

# Essays on Housing Market Search and Dynamics

by

Elliot Anenberg

Department of Economics

Duke University

Date: \_\_\_\_\_

Approved:

---

Patrick Bayer, Supervisor

---

Andrew Sweeting

---

James Roberts

---

Christopher Timmins

Dissertation submitted in partial fulfillment of the requirements for the degree of  
Doctor of Philosophy in the Department of Economics  
in the Graduate School of Duke University  
2011

ABSTRACT  
(Economics)

Essays on Housing Market Search and Dynamics

by

Elliot Anenberg

Department of Economics  
Duke University

Date: \_\_\_\_\_

Approved:

---

Patrick Bayer, Supervisor

---

Andrew Sweeting

---

James Roberts

---

Christopher Timmins

An abstract of a dissertation submitted in partial fulfillment of the requirements for  
the degree of Doctor of Philosophy in the Department of Economics  
in the Graduate School of Duke University  
2011

Copyright © 2011 by Elliot Anenberg  
All rights reserved except the rights granted by the  
Creative Commons Attribution-Noncommercial Licence

# Abstract

This dissertation comprises two papers on search in the housing market. The first paper looks at the effects of equity constraints and loss aversion on the home selling problem. The empirical findings are consistent with a search model where sellers with low equity in their homes and sellers subject to nominal losses have higher reservation prices because of downpayment constraints and loss aversion, and thus hold out for higher prices. The second paper formalizes the home selling problem with a theoretical model. In the model, sellers are uncertain about the distribution of buyer valuations for their particular house, and they learn endogenously about the parameters of this distribution in a Bayesian fashion. I use dynamic programming techniques, simulated method of moments, and a rich dataset on home listings and home transactions of single-family homes in Los Angeles to estimate the parameters of the model. Simulations of the model are informative about the value of information and housing price dynamics.

# Contents

<b>Abstract</b>	<b>iv</b>
<b>List of Figures</b>	<b>viii</b>
<b>Acknowledgements</b>	<b>x</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Loss Aversion, Equity Constraints and Seller Behavior in the Real Estate Market</b>	<b>4</b>
2.1 Introduction . . . . .	4
2.2 Related Literature . . . . .	8
2.3 Theory . . . . .	10
2.3.1 Equity Constraints . . . . .	10
2.3.2 Prospect Theory . . . . .	12
2.4 Data . . . . .	13
2.5 Empirical Model of Prices, Equity Position, and Loss Aversion . . . . .	15
2.6 Estimation . . . . .	17
2.6.1 Stage 1 . . . . .	18
2.6.2 Stage 2 . . . . .	20
2.7 Results . . . . .	22
2.7.1 Preliminary Results: Houses that Sold at Least Twice . . . . .	22
2.7.2 Foreclosures . . . . .	24
2.7.3 Main Results: Houses that Sold at Least Three Times . . . . .	25

2.7.4	Comparison to GM . . . . .	26
2.7.5	Implications for Repeat Sales Estimators . . . . .	28
2.8	Conclusion . . . . .	28
<b>3</b>	<b>Uncertainty, Learning and the Value of Information in the Residential Real Estate Market</b>	<b>40</b>
3.1	Data . . . . .	46
3.1.1	Expected Prices . . . . .	48
3.1.2	Summary Statistics . . . . .	48
3.2	A Dynamic Model of the Home Selling Problem . . . . .	50
3.2.1	Offer Process and Buyer Behavior . . . . .	50
3.2.2	Structure of Information . . . . .	53
3.2.3	Learning . . . . .	54
3.2.4	Seller's Optimization Problem . . . . .	57
3.2.5	Comparative Statics . . . . .	59
3.3	Estimation . . . . .	60
3.4	Identification . . . . .	63
3.5	Results . . . . .	66
3.5.1	Model Fit and Discussion . . . . .	67
3.6	Model Predictions . . . . .	69
3.6.1	Value of Information . . . . .	69
3.6.2	Alternative Demand Predictors . . . . .	71
3.6.3	Price Dynamics . . . . .	72
3.6.4	Suboptimal Behavior . . . . .	75
3.7	Expectation Bias: Theory and Evidence . . . . .	76
3.8	Conclusion . . . . .	78
<b>A</b>	<b>Appendix: Chapter 2</b>	<b>95</b>

<b>B Appendix: Chapter 3</b>	<b>98</b>
B.0.1 Detail on Calculation of Expected Prices . . . . .	98
B.0.2 Proof of Theorem 1 . . . . .	99
B.0.3 Detail on Alternative Demand Predictors . . . . .	100
B.0.4 Expectation Bias: Additional Results . . . . .	102
<b>Bibliography</b>	<b>104</b>
<b>Biography</b>	<b>108</b>

# List of Figures

2.1	Sample Means. Standard Deviations in Parenthesis . . . . .	31
2.2	LTV Ratios at Time of Purchase. Count of Observations by Year. . .	32
2.3	Monte Carlo Simulation Results. . . . .	33
2.4	Seller Characteristics and Selling Prices . . . . .	34
2.5	Seller Characteristics and Selling Prices. Fixed Rate Mortgages. . . .	35
2.6	Simulation of Average Price When an Offer is Accepted, T=20 . . . .	36
2.7	Prospect Theory . . . . .	37
2.8	Nominal Price Levels in San Francisco MSA 1985-2005 . . . . .	37
2.9	Premium Over Expected Selling Price by Transaction Type . . . . .	38
2.10	Nominal Price Levels 1990 - 1997 . . . . .	39
3.1	Real Log Price Index in Los Angeles 1987-2009 . . . . .	80
3.2	Timeline of Events in Model . . . . .	80
3.3	Simulation of Select Moments . . . . .	81
3.4	List Price Policy Function . . . . .	82
3.5	Home Value Indexes for Los Angeles . . . . .	82
3.6	Average Welfare Loss from Uncertainty . . . . .	83
3.7	Average Welfare Loss Under Suboptimal Behavior with 3 Percent Re- altor Fee Relative to Fully Optimal Behavior with 6 Percent Fee . . . .	83
3.8	Test for Expectation Bias . . . . .	84
3.9	Summary Statistics . . . . .	85



3.10	Percent of Sellers that Adjust List Price by Weeks Since Listing . . .	86
3.11	Parameter Estimates from the Structural Model . . . . .	87
3.12	Moments Used in Estimation . . . . .	88
3.13	Moments Used in Estimation Continued . . . . .	89
3.14	Additional Results from Model. All Results are Averages Across Sellers.	90
3.15	Seller Welfare Under Alternative Priors Relative to Welfare at Estimated Prior . . . . .	91
3.16	Sales Price Regressions from Simulated Model . . . . .	92
3.17	OLS Equations Illustrating Bias. <i>t</i> -statistics in italics. . . . .	93
3.18	OLS Equations Illustrating Effects of Uncertainty. <i>t</i> -statistics in italics.	94

# Acknowledgements

I am very grateful to my advisor, Pat Bayer, for continuous guidance and advice. I also thank my committee members, Jimmy Roberts, Andrew Sweeting, and Chris Timmins, for their support along the way and for providing helpful insights and suggestions on my research.

I presented parts of this dissertation at Carnegie Mellon, Duke, the ERID Conference on Housing Market Dynamics, the Federal Reserve Board, the Federal Reserve Bank of Cleveland, the Federal Reserve Bank of Philadelphia, and Wharton. I would like to thank various participants at these seminars for their helpful comments.

Finally, special thanks go to my family, and especially my wife, for constant support during graduate school.

# 1

## Introduction

If a homeowner wants to sell his home, he does not simply sell the home through an exchange at a market price as he would sell shares of a stock. Rather, the home selling problem involves searching through a thin pool of buyers, deciding which offers to accept or reject (and how the strategy should change over time), and bargaining. The two essays in this dissertation investigate how certain features of this search process affect housing market dynamics and welfare.

The first paper, “Loss Aversion, Equity Constraints and Seller Behavior in the Real Estate Market”, looks at the effects of loss aversion and equity constraints on the decision to sell a home. I find strong evidence that owners facing nominal losses on their housing investments and owners with high LTV ratios sell for higher prices, on average. This finding is consistent with a search model of the housing market, where sellers with low equity and sellers subject to nominal losses have higher reservation prices because of downpayment requirements and loss aversion.

Several previous studies have empirically studied the effects of these constraints on selling behavior. However, estimating the effects of potential losses and equity position on prices is difficult because these variables are non-linear functions of un-

observed house characteristics. I exploit the panel nature of a rich dataset on home transactions in San Francisco to develop an econometric model that can point identify the effects of loss aversion and equity constraints on actual selling prices in a diverse sample of housing transactions.

The results help explain why prices and transaction volumes are correlated in the housing market, and why selling prices adjust slowly to deteriorating fundamentals.

In the second paper, “Uncertainty, Learning and the Value of Information in the Residential Real Estate Market”, I formally model the home selling problem as a model of sequential search with Bayesian learning. In the model, sellers receive offers from an unknown offer distribution, and sellers try to maximize their sales price while minimizing the holding costs of keeping the home on the market. During this process, sellers endogenously learn about the parameters of this offer distribution in a Bayesian fashion. I estimate the structural model by matching the predictions of the model to a rich dataset on home listings and home transactions of single-family homes in Los Angeles. I identify parameters that summarize the latent beliefs of the seller, as well as parameters that measure buyer taste heterogeneity and holding costs.

At the estimated parameters, the model can match many of the key features of the housing data, including declining list prices over the selling horizon; and lower list prices, larger list price changes, and shorter TOM for homes that are sold versus homes that are withdrawn. I also find that a high withdrawal rate is perfectly compatible with rational behavior.

Simulations of the model reveal a number of interesting results about the value of information and house price dynamics. I find that realtors provide substantial value to less sophisticated sellers because uncertainty and the inability to process information have significant effects on welfare. At the estimated parameters, the model predicts a significant amount of persistence in price appreciation rates, which

has been difficult for existing studies to explain within an efficient markets framework, even when the fundamentals follow a random walk.

# Loss Aversion, Equity Constraints and Seller Behavior in the Real Estate Market

## 2.1 Introduction

This paper contributes to the literature on the role of loss aversion and equity constraints in determining how local housing markets operate (Genesove and Mayer, (1997 and 2001); Lamont and Stein (1999); Engelhardt (2003)). Sellers who are averse to selling their house for less than they initially paid face unique incentives when making decisions such as whether and when to sell, and what prices to accept. Those with little or negative equity in their house are also thought to evaluate housing decisions differently given the way that the mortgage market operates. In a market downturn, the effects of loss aversion and equity constraints become more pronounced because more homeowners see their homes depreciate in value. Thus, understanding their effects on homeowner behavior is essential for understanding key features of cold housing markets such as declines in sales volume, a relatively large inventory of unsold homes on the market at any given time, and gradually declining prices. While I discuss other studies that directly address the effects of these

constraints on sales volume, this paper revisits the challenge of estimating effects on prices with a richer dataset.

The effect of down payment constraints on selling behavior is best understood through the following example. Suppose a family has a house that is initially worth \$100,000 and an outstanding mortgage of \$85,000. The family wants to move for an exogenous reason, and the purchase of a new house requires a minimum down payment of 10 percent. If housing prices stay the same or increase, the family could sell the house and would likely have enough cash to make a down payment on a new house. However, if prices fall by 10 percent, the family would only have enough to make a down payment of \$5,000 (ignoring moving costs), and the family may be better off staying rather than moving.

Alternatively, the family could list their price at an above-average price (“fishing”) and hope to eventually match with a buyer who has a relatively high valuation of the house. Of course this strategy would tend to involve keeping the house on the market for longer than average. To the extent that sellers with low equity use this strategy, there should be a negative relationship between list prices and equity for low or negative levels of equity, but not necessarily for higher levels. Sellers with high equity have enough cash to make a downpayment, and so if the costs to keeping a house on the market are sufficiently high, these sellers have less of an incentive to fish.

I also discuss and simulate a simple model that generates the same prediction between equity constraints and price, but does not rely on downpayment constraints. The model shows that it is relatively less costly for sellers with low equity, as measured by loan-to-value (LTV) ratios, to wait for higher prices on average because the option to default on their mortgage is relatively more attractive.

Estimating the effects of potential losses and equity position on prices is difficult because these variables are non-linear functions of unobserved house characteristics.

In a seminal paper, Genesove and Mayer (henceforth, GM) address these identification issues using a clever estimation procedure that bounds the true effects of loss aversion and equity constraints. They find an effect of equity position on list and sales prices, and in the lower bound, GM find no statistically significant effect of loss aversion on the sales prices. They do, however, find a significant effect of loss aversion on the list price. Whether loss aversion carries through to the actual transaction prices remains in part an open question. It is possible, as GM note, that since loss aversion is a psychological reluctance to sell, its effect may quickly diminish with learning or exposure to market conditions.

The main contribution of this paper is to use a rich dataset to develop a closely related econometric model that is less parametric than GM, and can *point* identify the effects of loss aversion and equity constraints on actual selling prices in a more diverse sample of housing transactions. Whereas GM estimate their model on a sample of condominium sales, I use a dataset that provides details of every housing transaction that occurred in the San Francisco metropolitan area over a 18 year period. In a first stage, I restrict the sample to houses that sold at least two times during periods when prices were rising rapidly and it is reasonable to assume that sellers do not face potential losses or equity constraints. During these hot markets, the econometric model predicts that unobserved quality of a house only affects prices linearly because the equity constraint and potential loss variables are zeroed out. Thus, I can estimate unobserved quality for this sample of houses using simple panel data estimation methods, where I follow GM in treating unobserved quality as a fixed effect. The estimator that I use is a more flexible version of the repeat sales estimator described in Shiller (1991); I use locally linear regression to allow the time effect to vary by house.

In the second stage I restrict the first stage sample to houses that have an additional sale during the market downturn when equity constraints and loss aversion



may affect selling behavior. For the transaction price during the cold market, unobserved quality has the usual non-linear effect that complicates GM's estimation strategy. However, I can recover point estimates of the effects of loss aversion and equity constraints using least squares on the restricted sample, where I substitute the estimate of unobserved quality from the first stage into the model.

As a whole, my results largely support the findings in GM in a larger, more diverse sample of housing transactions. The key difference is that I find larger effects. I find that a seller facing a 10 percent prospective nominal loss receives a 3.55 percent higher price, on average, while a seller with an 100 percent LTV ratio receives a 3.3 percent higher price than a seller with an 80 percent LTV ratio, on average.

In addition, I present a number of new findings. I find that the effects of loss aversion and equity constraints are smaller for homes surrounded by similar houses, possibly because competition makes it more difficult for sellers to negotiate higher prices. I also find that transaction prices of foreclosed properties do not display sensitivity to the LTV ratio. This is expected if the theories discussed above are driving the results since the sellers of foreclosed properties do not face the same constraints as the delinquent owner. This result supports the claim that LTV is not proxying for some unobserved characteristic of the home.

I also find that failing to control for loss and LTV in a repeat sales estimator overstates prices that a non-credit constrained seller expects to receive. The results imply that selling prices do not adjust as quickly to deteriorating fundamentals because sellers facing equity constraints and nominal losses are reluctant to set lower prices. This is one explanation for the large inventory of unsold homes in markets where home prices are falling: buyers are unwilling to pay prices that include premiums for loss aversion and equity constraints. In addition, popular home price indexes like Case-Shiller do not capture changes in search behavior that accompany a market downturn, and so an analysis of selling prices alone can understate the severity of a

market downturn.

This paper proceeds as follows. Section 2 reviews the related empirical literature, and discusses GM's results. Section 3 discusses the theoretical literature that motivates my empirical strategy and also presents a new theory to motivate the effects of equity position. Section 4 describes my unique dataset and presents summary statistics. In section 5, I describe my empirical model and discuss how it differs from GM. Sections 6 and 7 present the estimation strategy and the results, as well as a discussion of how my estimates compare to GM, and why they differ in some cases. Finally, section 8 concludes by summarizing the results and presenting directions for future research.

## 2.2 Related Literature

The theoretical motivation for an effect of loss aversion on house prices comes from Tversky and Kahneman (1979), who develop prospect theory based on experimental evidence that losses relative to a reference point loom larger than gains. The effect of equity constraints is developed in a theoretical model by Stein (1995), which is the foundation for much of the empirical work. He lays out a model where down payment constraints make it more difficult for sellers to realize the gains from moving when prices are falling.

My empirical strategy is most closely related to GM (1997,2001), who use Stein's theory as motivation to look for reduced form relationships between high LTV ratios, loss aversion, list prices, transaction prices, time on the market, and probability of sale. In their earlier study, they find using data from the Boston condominium market that an owner with a higher LTV ratio sets a higher asking price and has a higher expected time on the market than an owner with proportionately less debt. They also find that if sold, a unit with an LTV of 1 has a sales price that is 4 percent higher than a unit with an LTV of 0.8, all else equal. The authors find no statistically

significant effect on selling prices for LTV values below 80 percent, consistent with the theory of a threshold effect.

Their findings suggest that equity constrained owners do indeed fish for better prices. Furthermore, they actually obtain higher selling prices, but at a cost: they need to keep their property on the market for longer. While these results are consistent with Stein's theory, they do not necessarily validate it. It is possible that LTV is endogenous and is proxying for unobserved characteristics of the house or unobserved characteristics of the seller that affect prices such as risk aversion or bargaining power<sup>1</sup>. For example, if more risk averse sellers tend to make larger down payments on average and tend to set lower list prices to reduce the risk of not receiving any offers, this would induce a positive relationship between LTV and price. In the empirical section below I discuss how I address this potential endogeneity.

In a follow-up paper, GM (2001) find that most of the effect of LTV ratios on selling prices, list prices, and time on market is actually explained by nominal loss aversion. They find that sellers who expect to receive less than they originally paid for their property set higher asking prices on average, controlling for the seller's equity position. As described in detail below, GM's estimation strategy cannot point identify the effects of loss on actual prices. In the lower bound, they find that loss has no statistically significant effect on actual selling prices. In the upper bound, they find that a 10 percent increase in loss is associated with a 1.8 percent increase in price, all else equal. In both the lower and upper bounds, they find that the effect of increasing LTV from 80 to 100 is only 1.5 percent.

Engelhardt (2003) investigates the effects on household mobility. Using data from the NLSY on household moves across multiple metropolitan areas in the U.S., Engelhardt finds that nominal loss aversion significantly restricts household mobility, while low equity because of fallen house prices does not. Chan (2001) uses actual

---

<sup>1</sup> See Arnold (1999) for a model where bargaining power affects prices.

mortgage data from a large commercial bank to estimate a similar specification. Unlike other studies, she can accurately calculate LTV ratios over time because she observes the exact terms of the mortgage. She finds evidence that low equity does constrain mobility for a sample of homeowners in the New York metropolitan area. Ferriera, Gyourko, and Tracy (2008) find similar results using data from the American Housing Survey.

## 2.3 Theory

Hedonic and repeat sales estimators typically specify that the transaction price of a house is a function of house characteristics and a time index that captures the average price in each time period, not on idiosyncratic characteristics of the seller. In this section, I briefly review how the theory implies that prices could depend on two characteristics that can vary across sellers of similar houses: LTV and return on housing investment. The framework for the discussion is a search model in the spirit of Chen and Rosenthal (1996), where sellers face a constant distribution of bid prices, and they can increase their chances of receiving a higher than average selling price by setting a high reservation price and waiting for a buyer with a high idiosyncratic valuation to arrive.

