markets is that sales volume and marketing time are positively and negatively correlated with sales price changes, respectively. The existing literature has convincingly emphasized the importance of search frictions and credit constraints as explanations.[3] Here, I investigate whether imperfect information is an additional friction that causes housing markets to display these unusual time series properties.

At a more micro level, the literature has also documented a set of stylized facts about the behavior of individual sellers (Merlo and Ortalo-Magne [2004]). For example, sellers tend to adjust their list prices downwards, even when market conditions do not change, and sales prices for observationally equivalent homes depend on time on market (TOM). These empirical patterns are inconsistent with the predictions of existing search, matching and bargaining theories of housing transactions, which do not accommodate duration dependence in seller behavior.[4] In addition to explaining the macro stylized facts discussed above, I will show that uncertainty and the gradual acquisition of information during the listing period is an explanation for these and many other dynamic features of the micro data.

I motivate the model with reduced form evidence that lack of information does affect selling behavior. While several studies show that homeowners misestimate their home values at various points during their ownership tenure[5], I am not aware of other studies that investigate the information set of sellers when the home is on the market for sale. Since most sellers hire a realtor when they are ready to sell their homes, this is potentially an important distinction. The reduced form evidence comes from a new micro dataset that I compile from two independent sources. For a large sample of single family homes listed for sale with a realtor in the two major California metropolitan areas from 2007-2009, the combined dataset describes the precise location of each listing, list prices each week that the home is listed for sale, TOM, and

---

[3]See Stein [1995], Ortalo-Magne and Rady [2006], Genesove and Mayer [1997], Anenberg [2011a], Krainer [2001], Ngai and Tenreyro [2010].
[4]See Carrillo [2010], Horowitz [1992], Novy-Marx [2009], Chen and Rosenthal [1996].
[5]See for example, Goodman and Ittner [1992], Kiel and Zabel [1999].

sales prices.

I exploit cross-sectional variation in the heterogeneity of the housing bust across neighborhoods to test whether sellers price their homes using the most up to date information about local market conditions. I find that they do not; initial list prices are overly sensitive to lagged market information. The regression results show that for two comparable homes in a given time period, the home in the neighborhood that experienced the greater amount of price depreciation in the previous four months will be listed at a *higher* price on average. Neighborhood price levels from longer than four months ago do not provide any additional explanatory power for initial list prices. This finding is consistent with anecdotal evidence that sellers, with the help of realtors, look to previous sales of similar houses when pricing their homes, presumably because comparable sales volume can be thin and sales prices become publicly available with a lag. Evidently, realtors do not fully adjust for the downward trend in prices during my sample period.

Using list prices to infer seller beliefs is complicated by the fact that many unobservables affect list prices, and some of them may be correlated with lagged market information. However, the correlation between *sales* prices and lagged neighborhood price depreciation confirms that expectation bias is indeed the explanation for the inflated list prices. In particular, I find that for the same homes that have higher list prices because of high lagged neighborhood price depreciation, sales prices are lower. Theoretically, higher list and lower sales prices can arise when sellers overstate their home value, but this crossing pattern, which is shown in Figure 1, cannot easily be explained by many alternative explanations for the high list prices including loss aversion (Genesove and Mayer [2001], Anenberg [2011a]), equity constraints (Stein [1995],Genesove and Mayer [1997], Anenberg [2011a]), high unobserved home quality, or low unobserved motivation to sell, among others. In other words, unobservables that increase list prices should also increase sales prices.[6] This

---

[6]In Figure 1, the two stylized neighborhoods trend at the same rate after the listing date. In the empirical specification, I control for heterogeneous trends across neighborhoods by normalizing list and sales prices by a time varying, neighborhood specific sales price index.

simple identification strategy for expectation bias could be useful in other settings where list prices or reserve prices are observed in addition to selling prices.

Having established this new evidence that information frictions are an important part of the home selling process, I incorporate them into a single-agent dynamic micro model of the home selling problem. I model seller uncertainty as a prior on the mean of the distribution of buyer valuations for their home, and this prior may be biased depending on the information available. Sellers set list prices to balance a trade-off: a high list price strengthens their bargaining position if a buyer arrives, while a low list price attracts more offers and increases the pace of learning. Conditional on the list price, buyers with idiosyncratic valuations arrive randomly. The house sells if the buyer's valuation is above the seller's reservation price, which depends on the value of declining the offer and continuing the dynamic process. Sellers in the model behave rationally and optimally given the available information.

I estimate the parameters of the model using simulated method of moments.[7] The parameter estimates themselves are informative about the amount of information that sellers have and the pace of learning. I find that the standard deviation of the typical seller's prior about their home value is \$38,000, which is about 7 percent of the average sales price. Learning over the course of the marketing period decreases this standard deviation by 37 percent by the time of sale, on average.

Simulations of the estimated model show that annual aggregate sales price appreciation rates persist even when changes in the market fundamentals do not. The model can account for over half of the persistence typically found in the data. To see the intuition behind this result, suppose that there is uncertainty about demand at time t, the expected value of demand is γ at time t, and the realization of a permanent demand shock is higher than expected at $\gamma + \epsilon$. Even if every seller receives an idiosyncratic signal at time t that

---

[7]In this respect, my paper contributes a new application to the growing literature on empirical learning models. See, for example, Crawford and Shum [2005], Ackerberg [2003], Hitsch [2006], Erdem and Keane [1996], Narayanan et al. [2007]. I highlight a couple unique features of my learning model in Section 4.1.

demand is high, in the absence of a mechanism (either formal or informal as in Grossman and Stiglitz [1976]) that publicizes private information, the reservation price of the average Bayesian updating seller will not fully adjust to the shock at time t. In subsequent periods, after more information about the positive shock becomes available, reservation prices, and thus sales prices, will fully adjust. It is this lag in the flow of information that gives rise to serial correlation in price changes.

The same lag in the updating of reservation prices to demand shocks generates a positive (negative) correlation between price changes and sales volume (TOM). For example, when there is a positive demand shock, reservation prices are too low relative to the fundamentals, which leads to higher sales volume and quicker sales. The model predicts that a 1 percent increase in quarterly prices leads to a 4.8 percent increase in volume and a 7.2 percent decrease in TOM. These predicted co-movements are comparable to what is observed in the data. Thus, uncertainty does appear to be important in explaining variation in transaction rates over the housing cycle as well.[8]

A related paper by Head et al. [2011], written recently and independently of this one, also explores the serial correlation of house price changes. Under some calibrations and functional form assumptions, their macro model, which is one of complete information, is also able to generate some, but not all, of the momentum observed in the data. Their model, like mine, does not rely on inefficiency or irrationality of the market. Matching frictions in the spirit of Mortensen and Pissarides [1994] and Pissarides [2000] interact with a lagged housing supply response to cause market tightness (i.e. the ratio of buyers to sellers), and thus prices, to gradually rise in response to an income shock. Future efforts to explain even more of the momentum in housing market conditions could try incorporating information frictions into this type of search and matching framework.

This paper proceeds as follow. Section 2 introduces the data. Section 3

---

[8]The idea that uncertainty about market conditions can lead to slow adjustment of prices and sales is also discussed in Berkovec and Goodman [1996]. However, the idea is introduced with a more stylized version of the model in this paper, and their model is not taken to the micro data.

motivates the model with reduced form evidence that information frictions are an important part of the home selling problem. To investigate the broader implications of seller bias and uncertainty for housing market dynamics, Section 4 develops a model where the flow of information has an endogenous effect on selling behavior. Section 5 and 6 discuss estimation details. Section 7 simulates the model to highlight the importance of information frictions in explaining the stylized facts discussed above. Section 8 discusses the robustness of the results to certain stylized features of the model and section 9 concludes the paper.

2 Data

I use home sale and listing data for the core counties of the San Francisco Bay Area and Los Angeles. These counties include Alameda, Contra Costa, Marin, San Francisco, San Mateo, and Santa Clara in San Francisco; and Los Angeles, Orange, Riverside, San Bernardino, and Ventura counties in the Los Angeles area.

The listing data come from Altos Research, which provides information on single-family homes listed for sale on the Multiple Listing Service (MLS) from January 2007 - June 2009. Altos Research does not collect MLS data prior to 2006. Since a seller must use a licensed real estate agent to gain access to the MLS, my sample only contains selling outcomes for sellers who use realtors.[9] Every Friday, Altos Research records the address, mls id, list price, and some characteristics of the house (e.g. square feet, lot size, etc.) for all houses listed for sale. From this information, it is easy to infer the date of initial listing and the date of delisting for each property. A property is delisted when there is a sale agreement or when the seller withdraws the home from the market. Properties are also sometimes delisted and then relisted in order to strategically reset the TOM field in the MLS. I consider a listing as new only if there was at least a 180 day window since the address last appeared in

---

[9]According to the National Association of Realtors, over 90 percent of non-arms length home sales were listed on the MLS in 2007.

the listing data.

The MLS data alone does not provide information on which listings result in a sale, and what the sales price is if a sale occurs. To obtain this information, I supplement the MLS data with a transactions dataset from Dataquick that contains information about the universe of housing transactions from 1988-2009. In this dataset, the variables that are central to this analysis are the address of the property, the date of the transaction, and the sales price.

Using the address, I attempt to merge each listing to a transaction record that is within 1 year of the date of delisting from the MLS.[10] I also attempt to merge each listing to a previous sale in the transaction dataset. The latter merge acquires information on the purchase price of each home, which I use to construct a predicted log selling price for each house:

$$\hat{p}_{ijt} : \quad \text{log predicted sales price for house } i \text{ located in neighborhood } j \text{ in month } t.$$

I calculate these prices by applying a zip code price change index to the previous log sales price. The price index is calculated using a repeat sales analysis following Shiller [1991]. I let the price index vary by zip code and month. The predicted price measures what the economist expects house i to sell for in time t, and it controls for time-invariant unobserved home quality and differences in neighborhood price appreciation rates. Appendix A.2 describes how I calculate these prices from the data in more detail.

Appendix A.1 describes more details of the data building process, including minor restrictions to the estimation sample (e.g. exclude listings where the ratio of the minimum list price to the maximum list price is less than the first percentile). I exclude listings where the initial listing date equals the first week of the sample and listings where the final listing data equals the last week of the sample to avoid censoring issues. I also drop all listings that do not merge to a previous transaction.[11]

---

[10]The sales date in the transaction data is the closing date, which lags the agreement date by a month on average.