### *2.3.1 Equity Constraints*

Two features of the housing market can help account for a relationship between price and LTV. The first, which is discussed in the introduction and used by GM to motivate their empirical findings, is down payment constraints. If sellers are moving into new houses, they need sufficient liquidity to secure an attractive mortgage. Often lenders will not approve a mortgage or will require private mortgage insurance without a sufficient down payment, and other lenders reserve their most attractive

financing packages for buyers who can put down more cash<sup>2</sup>. This feature of the mortgage market provides an additional incentive for equity constrained sellers to sell at a high price. Therefore, if the costs of trying to obtain a high sales price do not systematically depend on LTV, equity constrained sellers should set above-average reservation prices because their marginal benefit to receiving a higher sales price is relatively large.

A second feature of the housing market that leads to a similar prediction between LTV and price is that residential mortgages are often no recourse loans in many states<sup>3</sup>. That is, at any time, the owner of a house has the option to default on her mortgage and reset her LTV ratio to zero, but the creditor cannot take other assets from the owner or require future repayments. This outside options puts upward pressure on the reservation prices of equity constrained sellers. The extreme case can be illustrated in a simple search model. Suppose each period, a seller receives an offer from a buyer that is drawn independently from a stationary distribution,  $F$ . The seller has the option to accept – or reject the offer, pay a search cost, and move onto the next period. Different individuals likely face different search costs depending on how pressured they are to move, and how costly it is to advertise and show the house to prospective buyers<sup>4</sup>. Then, the value function of seller  $i$  in time  $t$  with offer  $x$  in hand and outstanding loan amount  $l$  is given by:

$$V_{it}(x, l_{it}) = \max_{\text{Accept or Reject}} \left\{ \max(x - l_{it}, 0), c_i + \beta \int V_{it'}(x', l_{it'}) dF(x') \right\} \quad (2.1)$$

where  $c_i$  denotes search cost and  $\beta$  is the discount factor. If the seller accepts offer  $x$ , they simply get utility from the difference between the offer amount and how much

---

<sup>2</sup> Given the data available to me, I focus on the early 1990's. The recent subprime crisis is largely a result of a relaxation of these requirements.

<sup>3</sup> This is the case in California for original home loans. In the empirical results below, I do not claim to be able to identify which of the two theories are driving the results.

<sup>4</sup> See, for example, Glower et al. 1998.

the seller needs to pay back to the mortgage lender. In this stylized model where there are no implications on credit rating, sellers never accept offers less than  $l$ .

Assuming a finite horizon, I simulate the average price received conditional on not exercising the default option for different types of sellers using backwards recursion<sup>5</sup>. This is equivalent to what we will observe in the data, except we will not observe  $c_i$ . Figure 2.6 shows how the premium above the average price varies with LTV, where the average price for this example is \$200,000. For low levels of LTV, there is no relationship between premium above the average price and LTV. However, for high values of LTV, on average, the results indicate that sellers tend to receive above average prices on average. The cost of drawing a new offer is just not as high for sellers with high LTV ratios because they always carry around an offer equal to  $l$ , and so they set higher reservation prices. Of course in this stylized model, the simulated wait times for these sellers are higher, which is consistent with the empirical findings of GM described above.

### 2.3.2 Prospect Theory

While low levels of equity can be a binding financial constraint for sellers, loss aversion can impose a behavioral constraint that has a similar effect on prices. Tversky and Kahneman (1979) develop prospect theory from the following insights gathered in experimental settings: gains and losses are examined relative to a reference point; the value of losses is steeper than for equivalently sized gains; and the marginal value of gains or losses diminishes with the size of the gain or loss. These insights are reflected in the value function in Figure 2.7, where the reference point is the price a seller initially paid for her house. Evidently, the marginal benefit to receiving a higher price is higher for sellers at risk of incurring a loss, but is diminishing. Again,

---

<sup>5</sup> I do the simulation separately for each value of  $l$  in a grid of 120 discrete points and for three different values of  $c_i$  reported in Figure 2.6. I let  $F$  be the normal distribution, with mean \$160,000 and standard deviation of \$2,700. I set the monthly discount factor at 0.99.

assuming that the costs to waiting for a higher price do not systematically depend on expected housing investment returns, prospect theory suggests that sellers facing losses set higher list prices.

## 2.4 Data

The data come from a national real estate company that provides detailed information about every housing transaction in the core San Francisco Bay Area counties from 1988-2005. The data report characteristics for every house sold such as square footage, lot size, year built, latitude and longitude, a unique property id, and the selling price<sup>6</sup>. The data also include specific information about the buyer and seller of each house sold such as the buyer's name, the seller's name, and central to the analysis of this paper, the buyer's first, second, and third loan amounts<sup>7</sup>. Variation in down payments allows me to separately identify the effects of equity constraints from loss aversion. Since I do not observe the terms of the mortgage, I calculate the home owner's outstanding loan at each quarter assuming that they took out a 30 year fixed rate mortgage at the prevailing interest rate<sup>8</sup>. The 30 year fixed rate assumption is reasonable given that I focus on mortgages originated in the late 80's/early 90's – which is before more creative financing became prevalent – in the main specification that I outline below. The same assumption is maintained in GM.

However, unlike in GM, for about 80 percent of the sales data, I also observe whether the interest rate on the loan is fixed or variable. I describe below how I use this information to show that my assumption about the terms of the mortgage is not driving my results.

---

<sup>6</sup> The unique id allows me to identify repeat sales. The data do not contain information on list prices. As a result, I focus on actual transaction prices unlike GM who primarily focus on list prices and time to sale.

<sup>7</sup> The LTV ratio that I use is thus a “combined” LTV ratio. Unfortunately, I do not observe secondary financing taken out after the purchase.

<sup>8</sup> The interest rate data come from <http://www.freddiemac.com/pmms/pmms30.htm>

In my analysis, I choose to focus on owner occupied and non-investor units as much as possible. It would be difficult to identify the effects of equity and loss aversion on investors' selling behavior in the Bay Area during my sample period because many investors take out high loans, renovate houses in unobserved ways, and "flip" the houses for large premiums. This type of behavior would bias my results because I would be attributing high prices to low equity or loss aversion, when in fact unobserved quality changes would cause me to underestimate the expected selling price. I avoid these types of transactions by excluding transactions that occurred within six months of each other<sup>9</sup>.

I make a few additional cuts to eliminate houses with potentially large changes in quality between transactions<sup>10</sup>. I eliminate potential re-builds, which are identified when the transaction date is prior to the year built. When a variable that indicates the year in which major improvements were made is populated, I delete all transactions of that house. This eliminates about 2 percent of the sample. And finally, I drop properties that experience appreciation/depreciation rates greater/less than 100 percent of the average rate<sup>11</sup>.

Figure 2.1 reports summary statistics of some key variables for the full sample, and for two different cuts of the data that are used in the analysis below. At closing, sellers put down 25 percent of the sale price and borrow the remaining 75 percent

---

<sup>9</sup> This eliminates about 4.5 percent of the sample. The official Case-Shiller price index also excludes transactions within six months of each other for the reasons cited above, as well as for the possibility that one of the transactions is non-arms-length (e.g., a transfer between family members before selling a property).

<sup>10</sup> In the data, the characteristics of the house (e.g. square feet, lotsize, etc) are not updated every time the house sells. Therefore, any kind of significant quality change is unobserved to the researcher. This is also an issue in GM's data.

<sup>11</sup> I calculate rough average yearly appreciation rates in San Francisco from a regression of log price on a set of house and year dummies. I also drop one percent of observations from each tail of the sales price distribution to minimize the effect of outliers. Finally, I remove non-arms-length transactions (e.g. property transfers between family members). In general, I followed the Case-Shiller home price indexes methodology used to eliminate outliers [http : //www2.standardandpoors.com/spf/pdf/index/SPCS\\_MetroAreaHomePrices\\_Methodology.pdf](http://www2.standardandpoors.com/spf/pdf/index/SPCS_MetroAreaHomePrices_Methodology.pdf)



on average<sup>12</sup>. There is heterogeneity in the housing stock across the San Francisco metropolitan area, as seen in the standard deviations of year built and square footage.

Figure 2.2 provides more details about the distribution of the initial LTV ratio, and how the distribution varies over time. The proportion of all transactions where the buyer makes a downpayment of 20 percent or less is increasing over time. By 2005, about 72 percent of transactions involve a downpayment of 20 percent or less.

Figure 2.8 shows the OFHEO nominal price index for the San Francisco metropolitan area from 1985 - 2005. Prices were generally rising leading up to and following the recession in the early 90's. Like much of the U.S., San Francisco experienced a slump in the housing market in the early to mid 90's. From the peak in the first quarter of 1990 to the trough in the last quarter of 1994, nominal prices declined by 12 percent. As in previous analyses, my empirical strategy exploits the variation in the market price levels over time, and relies heavily on the downturn in prices to get a sizable sample of sellers who face binding equity constraints and are at risk of incurring nominal losses on their housing investment.

## 2.5 Empirical Model of Prices, Equity Position, and Loss Aversion

The model follows GM (2001) closely. I specify that the expected log selling price of house  $i$  at time  $t$  is a linear function of observable attributes, the quarter of listing, and an unobservable component:

$$q_{it} = X_i\beta + \delta_t + v_i \tag{2.2}$$

where  $X_i$  is a vector of observable attributes (e.g. square feet, age),  $\delta_t$  is a time effect, and  $v_i$  is unobservable quality of the house.  $q_{it}$  should be interpreted as the average

---

<sup>12</sup> The notation  $LLTV$  will be used throughout to denote the initial LTV ratio, which is simply the loan amount divided by the purchase price. This differs from  $LTV$ , which is the outstanding loan at a given time divided by an estimated value of the house.

selling price sellers ultimately receive if they put their specific home on the market in quarter  $t$ , ignoring the effects of loss aversion and equity constraints.

I assume that the actual log selling price,  $p_{it}$ , differs from the expecting selling price in each period depending on both endogenous and exogenous idiosyncrasies of the transaction,  $w_{it}$ . That is,  $p_{it} = q_{it} + w_{it}$  where

$$w_{it} = \alpha_1 LTV_{it}^* + \alpha_2 LOSS_{its}^* + \epsilon_{it}. \quad (2.3)$$

$LTV_{it}^*$  measures the equity position of the seller; it is the difference between the LTV ratio and 80 percent, truncated from below at zero. Following the literature, I parameterize the effect with a spline function to only allow LTV to affect prices for values above 80 percent<sup>13</sup>. The LTV ratio at time  $t$  is simply the ratio of the seller's outstanding loan amount,  $l_{it}$ , to the expected selling price.  $LOSS_{its}^* = (p_{is} - q_{it})^+ = ((\delta_s - \delta_t) + w_{is})^+$ , which is the difference between the original log purchase price in some quarter  $s$  and the expected log selling price today, truncated from below at zero. This measures how much sellers expect to lose in percentage terms relative to the price they paid for the house if they put their house on the market today and receive the expected price. The final term,  $\epsilon_{it}$ , is an exogenous error that captures effects such as a buyer's unobserved taste for the particular house.

Substituting the definitions of  $LTV^*$  and  $LOSS^*$  into equation (2.3), we get an expression for the actual log selling price:

$$p_{it} = q_{it} + w_{it} = X_i\beta + \delta_t + v_i + \alpha_1 \left( \frac{l_{it}}{\exp(X_i\beta + \delta_t + v_i)} - 0.8 \right)^+ + \alpha_2 (\delta_s - \delta_t + w_{is})^+ + \epsilon_{it} \quad (2.4)$$

Note that both  $v_i$  and  $w_{is}$ , which is the amount the seller over or under paid relative to the expected price when they bought the house, are unobserved. Therefore, consistent estimates of the parameters using the relationship in equation (2.4)

<sup>13</sup> See for example GM(2001), Lamont and Stein (1999), Caplin et al. (1993), and Engelhardt(2003).

are not feasible because the unobservables enter the true model nonlinearly. GM address this issue by estimating two variations on equation (2.4) that bound the true effect. The first uses a noisy measure of *LOSS*, where  $X_i\beta + \delta_t$  is substituted for the expected selling price. This omits the unobserved quality term. The second also makes this substitution, but adds the residual from the previous selling price,  $p_{is} - X_i\beta - \delta_t = v_i + w_{is}$ , as an additional regressor. This essentially adds measurement error to the expected price. The intuition for why these specifications are biased in opposite directions is not clear as omitted variable bias and nonlinear errors in variable bias are both at play. They establish that their two estimates of  $\alpha_2$  are biased around the true effect of *LOSS* in opposite directions through simulation. However, the simulations impose the normal distribution on  $v$  and  $w$ , and assume that these two variables are independent of each other, which is a misspecification if the seller during the previous transaction was equity constrained or at risk of incurring a loss. They do not discuss simulation results for how their two estimates of  $\alpha_1$  compare to the true effect.

## 2.6 Estimation

I now present an estimation strategy that is less parametric than GM, and also can point identify *both* the effects of LTV and LOSS on selling prices. The main advantages relative to GM are that I get a precise estimate of the magnitudes of the effects and I avoid having to make distributional assumptions in simulations to argue that my empirical strategy is valid. The estimation proceeds in two stages.

### 2.6.1 Stage 1

In the first stage, I restrict the sample to houses that sold at least two times in the 1988-1989 and 1998-2005 intervals<sup>14</sup>. I assume that sellers during these periods were not equity constrained or at risk of a nominal loss on their housing investment. This assumption is reasonable given the large price increases for the Bay Area during these periods as shown in Figure 2.8. Leading up to the 90's, nominal prices appreciated in every quarter, at an average rate of 4 percent each quarter. From 1998 onwards, there was appreciation in every quarter except one, and the average rate was 3 percent. Given these high rates of average appreciation, it seems very reasonable to assume that homeowners did not face nominal losses during these time periods<sup>15</sup>. Homeowners who put little down at the time of purchase could still have LTV ratios above the 80 percent threshold at the time of sale if they move again quickly. As discussed above, I mitigate this potential bias by dropping multiple sales of the same property that occurred within a six month period. I also calculate a crude LTV ratio by applying a price index to the previous selling price, and then I exclude sales where the LTV ratio exceeds 65 percent<sup>16</sup>. It is also worth noting that a large percentage of the first stage sample involves sales in the early 2000's when low downpayment loans became more accessible, and so Stein's (1995) equity constraint theory may have less of an effect on selling behavior in this sample.

Thus, for quarters during these years,  $LTV^*$  and  $LOSS^*$  are zero and equation

---

<sup>14</sup> A house that sold twice between 1998-2005 and zero times in 1988-1989 would be included, as would a house that sold once between 1988-1989 and once between 1998-2005.

<sup>15</sup> One potential criticism is that Bay Area appreciation may be masking depreciation in some local areas. As a check to confirm that this is not a major issue, I calculate that only about 3 percent of all repeat sales in the 1988-1989 and 1998-2005 intervals were sold for prices less than their previous price. Many of these could be due to forced sales. During 1990-1997, the number is 33 percent.

<sup>16</sup> I discuss the calculation of this imperfect LTV ratio in more detail in Section 7 and Appendix A.

(2.4) reduces to

$$p_{it} = \delta_t + v_i + \epsilon_{it} \quad (2.5)$$

where the house fixed effect has absorbed  $X_i\beta$ .

In the first stage, I estimate equation (2.5) for the restricted sample to recover estimates of both the house fixed effects and the time effects. I use a more flexible version of the repeat sales price estimator described in Shiller (1991). A repeat sales estimator differences equation (2.5) for pairs of sales of the same house to remove the fixed attributes  $v_i$ , and then recovers estimates of the coefficients on the quarter dummy variables using OLS. The advantage of a repeat sales estimator is that I can control for both observed and unobserved characteristics of the house; the disadvantage is that  $v_i$  is assumed constant over time. GM also assume that the quality of the house remains constant over time in their specification. I try to minimize problems associated with the latter assumption by dropping houses suspected of large changes in quality, as described in Section 4.

I implement the repeat sales estimator, except I use locally linear regression to allow the time effects to vary by house. That is, I estimate equation (2.5) separately for each observation  $i$  using weighted least squares, where the weights for observations  $j \neq i$  are based on differences in the observables age, square feet, and location between  $j$  and  $i$ . Location is defined by latitude and longitude. Each individual kernel weight is formed using a standard normal kernel,  $N$ , and bandwidth,  $h$

$$W_{j(i)} = \left[ \prod_{k=1}^3 N\left(\frac{Z_i^k - Z_j^k}{h * std(Z^k)}\right) \right]^{1/2} \quad (2.6)$$

where  $Z_i$  is the 3x1 vector of the observables and  $W_{j(i)}$  is the weight for observation  $j$  when running least squares for observation  $i$ . Due to the symmetry of the normal distribution,  $W_{j(i)} = W_{i(j)}$ . In the results presented below, I set the bandwidth to

equal .75. The results are robust to a variety of bandwidths<sup>17</sup>.

Note the added flexibility from allowing  $\delta$  to vary by house. Whereas GM assume that all sizes of houses, new and old houses, and houses in different neighborhoods appreciate homogeneously over time, my specification allows different types of houses to appreciate differently in different time periods. Averaging this heterogeneity into a single parameter can lead to inaccurate predictions of the expected selling price since a single metropolitan area encompasses a variety of heterogeneous submarkets. This estimation approach also provides additional within quarter of sale variation to help identify the effects of *LTV* and *LOSS*.

### 2.6.2 Stage 2

In stage 2, I further restrict the sample to houses that transacted at least three times between 1988-2005, and had at least one transaction during 1990-1997. I require a house to have at least 2 transactions when the market was hot so that I can get an estimate of  $v_i$  for that house from the first stage<sup>18</sup>. I require an additional transaction when the market was cold so that I can get a sample of houses where equity constraints and loss aversion are potentially binding, *and* where I have an estimate of the house fixed effect. For now, I require the seller's name for the transaction during the early nineties to be the same as the buyer's name on the previous transaction for that house. This rules out foreclosures and cases where the seller may not face the same mortgage terms documented in the previous transaction.

For my second stage sample where *LTV\** and *LOSS\** may have an effect on

---

<sup>17</sup> In particular, I tried bandwidths ranging from as low as .15 (which is the optimal bandwidth according to Silverman's Rule of Thumb (Silverman (1986))) and as high as 1.5. A bandwidth of 1.5 produces little variation in the  $\delta$  estimates, and the results are essentially the same as restricting the  $\delta$ 's to be the same for every observation.  $h = .75$  led to reasonable variation in the estimates of  $\delta$  across houses.

<sup>18</sup> This is simply the average price for each house minus the average time effect for each house.

prices, I can rewrite equation (2.4) as

$$p_{it} - \hat{v}_i = \delta_t + \alpha_1 \left( \frac{l_{it}}{\exp(\delta_t + \hat{v}_i)} - 0.8 \right)^+ + \alpha_2 (p_{is} - \delta_t - \hat{v}_i)^+ + \epsilon_{it} \quad (2.7)$$

for years between 1990-1997 where I have moved the house fixed effect recovered in the first stage,  $\hat{v}_i$ , to the left hand side, and I substitute  $\hat{v}_i$  into the equations for  $LTV^*$  and  $LOSS^*$ . Note that the  $\delta$ 's in this equation still need to be estimated because the first stage regression recovers time effects for 1988-1989 and 1998-2005, but not 1990-1997. I also include a second order polynomial to control for tenure, which is the number of months since the prior transaction. I estimate equation (2.7) using non-linear least squares to get point estimates of  $\alpha_1$ ,  $\alpha_2$ , and  $\delta_t$ <sup>19</sup>.