[11]I compare summary statistics of the limited sample to the full sample to ensure that my sample is representative. The failure to merge here is because there is an idiosyncracy

8

2.1 Summary Statistics

Figure 2 shows the Case-Shiller home price index for Los Angeles and San Francisco from 2007 - 2009. During the years where the MLS and transactions data overlap, both cities experienced comparable and significant price declines. In Los Angeles, the market peaked in September 2006, and fell 37.5 percent in nominal terms through December 2009. The San Francisco market peaked in May 2006, and also fell by 37.5 percent by the end of 2009. The prolonged episode of falling prices, low sales volume, and long marketing time that my sample period covers is not an isolated event; Burnside et al. [2011] show that sustained booms and busts occur throughout housing markets around the world. For example, in real terms, Los Angeles experienced a comparable price decline during the housing bust of the early 90's.[12] Thus, my sample and my results like characterize market dynamics during cold housing markets more generally.

Table 1 presents summary statistics for the listings that sell. The median time to sale is about 3 months, and there is a lot of variation. Twenty five percent of listings sell in less than 5 weeks and 25 percent take more than 25 weeks to sell. Most sellers adjust their list price at least once before they sell. These list price changes tend to be decreases: only 6 percent of list price changes are increases.[13] Table 2 shows that list price changes occur throughout the selling horizon, and many occur in the first few weeks after listing. These stylized facts about list price changes seem challenging for models with complete information to explain.[14] Since some sellers will quickly adjust their

---

in the way the addresses are recorded, the house is new, or the current owner purchased the house prior to 1988.

[12]I also find evidence that the percent of transactions that are foreclosures during the recent downturn is comparable to the downturn during the 1990's. Campbell et al. [2011] also report that the foreclosure rate is not unusually high during the recent recession relative to the downturn during the 1990's in Massachusetts.

[13]50 percent of listings do not sell during the sample period. These withdrawals tend to have longer marketing times and higher list prices (normalized by predicted price) relative to listings that sell. I discuss reasons for these withdrawals in more detail in Section 8.

[14]For example, list prices could decline in a complete information framework if holding costs increase over time. But to generate list price increases during a market decline and to generate list price changes so quickly after listing likely requires a very flexible parameteri-

beliefs in response to new information, the learning model that I present below will also predict changing list prices in the first few weeks and some list price changes that are increases.

Few studies have had access to such a large dataset on home listings that includes the full history of list prices for each listing.[15] This feature of the data will be important for identifying the parameters of the non-stationary model of selling behavior presented below.

3 Motivating Empirical Facts

I begin by presenting strong evidence, which does not rely on my modeling assumptions below, that imperfect information does affect the home selling process. This is important because even though many features of the data are consistent with a model with uncertainty and learning as I show below and as documented in Knight [1996], alternative models may be able to explain these features as well.

3.1 Expectation Bias and List Prices

I test whether the initial list price choices of sellers reflect the most up to date market information, or whether they place undue weight on lagged information. Outdated information may affect list prices because the thinness of the market and the lag in which sales data become public make it difficult to assess current market conditions.[16] Conversations with a realtor suggest that lagged

---

zation of holding costs.

[15]A related study that uses a similar dataset from the Netherlands is de Wit and van der Klaauw [2010]. The authors find empirical evidence that list price changes affect selling outcomes such as the hazard rate of sale. Since list prices have no legal role, they interpret this as evidence that information frictions do exist in the housing market. My study is different from theirs in that they do not actually model the information frictions or address the implications of these frictions on market dynamics.

[16]Sales data become available only upon closing, which typically lags the date when the buyer and seller agree on price by months. In addition, home price indexes (e.g. Case-Shiller) that process sales data using econometric techniques lag the market by months.

comparable sales are often used as a proxy for the current market value.[17]

I implement this test by regressing the log list price in the initial week of listing, normalized by the log predicted sales price, on lagged price changes according to:

$$p_{i,j,t_0}^L - \hat{p}_{i,j,t_0} = \alpha_j + \gamma_t + \beta_1(\hat{p}_{i,j,t_0-1} - \hat{p}_{i,j,t_0}) + \beta_2 X_{i,j,t} + \epsilon_{i,j,t} \qquad (1)$$

where the 0 subscript denotes that it is $i's$ first month on the market and j denotes the neighborhood. Variation in the dependent variable across listings could be due to several factors, including heterogeneous motivation to sell, time-varying house characteristics, and heterogeneous beliefs about house values. The regressor of interest is the percentage change in average neighborhood prices from the previous month relative to the current month. The average value of $\hat{p}_{i,j,t_0-1} - \hat{p}_{i,j,t_0}$ is 1.7 percent and the standard deviation is 2.9 percent. I also include month fixed effects, zip code fixed effects, other controls X, and so β_1 is identified from heterogeneity in the variation in price declines across neighborhoods.

Column 1 of Table 3 reports the results. Standard errors are clustered at the zip code level. A 1 percent increase in the price depreciation rate leads to a 0.57 percent increase in the list price, all else equal. That this estimate is less than one suggests that realtors have some information, just not perfect information, that market conditions have deteriorated. In Columns 2-5, I continue to add lagged neighborhood price changes as regressors until the estimated coefficient becomes insignificant. One month price changes immediately before listing have the biggest effect on list price premiums, and the effect of 1 month price changes diminish as they occur further before the month of listing. A price change between month $t_0 - 5$ and $t_0 - 4$ does not affect the list price premium. It makes sense that the most recent price changes are the least capitalized into list prices because the least information is available about these

---

[17]My language suggests that sellers and realtors have the same objective function, even though the empirical evidence suggests otherwise (Levitt and Syverson [2008]). I do not feel that this distinction is important for my analysis since the results are identified off of cross-sectional variation and all of the sellers in my sample use realtors.

changes.[18]

In Table 4, I test how the list price premium varies over the entire distribution of lagged depreciation. The regression specification is

$$p_{i,j,t_0}^L - \hat{p}_{i,j,t_0} = \alpha_j + \gamma_t + \sum_{k=2}^{10} \alpha_k I[\Delta_{ijt}^4 < d_k] + \beta_2 X_{i,j,t} + \epsilon_{i,j,t} \qquad (2)$$

where

$$\Delta_{ijt}^4 = \hat{p}_{i,j,t_0-4} - \hat{p}_{i,j,t_0} \qquad (3)$$

is the local price change over the four months prior to initial listing, I is the indicator function, and d_k is the kth decile of the distribution of Δ^4 across all the listings in my sample.[19] As shown in Table 4, the higher the lagged price depreciation, the higher the list price, and the relationship is monotonic over the entire Δ^4 distribution. Coefficients in bold denote cases where the difference in the coefficient relative to the coefficient in the decile immediately below is statistically significant. The results are similar in column 2, where I restrict the sample only to listings that eventually sell.[20]

3.2   Expectation Bias and Selling Outcomes

The previous section showed evidence that sellers set higher list prices when their local market is declining at a faster than average rate. I interpret this as expectation bias. In this section, I test whether market deterioration affects other variables such as the sales price and marketing time. Here, I find patterns that are consistent with expectation bias, but not with other plausible explanations for the list price results.

Columns 3 and 4 of Table 4 substitute TOM as the dependent variable in equation (2) using the full sample and the sample of only sales, respectively.

---

[18]The results are similar when I restrict the sample to listings that sell.

[19]d_{10} denotes the largest price declines. The median value of Δ^4 is 6 percent.

[20]To test the robustness of these results to my assumption about off-market properties described in Section 2, I ran these regressions treating each listing where the address does not appear in the dataset in the week prior, as a new listing. The results are largely unchanged. In Table 3, the coefficient on $\hat{p}_{i,j,t_0-5} - \hat{p}_{i,j,t_0-4}$ becomes marginally significant.

TOM is increasing in Δ^4, although the extreme decile of the price change distribution appears to be an outlier. Column 5 shows that the propensity to withdraw is also increasing in Δ^4, and monotonic over the entire Δ^4 distribution. In this specification, I include in X an additional control for the change in price level during the marketing period, $\hat{p}_{i,j,T} - \hat{p}_{i,j,t_0}$, where T denotes the time period that the house sells or is withdrawn.

Column 6 reports results where the dependent variable is the log sales price normalized by the predicted log sales price in the month of sale, i.e. $p_{i,j,T}^{sale} - \hat{p}_{i,j,T}$. Sales prices are significantly decreasing in the lower deciles of Δ^4, but are flat or slightly increasing in the higher deciles.

3.3   Discussion

That higher lagged depreciation leads to longer marketing time and a higher propensity to withdraw is consistent with the expectation bias interpretation. Biased beliefs lead to higher reservation prices, which should increase marketing time in a standard search model. Biased beliefs may also draw sellers into the market, only to withdraw later once they realize that their home will not sell for what they initially expected.

The theoretical effect of inflated beliefs on sales price is ambiguous. For example, a high reservation price could cause sellers to stay on the market for longer, which allows them to sample more offers and ultimately receive a higher price at the expense of a longer time to sale. Inflated beliefs can also decrease sales prices if, for example, motivation to sell increases over time (e.g. there is a finite selling horizon). In this case, the seller is pricing too high, and potentially turning off potential buyers, exactly when he is most likely to accept higher offers.

For this reason, the sales price results alone do not tell us much about the existence of expectation bias. However, the fact that for some regions of the price decline distribution, list prices are significantly increasing while sales prices are significantly decreasing is an unusual pattern that is consistent with expectation bias but inconsistent with alternative explanations for the list

price, TOM and withdrawal results. A stylized case of this crossing property is illustrated in Figure 1. Finding a reasonable model where an omitted variable from equation (2) increases list prices above what is expected while decreasing sales price below what is expected is a challenge. In fact, finding a model where a variable increases list prices and leads to no change in sales price, which is the case for much of the Δ^4 distribution, is also difficult. For example, standard models of the home selling problem, including the model we present below, predict that unobservables such as high home quality, low motivation to sell, loss aversion, and equity constraints should all lead to higher list and higher sales prices.[21] Appendix A.3 presents additional results that are consistent with the conclusions established in this Section.

4 Model

4.1 Overview and Comparison with Related Literature

The heart of my model is similar to Carrillo [2010], Horowitz [1992], Salant [1991]. The seller's decision to list the home and sell it is taken as given, and the seller's objective is to maximize the selling price of the house less the holding costs of keeping the home on the market. My main contribution is to introduce uncertainty and Bayesian learning into this framework. This makes the home selling problem nonstationary; unlike in Carrillo [2010] and Horowitz [1992], sellers in my model will adjust their list prices over time and the hazard rate of selling varies over time as learning occurs.[22] I endogenize the effect of information on market dynamics while 1) only introducing parameters that are identified given the dataset described in Section 2 and 2) capturing the key features of the home selling process including search, a posting price mechanism, preference heterogeneity, and duration dependence in optimal seller behavior.

---

[21]Genesove and Mayer [2001] and Anenberg [2011a] present empirical evidence that loss aversion and equity constraints, which may be positively correlated with Δ^4, lead to higher list and higher sales prices

[22]Salant [1991] introduces nonstationarity into the search problem with a finite selling horizon, although sales prices in his model never differ from the asking price and he does not estimate his model.

I discuss extensions to the model and their implications for my conclusions in Section 8.