One concern is that a sample of houses that sold at least three times is not a representative sample of the housing market. However, since my dataset spans 18 years, it is not unusual for a house to be transacted at least three times. Figure 2.1 shows that 28 percent of houses sold are sold at least three times between 1988-2005. Figure 2.1 also shows that this sample is comparable in observables to the full sample. The first stage of the estimation strategy makes additional restrictions on the data to ensure that the first stage sample of prices is not affected by equity constraints. To ensure that the main results are not being driven by sample selection bias, I also present results using an alternative estimation procedure that follows GM (1997) and uses a much larger sample of home sales.

The above estimation strategy produces consistent estimates of  $\alpha_1$  and  $\alpha_2$  as the number of houses and the number of time periods gets large. In practice, since we observe a large number of houses but a small number of sales for each property,

---

<sup>19</sup> In the second stage, I revert to the assumption maintained in GM that all houses share the same time effect. Estimating equation (7) separately for each house using a form of weighted non-linear least squares is not computationally feasible. As a robustness check, I estimate equation (7) separately for different groups of houses (based on observables such as square feet). The estimates of  $\alpha_1$  and  $\alpha_2$  are not significantly different across all of the subsamples.

another concern is that the estimators of  $\alpha_1$  and  $\alpha_2$  have poor finite sample properties<sup>20</sup>. In a Monte Carlo simulation exercise, I use the above estimation procedure on a simulated sample of 1,150,497 homes where the number of sales for each house during the sample period is small. Figure 2.3 shows the results and Appendix B describes the details of the simulation procedure. I find almost no bias on the coefficient on  $LTV^*$  and a small positive bias on the coefficient on  $LOSS^*$ . In addition, the low root mean square errors imply that the estimators are precise for sample sizes comparable to the one used here. Under the null that there is no effect, the results suggest that an increase in loss of 10 percent leads to an increase in price of about .76 percent. The bias decreases by more than 50 percent under the alternative that there is an effect of loss on price. Given the 0.35 estimate of  $\alpha_2$  and the tight standard errors that I report below, I argue that the simulation results are evidence that the potential for bias does not affect any of my main conclusions.

## 2.7 Results

### 2.7.1 Preliminary Results: Houses that Sold at Least Twice

Before presenting the main results, I present results using an alternative estimation procedure that follows GM (1997). The advantage is that I can use the larger sample of houses that sold at least twice to estimate the effect of  $LTV$  semi-parametrically. The cost is that the estimates of expected selling price are biased.

The details of the estimation strategy are available in Appendix A. To summarize, I ignore the effects of  $LOSS^*$  and  $LTV^*$  in a first stage to get estimates of  $\delta$  using a repeat sales estimator on the entire sample. Then, I apply the estimated price indexes to the initial purchase price to compute the expected selling price. With an estimate of the expected selling price for each house in hand, the second stage

<sup>20</sup> For fixed  $T$  and large  $N$ , the estimators of  $\alpha_1$  and  $\alpha_2$  will be consistent if  $\sigma_\epsilon = 0$ . Thus, the use of local linear regression, which decreases  $\sigma_\epsilon$  relative to the standard fixed effect estimator, improves the properties of the estimator. I verify this in the simulations. I thank a referee for this insight.



reduces to an equation that is linear in the unknown parameters. This is one of the main specifications considered in GM (1997). Instead of parameterizing the effect of LTV as a spline function and using OLS, I use partially linear regression to get a semi-parametric estimate of how LTV affects prices<sup>21</sup>. However, it is clear that this estimation strategy introduces bias because I omit  $LOSS^*$  and  $LTV^*$  from the first stage. In addition, the estimates are confounded by measurement error because the expected selling price is calculated using the actual initial purchase price, which includes the idiosyncracies of the initial transaction.

Despite these caveats, it is useful to look at non-parametric plots to confirm that LTV does indeed have a threshold effect on prices. Figure 2.9 plots the premium of the actual selling price above the expected selling price as a function of LTV for properties sold during 1990-1997. I discretize LTV into bins corresponding to ten quantiles of the LTV distribution, and each dot represents the mean of the premium at the midpoint of each bin. The standard deviations of the premium in each bin are around 0.15. In this plot, I have imposed little parametric structure on the effect of LTV. I have only imposed that the tenure variables enter the model linearly, and are additively separable from LTV. For the Bay Area as a whole, the results are strikingly consistent with the simulated data shown in Figure 2.6. Sellers with high values of LTV tend to earn a higher premium above the expected selling price, and the premium is increasing in LTV. However, below a threshold, the effect flattens out: an increase in LTV over a relatively low value of LTV does not appear to significantly affect the premium. This helps to justify parameterizing the effect of LTV with a spline function in the estimation procedure described above.

---

<sup>21</sup> I do estimate the effects of LTV and LOSS using OLS here as another robustness check to make sure that my results are not being driven by using a sample of houses that sold at least three times. I find that the results using OLS are comparable to the results presented below.

### 2.7.2 Foreclosures

Sellers of foreclosed properties, usually banks or other institutions who represent the owners of the mortgage debt, neither face down-payment constraints in the sense of Stein (1995) nor do they typically have the outside option of resetting the LTV ratio to zero. If these constraints or the presence of the outside option are driving the relationship between LTV ratio and prices shown above, then in a sample of foreclosure transactions we would not expect to see a relationship between LTV and premium.

Although I cannot directly identify foreclosures in my transaction data, I parse the buyer and seller names for certain characters to create a sample that likely overstates the amount of foreclosures in the Bay Area. More specifically, I classify a transaction as a foreclosure if the seller name does not appear to be an individual (i.e. there is no comma separating a first and last name) but the buyer name on the previous transaction for the property does appear to be an individual<sup>22</sup>.

The semi-parametric plots for the foreclosure sample are shown in Figure 2.9. The standard deviations of the premium in each bin are around 0.2. In the sample of foreclosure transactions, the effects of LTV on premium for high levels of LTV is flatter. This helps to confirm the main results, because if LTV is correlated with unobservable changes to the house, the relationship between LTV and prices should show up here as well.

Furthermore, all else equal, sellers of foreclosed properties tend to receive lower prices<sup>23</sup>. This is consistent with beliefs that foreclosed properties are often sold for below market prices, possibly because the house is not properly maintained during

---

<sup>22</sup> I also create a sub-sample that may understate the number of foreclosures, where the seller's name has variations on the words "Bank", "Loan", etc. The results are similar, although there is more noise because this sample is quite small.

<sup>23</sup> The results for the more restrictive sub-sample lie below these prices, as expected if the main sample includes some transactions that are not foreclosures.

the foreclosure process or because it is especially costly for institutions to keep houses on the market. Campbell, Giglio, and Pathak (2009) find similar results.

### *2.7.3 Main Results: Houses that Sold at Least Three Times*

The first 2 columns of Figure 2.4 presents the main results of the estimation procedure described above. I calculate the standard errors by bootstrapping to account for the sampling variation in the first stage estimates. Sellers facing nominal losses do receive higher prices on average, holding fixed the quality of the house. A seller facing a nominal loss of 10 percent receives a 3.55 percent higher price on average, and the effect is statistically significant. Higher LTV values are also associated with higher prices. An increase in LTV from 0.8 to 1 is associated with a 3.3 percent increase in price.

Column 2 adds a quadratic loss term. There is no evidence that the effects of loss are diminishing as the value function in Figure 2.7 would predict. GM find a falling marginal response to the prospective loss for list prices, but they do not report results on selling prices.

The results in Figure 2.9 suggest that the effect of LTV on price might extend to LTV values below the 80 percent threshold used in GM. It is also possible that using the amortized original debt understates the number of households facing equity constraints because I do not observe secondary financing taken out after the purchase. To address these issues, column 3 changes the threshold to equal 0.6. The main results are robust to this alternative parameterization, although the effects of  $LOSS^*$  and  $LTV^*$  are slightly lower.

Column 4 addresses the potential endogeneity of  $LTV^*$  by adding the seller's LTV ratio at the time of initial purchase,  $ILLTV^*$ , as an additional regressor.  $ILLTV^*$  is included to proxy for unobservable seller characteristics, such as risk aversion or the propensity to make home renovations, that could affect future pricing and are

correlated with a high initial LTV ratio. The coefficients on  $LOSS^*$  and  $LTV^*$  remain largely unchanged when  $ILLTV^*$  is included. This result provides stronger evidence that the theories discussed in section 3 rather than correlated unobservables are driving the effects on pricing.

Figure 2.5 shows the results of the same specifications as in Figure 2.4, except now I remove sales during the 1990-1997 period where the interest rate type on the first, second, or third mortgage is listed as variable, graduated, or missing. For this sample, my assumption that the current loan balance is computed by amortizing the original mortgage amount is more reasonable. All of the main conclusions discussed above carry through to this sample, although the effect of LTV falls from 3.3 percent to 2.1 percent.

#### 2.7.4 Comparison to GM

As mentioned above, whether loss aversion affects actual selling prices is in part an open question because in the lower bound, GM find no statistically significant effect of  $LOSS^*$  on selling price. It is possible that loss aversion affects list price decisions early in the selling horizon, but then disappears with learning or exposure to market conditions. My results confirm that loss aversion is in fact associated with higher prices, and the .35 coefficient on  $LOSS^*$  in column 1 is about twice as high as GM's upper bound.

I also find a larger effect for  $LTV^*$  compared to GM, who find that an increase in LTV from 0.8 to 1 is only associated with a 1.2 percent increase in price<sup>24</sup>. There are several possible sources for this difference. As discussed above, unlike for the coefficient on  $LOSS^*$ , GM do not discuss how their two estimates of the effect of  $LTV^*$  compare to the true effect as their focus is on loss aversion. In addition, GM do not estimate  $LTV^*$  and the effect of  $LTV^*$  on price simultaneously as in equation

---

<sup>24</sup> In their 1997 paper that does not control for loss aversion, they find that the effect is 4.3 percent.

(2.4)<sup>25</sup>. Thus, one possibility is that their estimates for  $LTV^*$  are biased downwards.

Another potential reason for the larger estimates reported here is the different samples. GM estimate the effects for a sample of condos in Boston, whereas I use a broader sample that spans the entire Bay Area housing market. Since condos are typically less differentiated relative to other properties (i.e. there may be similar condos in the same building and condos typically have fewer distinguishing features such as lawns), it may be more difficult for condo sellers to earn a higher price because there is more competition. A seller of a more atypical house may be able to negotiate a higher price from a high valuation buyer because there are fewer substitute houses available to the buyer. To see if the discrepancy is the result of using a more heterogeneous sample of real estate, I investigate whether the effects of  $LOSS^*$  and  $LTV^*$  depend on whether the property is more homogeneous in that it is likely to be surrounded by similar units. The results are reported in column 5, where  $HETER_{it}$  is the standard deviation of the year built for all homes that sold within a .5 mile radius of house  $i$  within 4 years of quarter  $t$ <sup>26</sup>. The estimated effects of  $LOSS$  and  $LTV$  are indeed lower for more homogeneous properties, and the effect of  $LOSS$  appears to be more sensitive to  $HETER$  than  $LTV$ . A one standard deviation decrease in  $HETER_{it}$  (a .0765 decrease relative to the mean of .13) lowers the effect of a 10 percent increase in  $LOSS$  from 3.58 to 3.04 percent. For the effect of an increase in  $LTV$  from 0.8 to 1, the effect is decreased from 3.3 to 2.7 percent. One explanation for the higher sensitivity to  $HETER$  for  $LOSS$  is that while  $LTV$  represents an institutional constraint on sellers' behavior, loss aversion is a psychological effect that is more likely to be overcome if the seller is surrounded

---

<sup>25</sup> They do not explicitly describe how  $LTV^*$  is calculated, but it is probably through a procedure similar to the estimation described in Appendix A. This produces biased estimates for the reasons discussed above.

<sup>26</sup> I also run the regression where  $HETER_{it}$  is an indicator equal to zero if the property has a unit number (e.g. Apt. #, Suite #) in its address. The signs are similar but the estimates are imprecise because this measure of homogeneity does not have as much variation.

by competing sellers who do not demand a premium<sup>27</sup>.

### 2.7.5 Implications for Repeat Sales Estimators

My results have implications for the interpretation of home price indexes estimated using conventional methods, which do not control for seller characteristics. In repeat sales estimators for example, if equity position is not controlled for, then the estimated coefficients on the time dummies should be interpreted as an index of the selling price of a typical home in each time period. But the average selling price is inflated by sellers with losses and low equity who evidently keep their homes on the market for longer to wait for high prices. Thus, for example, a home builder looking to time her building decisions optimally would get a distorted estimate of average prices by looking at estimates from a traditional repeat sales analysis. To receive those average prices during a market downturn, the builder may have to keep the home on the market for an unusually long time.

The differences can be quite large. Figure 2.10 contrasts estimates of  $\delta$  when  $LTV^*$  and  $LOSS^*$  are included and excluded from equation (2.4). The market values estimated using traditional repeat sales estimators can be biased upwards by as much as 4 percent. One interpretation of these results is that, in a downturn, the selling prices of homes do not drop as fast as their fundamental values because equity constrained sellers and sellers facing nominal losses are reluctant to accept lower prices.

## 2.8 Conclusion

Using richer data, this paper has built upon the insights of GM to provide stronger evidence that loss aversion and equity constraints affect seller behavior. To my knowl-

---

<sup>27</sup> GM offer a similar explanation for why  $LTV$  has a similar effect on list price and selling price, whereas the effect of  $LOSS$  on the selling price is lower relative to its effect on the list price.

edge, my estimation strategy is the first to point identify the effects of loss aversion and equity constraints on actual transaction prices. I find that both loss aversion and equity constraints affect housing prices, and that the effects are larger than previously thought. I also provide stronger evidence that the relationship between these seller characteristics and prices is causal.

These results, when combined with the results in GM (2001), imply that sellers facing losses and equity constraints select higher reservation prices on average and realize a higher price conditional on a sale occurring<sup>28</sup>. Evidently, loss aversion is not a brief hope for a positive return that the market quickly corrects.

The empirical methodology developed in this paper, which relies on the panel nature of the data, could be applied to other housing applications such as discrimination. The methodology should become more useful as more years of transaction data become available. For example, if the demographics of the buyer and seller are observed, one could potentially look at houses that sold three times during the sample, twice when the demographics of the buyer and seller are the same and once when they are different. This would provide an estimate of the unobserved quality of the house that is pure of effects due to differences in demographics, which could be used in a second stage regression on the sample where the characteristics of buyer and sellers are different.

My findings combined with the results from other studies discussed in Section 2 suggest that during market downturns, sellers become locked-in to their homes because of loss aversion and equity constraints. This slows down the market. Selling prices do not drop as quickly because sellers are reluctant to accept lower prices; homes sit on the market for longer because many sellers are “fishing” for high prices, as shown in GM; and sales volume slows down because owners delay selling altogether

---

<sup>28</sup> The cost, as noted above, is that they must also be prepared for a longer time on market and a lower probability of sale.

to avoid nominal losses on their housing investments. Future work will try to generate these reduced form findings using a structural dynamic search model. Estimating this type of model could allow for interesting counterfactuals such as eliminating loss aversion and reducing outstanding mortgage balances.



Variable	All Transactions	Houses Sold 2x	Houses Sold >2x
Number of Observations	1,150,497	402,986	321,259
Sale Price	\$ 371,518 ( 238,111 )	\$ 374,350 ( 236,694 )	\$ 339,582 ( 218,439 )
Square Footage	1,612 ( 679 )	1,612 ( 657 )	1,456 ( 587 )
Year Built	1966 ( 23 )	1966 ( 23 )	1968 ( 22 )
Loan Amount	\$ 273,785 ( 193,563 )	\$ 277,766 ( 188,776 )	\$ 256,953 ( 181,230 )
ILTV <sup>a</sup>	0.75 ( 0.259 )	0.75 ( 0.254 )	0.77 ( 0.260 )

<sup>a</sup>The loan-to-value ratio at the time of sale.

FIGURE 2.1: Sample Means. Standard Deviations in Parenthesis

<u>Year</u>	<u>ILTV &lt; 0.8</u>	<u>ILTV &gt;= 0.8</u>
1988	36,118 ( 50.7 )	35,072 ( 49.3 )
1989	37,920 ( 55.7 )	30,117 ( 44.3 )
1990	25,575 ( 50.4 )	25,202 ( 49.6 )
1991	25,695 ( 49.9 )	25,777 ( 50.1 )
1992	24,523 ( 46.0 )	28,747 ( 54.0 )
1993	20,029 ( 40.7 )	29,195 ( 59.3 )
1994	19,867 ( 36.3 )	34,926 ( 63.7 )
1995	17,105 ( 36.6 )	29,669 ( 63.4 )
1996	21,625 ( 38.0 )	35,268 ( 62.0 )
1997	26,182 ( 39.3 )	40,476 ( 60.7 )
1998	26,598 ( 37.1 )	45,060 ( 62.9 )
1999	30,173 ( 38.6 )	48,041 ( 61.4 )
2000	30,846 ( 42.6 )	41,500 ( 57.4 )
2001	22,766 ( 39.9 )	34,255 ( 60.1 )
2002	27,734 ( 38.9 )	43,567 ( 61.1 )
2003	26,151 ( 34.9 )	48,729 ( 65.1 )
2004	25,320 ( 30.4 )	57,841 ( 69.6 )
2005	20,346 ( 27.9 )	52,482 ( 72.1 )

Note: Percentages are in parenthesis.  
All rows total 100 percent.

FIGURE 2.2: LTV Ratios at Time of Purchase. Count of Observations by Year.

N = 1,150,497 houses<sup>a</sup>

True Value alpha1,alpha2	Bias	RMSE
0.15;0.35	0.0033;0.0326	0.0101;0.0347
0;0	0.0061;0.0755	0.0125;0.0768

Simulation details are discussed in Appendix B.

The house fixed effects are drawn from a normal distribution with mean=12 and std=.445.

The error terms are drawn from a normal distribution with mean=0 and std=.099.

The true coefficients on the time dummy variables are assumed equal to the OFHEO repeat sales price indexes for San Francisco. I lower the coefficients on quarters during the years 1990-1997 by 5 percent.

<sup>a</sup>Since only houses that sell at least twice enter into the estimation routine, the number of houses used for estimation is substantially less than the sample size.

FIGURE 2.3: Monte Carlo Simulation Results.