The way that I model uncertainty and learning is similar to the existing empirical learning models (see cites in footnote 7)), but my model is unique in two ways. First, I allow the parameter, μ, that agents are learning about to change over time. Second, agents in my model receive direct signals about the unknown parameter (i.e. μ + noise) as in the existing studies, but also receive signals (when buyers do not inspect their house) that the unknown parameter is below a known threshold (i.e. μ + noise < T). The latter innovation introduces some new computational challenges that I discuss in Section 4.5.

4.2 Offer Process and Buyer Behavior

At the beginning of each week t that the house is for sale, seller/house combination i selects an optimal list price, p_{it}^L. This list price and a subset of the characteristics of the house are advertised to a single risk-neutral potential buyer. From now on, I refer to these potential buyers as simply buyers. The logarithm of each buyer j's willingness to pay (or valuation) v_{ijt} is parameterized as

$$v_{ijt} = \mu_{it} + \eta_{ijt} \tag{4}$$

where μ_{it} is common across all buyers and η_{ijt} represents buyer taste heterogeneity. The distribution of buyer valuations is exogenous to the model.

I assume that

$$\eta_{ijt} \sim N(0, \sigma_\eta^2). \tag{5}$$

That is, η is iid across time, houses, and buyers.

The advertisement only provides the buyer with a signal of their valuation. From the advertisement, the buyer forms beliefs about v that are assumed

$$\text{buyer beliefs: } v_{ijt} \sim N(\hat{v}_{ijt}, \sigma_{\hat{v}}^2) \tag{6}$$

where $\hat{v}_{ijt}$ is drawn from $N(v_{ijt}, \sigma_{\hat{v}}^2)$. Thus, buyers get an unbiased signal of

their true valuation from the advertisement.[23]

After observing $\hat{v}_{ijt}$, the buyer decides whether to inspect the house at some cost, κ.[24] If the buyer inspects, then v is revealed to both the buyer and the seller. If $v < p^L$, the seller has all the bargaining power and has the right to make a 'take it or leave it' offer to the buyer at a price equal to v (which I assume the buyer will accept). If $v > p^L$, then the buyer receives some surplus: the buyer has the right to purchase the house at a price equal to the list price. If the buyer chooses not to inspect or if the buyer's valuation lies below the seller's continuation value of remaining on the market, then the buyer departs forever and the seller moves onto the next period with her house for sale.

The proof of the following theorem, which characterizes the buyer's optimal behavior, appears in Appendix A.4.

Theorem 1 *The optimal search behavior for the buyer takes the reservation value form. That is, the buyer inspects when $\hat{v} > \bar{v}$, and does not inspect otherwise. $\bar{v} = T^* + p^L$ where T^* is a function of the parameters $(\kappa, \sigma_{\hat{v}}^2)$.*

This price determination mechanism delivers a closed form relationship between $\bar{v}$ and the list price, which is necessary to keep estimation tractable given that the list price choice will be endogenous. Since the buyer receives no surplus when $v < p^L$, $\bar{v}$ does not depend on the seller's reservation price or any other variable (like TOM) that provides a signal about the seller's reservation price.

This simple model of buyer behavior endogenizes the list price and leads to a trade-off (from the seller's perspective) when setting the list price between sales price and TOM.[25] The model also generates a mass point at the list price in the sales price distribution. These predicts are consistent with the empirical

---

[23]This specification of beliefs would arise if buyers had flat priors (i.e. prior variance = ∞) and processed the signal, $\hat{v}$, according to Bayes' rule.

[24]The inspection cost should be interpreted broadly as the cost of making a serious offer, including the inspection itself and the opportunity costs of lost time.

[25]The way the list price mechanism works in Carrillo [2010] is as follows. With exogenous probability θ, the list price serves an a take-it-or-leave-it offer to the buyer. With probability $(1 - \theta)$, the buyer can make an offer that extracts the entire surplus from the transaction.

evidence, and with the theoretical literature on the role of asking prices as a commitment device.[26]

4.3 The μ process

The underlying valuation process, μ_{it}, is exogenous to the model. It is not affected by the individual decisions of the buyers and sellers that I model.[27] I assume that it follows a random walk with drift, so that there is no predictability in changes in housing market fundamentals. In other words, in a frictionless environment, there should be no predictable returns to owning a house. The particular parametrization I use in estimation is

$$\mu_{it} - \mu_{it-1} = \gamma_0 + \epsilon_{it}. \tag{7}$$

where $\epsilon \sim N(0, \sigma_\epsilon^2)$.

4.4 Structure of Information

I assume that the seller knows all of the parameters that characterize the search problem except for the mean of the valuation distribution, μ_{it}. When sellers receive an offer, they cannot separately identify η from μ. That is, sellers have difficulty distinguishing a high offer due to high average demand from a high offer due to a strong idiosyncratic taste for the house.

Sellers have rational expectations about the μ process in (7). Sellers do not observe the realizations of ϵ, but they observe an unbiased signal z parameterized as

$$z_{it} \sim N(\mu_{it} - \mu_{it-1}, \sigma_z^2). \tag{8}$$

The source of this signal is exogenous to the model, but we can think about it as idiosyncratic information about real-time market conditions that realtors can collect as professional observers of the market.

---

[26]See for example Glower et al. [1998], Merlo and Ortalo-Magne [2004], Chen and Rosenthal [1996]. 15 percent of sales occur at the list price in my data.

[27]In other words, sellers are price takers in this model.

To summarize, there are three sources of information that sellers receive during the selling horizon. Sellers observe whether or not a buyer inspects. This reveals whether or not a noisy signal of a buyer's valuation exceeds a known threshold. Secondly, sellers observe the buyer's valuation if the buyer inspects. Thus, inspections are more informative to the seller than non-inspections. Since the choice of list price affects whether or not a buyer inspects, the list price has an endogenous effect on the flow of information. I am not aware of other models where the optimal list price will depend on the amount of information that the buyer response to the price is likely to provide. Finally, each period sellers observe the exogenous signal about changes in the valuation process.

Appendix A.5 shows how sellers update their beliefs with this information using Bayes' rule. The final piece of the information environment is the seller's initial prior, which is assumed to be:

$$\text{initial prior: } \mu_{it_0} \sim N(\hat{\mu}_{it_0}, \hat{\sigma}^2). \tag{9}$$

The mean of the prior is given by

$$\hat{\mu}_{it_0} = \theta_1 \Big(\frac{\sum_{j=1}^{4} \mu_{t_0-j}}{4} - \mu_{t_0} \Big) + \theta_2 \Big(\frac{\sum_{j=5}^{8} \mu_{t_0-j}}{4} - \frac{\sum_{j=1}^{4} \mu_{t_0-j}}{4} \Big) +$$
$$\theta_3 \Big(\frac{\sum_{j=9}^{12} \mu_{t_0-j}}{4} - \frac{\sum_{j=5}^{8} \mu_{t_0-j}}{4} \Big) + \theta_4 \Big(\frac{\sum_{j=13}^{16} \mu_{t_0-j}}{4} - \frac{\sum_{j=9}^{12} \mu_{t_0-j}}{4} \Big) + \nu_{it_0} \tag{10}$$

where

$$\nu_{it_0} \sim N(\mu_{it_0}, \hat{\sigma}^2). \tag{11}$$

The parameters θ allow the initial beliefs to be sensitive to market conditions from the previous 4 months, as the evidence in Section 3 suggests. If $\theta_j = 0$ for $j = 1, ..., 4$, then the average seller will have unbiased initial beliefs, although there will still be heterogeneity due to ν. Although I do not explicitly model how this initial prior is generated, I show in Appendix A.7 that if a similar Bayesian learning framework applies prior to the beginning of the

18

selling horizon, then initial priors will depend on lagged information.

4.5 Seller's Optimization Problem

The timing of the model is summarized in Figure 3. Each period begins with the realization of z. The seller updates his beliefs, and then chooses an optimal list price. The list price is set to balance the tradeoffs that emerge from Theorem 1. Once the list price is advertised, the buyer decides whether to inspect, the seller updates the reservation price with the information from buyer behavior, and then the seller chooses to either sell the house (if an offer is made) and receive a terminal utility equal to the log sales price or to move onto the next period with the house for sale. Each period that the home does not sell, the seller incurs a cost c, which reflects discounting and the costs of keeping the home presentable to show to prospective buyers. I impose a finite selling horizon of 80 weeks.

The following Bellman's equation, which characterizes selling behavior at the third hash mark on the timeline in Figure 3, summarizes the seller's optimization problem:

$$
\begin{aligned}
V_t(\Omega_t) = \max_{p_t^L} \int_z \Bigg(& (c + V_{t+1}(\Omega_{t+1}|\hat{v}_t < T^* + p^L, z_{t+1}))Pr(\hat{v}_t < T^* + p^L) \\
& + \Big[\int_{-\infty}^{p^L} \max\{v_t, c + V_{t+1}(\Omega_{t+1}|v_t, z_{t+1})\} \\
& + \int_{p^L}^{\infty} p_t^L \Big] Pr(\hat{v}_t > T^* + p^L)g(v_t|\hat{v}_t > T^* + p^L)dv_t \Bigg) f(z_{t+1})dz_{t+1}
\end{aligned}
\tag{12}
$$

where Ω_t denotes the state variables. The normality assumptions imposed throughout, in addition to an approximation method borrowed from a statistics paper by Berk et al. [2007] on Bayesian learning with censored demand, imply that Ω_t is comprised of a single mean and variance. The self-conjugacy of the normal distribution is critical in avoiding the curse of dimensionality

that can make dynamic models infeasible to estimate.[28] This is discussed in more detail in Appendix A.5. Expectations are with respect to whether $\hat{v}$ will exceed $T^* + p^L$, the realization of v, which is correlated with $\hat{v}$, and the realization of the signal about demand changes, z. The top line of equation (12) reflects the case when a buyer does not inspect. In this case, the seller updates his beliefs to Ω_{t+1} and moves onto the next period. If $\hat{v}_t > T^* + p^L$, the seller receives an offer, v. If the offer is above p^L, the payoff is p^L. If the offer is below p^L, the seller Bayesian updates and then decides whether to accept, or reject and move onto the next period. p^L is chosen by the seller to maximize this expected utility. In the Appendix A.5, I show how Ω_t transitions to Ω_{t+1} given the realizations of v, z, and $\hat{v}$.

5 Estimation and Identification

Table 5 summarizes the notation of all of the model parameters. I estimate the parameters using simulated method of moments. That is, I minimize a weighted average distance between sample moments and simulated moments with respect to the model's parameters. The weights are the inverses of the estimated variance of the moments. The target moments are listed in Table 6. I calculate the empirical moments using the subset of listings that sell (which introduces potential sample selection issues that I discuss in Section 8). I describe the dynamic programming techniques used to estimate the model in Appendix A.6.

The parameters that are not estimated are the time invariant holding cost ($c = 0$) and the buyer's inspection cost ($\kappa = -.005$ or .5 percent of the list price). The conclusions that follow are related to the parameters that dictate the flow of information, and so I find that my results are not sensitive to the choices for these parameters. I calibrate the mean ($\gamma_0 = -.0033$ or .33 percent) of the μ process using the average monthly change in the Case-Shiller index

---

[28]The same normality assumptions are made in most empirical learning models. See, for example, Crawford and Shum [2005] and Ackerberg [2003]. I am not aware of any empirical learning models that allow for the type of censored demand that arises when the buyer's valuation exceeds the list price.

for San Francisco and Los Angeles during my sample period. I set $\theta_1, \theta_2, \theta_3, \theta_4$ equal to the coefficients on lagged depreciation estimated in Column 4 of Table 3. Since $\partial p^L / \partial \hat{\mu} = 1$ in the model (proof not reported), this implies that the initial list price will display the same level of sensitivity to lagged market conditions as found in the data.

The variance of the offer distribution, σ_η^2, is identified by the distribution of sales prices relative to the list price. The level of initial uncertainty, $\hat{\sigma}^2$, is identified by the size of list price changes, especially in the first couple weeks after listing before depreciation in μ increases the variance in list price changes. Both variances also have different predictions for TOM. More variance in the offer distribution increases TOM because the higher incidence of very good offers increases the value of searching. More uncertainty distorts the choice of list price and reservation price, which decreases the returns to staying on the market. The average premium of the list price relative to the sales price helps to identify $\sigma_{\hat{v}}^2$, as does the propensity for prices to occur at the list price. For example, if buyers have a lot of information about their valuation prior to inspection (low $\sigma_{\hat{v}}^2$), sellers need to set low list prices to induce inspections. σ_z^2 is identified by the correlation between list price changes and changes in $\hat{p}$. A high correlation suggests that σ_z^2 is low because sellers can quickly and fully internalize changes in market conditions into their list price decisions.[29] The variance of the μ process, σ_ϵ^2, is identified by the variance of changes in average prices over time. In the data, I calculate this moment by taking the standard deviation of monthly price changes in the Case-Shiller index for San Francisco and Los Angeles during my sample period.

---

[29]If I used weekly list price changes, the model would generate a high σ_z^2 because in most cases list prices do not change week by week. Some of this may be due to high σ_z^2, but some of it may be due to menu costs, which I do not model. As a result, I use list price changes in the initial week of listing relative to the final week of listing. Menu costs should have less of an effect on average changes over longer horizons.

6 Estimation Results and Model Fit

The learning model matches the data well (see Table 6 for the simulated moments at the estimated parameter vector). Even when agents have rational expectations about the severe market decline during my sample period, the model matches the lengthy TOM observed in the data. At the estimated parameters, more uncertainty raises the list price because sellers want to test demand before dropping the price, which will attract more buyers but will also transfer more of the bargaining power to the buyer. Since uncertainty decreases over the selling horizon, the model generates declining list prices, and would do so even if market conditions (i.e. the μ process) were constant. However, as in the data, a minority of list price changes are increases (6 percent in the data versus 4 percent in the model). The fraction of list price changes that are increases is not a moment that I directly target in estimation. In the model, list price increases primarily occur from sellers with low draws of ν in equation (10).

Interestingly, the model fits the fact that list to sales price premiums increase with TOM. This is true even though list prices are increasing in the level of uncertainty, and the level of uncertainty decreases over the selling horizon. The reason for the increasing wedge between list and sales prices is selection: sellers with low reservation prices tend to sell quickly and post lower list prices. The model overpredicts the average list price change at delisting relative to listing. This is partly because I do not model menu costs. In Section 8 I discuss the robustness of my results to this and other simplifications.

Table 5 reports the parameter estimates and their standard errors.[30] The results suggest that sellers typically accept offers that are 12 percent above the mean of the valuation distribution, which is the 92nd percentile. Given that the average sales price is \$ 628,000, this implies that the mean of the offer distribution is about \$ 561,000 (628000/1.12) for the typical house. Thus, the standard deviation of the seller's prior for the typical house is about \$

---

[30]The variance-covariance matrix of the parameter estimates is given by $(G'W^{-1}G)^{-1}$, where G is the matrix of derivatives of the moments with respect to the parameters and W is the variance-covariance matrix of the moments. Off-diagonal elements are ignored.

38,000 and the standard deviation of the offer distribution is about \$47,000. I calculate that Bayesian learning reduces the standard deviation (variance) of the seller's initial prior by 37 percent (60 percent) over the course of the selling horizon.

I also relate the predictions of the model to the reduced form results presented in Section 3. As mentioned above, the specification of the initial prior in equation (10) ensures that the model replicates the correlation between lagged price depreciation and list prices observed in the data. The model also predicts that lagged price depreciation is positively correlated with TOM.[31] The stronger is perceived demand, the higher is the reservation price, which increases TOM, all else equal. The model predicts a small effect of lagged price depreciation on sales price.[32] Just as in the data, for some parts of the distribution of initial bias, more bias leads to lower sales prices. The model predicts that two alternative explanations for the high list prices found in Section 3 – high unobserved home quality (a higher μ) and low unobserved motivation to sell (a higher $|c|$) – lead to unambiguously *higher* sales prices as well higher list prices (proof not reported). Thus, the model illustrates how these explanations are inconsistent with the evidence from Section 3.

7   Simulations of Market Dynamics

7.1   Price Dynamics

It has been well documented that house price appreciation rates are persistent in the short-run. An important question is whether this predictability can be supported in an equilibrium where market participants are behaving optimally. My model of rational behavior conditional on an exogenous level of information suggests that it can. I show this by simulating average weekly sales prices using the model for $T = 48000$ and $N = 20000$ new listings each week. The parameters of the model are set at their estimated values. Following

---

[31]A 1 percent increase in lagged price depreciation increases TOM by 5.1 percent.

[32]A 1 percent increase in lagged price depreciation increases sales price by .1 percent.

the literature, I run the following regression on the simulated price series

$$p_t - p_{t-52} = \rho_0 + \rho_1(p_{t-52} - p_{t-104}) + \nu_t. \tag{13}$$

where p_t is the log average price over all simulated sales in week t.[33]

Table 7 shows the results. The level of sales price persistence generated by the model is .124. The information frictions are completely responsible for the persistence. Column 1 shows that when the average seller has unbiased beliefs at the time of initial listing and when $\sigma_z = 0$, $\rho_1 = 0$.[34] The third and fourth columns show results when I aggregate weekly prices to the quarterly level. In this case, the dependent variable is $p_t - p_{t-4}$ where t is a quarter and p_t is the simple average of all sales in quarter t. I present these results because in practice sales do not occur frequently enough to compute price indexes at the weekly level. Case and Shiller [1989] and Cutler et al. [1991], for example, run their regressions at the quarterly level. The aggregation alone introduces persistence, and the AR(1) coefficient rises to .152.

7.1.1 Discussion, Robustness, and Further Results

At the parameter estimates, the model generates persistence that is over half the level typically found in the data. The intuition for the result is as follows. Sellers do not fully adjust their beliefs in time t to a shock to μ in time t, on average. The optimal Bayesian weighting places some weight on the signal about the shock and some weight on the seller's prior expectation. Then, for example, when there is a positive shock, the average reservation price in the population rises, but is too low relative to the perfect information case. As time progresses, however, learning from buyer behavior provides more information about the shock, and reservation prices eventually fully adjust. The same intuition holds for a negative demand shock. Thus, serial correlation in

---

[33]By simulating a large number of sales each week, I avoid measurement errors that affect the estimation of these regressions in practice. See Case and Shiller [1989].

[34]Recall that the fundamental determinant of house values in the model, μ_t, follows a random walk (see equation 7). So by construction, there is zero persistence in changes in the fundamentals.

price changes arises because 1) persistent demand shocks are not immediately capitalized into reservation prices and 2) there exists a mechanism through which additional information about these shocks arrives with a lag.

Over shorter frequencies, the persistence is even higher, as shown in the right-most columns of Table 7. We can see this through the equation for the OLS estimate of ρ_1:

$$\hat{\rho}_1 = \frac{cov(p_t - p_{t-L}, p_{t+L} - p_t)}{var(p_t - p_{t-L})}. \tag{14}$$

As the lag length, L, gets smaller, the numerator stays approximately the same and the denominator gets smaller because there are fewer shocks between time t and $t - L$. By the same logic, the persistence dies out as L increases. Thus, the short-run persistence generated here does not preclude long-run mean reversion in price changes, which is an additional stylized fact about house price dynamics. The model can generate long-run mean reversion with the addition of a mean reverting shock to the μ process.[35]

The persistence generated by the model could be arbitraged away if some traders have superior access to information about the current period fundamentals. However, given that realtors already provide the typical seller with information based on their professional insights and access to data in the MLS, this seems like a difficult arbitrage. In addition, the difficulty in taking short positions and the large transaction costs involved in trading homes complicates any potential trading strategy (Meese and Wallace [1994]).

σ_z^2 potentially plays a large role in determining the amount of persistence because it affects how much of a demand shock is immediately capitalized into reservation prices. When σ_z^2 is high, there is a lot of scope for persistence because most of the information about the demand shock will arrive with a lag. To test the sensitivity of the results in Table 7 to the point estimate of σ_z^2, I re-simulate the model at the upper and lower limits of the 95 percent confidence interval for the estimate of σ_z. The annual persistence (weekly

---

[35]Glaeser and Gyourko [2006] cannot generate short-run momentum, but generate long-run mean reversion with a mean reverting shock to local productivity and a slow construction response.

prices) always lies between 0.12 and 0.13. Given that $\sigma_z = .018$ (1.8 percent) at the lower end of the confidence the interval, the model does not require much signal noise at all to generate a significant amount of persistence.

I also test the sensitivity of the results to the assumption that the mean of the valuation distribution, μ_t, changes each period (i.e. each week). I simulate a version of the model where μ_t only changes every four periods; I multiply σ_ϵ^2 by four so that the variance of $\mu_t - \mu_{t-4}$ is the same as in the baseline model. In this version of the model, the annual persistence is slightly higher at .131 for weekly prices and .162 for quarterly prices.

7.2 Sales Volume and TOM Dynamics

The existing literature has identified frictions related to search and credit constraints as explanations for the positive price-volume correlation in the housing market. In this section, I show that an information friction is an additional contributor to this correlation.

Table 8 shows the results when I regress $log(Volume)$ and $log(TOM)$ on quarterly price changes using data simulated from the model at the estimated parameters. Column 1 shows that a 1 percent increase in quarterly prices leads to a 6.7 percent increase in sales volume. The estimate from running the same specification on the actual data for Los Angeles and San Francisco is 5 percent, which is close to the estimate reported in Stein [1995] who uses data from the entire US housing market.[36] Column 2 shows that absent the information friction, the model does not predict a relationship between price changes and volume.