Dependent Variable: Log (Selling Price) - House Fixed Effect  
 NLLS equations, Standard Errors are in parentheses  
 1990 - 1997

Variable	All Mortgage Types				
	(1)	(2)	(3)	(4)	(5)
<b>Loss</b>					
LOSS*	0.355 ** ( 0.025 )	0.350 ** ( 0.029 )	0.326 ** ( 0.023 )	0.354 ** ( 0.030 )	0.266 ** ( 0.038 )
LOSS*-Squared	--	0.016 ( 0.100 )	--	0.016 ( 0.100 )	--
LOSS* x HETER	--	--	--	--	0.706 ** ( 0.292 )
<b>Loan-to-Value Ratio</b>					
LTV*	0.163 ** ( 0.015 )	0.163 ** ( 0.015 )	0.139 ** ( 0.008 )	0.157 ** ( 0.021 )	0.113 ** ( 0.038 )
ILTV*	--	--	--	0.011 ( 0.023 )	--
LTV* x HETER	--	--	--	--	0.417 * ( 0.247 )
HETER	--	--	--	--	0.112 ** ( 0.019 )
T*	0.8	0.8	0.6	0.8	0.8
Controls for Tenure	Yes	Yes	Yes	Yes	Yes
Quarter of Sale Dummies	Yes	Yes	Yes	Yes	Yes
Number of Obs.	27,467	27,467	27,467	27,467	27,467

\*\* Indicates significance at the 5 percent level.  
 \* Indicates significance at the 10 percent level.

Standard Errors are calculated by bootstrapping to account for the generated regressors in the second stage.

LOSS is the greater of the difference between the previous selling price and the estimated value in the quarter of sale, and zero. LTV is the greater of the difference between the ratio of outstanding loan to estimated value in the quarter of sale and T\*, and zero. ILTV is the greater of the difference between the ratio of initial loan to previous selling price and T\*, and zero. HETER is the standard deviation of the year built for all homes that sold within a .5 mile radius of house *i* within 4 years of quarter *t*.

FIGURE 2.4: Seller Characteristics and Selling Prices

Dependent Variable: Log (Selling Price) - House Fixed Effect  
 NLLS equations, Standard Errors are in parentheses  
 1990 - 1997

Variable	Fixed Rate Mortgages				
	(1)	(2)	(3)	(4)	(5)
<b>Loss</b>					
LOSS*	0.364 ** ( 0.034 )	0.312 ** ( 0.043 )	0.350 ** ( 0.036 )	0.309 ** ( 0.038 )	0.278 ** ( 0.060 )
LOSS*--Squared	--	0.160 ( 0.158 )	--	0.161 ( 0.156 )	--
LOSS* x HETER	--	--	--	--	0.798 ** ( 0.407 )
<b>Loan-to-Value Ratio</b>					
LTV*	0.105 ** ( 0.033 )	0.102 * ( 0.055 )	0.087 ** ( 0.026 )	0.105 * ( 0.057 )	0.052 ( 0.064 )
ILTV*	--	--	--	-0.005 ( 0.028 )	--
LTV* x HETER	--	--	--	--	0.455 ( 0.379 )
HETER	--	--	--	--	0.123 ** ( 0.022 )
T*	0.8	0.8	0.6	0.8	0.8
Controls for Tenure	Yes	Yes	Yes	Yes	Yes
Quarter of Sale Dummies	Yes	Yes	Yes	Yes	Yes
Number of Obs.	12,248	12,248	12,248	12,248	12,248

\*\* Indicates significance at the 5 percent level.

\* Indicates significance at the 10 percent level.

Standard Errors are calculated by bootstrapping to account for the generated regressors in the second stage.

LOSS is the greater of the difference between the previous selling price and the estimated value in the quarter of sale, and zero. LTV is the greater of the difference between the ratio of outstanding loan to estimated value in the quarter of sale and T\*, and zero. ILTV is the greater of the difference between the ratio of initial loan to previous selling price and T\*, and zero. HETER is the standard deviation of the year built for all homes that sold within a .5 mile radius of house *i* within 4 years of quarter *t*.

FIGURE 2.5: Seller Characteristics and Selling Prices. Fixed Rate Mortgages.

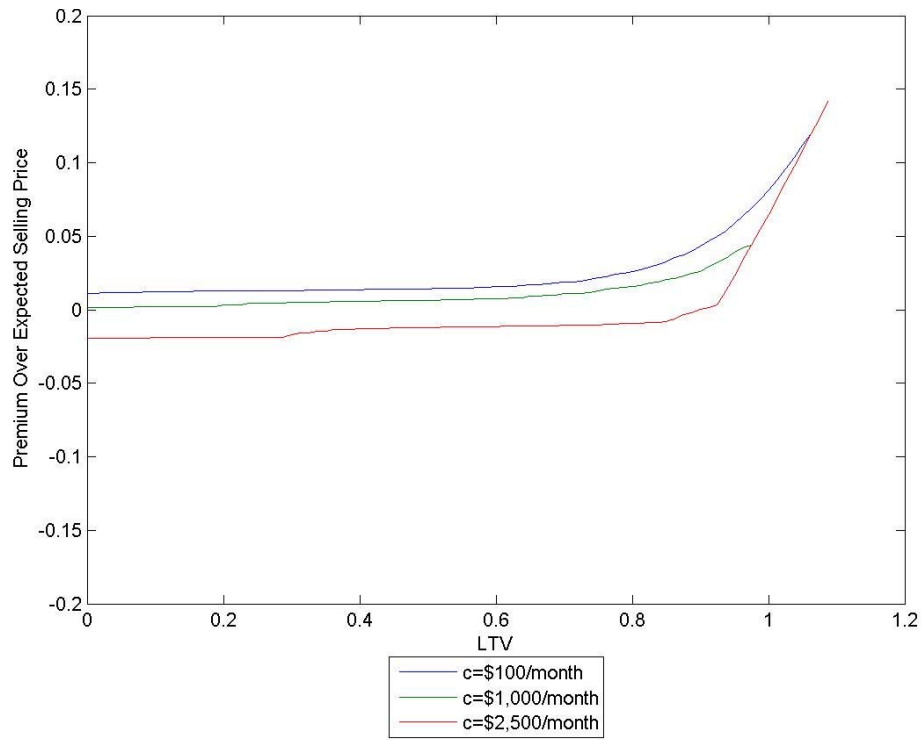


FIGURE 2.6: Simulation of Average Price When an Offer is Accepted,  $T=20$

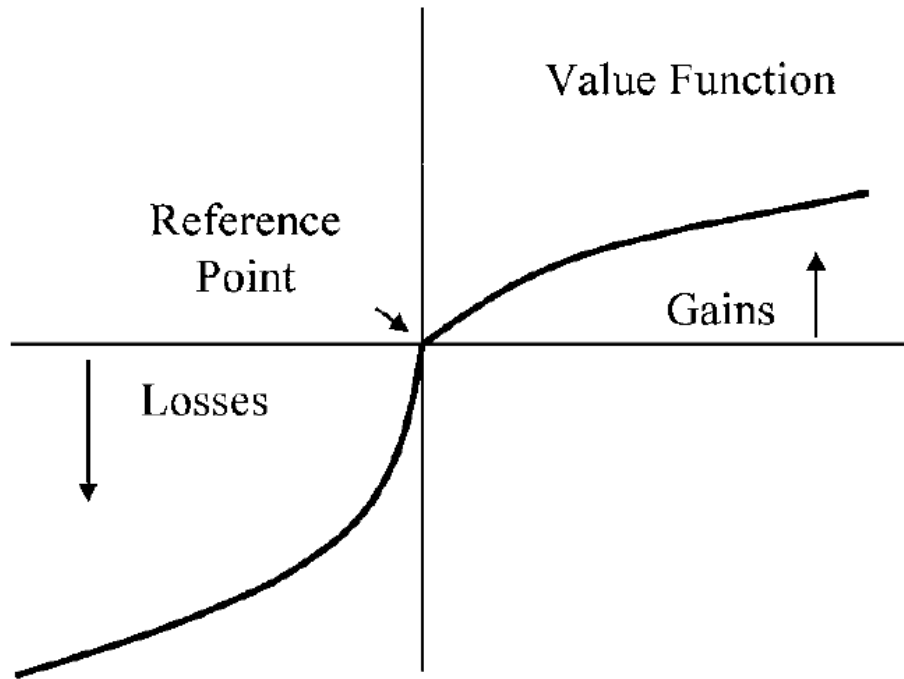


FIGURE 2.7: Prospect Theory

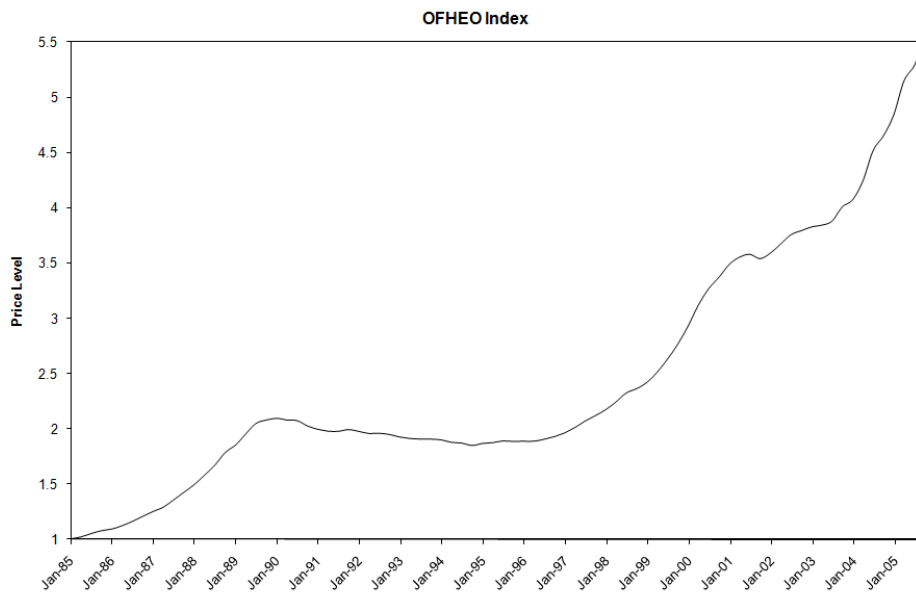


FIGURE 2.8: Nominal Price Levels in San Francisco MSA 1985-2005

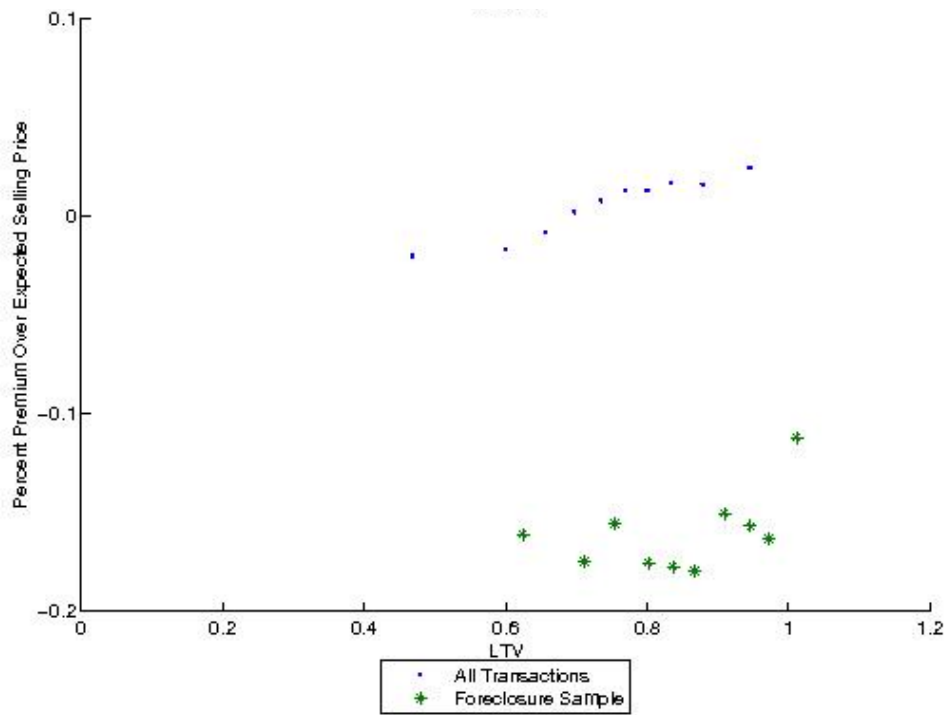


FIGURE 2.9: Premium Over Expected Selling Price by Transaction Type



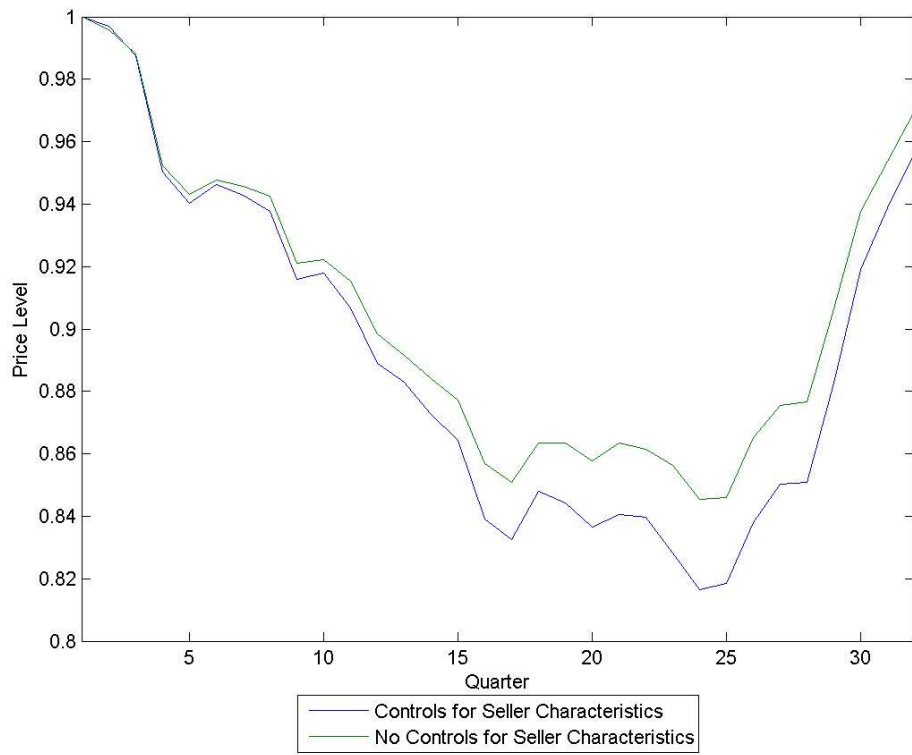


FIGURE 2.10: Nominal Price Levels 1990 - 1997

## Uncertainty, Learning and the Value of Information in the Residential Real Estate Market

The housing market is a classic example of a thin market. Even though there may be a large stock of homes in a given neighborhood, the number of houses that share a particular physical characteristic or amenity in that location can be small. This large degree of differentiation, combined with the low turnover rate of homes and the volatility in prices over time, makes it difficult for sellers to assess home values based on comparable sales. This uncertainty may be especially severe in the housing market relative to other thin markets because it is more difficult to survey demand given that buyers must visit a home to evaluate its characteristics. In addition, sellers tend to be inexperienced because home owners typically do not sell many homes over the course of their lifetime.

Given the potential for uncertainty in the housing market, the primary goal of this paper is to estimate how uncertainty affects seller welfare and seller behavior. Understanding these relationships is important for two main reasons. First, it provides an estimate of the value of information in the housing market. This is becoming

increasingly important as rich micro housing datasets are becoming available<sup>1</sup> and firms (e.g. zillow.com) want to know whether there is demand for more sophisticated home appraisal techniques. Estimating the cost of less information contributes to the literature on the value of real estate agents (Hendel, Nevo, and Ortalo-Magne (2009); Bernheim and Meer (2008)). Despite high commission rates, these existing studies present convincing evidence that sellers who select to sell by owner would have had similar selling outcomes if they had used a realtor. This begs the question: why are sellers willing to pay such high commissions, which are typically 3 percent of the sales price? Are they paying a high price for convenience, as Hendel, Nevo, and Ortalo-Magne (2009) suggest? Suppose that compared to sellers that sell by owner, the types of sellers who select realtors have less access to information or are less able to process information on their own. Then, if the value of information is large, the realtor's ability to provide and process information is a valuable service to most sellers.<sup>2</sup>

Secondly, since previous empirical housing models abstract from seller uncertainty, incorporating it may help explain some previously unexplained stylized facts about the housing market, including the predictability of short-run price appreciation rates that has been well-documented in the literature.<sup>3</sup> This is an important contribution because housing data are used in a variety of important economic settings. For example, movements in the widely reported Case-Shiller home price index affect stock prices, banks and mortgage providers use sales data in determining whether to extend credit, and researchers, especially in public economics, use home price data to value public goods and amenities.

---

<sup>1</sup> An antitrust suit against the National Association of Realtors (NAR), which settled in 2008, made it easier to collect data on home listings from the internet.

<sup>2</sup> 84 percent of sellers use realtors according to the NAR (2007).

<sup>3</sup> Lewis (2009) and Yang and Ye (2008) also use models with search, uncertainty, and learning to explain equilibrium price dynamics in other product markets, although their focus is on asymmetric price adjustments (also known as rockets and feathers).

The framework I develop for investigating these ideas combines a dynamic model of the home selling problem with a rich micro housing dataset on home listings and home transactions. In the model, sellers face a time-varying distribution of buyer valuations for their home, and their objective is to maximize the selling price less the holding costs of keeping the home on the market. As in the existing empirical search models of the real estate market (Horowitz (1992), Carrillo (2010)), I endogenize the seller’s decision over what list price to set and when to sell the house. The first key contribution of my model is that I allow sellers to withdraw their homes from the market without sale. This is critical because close to 50 percent of homes listed for sale are eventually withdrawn in my sample, and excluding this segment of the market introduces sample selection problems. The second main contribution of my model is that I allow for seller uncertainty about the distribution of buyer valuations. This is modeled through a prior on the mean of this distribution. Over the course of the selling horizon, buyer behavior provides (noisy) information about demand, and sellers update their priors using Bayes’ rule as in Lazear (1986).<sup>4</sup> This learning process introduces duration dependence into the model, which is necessary to explain why list prices typically decline over the selling horizon and why the time on market (TOM) distribution is not geometric.<sup>5</sup> Since I do not observe rich enough data on the buyer side to also identify the parameters of a dynamic model of buyer behavior, I present a simple, static model of buyer behavior where the seller influences the buyer’s search problem through the choice of a list price. Nonetheless, the full model is rich enough to capture the key features of the home selling process including sequential search, a posting price mechanism, preference heterogeneity, heterogeneity

---

<sup>4</sup> My paper contributes a new application to the growing literature on empirical learning models. See, for example, Akerberg (2003), Crawford and Shum (2005), Goettler and Clay (2009), Narayanan, Chintagunta, and Miravete (2007), Hitsch (2006), Erdem and Keane (1996).

<sup>5</sup> Seller optimal strategies are time invariant in Horowitz (1992) and Carrillo (2010), presumably because they lack data on list price changes to identify a richer model.

in the motivation to sell, and observed and unobserved heterogeneity in the stock of homes.

I estimate the model using dynamic programming techniques and simulated method of moments to recover the structural parameters. I match the predictions of the model to a dataset on single-family homes listed for sale on the Multiple Listing Service (MLS) in Los Angeles from 2006-2009. I merge the listing data with a transaction dataset to get detailed information on weekly list prices, sales prices (if the home sells), and TOM. I am not aware of other studies that have access to such a large dataset representing the full inventory and pricing histories of homes on the market.