When the dependent variable is $log(TOM)$, the model generates a $\beta_1 = -7.6$: a 1 percent increase in quarterly prices leads to a 7.6 percent decline in TOM.[37] To compare this prediction to the data, I collect a TOM time-

---

[36]To estimate this regression on the data, I first run a repeat sales regression with quarter dummies on the transaction data from 1988-2009. The change in the quarter dummies (adjusted for inflation) is then the explanatory variable in an OLS regression where volume is the dependent variable. I run the regressions separately for Los Angeles and San Francisco, but the estimates are similar.

[37]In the version of the model where shocks arrive only once a month, $\beta_1 = -6.3$ for TOM

series from the Annual Historical Data Summary produced by the California Association of Realtors. The TOM data reflects averages for the entire state of California, while the quality adjusted price data I have is from Los Angeles and San Francisco so the comparison is quite rough. The estimate of β_1 is -5.6 percent using LA prices and -4.1 percent using SF prices, suggesting that the model is generating predictions that are of the same order of magnitude as the empirical price-TOM relationship.

The results suggest that information frictions are important for explaining variation in transaction rates over the housing cycle. The intuition for these results is that positive shocks to home values are accompanied by reservation prices that are too low relative to the perfect information case. Lower reservation prices relative to the fundamentals leads to more and quicker transactions.

8   Discussion of Model Assumptions

In this section, I discuss the robustness of the results to a few stylized features of the model.

8.1   No Buyer Learning

In the model, changes in the offer, or willingness to pay, distribution are exogenous. A model that endogenizes buyer willingness to pay from the fundamental demand and supply conditions in the economy could also include a dynamic learning process, as the thinness and volatility of the market probably make it difficult for buyers to observe market conditions as well. We do not model such a dynamic process because it would be difficult to identify without data on buyer behavior and it would significantly complicate the seller's problem. However, I suspect that including a buyer learning process would increase the level of price persistence. In the current setup, reservation prices adjust to market shocks with a lag, but offers adjust immediately. If offers adjust with a lag as well, then the adjustment of prices to market shocks would be

_____

and $\beta_1 = 7.3$ for volume.

even slower. The correlations between volume and TOM with price changes, however, may be attenuated because sluggish reservation prices are met with sluggish demand.

8.2 Abstraction from Features of the Micro Data

The model presented above is the simplest version of an empirical model needed to highlight the effects of information frictions on market dynamics. As a result, the current version of the model does not explain some features of the micro data such as withdrawals, sales prices above list prices, and sticky list prices.[38] In a working paper version of the paper Anenberg [2011b], I experimented with more detailed versions of the model to address each of these features. The main conclusions are unchanged. In this section, I summarize these adjustments to the model.

To accommodate withdrawals, I allow sellers to withdraw at any time and receive an exogenous and heterogenous termination utility, v^w. The parameter estimates from that model suggest that there is a group of motivated sellers, with very low v^w as modeled above, and a group of unmotivated sellers with high v^w. Hardly any of the unmotivated sellers end up selling given the decline in the market. Thus, the predictions of the model with respect to sales price and volume dynamics are similar.[39] This version of the model requires positive holding costs, c, to explain why sellers do not stay on the market indefinitely. The estimated holding costs are small.

To accommodate sales prices above the list price, I assume that when $v > p^L$, there is some exogenous probability that the price gets driven up above p^L. This addition to the model does not affect the main parameter estimates or conclusions.

---

[38]The models of Carrillo [2010] and Horowitz [1992] do not accommodate these features of the data either.

[39]I also note here that uncertainty is able to rationalize the high withdrawal rate observed in the data. The estimated amount of uncertainty is high enough and the holding costs of keeping a home on the market are low enough that unmotivated sellers find it optimal to test the market even though they fully anticipate withdrawing if they learn that demand for their house is insufficient.

The current model predicts that sellers should adjust their list price, oftentimes by an ϵ amount, each period. This is one reason that the model overpredicts the average list price change at delisting relative to listing. In practice, list prices are sticky, presumably due to menu costs. In the working paper version of this paper, I show that very small menu costs can rationalize sticky list prices. Thus, I do not expect the addition of a menu cost, which would significantly increase the computational burden, to affect the conclusions.

8.3 One Offer per Period

In the model I assume that the expected amount of information that the seller receives does not vary much over the selling horizon. In practice, the arrival of buyers may be especially strong in the first several periods while the listing is fresh. Thus, learning in the initial weeks may be higher than the model allows for. Modeling multiple offers significantly complicates estimation, and it is not clear how the arrival rate would be identified without information on buyer behavior. Instead, I test the robustness of my results to stronger learning in the initial weeks by allowing sellers to observe an additional draw from the offer distribution, v_{it}, during each period in the first month after listing. The annual price persistence declines from .124 to .092. The effect of price change on volume decreases from 6.7 percent to 3.3 percent, and the effect of price change on TOM increases from -7.6 percent to -5.8 percent.

9 Conclusion

This paper shows that information frictions play an important role in the workings of the housing market. Using a novel and robust identification strategy for expectation bias, I find direct evidence that imperfect information affects the micro decisions of individual sellers. Then, I show that a search model with uncertainty and Bayesian learning fits the key features of the micro data remarkably well, suggesting that information frictions are important in ex-

plaining the distribution of marketing times, the role of the list price, and the microstructure of the market more generally. I also use the model to highlight how micro-level decision making in the presence of imperfect information affects aggregate market dynamics. Most notably, I find a significant microfounded momentum effect in short-run aggregate price appreciation rates.

The analysis here raises several interesting directions for future research. Given the sample period that I have access to, I have argued that my results likely characterize market dynamics during cold housing markets. Since the basic mechanisms generating the main results do not depend on the market being in decline, I suspect that my model, estimated using a sample with rising prices, would be successful in explaining momentum and correlations between price, volume, and marketing time in rising markets as well. However, the magnitude of the results may differ as the pace of learning may change over the housing cycle. This paper does not discuss the welfare implications of uncertainty. In a working paper version, I use a similar model to show that the value of information to sellers is large, which helps to explain the demand for realtors that typically charge 3 percent of the sales price. On the modeling side, extending micro models of the home selling problem to the multi-agent setting, so that the pricing and selling outcomes of neighboring listings has an endogenous effect on the flow of information, is an interesting direction for future research.

References

Daniel A. Ackerberg. Advertising, learning, and consumer choice in experience good markets: An empirical examination. *International Economic Review*, 44(3):1007–1040, 2003.

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

Elliot Anenberg. Uncertainty, learning, and the value of information in the residential real estate market. *mimeo Duke University*, 2011b.

Emre Berk, lk Grler, and Richard A. Levine. Bayesian demand updating in the lost sales newsvendor problem: A two-moment approximation. *European Journal of Operational Research*, 182(1):256 – 281, 2007.

James A. Berkovec and John L. Goodman. Turnover as a measure of demand for existing homes. *Real Estate Economics*, 24(4):421–440, 1996.

Craig Burnside, Martin Eichenbaum, and Sergio Rebelo. Understanding booms and busts in housing markets. Working Paper 16734, National Bureau of Economic Research, January 2011.

John Campbell, Stefano Giglio, and Parag Pathak. Forced sales and house prices. *American Economic Review*, 2011.

Paul Carrillo. An empirical stationary equilibrium search model of the housing market. *forthcoming International Economic Review*, 2010.

Karl E. Case and Robert J. Shiller. The efficiency of the market for single-family homes. *The American Economic Review*, 79(1):125–137, 1989.

Yongmin Chen and Robert W. Rosenthal. Asking prices as commitment devices. *International Economic Review*, 37(1):129–155, 1996.

Man Cho. House price dynamics: A survey of theoretical and empirical issues. *Journal of Housing Research*, 7(2):145–172, 1996.

Gregory S. Crawford and Matthew Shum. Uncertainty and learning in pharmaceutical demand. *Econometrica*, 73(4):1137–1173, 2005.

David M. Cutler, James M. Poterba, and Lawrence H. Summers. Speculative dynamics. *The Review of Economic Studies*, 58(3):pp. 529–546, 1991.

Erik de Wit and Bas van der Klaauw. Asymmetric information and list price reductions in the housing market. *CEPR Discussion Paper*, 2010.

Tlin Erdem and Michael P. Keane. Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing Science*, 15(1):pp. 1–20, 1996.

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

David Genesove and Christopher J. Mayer. Equity and time to sale in the real estate market. *The American Economic Review*, 87(3):255–269, 1997.

Edward L. Glaeser and Joseph Gyourko. Housing dynamics. Working Paper 12787, National Bureau of Economic Research, December 2006.

M Glower, D.R. Haurin, and P.H. Hendershott. Selling time and selling price: the influence of seller motivation. *Real Estate Economics*, 26(4):719–740, 1998.

John L. Goodman and John B. Ittner. The accuracy of home owners' estimates of house value. *Journal of Housing Economics*, 2(4):339 – 357, 1992.

Sanford J. Grossman and Joseph E. Stiglitz. Information and competitive price systems. *The American Economic Review*, 66(2):pp. 246–253, 1976.

Allen Head, Huw Lloyd-Ellis, and Hongfei Sun. Search, liquidity and the dynamics of house prices and construction. *mimeo Queens University*, 2011.

Gunter J. Hitsch. An Empirical Model of Optimal Dynamic Product Launch and Exit Under Demand Uncertainty. *Marketing Science*, 25(1):25–50, 2006.

Joel L. Horowitz. The role of the list price in housing markets: Theory and an econometric model. *Journal of Applied Econometrics*, 7(2):115–129, 1992.

Michael P. Keane and Kenneth I. Wolpin. The solution and estimation of discrete choice dynamic programming models by simulation and interpolation: Monte carlo evidence. *The Review of Economics and Statistics*, 76(4): 648–672, 1994.

Katherine A. Kiel and Jeffrey E. Zabel. The accuracy of owner-provided house values: The 19781991 american housing survey. *Real Estate Economics*, 27 (2):263–298, 1999.

John R. Knight. Listing price, time on market, and ultimate selling price: Causes and effects of listing price changes. *Real Estate Economics*, 30(2): 213–237, 1996.

John Krainer. A theory of liquidity in residential real estate markets. *Journal of Urban Economics*, 49(1):32 – 53, 2001.

Edward P. Lazear. Retail pricing and clearance sales. *The American Economic Review*, 76(1):14–32, 1986.

Steven D. Levitt and Chad Syverson. Market distortions when agents are better informed: The value of information in real estate transactions. *Review of Economics and Statistics*, 90(4):599–611, 2008.

Richard Meese and Nancy Wallace. Testing the present value relation for housing prices: Should i leave my house in san francisco? *Journal of Urban Economics*, 35(3):245 – 266, 1994.

Antonio Merlo and Francois Ortalo-Magne. Bargaining over residential real estate: evidence from england. *Journal of Urban Economics*, 56(2):192–216, September 2004.