At the estimated parameters, the model can match many of the key features of the housing data, including declining list prices over the selling horizon; and lower list prices, larger list price changes, and shorter TOM for homes that are sold versus homes that are withdrawn. I also find that a high withdrawal rate is perfectly compatible with rational behavior. The estimated amount of uncertainty is high enough and the holding costs of keeping a home on the market are low enough that a pool of risk-neutral, 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.<sup>6</sup>

Since the sellers in my sample use real estate agents, the estimated parameters reflect the beliefs of sellers after incorporating advice from realtors. I find that these beliefs are relatively precise, and the precision is increasing in a measure of the similarity of surrounding homes. The standard deviation of buyer preference heterogeneity is about three times the size of the standard deviation of the seller's prior. As a result, learning occurs relatively slowly because it is difficult for sellers to dif-

---

<sup>6</sup> Hitsch (2006) finds a similar result among potential entrants with uncertain product quality in the ready-to-eat breakfast cereal industry.

ferentiate between, for example, high average demand and a high idiosyncratic taste for their house. Despite their shorter TOM, motivated sellers learn more than unmotivated sellers because their aggressive pricing attracts more buyers, who provide more information.

One of the most interesting findings to emerge is that sellers who put their homes on the market after a period where prices declined by 1 percent tend to overstate the value of their home by .65 percent.<sup>7</sup> This method of price discovery is consistent with anecdotal evidence that sellers look to previous sales of similar houses when pricing their homes. Evidently, sellers (or their realtors) do not fully adjust their comparable sales analysis for the downward trend in prices during my sample period. My estimate is identified by comparing average list and sales prices for sellers entering the market under varying previous market conditions, and the identification is robust to unobserved heterogeneity. The identification strategy for expectation bias could be useful in other settings where list prices or reserve prices are observed in addition to selling prices.

I use simulations of the model to show that biases and uncertainty have large welfare consequences for sellers, and the welfare consequences are larger when the market is more volatile. Sellers are \$22,482 better off (5.06 percent of the expected sales price) on average under the beliefs estimated from the model relative to beliefs that would arise if sellers did not use a realtor and simply formed price expectations based on the FHFA city-wide house price index that is publicly available. This comparison suggests that realtors provide a lot of value to less sophisticated sellers through appraisals alone.

Although I find that sellers with realtors have better information relative to what zillow.com and aggregated price indexes can provide, there is the potential for large

---

<sup>7</sup> Kuzmenko (2010) finds a similar result when comparing owner estimates of home value from the Census with econometric estimates of home value.

welfare gains from more accurate expectations. The average seller in my sample would be \$8,658 (1.95 percent of the expected sales price) better off with complete information. There is heterogeneity across types of sellers in the willingness to pay for information, and over 25 percent of sellers would be willing to pay over \$12,000 for full information. These results suggest that there is a market for firms that can offer more sophisticated home appraisal techniques. At levels of uncertainty that are slightly lower than the estimated level of uncertainty, simulations show that suboptimal behavior by sellers does not have a devastating effect on seller welfare. This result helps to explain the emergence and appeal of discount realtors that provide convenience, but offer less advice and strategy.

I use simulations of the model to show that uncertainty and learning is one explanation for the persistence of sales price appreciation rates in the short-run (annually). Previous research has consistently documented the predictability of price appreciation rates, and this stylized fact has been a challenge for researchers to explain within an efficient markets framework.<sup>8</sup> When sellers have rational expectations, simulations of the model show that annual sales price appreciation rates persist even when the market fundamentals follow a random walk (the AR(1) coefficient is 0.24) because reservation prices slowly adjust to market shocks as learning occurs. The persistence is even stronger at shorter frequencies. Of course, we cannot rule out irrationality or inefficiency as additional sources for the observed predictability. However, we can show that a significant amount of predictability can arise from a purely rational model of the home selling problem due to lack of information.

This paper proceeds as follows. I begin by introducing the dataset that I assemble because it motivates the modeling assumptions. Sections 3, 4, and 5 introduce the structural model, and the identification and estimation issues that arise. Section 6

---

<sup>8</sup> See, for example, Case and Shiller (1989) and Glaeser and Gyourko (2007). Cho (1996) provides a survey of the literature on house price dynamics.

presents the parameter estimates and discusses model fit. Section 7 uses simulations of the model to generate the predictions about price dynamics and seller welfare described above. Section 8 discusses identification of expectation bias in detail, and Section 9 concludes.

### 3.1 Data

The data come from two sources. The listing data come from Altos Research, which provides information on single-family homes first listed for sale on the MLS in the Los Angeles metropolitan area from October 2006 - February 2009. For each property, I observe the address, the date when the property is listed for sale, the date when the property is taken off the market, the list price each week, and some characteristics of the house (e.g. square feet, lot size, etc.). Most existing studies neither have access to weekly variation in list price nor the listing history of all homes on the market (i.e. both sales and withdrawals). I exploit this new source of variation to identify a more detailed model of seller behavior, as described below.

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.<sup>9</sup> According to the NAR, agent-assisted sales accounted for 84 percent of all home sales in 2007.<sup>10</sup>

Since the MLS data does not provide information about whether a particular property actually sells, I supplement the MLS data with a transactions dataset that contains information about the universe of housing transactions in the LA metro area from 1988-2009. In this dataset, the variables that are central to this analysis

---

<sup>9</sup> My dataset does not report the real estate company that each seller hires. However, the services that realtors provide and the fees that realtors charge are fairly homogeneous (Hsieh and Moretti (2003)). An exception is flat-fee realtors, who provide access to the MLS for sellers selling by owner. However, Jia and Pathak (2010), who observe information about realtors and commissions, find that the market share of flat-fee contracts or other non-percentage commissions in their MLS dataset for Boston is less than 4 percent.

<sup>10</sup> Of the percentage of homes sold without a realtor, 40 percent involved sales to buyers who were known to the seller prior to the transaction.



are the address of the property, the date of the transaction, and the sales price.

Using the address, I merge the listing data with the transaction data. The failure of a listing in the MLS data to merge with an entry in the transaction dataset does not necessarily mean that the property was not sold. Differences in how an address is coded in the two datasets can be responsible for a failed merge. Even after making a considerable effort to standardize common abbreviations and eliminate inconsistencies by hand and visual inspection of the data, the resulting dataset after the initial merge overstates the proportion of withdrawals. To fix this, I restrict the MLS dataset to addresses that merge to a previous sale in the transaction dataset.<sup>11</sup> The failure to merge here is because the addresses are inconsistent, the house is new, the current owner purchased the house prior to 1988, or the listing data cover some neighborhoods that are not included in the transaction data. However, since the addresses in the transaction data are identical for every sale of a given property, I can be confident that a listing that merges with a previous sale but does not merge with a recent sale is a withdrawal.

Some listings correspond to the same property being listed more than once in a short amount of time.<sup>12</sup> I consider a listing as new only if there was at least a 180 day window since the address last appeared in the listing data. If the window is less than 180 days, I assume the property remained on the market during that interval at a list price equal to the list price in the final week before the gap begins.<sup>13</sup> To avoid confusion during the merge that can arise from multiple sales occurring close together, I also drop listings that match to more than 1 transaction during the years

---

<sup>11</sup> This cut of the data drops approximately 40 percent of the listings in the MLS data. I compare summary statistics of the limited sample to the full sample to ensure that my sample is representative.

<sup>12</sup> Some common reasons for relistings include price reductions, sellers changing agents, or to try to game the TOM measure.

<sup>13</sup> Levitt and Syverson (2008) make the same assumption. 3.65 percent of listings involve a property exiting the market and then returning after more than 180 days but less than 365 days

2007-2009 that are less than 1.5 years apart ( $< 1$  percent of listings).

I adjust all sales prices within \$ 500 of the list price to equal the list price. The final restrictions remove extreme/miscoded prices: I 1) exclude listings where the ratio of the minimum list price to the maximum list price is less than the first percentile of the distribution; 2) exclude listings where the final list price is greater than \$3,000,000 or less than \$50,000; 3) exclude listings where the difference between the ratio of the final list price to the expected price (to be defined below) is less than the 1st or greater than the 99th percentile of the distribution; and 4) exclude listings where the ratio of the sales price to the final list price is greater than 1.15.

### *3.1.1 Expected Prices*

For each house in the final sample of listings, I construct two expected log prices that I will refer to throughout the paper:

$$\begin{aligned}\hat{p}_{it} &: \text{log expected sales price for house } i \text{ in week } t \\ \hat{p}_{it}^L &: \text{log expected list price if the house is put on the market in week } t\end{aligned}$$

Appendix A.1 describes how I calculate these prices from the data. In short, they are calculated by applying a neighborhood level of list/sales price appreciation to the previous log list/sales price as in Shiller (1991).

### *3.1.2 Summary Statistics*

Figure 1 shows the Case-Shiller home price index adjusted for inflation in Los Angeles from 1987 - 2009 on a log scale. During the years where the MLS and transactions data overlap, prices fell significantly.<sup>14</sup> From the peak in July 2006 to the trough in May 2009, real house prices fell 45 percent. While this time period was unique in many ways, the severity of the price changes is not unprecedented. Figure 1 shows

<sup>14</sup> I do not have access to the rich MLS data that I use here, which contains list price changes and all homes listed for sale, before 2006.

that the Los Angeles housing market has historically been quite volatile, and during the housing bust of the early 90's, real house prices fell 40 percent from peak to trough. Many other metro areas exhibit similar boom and bust cycles. I also find evidence that the percent of transactions that are foreclosures in Los Angeles during the recent downturn is comparable to the downturn during the 1990's.<sup>15</sup> Thus, I argue that the main results of this paper are not only informative about selling behavior during this recent housing crisis; they are also likely to provide a useful characterization of selling behavior during cold markets in general.

Table 1 shows summary statistics of the final sample for Los Angeles, which is considerably larger than those of previous studies. One of the most striking statistics from Table 1 is that 55 percent of the properties are withdrawn without sale. Genesove and Mayer (1997) find a similar result, and their sample also covers a period of declining prices. In a sample of Chicago home sales during a period of rising prices, Levitt and Syverson (2008) find that 22 percent of properties are withdrawn. Thus, one stylized fact of the housing market appears to be that homes are more likely to be withdrawn without sale during cold markets. Properties that are withdrawn tend to have higher list prices, longer TOM, and smaller list price changes even though they are on the market for longer. All of the differences in means between the variables summarized in Table 1 are statistically significant. Most properties experience at least one list price change before they are withdrawn or sold, and conditional on having a list price change, the average change is about -4 percent. However, about 8 percent of list price changes are increases. Table 2 shows that list price changes occur throughout the selling horizon, and many occur early in the selling horizon. Since some sellers will quickly adjust their beliefs in response to new information,

---

<sup>15</sup> This calculation is imperfect since foreclosures are identified by parsing the seller's name for revealing words such as "Bank". However, Campbell, Giglio, and Pathak (2009) also report that the foreclosure rate is not unusually high during the recent recession relative to the downturn during the 1990's in Massachusetts.

the learning model that I present below will also predict changing list prices in the first few weeks and some list prices changes that are increases. Knight (2002) and Endgelberg and Parsons (2010) present additional reduced-form evidence that sellers learn over the selling horizon in the spirit of Lazear (1986).

## 3.2 A Dynamic Model of the Home Selling Problem

The objective of the seller is to maximize the sales price less the holding costs of keeping the house on the market for sale. However, sellers always have the option to withdraw the home from the market and receive a terminal utility,  $v^w$ , to be specified below. Thus, I model the seller's problem as an optimal stopping decision.

### 3.2.1 Offer Process and Buyer Behavior

Before presenting the core of the model, I present the simple, static model of buyer behavior that will make it necessary for the seller to post a list/asking price in equilibrium. To preview the results of Theorem 1, I find that high list prices discourage buyers from visiting homes, but high list prices also result in higher sales prices conditional on a buyer arriving. Therefore, in selecting the list price, the seller faces a trade-off between TOM and the sale price. This is consistent with my empirical evidence as well as with results from other empirical and theoretical studies.<sup>16</sup>

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

---

<sup>16</sup> For empirical work, see Glower, Haurin, and Hendershott (1998) and Merlo and Ortalo-Magne (2004). Chen and Rosenthal (1996) have a much more general, theoretical model of the home selling problem, and they find that list prices serve as commitment devices in equilibrium to induce buyers to undergo costly inspections.

$$v_{ijt} = \mu_{it} + \eta_{ijt}. \quad (3.1)$$

$\mu_{it}$  is common across all buyers, whereas  $\eta_{ijt}$  represents buyer taste heterogeneity. Of course,  $\mu_{it}$  will vary by  $i$  because, for example, buyers will be willing to pay more for larger houses. I also allow  $\mu_{it}$  to vary over time. More details about  $\mu$  and the transition of  $\mu$  will be discussed in Section 4; for now, note that the level of  $\mu$  and the  $\mu$  process are exogenous. I assume that

$$\eta_{ijt} \sim N(0, \sigma_\eta^2) \quad (3.2)$$

and is *iid* over buyers, houses and time.

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

$$\text{buyer beliefs: } v_{ijt} \sim N(\hat{v}_{ijt}, \sigma_{\hat{v}}^2) \quad (3.3)$$

where  $\hat{v}_{ijt}$  is drawn from  $N(v_{ijt}, \sigma_v^2)$ . Thus, buyers get an unbiased signal of their true valuation from the advertisement.<sup>17</sup>

After observing  $\hat{v}_{ijt}$ , the buyer decides whether or not to inspect the house at some cost,  $\kappa$ . 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 we 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.<sup>18</sup> If the buyer chooses not to

<sup>17</sup> This specification of beliefs would arise if buyers had flat priors (i.e. prior variance =  $\infty$ ) and processed the signal,  $\hat{v}$  according to Bayes' rule. See Hitsch (2006) and Narayanan, Chintagunta, and Miravete (2007) for a similar specification of initial priors.

<sup>18</sup> Assuming that the seller must sell if  $v > p^L$  should not affect the results. At the estimated parameters, simulations show that sellers want to sell in over 99 percent of the cases when  $v > p^L$ . In some cases, real estate agent contracts explicitly require sellers to accept offers above the list price. Albrecht, Gautier, and Vroman (2010) note that realtor concerns about reputation can also lead to some limited commitment.

inspect or if the buyer's valuation lies below the seller's reservation price, then the buyer departs forever and the seller chooses to either withdraw the home from the market or to move onto the next period with her house for sale.

This setup can account for the significant percentage of homes that sell for the list price. However, it does not allow for prices above the list price, which can occur if competition between buyers drives up the price. I accommodate prices above the list price with a reduced form. In particular, I assume that when  $v > p^L$ , the price gets driven up to  $v$  with exogenous probability  $\lambda$ . This reduced form preserves the incentives of sellers, whose welfare and behavior is the focus of this paper: the optimal list price and reservation price will reflect the possibility that offers can be above the list price, and that there is a stochastic component to when this occurs.<sup>19</sup>

To summarize, conditional on a buyer arriving with a valuation that exceeds the seller's reservation price,  $p^R$  to be defined below, a transaction will occur with log price equal to

$$p^* = \begin{pmatrix} v & \text{if } v \leq p^L \text{ and } v \geq p^R \\ p^L & \text{if } v \geq p^L \text{ and } \Lambda = 0 \\ v & \text{if } v \geq p^L \text{ and } \Lambda = 1 \end{pmatrix} \quad (3.4)$$

where  $\Lambda$  is a Bernoulli random variable with parameter  $\lambda$ .

The proof of the following theorem, which characterizes the buyer's optimal behavior, appears in Appendix A.2.

**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  $(\kappa, \sigma_{\hat{v}}^2, \lambda)$ .*

---

<sup>19</sup> A more general model of multiple bidders significantly complicates the computation of the model because the buyer's optimal behavior would depend on their beliefs about the population distribution of  $v$ . In addition, it is not clear how these beliefs would be identified without more data on buyer behavior. This is left for future research.

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. Theorem 1 shows that under this model of buyer behavior, there is a closed form relationship between  $\bar{v}$  and the list price. As I show below, this is important to keep estimation tractable because the list price will be endogenous.

### 3.2.2 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,  $\mu_{it}$ . In the week when the home is first put on the market, the initial prior is given by

$$\text{initial prior: } \mu_{it} \sim N(\hat{\mu}_{it_0}, \hat{\sigma}_{it_0}^2) \quad (3.5)$$

and  $\hat{\mu}_{it_0}$  is drawn from  $N(\mu_{it} + \alpha\Delta_6, \hat{\sigma}_{it_0}^2)$ , where

$$\Delta_6 \equiv \frac{\sum_{k=1}^{24} \hat{p}_{it_0-k}}{24} - \hat{p}_{it_0} \quad (3.6)$$

and  $t_0$  denotes the week of initial listing.  $\Delta_6$  simply measures the percentage change in prices between the average expected price in the 6 months before the seller puts the home on the market and the expected price in the week that the seller lists the home for sale.<sup>20</sup>  $\alpha$  is a parameter to be estimated. If  $\alpha \neq 0$ , then falling price levels prior to the time of listing will bias seller beliefs.

Sellers may have upward-biased beliefs when prices are falling because the thinness in the housing market makes it difficult or costly for the seller to acquire information about the most up to date market conditions. Sellers (or their realtors) may find it optimal to use recent sales prices as a proxy for current market conditions. This theory is supported by conversations with realtors.<sup>21</sup>

<sup>20</sup> The mean and standard deviation of  $\Delta_6$  in the sample are 5.6 percent and 5.1, respectively.

<sup>21</sup> Bias during a particular time period does not necessarily imply irrationality or suboptimality.

I allow the standard deviation of the prior to depend on how similar seller  $i$ 's house is relative to nearby houses. I parameterize this as

$$\hat{\sigma}_{it_0} = h_0 + h_1 * \log(1 + S_i) \quad (3.7)$$

where  $S$  is the standard deviation of the year built for all homes that sold within a .5 mile radius of house  $i$  in the 6 years prior to  $t$ .<sup>22</sup> We should expect  $h_1 > 0$  since the ability to observe prices of close substitutes should reduce uncertainty.<sup>23</sup>

During the selling horizon, market conditions ( $\mu_{it}$ ) can change. This is not standard in the existing literature on empirical learning models where the variable that the agent is learning about typically remains constant over time. To accommodate this extra layer of complexity, I assume that all sellers, regardless of their neighborhood or the particular time during my sample period, share the expectations that

$$\text{seller beliefs: } \mu_{it} - \mu_{it-1} \sim N(\hat{\mu}_p, \hat{\sigma}_p^2). \quad (3.8)$$

I now discuss how agents learn both about transitions in  $\mu$  and the level of  $\mu$ .

### 3.2.3 Learning

Sellers process all information optimally using Bayes' rule. At the beginning of week  $t$ , before the list price is set and before the buyer receives their signal about their valuation, I assume that the seller observes an exogenous signal,  $z_{it}$  that is

$$z_{it} \sim N(\mu_{it} - \mu_{it-1}, \sigma_z^2). \quad (3.9)$$

---

Whether sellers acquire the optimal amount of information is beyond the scope of this paper since we do not estimate or observe the cost of acquiring information.

<sup>22</sup> I choose year built because it is exogenous and it does not vary over time. The latter is important because home characteristics are not updated in the transaction data to account for renovations. Nonetheless, the results when I instead use square feet are comparable to the results presented below.

<sup>23</sup> Levitt and Syverson (2008) find that real estate agents representing sellers on more homogeneous blocks have less of an information advantage, presumably because sellers learn a lot from nearby sales.



This signal about demand shifts could come from research by the realtor or from observing neighbor behavior.  $\sigma_z^2$  should be greater than zero due to the thinness of the market and the fact that closing dates (when sales price data become public) lag agreement dates.

Before continuing, it is useful to define the following means and variances of seller beliefs over  $\mu_{it}$ :

$\hat{\mu}_{it}^{pre}, \hat{\sigma}_{it}^{pre}$  : 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<sup>24</sup> where

$$\begin{aligned}\hat{\mu}_{it}^{pre} &= \hat{\mu}_{it-1} + \frac{\sigma_z^2 \hat{\mu}_p + \hat{\sigma}_p^2 z_{it}}{\sigma_z^2 + \hat{\sigma}_p^2} \\ \hat{\sigma}_{it}^{pre} &= \hat{\sigma}_{it-1}^2 + \frac{\sigma_z^2 \hat{\sigma}_p^2}{\sigma_z^2 + \hat{\sigma}_p^2}.\end{aligned}\tag{3.10}$$

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  $\mu_{it}$ . The posterior distribution of  $\mu$  after the seller processes the information in  $v_{it}$  remains normal with mean and variance at time  $t$  given respectively by:

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

---

<sup>24</sup> See DeGroot (1970), and Ljungqvist and Sargent (2004) for a presentation of the updating formulas when the hidden state transitions over time.

$$\hat{\sigma}_{it}^2 = \frac{\hat{\sigma}_{it}^{pre} \sigma_{\eta}^2}{\hat{\sigma}_{it}^{pre} + \sigma_{\eta}^2}. \quad (3.11)$$

The initial conditions are given in equation (3.5).

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

$$f(\hat{\mu}_t | \hat{v} < T^* + p^L) = \frac{\Phi\left(\frac{T^* + p^L - \hat{\mu}_t}{\sqrt{\sigma_{\eta}^2 + \sigma_{\hat{v}}^2}}\right) \phi\left(\frac{\hat{\mu}_t - \hat{\mu}_t^{pre}}{\hat{\sigma}_t^{pre}}\right) \frac{1}{\hat{\sigma}_t^{pre}}}{\Phi\left(\frac{T^* + p^L - \hat{\mu}_t^{pre}}{\sqrt{\hat{\sigma}_t^{pre} + \sigma_{\eta}^2 + \sigma_{\hat{v}}^2}}\right)}. \quad (3.12)$$

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, Gurler, and Levine (2007) shows that a normal distribution with mean and variance equal to the mean and variance of the distribution in equation (3.12) 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:

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

$$\begin{aligned} \hat{\sigma}_{it}^2 = & \frac{1}{\tau^2} ((\hat{\mu}_{it}^{pre})^2 \sigma^4 + \tau \hat{\sigma}_{it}^{pre} \sigma^2 + 2 \hat{\sigma}_{it}^{pre} \sigma^2 (\hat{\mu}_{it}^{pre})^2 + (\hat{\sigma}_{it}^{pre})^2 \tau + (\hat{\mu}_{it}^{pre})^2 (\hat{\sigma}_{it}^{pre})^2) \\ & + (2 \hat{\mu}_{it}^{pre} \hat{\sigma}_{it}^{pre} \sigma^2 + (\hat{\sigma}_{it}^{pre})^2 (T^* + p^L + \hat{\mu}_{it}^{pre})) - h(T^* + p^L) / \tau - (\hat{\mu}_{it})^2 \end{aligned} \quad (3.13)$$

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  $\tau$ .

Since there is less information in the signal that the seller receives when a buyer does not arrive relative to when a buyer does arrive, for a given prior variance, the posterior variance is relatively higher if a buyer does not arrive. This can be shown by manipulating the variance expressions in equations (3.11) and (3.13).

### 3.2.4 Seller's Optimization Problem

The timing of the model is summarized in Figure 2. Each period begins with the realization of  $z$ . The seller processes this signal as described above, and then chooses an optimal list price (a continuous variable). The list price is set to balance the tradeoffs that emerge from Theorem 1. Once the list price is advertised, the buyer moves, the seller processes 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 sales price, move onto the next period with her house for sale, or withdraw the home from the market and receive a terminal utility  $v^w$ . Each period that the home is on the market, the seller incurs a time-invariant holding cost,  $c$ . Without some cost of keeping the home on the market, the model cannot rationalize any withdrawals. These costs include keeping the home presentable and showing the house to prospective buyers. I assume a finite-horizon ( $T = 80$  weeks) for the selling problem.

The following Bellman's equation, which characterizes selling behavior at the third hash mark on the timeline in Figure 2, summarizes the seller's optimization problem:

$$\begin{aligned}
V_t(\Omega_t) = \max_{p_t^L} & \\
& \int_z (\max \{ (1 - \beta)v^w + \beta(c + V_{t+1}(\Omega_{t+1}|\hat{v}_t < T^* + p^L, z_{t+1})), v^w \} Pr(\hat{v}_t < T^* + p^L) \\
& \quad + (\int_{-\infty}^{p^L} \max \{ v_t, (1 - \beta)v^w + \beta(c + V_{t+1}(\Omega_{t+1}|v_t, z_{t+1})), v^w \} \\
& \quad + \int_{p^L}^{\infty} ((1 - \lambda)p_t^L + \lambda v_t))g(v_t|\hat{v}_t > T^* + p^L)dv_t Pr(\hat{v}_t > T^* + p^L))f(z_{t+1})dz_{t+1}
\end{aligned} \tag{3.14}$$

where  $\beta$  is the weekly discount factor and  $\Omega_t$  denotes states that vary across sellers or time  $(\hat{\mu}_t^{pre}, \hat{\sigma}_t^{pre}, v^w)$ . Without the normality assumptions made throughout the model, the state space expands beyond a single mean and variance, and computational resources quickly become binding. This is why the result by Berk et al. is so useful. It allows me to introduce a second type of learning, which is new to the empirical learning literature in economics, without compromising the manageable dimension of the state space. While not shown explicitly in the Bellman's equation, I do take out a 6 percent realtor commission, which is typically 3 percent to the seller agent and 3 percent to the buyer agent, from the sales price.

I parameterize the terminal utility from withdrawing as

$$v_{it}^w = w_i + \mu_{it_0} \quad (3.15)$$

where  $w$  is time-invariant, *iid* over potential sellers, and  $N(\mu_w, \sigma_w^2)$ . I allow  $\mu_w$  to take on one of two values, low(L) or high(H), with probabilities  $\gamma$  and  $1 - \gamma$ , respectively. I assume that the seller knows  $v^w$ , but is unaware of the decomposition in (3.15).

Note from equation (3.14) that when sellers opt to stay on the market, they receive a flow utility equal to  $(1 - \beta)v^w$  in addition to  $c$ . Thus, the cost of being on the market is relatively higher for sellers who receive low utility from living in their existing homes. In this way, the model endogenously generates the prediction that more motivated sellers will have shorter TOM on average. This is consistent with Glower et al. (1998), who find evidence that sellers with high motivations to sell due to, for example, a job change have shorter TOM.

When the seller decides to stay on the market for an additional week, the seller updates  $\Omega_t$  to  $\Omega_{t+1}$  using equation (3.10), and using equations (3.11) and (3.13) depending on whether  $\hat{v} < T^* + p^L$  or  $\hat{v} > T^* + p^L$ , respectively.

### 3.2.5 Comparative Statics

The value function and the optimal list price are non-decreasing in  $\hat{\mu}$ . The value function and the optimal list price are also non-decreasing in  $v^w$ . Since the likelihood of withdrawing is also increasing in  $v^w$ , the model predicts that sellers who ultimately withdraw from the market set higher list prices on average. Table 1 shows that this is true empirically as well.

I cannot formally derive how  $\hat{\sigma}_{it}$  affects the list price policy function and the value function. However, for numerous simulations at reasonable parameter values, I find that the list price is non-decreasing in  $\hat{\sigma}_{it}$  for all values of  $v^w$  and the relationship between the value function and  $\hat{\sigma}_{it}$  is non-monotonic and depends on  $v^w$ . Figure 4 is typical of the relationship between  $p^L$  and  $\hat{\sigma}_{it}$ . More uncertainty raises the log list price (on the vertical axis) 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. Thus, the model can generate declining list prices during the selling horizon even if market conditions (i.e. the  $\mu$  process) process remain constant. This is consistent with anecdotal evidence as well as with Merlo and Ortalo-Magne (2004), who find that even in booming markets, most list price changes tend to be declines.

For sufficiently low levels of  $v^w$ , the value function is decreasing in  $\hat{\sigma}_{it}$ ; for higher values of  $v^w$ , the value function is increasing in  $\hat{\sigma}_{it}$ . When  $v^w$  is high, uncertainty increases the value function for the same reason that a mean preserving spread increases expected utility in the classic labor search models (see McCall (1970)). The seller targets the list and reservation prices for the high demand state, while fully expecting to withdraw if demand is low. The positive effects of more variance evidently outweigh the negative effects of less information. When  $v^w$  is sufficiently low and the seller needs to move, the seller does not have an outside option to hedge against low

demand. In this case, the distortion of the list and reservation price relative to the full information case outweighs the benefits of move variance.

This feature of the model can explain why uncertainty leads to a large proportion of withdrawals. More uncertainty encourages high  $v^w$  sellers to test the market, which increases the average  $v^w$  in the pool of sellers in the market.

### 3.3 Estimation

Table 3 summarizes the notation of all the parameters to be estimated. I estimate the parameters using simulated minimum distance estimation. The objective function that I minimize is of the form

$$[m - m_S(\Theta)]'W[m - m_S(\Theta)] \tag{3.16}$$

where  $m$  denotes a vector of moments from the data,  $m_S$  are the simulated counterparts generated by the model (which depend on the parameter vector), and  $W$  is a weighting matrix.<sup>25</sup> The moments and the corresponding weights used in the estimation are listed in Table 4, and are based on the discussion of identification that follows. In practice, there are several computational issues that arise. I discuss each in turn.

It is well known that in these types of dynamic discrete choice problems,  $V$  from equation (3.14) needs to be calculated for each point in the state space for every trial parameter vector. 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. In practice, I use 1300 points for each time period ( $T=80$ ), which results in 104,000 evaluations of  $V$  per parameter vector.

---

<sup>25</sup> I use the 2-step procedure described in Lee and Wolpin (2010) to calculate the weighting matrix, which is assumed to be diagonal. The weighting matrix adjusts for the scale of the moments, the precision of each of the estimated moments from a first stage, and the number of observations that comprise each moment.

A second issue that typically arises relates to the calculation of the integrals in equation (3.14). Simulation methods preserve consistency if the number of simulation draws rises with the sample size.<sup>26</sup> However, since the value function typically needs to be calculated at a large number of points, a large number of simulation draws is often not computationally feasible. I avoid these issues altogether as the term inside the max operator in equation (3.14) has a closed form. 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 (3.10) and (3.11), and linearity in the interpolating function.

The optimal list price, however, does not have a closed form. This raises computational demands in two areas of the estimation routine. First of all, approximating  $V$  involves calculating the optimal list price. I do not make any additional simplifications here. For each point in the discretized state space and for every trial parameter vector, 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 in the data. This would involve calculating the optimal list price  $N * NSIM * STOM$  times, where  $N$  is the number of observations,  $NSIM$  is the number of simulations, and  $STOM$  is the simulated TOM for each observation. To overcome this computational obstacle, I approximate the list price policy function using linear (in parameters) interpolation. This is done using the 104,000 discrete points used to approximate the value function. To see how well this approximation works in practice, Figure 4 shows the actual list price policy for a grid of  $\hat{\sigma}$  and  $v^w - \hat{\mu}$  at the estimated parameters. The list price policy function is monotonic, and as a result, is well-approximated by simple interpolation.

<sup>26</sup> This is true for simulated maximum likelihood, but also usually true for simulated method of moments because the value function enters non-linearly into the simulated moments. See Keane and Wolpin (1994) for a more detailed discussion.

$\mu_{it}$  is unobserved to the econometrician  $\forall i, t$ . We can write  $\mu_{it}$  as

$$\mu_{it} = \hat{p}_{it}^L + e_{it} \tag{3.17}$$

where  $\hat{p}_{it}^L$  is observed (as described in Section 2.1) and  $e_{it}$  is simply the unobserved residual between  $\mu_{it}$  and  $\hat{p}_{it}^L$ .  $e_{it}$  reflects time varying characteristics of the house, such as the market value of renovations since the initial purchase and how much the seller invests in the presentation of the house to prospective buyers. For each observation and for each simulation of the other unobservables that affect the choice of list price in the initial week of listing, I recover a value of  $e_{it}$  that is consistent with the observed list price in the initial week.<sup>27</sup> This inversion of the list price policy function is valid because the optimal list price is monotonic in  $\hat{\mu}$  and  $w$ , as discussed in Section 3.5.

One assumption that I make is that all changes in  $\mu_{it}$  during the selling horizon are reflected in changes in  $\hat{p}_{it}^L$ ;  $e_{it}$  remains fixed over the selling horizon. Under this assumption, even though  $\mu$  is unobserved, I can recover *changes* in  $\mu$  from the observed changes in  $\hat{p}^L$ , which I allow to be neighborhood specific as described in Appendix A.1. It is only the changes in  $\mu$  that are needed to simulate draws of  $z$  in estimation (see equation (3.9)). Since the time series of  $\hat{p}^L$  can misrepresent the time series of  $\mu$  due to time-varying biases in the initial list prices, for each trial parameter vector, when I feed transitions of  $\hat{p}^L$  into the estimation routine, I undo the effect of expectation biases on the initial list price of each observation using the trial value of  $\alpha$ ,  $\Delta_6$ , and the list price policy function.

I choose not to include a menu cost to explain why list prices do not change continuously from week to week as shown in Table 2. However, Section 7 presents evidence that very small menu costs can rationalize sticky list prices. Thus, I would

---

<sup>27</sup> This procedure effectively treats the initial list price choices for every observation as moments that the model must fit exactly. Each simulation will result in a different value of  $e_{it}$ .



not expect the addition of a menu cost, which would increase the computational burden, to affect the conclusions.

Finally, note that the distributions of  $w_i$  and  $\hat{\mu}_{it_0}$  in the sample of observed listings are not normal. Some sellers with, for example, a high  $v^w$  relative to  $\hat{\mu}_{t_0}$  will not find it optimal to enter the market and pay  $c$  in the first period. I account for this selection in the simulation of  $w_i$  and  $\hat{\mu}_{it_0}$  by only accepting draws that are consistent with the decision to list.

I minimize the objective function using the BCPOLE command from the Fortran IMSL library, which is a nonsmooth, direct search algorithm. I use a variety of starting values to make sure that the results do not reflect a local minimum.

### 3.4 Identification

In this section I provide a description of the key variation in the data that separately identifies each of the parameters to be estimated. When the intuition is less clear, I provide simulations of select moments under various parameter values. This discussion is incomplete since in practice all 88 moments listed in Table 4 contribute to the identification of the parameters.

$\lambda$  is simply identified by the proportion of sales that occur at versus above the list price.  $\kappa$  and  $\sigma_{\hat{v}}$  are identified from the ratio of list prices to expected prices and the percentage of sales that occur at or above the list price. The reason is that when  $T^*$  (which is only a function of the parameters  $\kappa, \sigma_{\hat{v}}^2, \lambda$ ) is high, more sales will occur at or above the list price and sellers need to set lower list prices, all else equal, to attract buyers.  $\sigma_{\hat{v}}^2$  is separately identified because it has a direct effect on the distribution of  $p^* - p^L$  conditional on  $p^* > p^L$ . Simulations show that heterogeneity in  $v^w$  is identified from the percentage of withdrawals, the full TOM distribution for both sales and withdrawals, and the differences in list price changes for withdrawals versus sales.

In Figure 3, I present simulations that illustrate how  $\sigma_\eta$ ,  $\hat{\sigma}_{t_0}$ , and  $c$  move around a few select moments. For each parameter of interest, I run simulations in a neighborhood around the estimated value, which is highlighted with a vertical line. I fix the remaining parameters at their estimated values. More initial uncertainty increases the number of withdrawals for the reasons discussed in Section 3.5. It also increases the size of list price changes over the selling horizon.<sup>28</sup> This is because the list price policy function is increasing in the amount of uncertainty and  $\hat{\sigma}^{pre} - \hat{\sigma}^2$  is convex with respect to  $\hat{\sigma}^{pre}$ , which can be shown using equation (3.11). The TOM for homes that are withdrawn is non-monotonic in  $\hat{\sigma}_{t_0}$  because even though the value function for unmotivated sellers is increasing in  $\hat{\sigma}_{t_0}$ , higher  $\hat{\sigma}_{t_0}$  induces less motivated sellers into the market, who are more likely to withdraw quickly as the market declines. Higher  $\sigma_\eta$  has a similar effect on the number of withdrawals, but has different predictions for the TOM distribution and the size of list price changes. When  $c$  is more negative, a larger percentage of listings result in a sale mainly because the pool of potential sellers becomes more motivated on average.  $c$  does not have much of an effect on list price changes of homes that sell. The variances of the two signals,  $\sigma_z^2$  and  $\sigma_\eta^2$ , are separately identified because  $\sigma_\eta^2$  affects the pace of learning and has a direct effect on utility (i.e. it is also the variance of the offer distribution), while  $\sigma_z^2$  only affects the pace of learning.

I identify  $\alpha$  and  $h_1$  using the parameter estimates of two regressions as moments. The first is

$$(p_{it_0}^L - \hat{p}_{t_0}) - (p_{iT}^* - \hat{p}_T) = \gamma_0 + \gamma_1 \Delta_6 + \gamma_2 S_i + \epsilon_{it} \quad (3.18)$$

where  $T$  denotes the time period of sale. I also include quarter of listing dummies in equation (3.18). A one standard deviation increase in  $\Delta_6$  and  $S_i$  is predicted to increase the premium of list price over sales price by 1 and .8 percent, respectively.

---

<sup>28</sup> The list prices in these simulations are normalized by  $\mu_{it}$ .

The estimates are statistically significant. Many simulations of the model under alternative parameterizations show that this is consistent with positive values for  $\alpha$  and  $h_1$ .

The second regression is similar to (3.18), except the propensity to withdraw is substituted as the dependent variable. I also include the trajectory of average neighborhood prices during the selling horizon as an additional explanatory variable. A one standard deviation increase in  $\Delta_6$  and  $S_i$  is predicted to increase the propensity to withdraw by .01 and .031, respectively. The estimates are statistically significant. Many simulations of the model under alternative parameterizations show that this is consistent with positive values for  $\alpha$  and  $h_1$ .<sup>29</sup>

One weakness of this identification strategy for expectation bias is that unobservables (e.g. motivation to sell, home quality) can bias estimates of (3.18) if these unobservables are correlated with  $\Delta_6$ . Since the model will not capture these correlations, we may be matching a moment of the data through the incorrect mechanism. Section 8 presents even stronger evidence, which is robust to this criticism, that sellers who put their homes on the market after a period of declining prices tend to overstate the value of their homes.