Dale T. Mortensen and Christopher A. Pissarides. Job creation and job destruction in the theory of unemployment. *The Review of Economic Studies*, 61(3):397–415, 1994.

Sridhar Narayanan, Pradeep Chintagunta, and Eugenio Miravete. The role of self selection, usage uncertainty and learning in the demand for local telephone service. *Quantitative Marketing and Economics*, 5:1–34, 2007. ISSN 1570-7156.

Rachel Ngai and Silvana Tenreyro. Hot and cold seasons in the housing market. *mimeo London School of Economics*, 2010.

Robert Novy-Marx. Hot and cold markets. *Real Estate Economics*, 37(1): 1–22, 2009.

Francois Ortalo-Magne and Sven Rady. Housing market dynamics: On the contribution of income shocks and credit constraints. *Review of Economic Studies*, 73(2):459–485, 2006.

Christopher Pissarides. Equilibrium unemployment theory. *MIT Press*, 2000.

Stephen W. Salant. For sale by owner: When to use a broker and how to price the house. *The Journal of Real Estate Finance and Economics*, 4:157–173, 1991. ISSN 0895-5638.

Robert J. Shiller. Arithmetic repeat sales price estimators. *Journal of Housing Economics*, 1(1):110–126, 1991.

Jeremy C. Stein. Prices and trading volume in the housing market: A model with down-payment effects. *The Quarterly Journal of Economics*, 110(2): 379–406, 1995.

George J. Stigler. The economics of information. *Journal of Political Economy*, 69(3):pp. 213–225, 1961.

Curtis R Taylor. Time-on-the-market as a sign of quality. *Review of Economic Studies*, 66(3):555–78, July 1999.

A Appendix: For Online Publication

A.1 Data Appendix

I first describe how I combine the listing data from Altos Research with the transaction data from Dataquick. I begin by cleaning up the address variables in the listing data. The address variables in the transaction data are clean and standardized because they come from county assessor files.

The listing data contains separate variables for the street address, city, and zip code. I ignore the city variable since street address and zip code uniquely characterize houses. The zip code variable does not need any cleaning. In a large majority of cases, the address variable contains the house number, the street name, the street suffix, and the condo unit number (if applicable) in that order. We alter the street suffixes to make them consistent with the street suffixes in the transaction data (e.g. change "road" to "rd", "avenue" to "ave", etc). In some cases, the same house is listed under 2 slightly different addresses (e.g. "123 Main" and "123 Main St") with the same MLSIDs. We combine listings where the address is different, but the city and zip are the same, the MLSids are the same, the difference in dates between the two listings is less than 3 weeks, and at least one of the follow conditions applies:

1. The listings have the same year built and the ratio of the list prices is greater than 0.9 and less than 1.1.

2. The listings have the same square feet and the ratio of the list prices is greater than 0.9 and less than 1.1.

3. The listings have the same lotsize and the ratio of the list prices is greater than 0.9 and less than 1.1.

4. The first five characters of the address are the same.

We merge the listing data and the transaction data together using the address. If we get a match, we keep the match and treat it as a sale if the difference in dates between the transaction data (the closing date) and the date the listing no longer appears in the MLS data (the agreement date) is greater than zero and less than 365 days. If the match does not satisfy this timing criteria, we keep the most recent transaction to record the previous selling price. Before we do the merge, we flag properties that sold more than once during a 1.5 year span during our sample period. To avoid confusion during the merge that can arise from multiple sales occurring close together,

we drop any listings that merge to one of these flagged properties (< 1 percent of listings).

I drop listings where the ratio of the minimum list price to the maximum list price is less than the first percentile. I drop listings where the TOM is greater than the 99th percentile. I drop listings where the list to predicted price ratio is less than the 1st or greater than the 99th percentile. I drop listings where the predicted price is less than the 1st or greater than the 99th percentile. I drop listings where the sales to predicted price ratio is less than the 1st or greater than the 99th percentile.

A.2 Detail on Calculation of Predicted Prices

$\hat{p}_{it}$ is the log expected sales price for house i in month t. This expected price is simply equal to the previous log price paid for the house plus some neighborhood (zip code in this analysis) level of appreciation or depreciation. To calculate the level of appreciation, I follow Shiller [1991], who estimates the following model

$$p^*_{ijt} = v_i + \delta_{jt} + \epsilon_{ijt} \tag{15}$$

where v is a house fixed effect, δ_{jt} is a neighborhood specific time dummy, and ϵ_{ijt} is an error term. We can estimate the coefficients on the neighborhood-specific time dummies, which form the basis of a quality adjusted neighborhood index of price appreciation, through first-differencing and OLS using a sample of repeat-sales. In practice, when I estimate the time-dummy coefficients for a particular zip code j, I use the entire sample of repeat sales from 1988-2009, except I weight the observations for zip code i using

$$W_{i(j)} = \left[\frac{1}{h} \phi \left(\frac{dist_{ij}}{h * std(dist_{ij})} \right) \right]^{1/2} \tag{16}$$

where ϕ is the standard normal pdf, $dist$ is the distance between the centroid of the zip codes i and j, and h is a bandwidth.[40] I use this weighting scheme because sometimes the number of sales in a particular zip code in a particular month is not large.

---

[40]I set the bandwidth equal to 0.25. This choice of bandwidth implies that the weights decline about 40 percent as we move 10 miles away from the centroid of a neighborhood. The main results of the paper are robust to alternative choices of bandwidth.

A.3 Further Results on Expectation Bias

Appendix Table 1 presents additional results that are consistent with the conclusions established Section 3. I test whether sellers in neighborhoods where there have been a lot of recent sales are better able to detect recent price trends. In the context of the model, we could think of these sellers as receiving signals, z, with a tighter variance because there is more information about recent price trends. We run the following variation of specification (2)

$$y_{i,j,t} = \gamma_t + \beta_1 \Delta_{jt}^4 + \beta_2 \Delta_{jt}^4 * sales_{jt} + \beta_3 X_{i,j,t} + \epsilon_{i,j,t} \qquad (17)$$

where $sales_{jt}$ is the total number of sales in neighborhood j in the previous 4 months. Column 1 of Appendix Table 1 shows that a 1 standard deviation increase in $sales_{jt}$ lowers the effect of Δ_{jt}^4 by 0.12, or 22 percent. Columns (2)-(4) show the results when we substitute TOM, $I[Withdraw]$, and the sales price premium as the dependent variable.[41] More thickness decreases the positive (negative) effects of lagged depreciation on TOM (sales price). The effects on the propensity to withdraw are economically insignificant.

Column (5) shows that the effects of Δ_{jt}^4 on list prices diminish as we move later in the sample period. The time series of prices in Figure 2 provides a likely explanation. As the housing decline deepened and sellers learned that prices were depreciating rapidly, they did a better job of adjusting the prices of recent comparable sales for the downward trend.

A.4 Proof of Theorem 1

Buyers will inspect house i when the expected surplus from visiting exceeds the expected cost, i.e. when

$$\int_{p^L}^{\infty} (v - p^L) \frac{1}{\sigma_{\hat{v}}} \phi(\frac{v - \hat{v}}{\sigma_{\hat{v}}}) dv \geq -\kappa \qquad (18)$$

where ϕ is the standard normal distribution. The lower limit of integration is p^L because the buyer receives no surplus when her valuation is below the list price.

To show that the optimal buyer behavior takes the reservation value form, it is sufficient to show that the term in the integral in equation (18) is increasing in $\hat{v}$. Using properties of the truncated normal distribution, we rewrite the

---

[41]The adjustments to the regressors in (17) depending on the dependent variable follow the discussion/specifications in Section 3. In these specifications where we have no controls for neighborhood, we also include an additional control for the level of $sales_{jt}$.

integral as

$$(\hat{v} - p^L)(1 - \Phi(\frac{p^L - \hat{v}}{\sigma_{\hat{v}}})) + \sigma_{\hat{v}}\phi(\frac{p^L - \hat{v}}{\sigma_{\hat{v}}}) \tag{19}$$

Taking the derivative of this expression with respect to $\hat{v}$ gives

$$(1 - \Phi(\frac{p^L - \hat{v}}{\sigma_{\hat{v}}})) + (\hat{v} - p^L)\phi(\frac{p^L - \hat{v}}{\sigma_{\hat{v}}}) + (p^L - \hat{v})\phi(\frac{p^L - \hat{v}}{\sigma_{\hat{v}}}) = 1 - \Phi(\frac{p^L - \hat{v}}{\sigma_{\hat{v}}}) > 0. \tag{20}$$

To show the particular form of $\bar{v}$, using properties of the truncated normal distribution, we rewrite equation (18) for $\hat{v} = \bar{v}$ as

$$(\bar{v} - p^L)(1 - \Phi(\frac{p^L - \bar{v}}{\sigma_{\hat{v}}})) + \sigma_{\hat{v}}\phi(\frac{p^L - \bar{v}}{\sigma_{\hat{v}}}) + \kappa = 0. \tag{21}$$

Let z be the left hand size of (19). It is clear from (19) that $\frac{\partial z}{\partial p^L} = -\frac{\partial z}{\partial \bar{v}}$. Then, using the implicit function theorem, $\frac{\partial \bar{v}}{\partial p^L} = 1$. Thus, the remaining determinant of $\bar{v}$ will be an additively separable term, T^*. To get an expression for T^*, plugging the solution for $\bar{v}$ into (19), we get

$$(T^*)(1 - \Phi(\frac{-T^*}{\sigma_{\hat{v}}})) + \sigma_{\hat{v}}\phi(\frac{-T^*}{\sigma_{\hat{v}}}) + \kappa = 0. \tag{22}$$

Given values for $(\sigma_{\hat{v}}, \kappa)$, we can solve for T^* using fixed-point iteration.

A.5 Model Details on Learning

In this section, we detail how sellers update their beliefs in response to information that arrives during the selling horizon.

Define the following means and variances of seller beliefs over μ_{it}:

$\hat{\mu}_{it}^{pre}, \hat{\sigma}_{it}^{pre}$: Beliefs after observing z but before observing buyer behavior in week t.
$\hat{\mu}_{it}, \hat{\sigma}_{it}^2$: Beliefs after observing buyer behavior in week t.

Suppose that $\hat{\mu}_{it}$ and $\hat{\sigma}_{it}^2$ are the mean and variance of a normal distribution at any time t. Given the assumptions made in the model, I show below that this will be the case. Then, Bayes' rule implies that the posterior after processing z is also normal where

$$\hat{\mu}_{it}^{pre} = \hat{\mu}_{it-1} + \frac{\sigma_z^2 \hat{\mu}_p + \hat{\sigma}_p^2 z_{it}}{\sigma_z^2 + \hat{\sigma}_p^2}$$

$$\hat{\sigma}_{it}^{pre} = \hat{\sigma}_{it-1}^2 + \frac{\sigma_z^2 \hat{\sigma}_p^2}{\sigma_z^2 + \hat{\sigma}_p^2}. \tag{23}$$

The best case scenario for the seller is that $\sigma_z^2 = 0$; in this case, weekly changes to the mean of the valuation distribution do not increase uncertainty.