I set  $\beta = 0.999$ , which corresponds to an annual discount factor of about 0.95. I set  $\hat{\mu}_p = -.003$  (-.3 percent), which is the average weekly change in  $\hat{p}_{it}^L$  during my sample period.<sup>30</sup> I set  $\hat{\sigma}_p^2 = .008^2$  (std = .8 percent), which is the average weekly variation in  $\hat{p}_{it}^L$  faced by a seller in my sample.

---

<sup>29</sup> Section 3.5 discusses why this is the case for  $h_1$ . When  $\alpha > 0$ , the average value of  $w_i$  that is consistent with the observed decision to list is higher, and the likelihood of withdrawal is increasing in  $w_i$ .

<sup>30</sup> As a robustness check, I also run the model assuming sellers are more optimistic ( $\hat{\mu}_p = -.002$ ). This does not affect the main results.

### 3.5 Results

Table 3 reports the parameter estimates and asymptotic standard errors. Recall from Section 3.1 that prices enter the model in logs. Thus, the estimate of  $\sigma_\eta$  implies that a buyer with a valuation that is one standard deviation higher than the mean valuation has a 24.5 percent ( $\exp(.22)-1$ ) higher valuation relative to the mean, or \$79,655 for the average home in the data.

The amount of uncertainty,  $\hat{\sigma}_{t_0}$ , is small relative to the size of buyer preference heterogeneity. Equation (3.11) implies that this decreases the pace of learning because each offer is less informative. As predicted by the theory,  $h_1 > 0$ : the amount of uncertainty is increasing in the heterogeneity of nearby houses. At the mean level of  $S$ , the standard deviation of the prior is 7.4 percent or \$24,605 for the average home in the data. At the 10th percentile of the  $S_i$  distribution, the standard deviation of the prior is 6.5 percent; at the 90th percentile, it is 8.5 percent.

The estimated value of  $\alpha$  implies that sellers who put their homes on the market after a period of declining prices tend to overstate the value of their home by .65 percent for each 1 percent decline in prices. Evidently, realtors do not fully adjust their comparable sales analysis for the downward trend in prices during my sample period.

The estimates of  $\mu_{wL}$  and  $\mu_{wH}$  suggest that there are two types of sellers who contact real estate agents, motivated and unmotivated, with heterogeneity around each type. We can think of motivated sellers as having low  $v^w$  because they need to move houses for a new job or because it is a forced sale. Unmotivated sellers, on the other hand, are still relatively well-matched with their existing house, and will only sell for a high price.

$c$  is estimated to reduce the seller's weekly flow utility from living in the home by .2 percent. For the average seller, this translates into weekly costs of about \$740

dollars.

Using the estimates of  $(\kappa, \sigma_v^2, \lambda)$ , I calculate that  $T^* = -.181$  (see Appendix A.2 for the derivation). This implies that if the list price is \$800,000, then a buyer with an expected valuation of \$667,548 is indifferent over the decision to inspect.<sup>31</sup>

### 3.5.1 Model Fit and Discussion

Table 4 compares the actual moments to the simulated moments and the weights placed on each of the moments in the estimation. The learning model performs well. Even when agents are correct on average in their expectations about the severe market decline during my sample period, the model predicts lengthy TOM for homes that are both sold and withdrawn. The model predicts a large percentage of withdrawals with a very reasonable amount of uncertainty because the cost of testing the market,  $c$ , is low. If I assume that the seller has full information about the distribution of buyer valuations, only about 2 percent of listings would result in withdrawals.<sup>32</sup> The model also matches the effects of  $\Delta_6$  and  $S$  on prices and withdrawal rates. As an additional check on model fit, I find that the model predicts that 7 percent of list price changes are increases compared to 7.9 percent in the data even though I do not require the model to match this particular moment in estimation.

In Table 5, I show additional results of the model for my sample. 46 percent of the listings in my sample are by motivated sellers.<sup>33</sup> These motivated sellers are more likely to sell rather than withdraw and receive lower sales prices conditional on selling due to their lower reservation prices. All types of sellers are being matched with buyers who have high idiosyncratic tastes for the house. Sellers accept offers

---

<sup>31</sup>  $667,548 = \exp(\ln(800000) - .181)$ .

<sup>32</sup> Hitsch (2006) finds comparable sensitivity of the entry and exit rates to the level of uncertainty in the ready-to-eat breakfast cereal industry.

<sup>33</sup> This percentage differs from the value of  $\gamma$  reported in Table 3 because of the selection described in Section 4.

from the 97th percentile of the valuation distribution on average. I also find that the discount in sales price for a highly motivated seller relative to the average is comparable to the forced sale discount estimated in Campbell et al. (2009). Sellers can wait for high-valuation buyers because holding costs are low.

List prices for motivated sellers decline over the selling horizon, but actually increase relative to  $\mu$  for unmotivated sellers. Even though  $\mu$  is declining, unmotivated sellers do not drop their list prices as much because their outside option remains constant in the model. On the other hand, for motivated sellers, the outside option is not as attractive and so they drop their list prices as the market declines and uncertainty gets resolved. This explains how the model generates sticky list prices when price levels are falling over the selling horizon and why sellers who sell drop their list prices more relative to sellers who withdraw.

Bayesian updating reduces the standard deviation of the seller's initial prior by about 26 percent over the course of the selling horizon for motivated sellers even though uncertainty continues to accrue from changes in  $\mu$ . Unmotivated sellers learn less because they price high, which attracts fewer buyers.

The results in Table 5 also show why Case-Shiller sales price indexes can be misleading. Sales prices are selected prices that tend to overstate how an average buyer values a house, and prices vary significantly for identical houses depending on the motivation of the seller. Figure 5 illustrates how this selection can cause trends in the Case-Shiller price index over time to misrepresent trends in actual home values (or  $\mu$  using the notation of the model). Using my data, I plot trends in  $\hat{p}$  ('Sales Price/Case-Shiller Index') versus trends in  $\hat{p}^L$  ('Initial List Price Index') for the typical home in Los Angeles. Section 2.1 and Appendix A.1 describe exactly how these indexes are calculated; the important difference is that the  $\hat{p}$  index is based on the sales prices of homes that sell whereas the  $\hat{p}^L$  index is based on the initial list prices of all homes listed for sale regardless of their ultimate outcome. Since

the model provides an estimate of how initial list prices are affected by biases from previous comparable sales, I also present a third line (the dashed line) in Figure 5 that adjusts the  $\hat{p}^L$  index for these biases using the estimate of  $\alpha$  and the list price policy function. This latter index is the most preferred measure of trends in  $\mu$  of the three because it is not affected by biases or by selection on the types of sellers that choose to sell. The Case-Shiller index appears to overstate the decline in home values by as much as 9 percent during my sample period. A likely explanation is that only the most motivated sellers select to sell when prices are falling, and my model predicts that these sellers accept offers from the lower-end of the offer distribution.

## 3.6 Model Predictions

### 3.6.1 Value of Information

To calculate the value of information, for various levels of initial uncertainty  $\hat{\sigma}$ , I simulate optimal selling behavior and calculate the present discounted utility for many draws of  $v^w - \mu = w_i$ .<sup>34</sup>

Figure 6 shows the average welfare cost of uncertainty relative to the full information case.<sup>35</sup> Uncertainty lowers welfare because it distorts the entry decision and the choice of list and reservation prices. The welfare losses are significant, especially if the amount of uncertainty is high enough. At  $\hat{\sigma} = .06$ , the welfare costs of uncertainty are \$2,331; however, at  $\hat{\sigma} = .1$ , the costs jump to \$17,501.<sup>36</sup> The lighter bar in Figure 6 shows the value of information when there is no weekly volatility in valuations ( $\hat{\sigma}_p = 0$ ). The costs of uncertainty are lower in this case. This is because it is easier for sellers to resolve initial uncertainty when no additional uncertainty

<sup>34</sup> I assume that the  $\mu$  process behaves over time according to the seller's expectations in (3.8)

<sup>35</sup> In the full information case ( $\hat{\sigma} = 0$ ), I also assume  $\sigma_z = 0$ . I find that the value of having  $\sigma_z = 0$  is highest when  $\hat{\sigma} = .09$ . The improvement in welfare at this level of  $\hat{\sigma}$  is \$1250.

<sup>36</sup> The dollar figures are calculated using the average  $\mu$  in the sample for Los Angeles at the mean level of  $S$ . For lower priced cities, the dollar figures would be lower and the percentages would be the same.

accrues during the selling horizon.

Table 6 shows that the average seller is willing to pay \$8,658 or 2 percent of the average sales price to move from the estimated prior to the full information case. About 1/4 of sellers (i.e. 1/4 of the draws of  $w_i$ ) would be willing to pay over \$12,000 or 2.7 percent of the average sales price for full information. I also find that the welfare costs of the positive bias about home values is about \$2000 on average. This relatively small effect compared to the welfare costs of uncertainty shown in Figure 6 is one possible explanation for why sellers (and their agents) do not make the effort to completely account for trends in local market conditions.

In this static model of buyer behavior, seller welfare losses are not simply transfers to buyers. Buyers only obtain surplus when their valuation exceeds the list price, and list prices are increasing in uncertainty. In addition, more uncertainty attracts unmotivated sellers into the market, which increases the likelihood that a costly inspection will result in no surplus for the buyer.

Since I model the seller's decision as a single-agent problem, the welfare results do not account for changes in competition or changes in market tightness due to competitive entry as the information structure changes. However, changing uncertainty primarily affects the entry and exit decisions of unmotivated sellers, and Table 5 shows that these sellers price so highly (they attract less than 1 buyer over the course of the selling horizon) that they probably do not have a large effect on the market. I also find reduced-form evidence that competition, as measured by the number of listings within a 1/2 mile radius, does not have a large effect on pricing. This result is not surprising given that single-family homes tend to be highly differentiated even within the same neighborhood.<sup>37</sup>

---

<sup>37</sup> The results are available upon request. Measurement error in listings and endogeneity of the number of listings due to high inventory being correlated with unobservably worse market conditions should lead me to overstate the effect (in absolute value).



### 3.6.2 Alternative Demand Predictors

In the bottom panel of Table 6, I calculate welfare assuming that the seller forgoes a realtor and uses publicly available data sources to form demand expectations. The Appendix A.3 describes how I generate these alternative priors. In these simulations, I continue to assume fully optimal behavior conditional on a particular prior. I consider the effects of suboptimal behavior in Section 7.4.

In specification (6), I assume that sellers form price expectations using *Zestimates* from the popular website *zillow.com*.<sup>38</sup> Comparing the standard deviation of the prior in (1) versus (6), we see that realtors provide more information than *zillow.com*. This is not surprising given that realtors can observe information that is not available in transaction data such as renovations, decorations, how similar in unobservables homes are to the comparables, characteristics of the seller, etc. Even after accounting for savings in seller realtor commissions, the average seller is at least \$4,135 worse off using *zillow.com* instead of a realtor.<sup>39</sup>

The benefits of observing the prices of close substitutes can be seen from the large welfare losses in specification (9) relative to specification (6). A demand predictor based on the publicly available Case-Shiller price index lowers uncertainty significantly relative to specification (9) because it controls for all time-invariant home characteristics (both observed and unobserved). However, it performs worse than specification (6) because it measures metro-level price changes, and appreciation rates vary significantly across neighborhoods within a metro area.<sup>40</sup> Since the Case-

---

<sup>38</sup> Zillow.com does not make its estimation methodology publicly available. However, the information provided on their website suggests that they apply econometric models to the type of transaction data used in this paper. Zillow.com does provide summary statistics on the accuracy of their estimates, which I use to calculate the prior in (6). See Appendix A.3 for details.

<sup>39</sup> This understates the welfare loss because I do not have information to calculate biases in *Zestimates* during my sample period. Thus, I conservatively set the bias to zero.

<sup>40</sup> This index is also reported with a 2-month lag, which explains why the bias is positive. If I assume that the seller is sophisticated and forecasts the current index using an OLS regression, welfare is only slightly higher.

Shiller price index is only available in select cities, specification (8) reports welfare using the FHFA index, which covers more metro areas. The disadvantage of FHFA relative to Case-Shiller is that it is reported quarterly rather than monthly and it is only based on homes sold with conforming loans. After accounting for savings in realtor commissions, sellers would still be willing to pay \$22,482 or 5.06 percent of the expected sales price on average to move from specification (8) to specification (1). A significant fraction of sellers would be willing to pay over \$32,000.

If the alternative to using a realtor is to form price expectations without accounting for the heterogeneity of the housing stock and the heterogeneity in appreciation rates within the metro area, then these results suggest that realtors provide enough value on average to justify their commissions through information alone. This may be the reality for less sophisticated sellers.

Table 6 also shows that there is a large amount of demand for more sophisticated forecasting techniques. If zillow.com is indeed applying the best statistical tools to the type of transaction data used in this paper, then a technology that could approach the full information case would need to solicit information on time varying home characteristics that are not typically recorded in transaction datasets.<sup>41</sup> Given the significant heterogeneity in sales price conditional on quality as shown in Table 5, another possibility is to adjust comparable sales for heterogeneity in the motivation of the seller using TOM or equity position (see Genesove and Mayer (2001); Anenberg (2011); Ortalo-Magne and Rady (2006)).

### *3.6.3 Price Dynamics*

The predictability of house price appreciation rates is a key stylized fact of housing markets. In their pioneering work, Case and Shiller (1989) find that a 1 percent

---

<sup>41</sup> Zillow.com appears to be moving in this direction. Home owners can update and correct information about their home that is not in the public records.

increase in real house prices in one year is associated with a .2-.5 percent increase the next year. Glaeser and Gyourko (2007) calibrate a dynamic rational expectations model of house price formation to try to explain the persistence. Their model succeeds in matching many of the features of the data; however, it cannot explain the persistence in house prices “under virtually any reasonable parameterization” of the model. Like Case and Shiller, they cite inefficiency in the housing market as a likely explanation.

As shown above, however, sales prices are not a perfect proxy for the fundamentals because prices are determined in a search and bargaining environment. In fact, uncertainty generates persistence in price appreciation rates even when the fundamentals follow a random walk. The reason is that sellers with rational expectations but with uncertainty over current demand shocks do not instantly adjust the mean of their beliefs to such demand shocks.<sup>42</sup> For example, when there is a positive shock, reservation prices rise slightly, but are too low relative to the perfect information case. As time progresses, however, learning from buyer behavior will allow sellers to fully adjust their beliefs to that initial shock. The same intuition holds for a negative demand shock. If sellers have biased beliefs at the beginning of the selling horizon from using previous comparable sales, then the persistence is even stronger.

I formalize this intuition using simulations of the model. I assume that all houses are identical and

$$\mu_t - \mu_{t-1} \sim N(\hat{\mu}_p, \hat{\sigma}_p^2) \quad (3.19)$$

where  $\hat{\mu}_p = -0.003$  and  $\hat{\sigma}_p = 0.008$  so that the simulated  $\mu$  process behaves similarly to the  $\mu$  process used to estimate the model. The parameters of the model are set at their estimated values and I continue to assume that (3.8) holds so that sellers have rational expectations about the  $\mu$  process.

---

<sup>42</sup> Recall that the formula for the posterior mean places weight on an unbiased signal of the current period demand shock, but also on the seller’s prior expectation of the price change.

Using the estimates in Table 3, I simulate sales prices for 20,000 sellers over 48,000 weeks.<sup>43</sup> Then, following the literature, I run the following regression

$$p_t - p_{t-52} = \rho_o + \rho_1(p_{t-52} - p_{t-104}) + \nu_t. \quad (3.20)$$

where  $p_t$  is the log average price over all simulated sales in week  $t$ . If we ran this regression on the simulated  $\mu$  process,  $\rho_1 = 0$ .

The first row of Table 7 shows that when 1)  $\alpha = 0$  (i.e. no initial biases) and 2)  $\sigma_z = 0$  (i.e. sellers can perfectly observe price changes during the selling horizon),  $\rho_1$  is close to zero.<sup>44</sup> Thus, search frictions alone cannot generate persistence. In the second row, when I relax assumption 2),  $\rho_1$  turns positive. The third row shows that the persistence rises to 0.204 when assumption 1) is also relaxed.

The right panel of Table 7 shows the equivalent set of results when I aggregate prices to the quarterly level. In this case, the dependent variable is  $p_t - p_{t-4}$  where  $t$  is a quarter instead of a week. 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), for example, run their regressions at the quarterly level. The aggregation alone introduces persistence, and the AR(1) coefficient rises to 0.235.

When I run these same regressions over shorter frequencies, the persistence is even higher. We can see this through the equation for the OLS estimate of  $\rho_1$ :

$$\hat{\rho}_1 = \frac{cov(p_t - p_{t-L}, p_{t+L} - p_t)}{var(p_t - p_{t-L})}. \quad (3.21)$$

As  $L$  gets smaller, the numerator stays approximately the same and the denominator gets smaller because there are fewer shocks between time  $t$  and  $t - L$ .

<sup>43</sup> The results are robust when there is less volatility in the market. For example, when  $\hat{\sigma}_p = 0.0043$ , which gives variation in monthly price changes that is equal to the variation in monthly price changes in the historical Case Shiller index for Los Angeles, the persistence results are even stronger. Because I simulate a large samples of houses, I avoid complications that arise in practice from measurement error.

<sup>44</sup> The coefficient is slightly negative because when prices rise, unmotivated sellers are more likely to sell rather than withdraw, and unmotivated sellers tend to sell at high prices.

To exploit the source of the predictability generated here, an investor/trader would need access to better information about the current period fundamentals. This information is difficult to obtain given the thinness in the housing market, and given that sales data typically become available with a lag because closing dates lag agreement dates. The large transaction costs involved in selling homes also complicates any potential trading strategy. It is worth emphasizing that we cannot rule out irrationality or inefficiency as sources for the observed predictability. Rather, we can show that a significant amount of predictability arises from a purely rational model of the home selling problem because of information problems.

#### *3.6.4 Suboptimal Behavior*

In the simulations in Sections 7.1 and 7.2, we assumed fully optimal behavior while varying the initial prior to isolate the value of information. In this section, I consider two different kinds of suboptimal behavior to get an estimate of the value of some of the other services that real estate agents provide. The two lighter bars in Figure 7 show the implications of setting a suboptimal list price, but not having to pay a 3 percent realtor fee upon sale. Realtor fees exceed the welfare loss on average when the list price optimization error is normal with mean zero and a standard deviation of 10 percent. However, when the optimization error is large enough, setting the wrong list price can significantly affect welfare. The black bar shows relative welfare assuming that sellers do not update their priors in response to new information. For low levels of uncertainty, failing to update priors does not have a large effect on welfare, but the effects rise quickly as uncertainty increases.

The results suggest that low cost alternatives to full service agents that provide advertising and paper work services, but provide less advice and strategy, may be attractive even to less sophisticated sellers, especially if the initial level of uncertainty is low enough. This helps to explain the emergence of no frills services such as Assist-

2-Sell and Redfin, as the internet has made price estimates and information about the housing market more available to potential sellers. In addition, the alternative list price policy simulations suggest that very small menu costs can rationalize the unchanging list prices that we observe throughout the selling horizon, as shown in Table 2. In fact, in a simulation not shown, I find that seller welfare is only slightly lower at the estimated prior when I only allow sellers to adjust their list price every two months.