The source of learning that decreases uncertainty in week t is buyer behavior. If a buyer arrives, recall that the seller observes v_{it}, which is a noisy signal of μ_{it}. The posterior distribution of μ after the seller processes the information in v_{it} remains normal with mean and variance at time t given respectively by:

$$\hat{\mu}_{it} = \frac{\sigma_\eta^2 \hat{\mu}_{it}^{pre} + \hat{\sigma}_{it}^{pre} v_{it}}{\sigma_\eta^2 + \hat{\sigma}_{it}^{pre}}$$

$$\hat{\sigma}_{it}^2 = \frac{\hat{\sigma}_{it}^{pre} \sigma_\eta^2}{\hat{\sigma}_{it}^{pre} + \sigma_\eta^2}. \tag{24}$$

The initial conditions are given in equation (9).

If a buyer does not arrive, the seller observes that $\hat{v}_{it} < T^* + p_{it}^L$ and the density function of the posterior is

$$f(\hat{\mu}_t | \hat{v} < T^* + p^L) = \frac{\Phi\left(\frac{T^* + p^L - \hat{\mu}_t}{\sqrt{\sigma_\eta^2 + \sigma_{\hat{v}}^2}}\right) \phi\left(\frac{\hat{\mu}_t - \hat{\mu}_t^{pre}}{\hat{\sigma}_t^{pre}}\right) \frac{1}{\hat{\sigma}_t^{pre}}}{\Phi\left(\frac{T^* + p^L - \hat{\mu}_t^{pre}}{\sqrt{\hat{\sigma}_t^{pre} + \sigma_\eta^2 + \sigma_{\hat{v}}^2}}\right)}. \tag{25}$$

This is not a normal distribution because of the $\hat{\mu}_t$ term in the normal *cdf* in the numerator. A statistics paper by Berk et al. [2007] shows that a normal distribution with mean and variance equal to the mean and variance of the distribution in equation (25) is a good approximation for the true posterior when demand is censored in exactly this way. I use this approximation method here, noting that simulations show this approximation to work extremely well for my application. Then, when a buyer does not arrive, the posterior distribution after processing that $\hat{v}_{it} < T^* + p_{it}^L$ is normal with mean and variance at given respectively by:

$$\hat{\mu}_{it} = \hat{\mu}_{it}^{pre} - \hat{\sigma}_{it}^{pre} h(T^* + p^L)$$

$$\hat{\sigma}_{it}^2 = \frac{1}{\tau^2}\left((\hat{\mu}_{it}^{pre})^2 \sigma^4 + \tau \hat{\sigma}_{it}^{pre} \sigma^2 + 2\hat{\sigma}_{it}^{pre} \sigma^2 (\hat{\mu}_{it}^{pre})^2 + (\hat{\sigma}_{it}^{pre})^2 \tau + (\hat{\mu}_{it}^{pre})^2 (\hat{\sigma}_{it}^{pre})^2\right)$$
$$+ (2\hat{\mu}_{it}^{pre} \hat{\sigma}_{it}^{pre} \sigma^2 + (\hat{\sigma}_{it}^{pre})^2 (T^* + p^L + \hat{\mu}_{it}^{pre})) - h(T^* + p^L)/\tau - (\hat{\mu}_{it})^2 \quad (26)$$

where $\tau = \hat{\sigma}_{it}^{pre} + \sigma_\eta^2 + \sigma_{\hat{v}}^2$, $\sigma^2 = \sigma_\eta^2 + \sigma_{\hat{v}}^2$, and h is the hazard rate corresponding

to the normal distribution with mean $\hat{\mu}_{it}^{pre}$ and variance τ.

A.6   Dynamic Programming Methods

I assume a finite horizon of 80 weeks for the selling horizon. It is well known that in these types of dynamic programming problems, V from equation (12) needs to be calculated for each point in the state space. I calculate V for a discrete number of points and use linear (in parameters) interpolation to fill in the values for the remainder of the state space.

The integrals in equation (12) have a closed form. No simulation is required. This avoids a source of bias that often arises in practice when the number of simulations required to preserve consistency is not feasible to implement. See Keane and Wolpin [1994] for a more detailed discussion. The closed form arises due to the normal approximation for the pdf, g, described in Berk et al. [2007], properties of the truncated normal distribution, the absence of idiosyncratic choice specific errors from the model, linearity in equations (23) and (24), and linearity in the interpolating function.

The optimal list price, however, does not have a closed form. For each point in the discretized state space, I solve for the optimal list price using a minimization routine. The optimal list price also needs to be calculated when simulating selling outcomes for each seller. I approximate the list price policy function using linear (in parameters) interpolation. This is done using the discrete points used to approximate the value function.

A.7   Bayesian Learning Prior to Listing

In this section, I show that if a similar Bayesian learning framework applies prior to the beginning of the selling horizon, the initial priors will depend on lagged information. To see how, consider a simplified information structure where μ_t follows a random walk with a drift equal to zero (and normally distributed shocks). Furthermore, assume that μ_{t_0-1} is observable, but the seller only gets a signal z about $\mu_{t_0} - \mu_{t_0-1}$. Then, the seller's beliefs about μ_{t_0} will be

$$\mu_{t_0-1} + \theta z \tag{27}$$

where $\theta = \frac{\sigma_\epsilon^2}{\sigma_\epsilon^2 + \sigma_z^2}$ is the optimal Bayesian weight that sellers put on the signal. For a realization of $\mu_{t_0} < 0$, sellers will tend to overstate μ_{t_0} on average. This will lead to high list prices because the optimal list price is monotonic in $\hat{\mu}$ as discussed above. The noisier the signal z, the lower is θ, and the more sellers will overstate μ_{t_0} for low realizations of μ_{t_0}.

Figure 1: Test for Expectation Bias
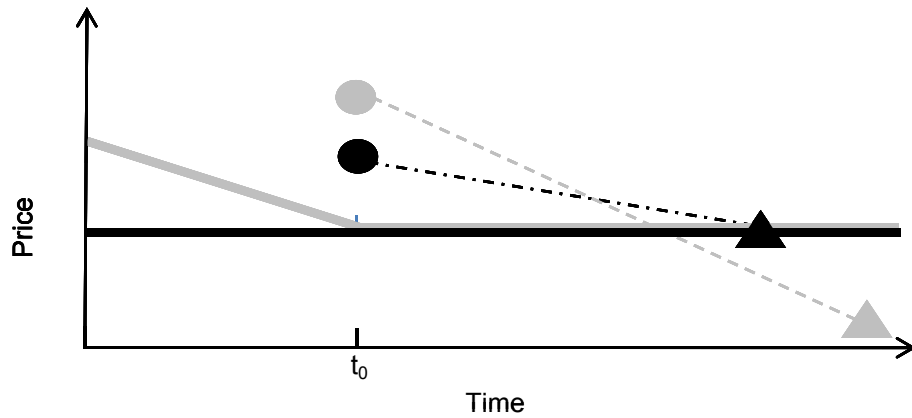
Note: Black and grey denote 2 different neighborhoods. The solid lines denote the average sales prices among all homes in the neighborhood that sell at each point in time.

○ Average list price for homes first listed for sale at month t_0
△ Average sales price for homes first listed for sale at month t_0

Figure 2: Case Shiller Home Price Index 2007-2009
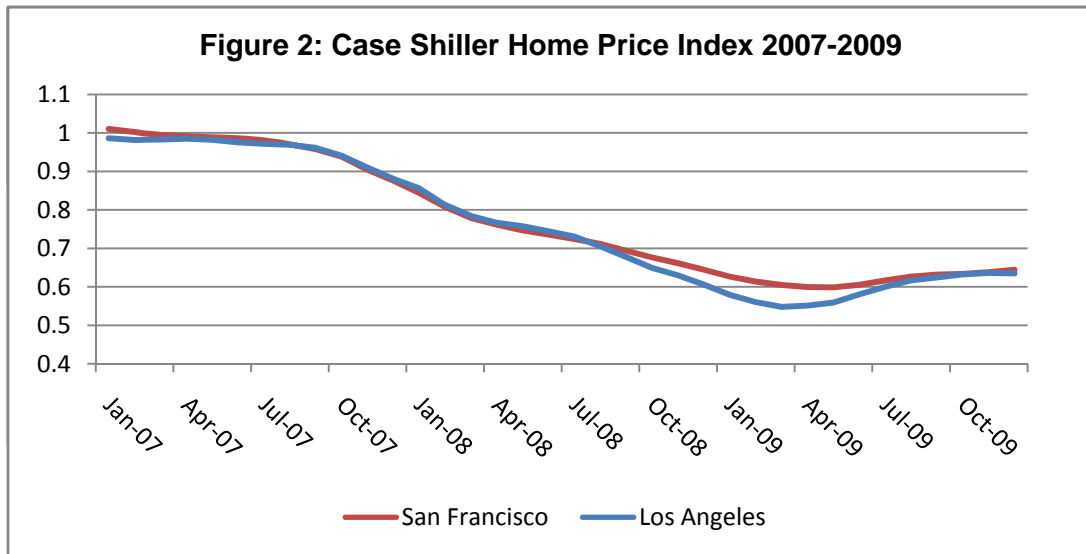
Figure 3: Timeline of Events in Model

| Pay cost c | Seller gets signal, z, about $\mu_t - \mu_{t-1}$; seller updates beliefs | Seller chooses p^L | Buyer gets signal, chooses whether to inspect | Seller updates beliefs from buyer behavior | If offer made, seller decides to accept or reject | Seller repeats process if reject chosen in time t |

t   t+1
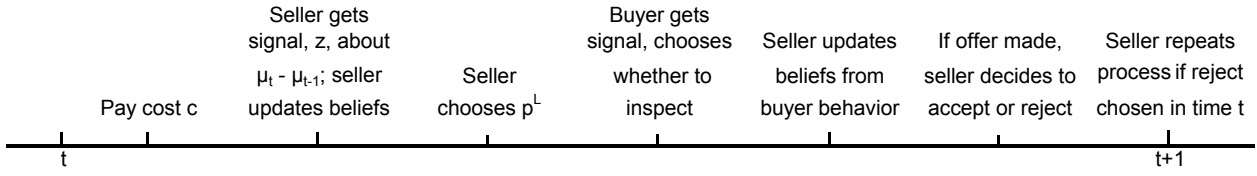
Table I : Summary Statistics

		List Price - Predicted Price (%)	Square Feet	Year Built	Time on Market (Weeks)	Change in List Price over Selling Horizon	Sales Price
Sell N= 87,879	Mean	8%	1683	1961	18	-10%	628372
	p25	-5%	1188	1950	5	-14%	360000
	p50	6%	1513	1960	12	-2%	530000
	p75	18%	1999	1980	25	0%	765000
Withdraw N=88,060	Mean	17%	1673	1960	20	-7%	
	p25	2%	1156	1949	8	-9%	
	p50	14%	1495	1958	16	-1%	
	p75	28%	2032	1979	26	0%	
Total N=175,939	Mean	12%	1678	1960	19	-9%	628372
	p25	-2%	1172	1949	6	-11%	360000
	p50	9%	1504	1959	14	-2%	530000
	p75	23%	2014	1979	26	0%	765000

Notes: The Predicted Price is calculated by applying a neighborhood level of sales price appreciation to the previous log selling price.

Table 2: Percent of Sellers on Market that Adjust List Price by Week Since Initial Listing

Weeks Since Listing	% Adjusting List Price
1	4.47%
2	7.72%
3	9.91%
4	11.21%
5	11.42%
6	10.82%
7	10.20%
8	10.27%
9	10.47%
10	10.26%
11	10.04%
12	9.78%
13	9.88%
14	9.86%
15	8.91%
16	8.97%
17	8.95%
18	9.24%
19	8.99%
20	8.92%
21	8.83%
22	8.69%
23	8.68%
>=24	9.54%

Table 3: Effects of Lagged Market Conditions on List Price

Dependent Variable	(1) Log List Price- Log Predicted Price	(2) Log List Price- Log Predicted Price	(3) Log List Price- Log Predicted Price	(4) Log List Price- Log Predicted Price	(5) Log List Price- Log Predicted Price
Log Predicted Price$_{t-1}$ - Log Predicted Price$_t$	0.5631***	0.7052***	0.7298***	0.7439***	0.7475***
	(0.0266)	(0.0327)	(0.0347)	(0.0367)	(0.0381)
Log Predicted Price$_{t-2}$ - Log Predicted Price$_{t-1}$		0.3933***	0.4858***	0.5167***	0.5298***
		(0.0321)	(0.0415)	(0.0473)	(0.0534)
Log Predicted Price$_{t-3}$ - Log Predicted Price$_{t-2}$			0.2262***	0.3071***	0.3238***
			(0.0366)	(0.0548)	(0.0644)
Log Predicted Price$_{t-4}$ - Log Predicted Price$_{t-3}$				0.1583***	0.1999***
				(0.0478)	(0.0741)
Log Predicted Price$_{t-5}$ - Log Predicted Price$_{t-4}$					0.0771
					(0.0570)
Month fixed effects	X	X	X	X	X
Zip code fixed effects	X	X	X	X	X
Observations	175939	175939	175939	175939	175939
Adjusted R-squared	0.111	0.113	0.113	0.114	0.114

*** p<0.01, ** p<0.05, * p<0.1

Notes: The Predicted Price is calculated by applying a neighborhood level of sales price appreciation to the previous log selling price. The t in the subscript of the explanatory variables denotes the month of initial listing. Regressions also include a Real Estate Owned dummy. Standard errors clustered at the zip code level are in parentheses.

Table 4: Effects of Lagged Market Conditions on Selling Outcomes

Dependent Variable	(1) Log List Price- Log Predicted Price	(2) Log List Price- Log Predicted Price	(3) TOM	(4) TOM	(5) Withdraw	(6) Log Sales Price- Log Predicted Price
Distribution of Price Depreciation Rates:						
2nd decile	**0.0196***	**0.0111**	**2.8823***	**3.0797***	**0.0499***	**-0.0171***
	(0.0048)	(0.0047)	(0.3819)	(0.4404)	(0.0099)	(0.0033)
3rd decile	0.0226***	0.0121**	**3.7888***	**3.9170***	**0.0828***	**-0.0245***
	(0.0046)	(0.0049)	(0.4001)	(0.4555)	(0.0109)	(0.0034)
4th decile	**0.0279***	0.0132***	3.9401***	**4.4235***	**0.0983***	**-0.0291***
	(0.0048)	(0.0051)	(0.4131)	(0.4910)	(0.0111)	(0.0034)
5th decile	**0.0447***	**0.0308***	**4.5446***	**5.4883***	**0.1151***	**-0.0227***
	(0.0052)	(0.0054)	(0.4038)	(0.4982)	(0.0104)	(0.0035)
6th decile	0.0490***	**0.0371***	4.7441***	5.5356***	0.1199***	-0.0197***
	(0.0054)	(0.0059)	(0.4716)	(0.6070)	(0.0112)	(0.0036)
7th decile	0.0474***	0.0364***	4.6186***	5.7616***	0.1250***	-0.0224***
	(0.0056)	(0.0059)	(0.5134)	(0.6659)	(0.0119)	(0.0037)
8th decile	**0.0588***	**0.0474***	4.4823***	4.7369***	**0.1505***	**-0.0172***
	(0.0057)	(0.0061)	(0.5507)	(0.6548)	(0.0118)	(0.0038)
9th decile	**0.0657***	**0.0553***	4.9347***	5.2629***	**0.1621***	-0.0157***
	(0.0063)	(0.0065)	(0.5504)	(0.6645)	(0.0124)	(0.0040)
10th decile	**0.0927***	**0.0815***	**3.5064***	**2.9475***	0.1672***	**0.0088**
	(0.0087)	(0.0091)	(0.5506)	(0.6505)	(0.0135)	(0.0042)
Change in Predicted Price over Selling Horizon					-0.5106***	
					(0.0369)	
Month fixed effects	X	X	x	x	x	x
Zip code fixed effects	X	X				
Only Listings that Sell		x		x		x
Observations	175939	87879	175939	87879	175939	87879
Adjusted R-squared	0.111	0.104	0.100	0.127	0.207	0.101

*** p<0.01, ** p<0.05, * p<0.1

Notes: The price depreciation rate is the the predicted price from 4 months prior to the initial listing date minus the predicted price at the initial listing date. d1 is the excluded group. TOM is measured in weeks. Withdraw is a dummy variable equal to one if the listing is ultimately withdrawn. Regressions also include a Real Estate Owned dummy. Coefficients in bold denote cases where the difference in the coefficient relative to the coefficient in the decile immediately below is statistically significant at the 10 percent level. Standard errors clustered at the zip code level in parentheses. In specification 6, robust standard errors are reported.

Table 5: Parameter Estimates of the Structural Model

Variable	Description	Estimate	Std. error
σ	St. dev. of Initial Prior	0.0679	0.0054
σ_η	St. dev. of buyer valuations	0.0841	0.0022
σ_v	St. dev. of buyer uncertainty over their valuation prior to inspection	0.0606	0.0016
σ_z	St. dev. of signal about weekly decline in mean valuations.	0.0322	0.0073
σ_ε	St. dev. of Belief about weekly decline in mean valuations.	0.0094	0.0007
κ	Buyer inspection cost.	-0.0050	--
c	Weekly holding cost.	0.0000	--
γ_0	Mean of Belief about weekly decline in mean valuations.	-0.0033	--

Note: The parameters without standard errors are fixed in estimation.

Table 6: Moments Used in Estimation

		Moment	Simulated Moment
1	% of homes that sell 5 weeks after initial listing	4.30%	4.65%
2	% of homes that sell 10 weeks after initial listing	3.30%	3.89%
3	% of homes that sell 15 weeks after initial listing	2.20%	2.92%
4	% of homes that sell 20 weeks after initial listing	1.55%	2.23%
5	Median Time on Market	12	12
6	25th pctile of List Price - Sales Price	0.00%	0.00%
7	50th pctile of List Price - Sales Price	1.90%	2.86%
8	75th pctile of List Price - Sales Price	5.40%	5.42%
9	Corr(Change in List Price, Change in Predicted Price)	0.76	0.76
10	5th percentile of list price change in week 3	-4.40%	-3.32%
11	Average Change in List Price	-10.00%	-12.23%
12	Median (List Price - Sales Price in Week 10 - (List Price - Sales Price in Week 5))	0.95%	0.02%
13	Median (List Price - Sales Price in Week 15 - (List Price - Sales Price in Week 10))	0.42%	0.13%
14	Stdev. of Monthly Price Changes	1.20%	0.67%

Note: All prices are in logs. The changes in moments 9 and 11 are at the week of sale relative to the week of initial listing. In moment 14, the empirical moment is the standard deviation of monthly price changes in the Case-Shiller index for San Francisco and Los Angeles during my sample period. The model equivalent is the standard deviation of monthly average sales price changes.

Table 7: Sales Price Dynamics from the Simulated Model

	Dependent Variable			
	Annual Price Change	Annual Price Change Prices Aggregated	Semi-Annual Price Change	Semi-Annual Price Change Prices Aggregated
OLS Estimates of AR(1) Coefficient	-0.003 0.124	0.040 0.152	-0.013 0.279	0.079 0.346
Assumptions				
Uncertainty Over Changes in Market Conditions	x	x	x	x

Average weekly sales prices are simulated for T=48,000 weeks and 20,000 new listings each week. Columns 2 and 4 show the results when average quarterly prices are used instead of average weekly prices. When there is no uncertainty over changes in market conditions, sellers beliefs at the time of listing are not sensitive to lagged market conditions, and any change in the distribution of buyer valuations during the listing period is perfectly observable.

Table 8: Volume and Time-on-Market Dynamics from the Simulated Model

	Dependent Variable	
	Log(Sales Volume)	Log(TOM)
Quarterly Change in Price	0.000 6.732	0.000 -7.633
Assumptions		
Uncertainty Over Changes in Market Conditions	x	x

Average weekly sales prices are simulated for T=48,000 weeks and 20,000 new listings each week. When there is no uncertainty over changes in market conditions, sellers beliefs at the time of listing are not sensitive to lagged market conditions, and any change in the distribution of buyer valuations during the listing period is perfectly observable.

Appendix Table 1: Robustness Specifications

Dependent Variable	(1) List Price	(2) TOM	(3) Withdraw	(4) Sales Price	(5) List Price
Lagged Depreciation	0.6022***	21.6028***	0.7018***	0.1370***	0.7818***
	(0.0482)	(1.0780)	(0.0315)	(0.0247)	(0.1185)
Lagged Depreciation*Lagged Num. Sales	-0.0018***	-0.0590***	0.0002	0.0000	
	(0.0003)	(0.0062)	(0.0002)	(0.0002)	
Lagged Num. Sales		-0.0055***	-0.0000	-0.0001***	
		(0.0009)	(0.0000)	(0.0000)	
Change in Expected Price over Selling Horizon			-0.5124***		
			(0.0096)		
Lagged Depreciation*Months Since Beginning of Sample Period					-0.0187**
					(0.0078)
Month fixed effects	X	X	x	x	x
Zip code fixed effects	X				
Observations	171066	171066	171066	85603	172460
Adjusted R-squared	0.111	0.096	0.202	0.100	0.112

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1