### 3.7 Expectation Bias: Theory and Evidence

In this section, I supplement Section 5 to provide strong reduced form evidence that the estimate of  $\alpha$  (which determines the amount of bias in initial beliefs) from the structural model is due to biased beliefs rather than unobserved heterogeneity that I do not model. My approach to identifying bias contrasts with other studies in the literature that compare stated home values from various surveys to actual transaction prices to identify whether home owners have biased beliefs.<sup>45</sup> I avoid a host of problems associated with selection, inconsistent timing, measurement error and misreported information in survey data by using data on an endogenous choice variable that has significant consequences for seller utility.

Suppose we observe the situation depicted in Figure 8 in the data. Sellers in a blue neighborhood and a red neighborhood list their homes for sale in the same time period. Price levels in the red neighborhood have been falling more severely relative to prices in the blue neighborhood before the time of listing. In the first week of listing, sellers in the red neighborhood set higher list prices on average, controlling for time-invariant characteristics of the house and the timing of the initial listing. However, sellers in the red neighborhood also receive *lower* sales prices and take longer to sell on average. A likely explanation for this pattern given that

<sup>45</sup> See, for example, Benitez-Silva, Eren, Heiland, and Jimenez-Martin (2011) and citations within.

sellers (or their realtors) rely on previous comparable sales is that sellers in the red neighborhood have biased beliefs relative to sellers in the blue group. In most models, positive bias about the quality of the house should increase the initial list price. But a higher than optimal list price should have negative consequences for seller utility, either through a lower sales price, a higher TOM, a lower probability of sale, or a combination of any of the three. Since in theory, inflated beliefs can have negative consequences through mechanisms other than the sales price, this test for expectation bias has low power.<sup>46</sup> However, the appeal of the test is that the probability of making a Type I error is very low. For example, lower unobserved motivation to sell or higher unobserved quality of the homes in the red neighborhood should be associated with both higher list *and* sales prices.

To get an even larger sample to run the test described above, I supplement the Los Angeles dataset with an identical dataset for San Francisco. The results are similar when I run the tests separately for each city. Table 8 presents the tests for bias described above. In the first column, I run the following regression on the sample of sales

$$p_{it_0}^L - \hat{p}_{it_0} = \alpha_0 + \sum_{k=2}^{10} \alpha_k I[\Delta_6 < d_k] + \epsilon_{it} \quad (3.22)$$

where  $I$  is the indicator function and  $d_k$  is the  $k$ th decile of the  $\Delta_6$  distribution. I also include a dummy variable for whether the listing is an REO property (determined by parsing the seller name for revealing words) and quarter of listing dummies. Column 2 shows the results from the same specification as (3.22), except the sales price is substituted for the list price in the dependent variable and the time period of sale is substituted for  $t_0$ .

---

<sup>46</sup> That is, if the null hypothesis is that there is no expectation bias, we may often fail to reject the null when the null is false.

The list price specifications in Table 8 show that the more prices have fallen, the higher is the list price on average. The effect is the opposite when we look at sales prices, and there are even adjacent regions of the price change distribution where list prices are significantly increasing but sales prices are significantly decreasing (highlighted in bold). This evidence is consistent with sellers having biased beliefs when price levels are falling, and as discussed above, correlated unobservables cannot generate this relationship.

The next two specifications in Table 8 are consistent with the hypothesis that declining prices lead to biased beliefs. TOM and the propensity to withdraw are increasing in  $\Delta_6$ , although the extreme decile of the price change distribution appears to be an outlier. The structural model presented above matches both of these moments with a positive value of  $\alpha$ ; the intuition is presented in Section 3.5.<sup>47</sup> Appendix A.4 presents additional robustness specifications.

### 3.8 Conclusion

This paper incorporates uncertainty and Bayesian learning into a search model of the residential real estate market. Using the model and a new dataset that combines information on all housing transactions with information on the complete listing and list price history of each property, I find that sellers' lack of information about home values matters. It affects listing and pricing decisions enough to have a significant impact on welfare, and it has important effects on equilibrium house price dynamics. In particular, uncertainty and learning can explain a significant amount of the predictability of short-run price appreciation rates that has been well-documented in the literature and has led researchers to question the efficiency of the housing market.

<sup>47</sup> In results not reported, I find that the effects of  $\Delta_6$  on prices diminish as we move later in the sample period. The time pattern of prices in Figure 1 provides a likely explanation. As the recession deepened, sellers learned that prices were on a downward trajectory, and did a better job of adjusting the prices of recent comparable sales for the downward trajectory of prices. I also find that the effects of  $\Delta_6$  are stronger for houses with higher  $S$ .



The parameter estimates themselves also reveal several new findings about the nature of seller information at the time of listing and during the listing period. Seller beliefs are relatively precise: the variance of the seller's prior is low relative to the variance that would arise if sellers used alternative sources of market information. The estimates of the mean of the seller's initial prior show that sellers place too much weight on lagged information. Thus, when prices are falling, sellers overestimate the value of their homes on average, leading to higher list prices, longer TOM, a higher withdrawal rate, and sometimes lower sales prices.

The data set I use in this paper comes from one market over a short time period. I selected this market and time period because of data availability. It would be interesting to test how the predictions of the model are affected during hot markets, and also how the predictions change going forward as the internet continues to improve the availability of information at lower costs. I have presented evidence that abstracting from competition would not significantly affect the conclusions reached in this paper, but extending housing search models to the multi-agent setting is important for addressing a number of other interesting topics, including endogenous cycling between housing booms and busts, and the effects of foreclosures externalities.

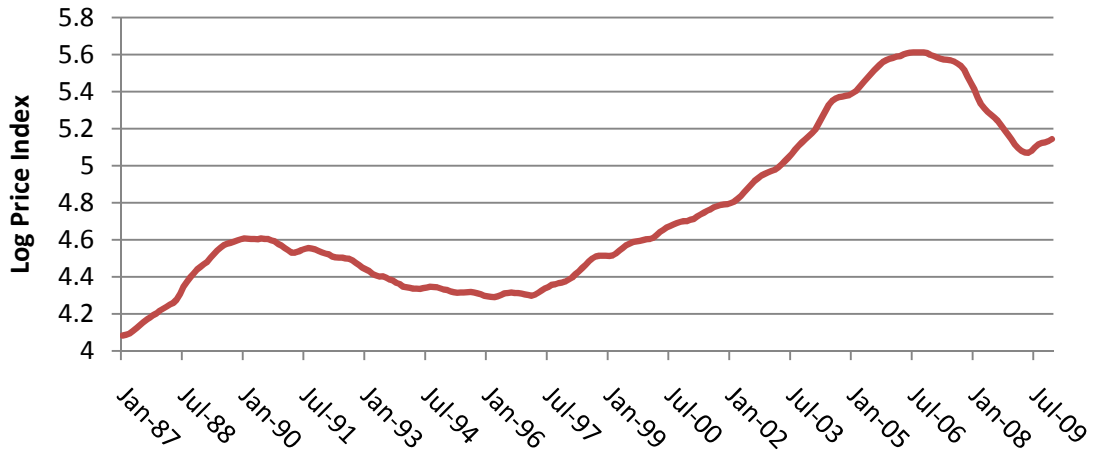
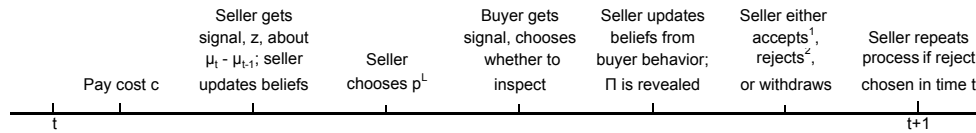


FIGURE 3.1: Real Log Price Index in Los Angeles 1987-2009



<sup>1</sup>This option is only available if the buyer inspects.  
<sup>2</sup>Seller receives a flow utility equal to  $(1-\beta)v^w$ .

FIGURE 3.2: Timeline of Events in Model

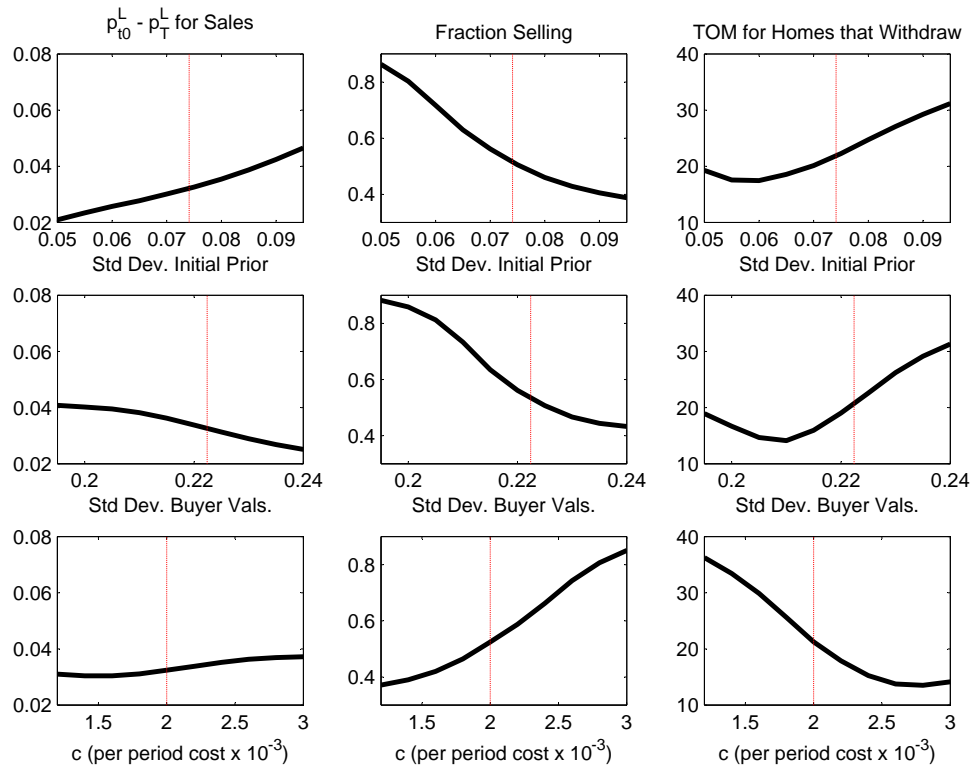


FIGURE 3.3: Simulation of Select Moments

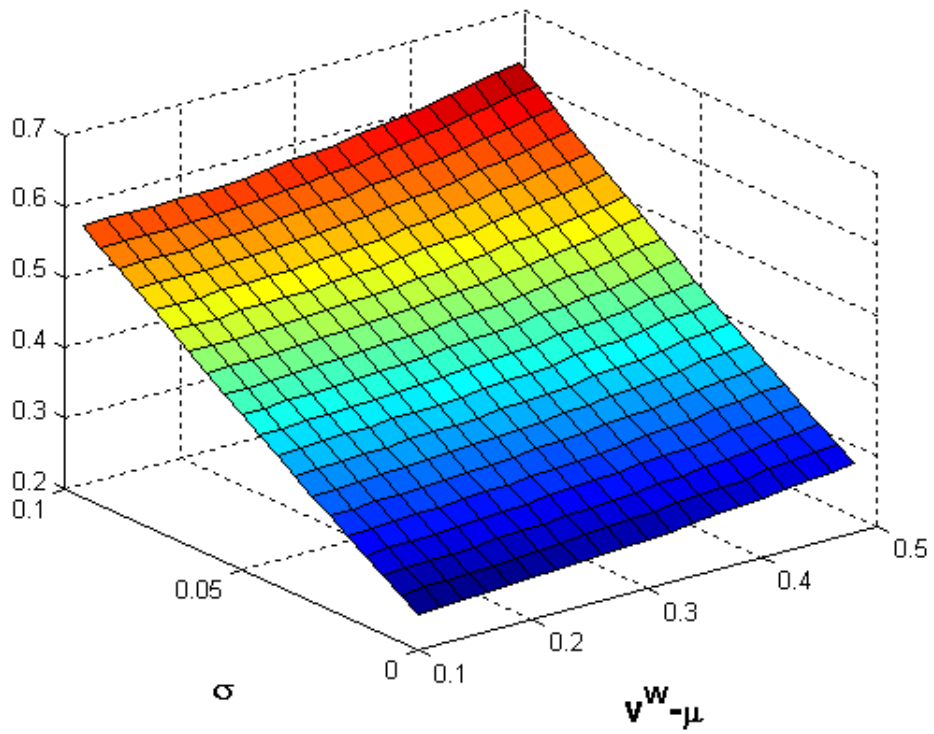


FIGURE 3.4: List Price Policy Function

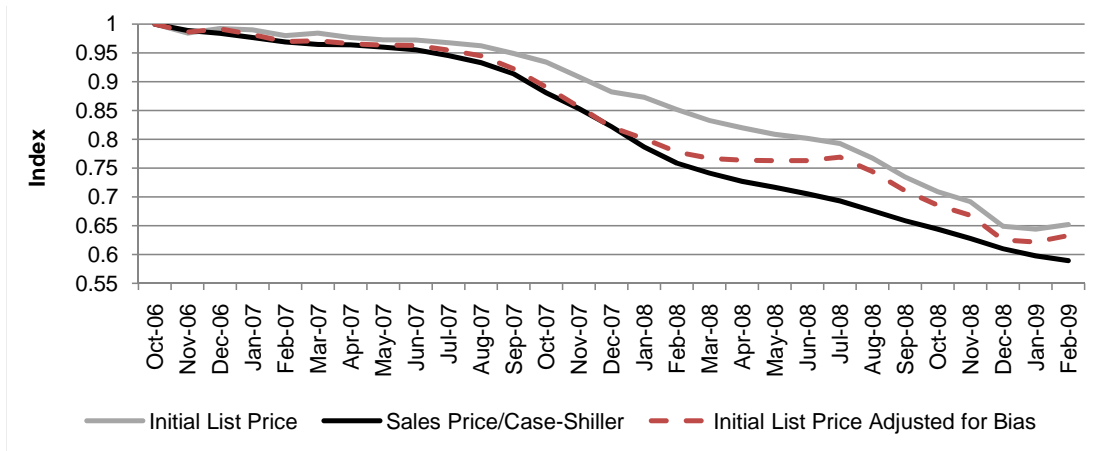


FIGURE 3.5: Home Value Indexes for Los Angeles

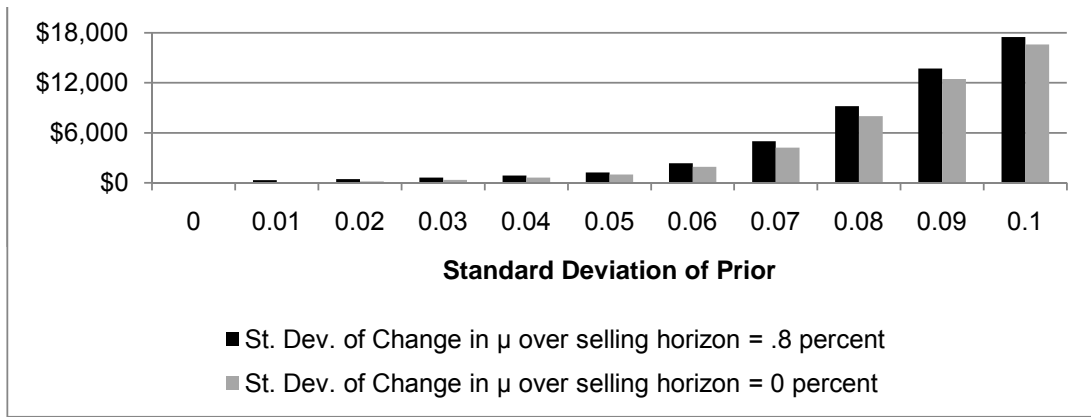


FIGURE 3.6: Average Welfare Loss from Uncertainty

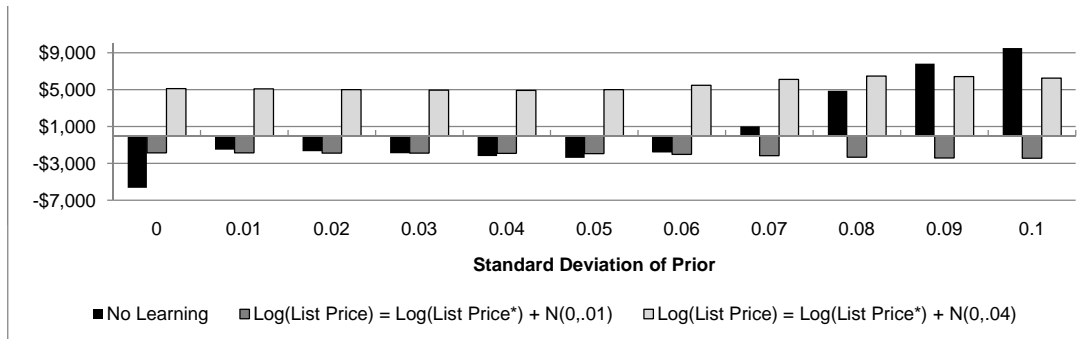
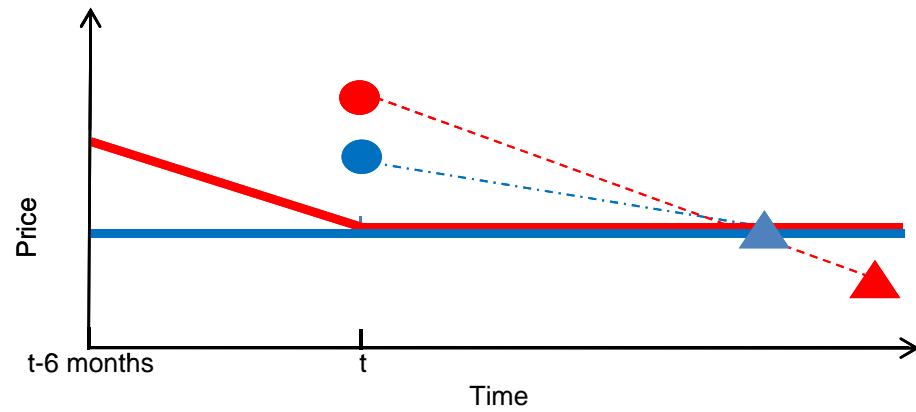


FIGURE 3.7: Average Welfare Loss Under Suboptimal Behavior with 3 Percent Realtor Fee Relative to Fully Optimal Behavior with 6 Percent Fee





-  Average list price for homes listed at month t
-  Average sales price for homes listed at month t
-  Average sales price

FIGURE 3.8: Test for Expectation Bias

	Sold (N = 47799)			Withdrawn (N = 59792)		
	25th Pctile	Median	75th Pctile	25th Pctile	Median	75th Pctile
List Price/Expected Sales Price	0.94	1.05	1.18	1.02	1.16	1.34
Weeks on Market	7	14	27	11	20	31
Weeks Until First List Price Change <sup>1</sup>	4	6	11	4	7	13
Change in List Price <sup>1</sup>	-7.23%	-4.08%	-2.11%	-7.25%	-3.87%	-1.87%
Change in List Price at Delisting Relative to Listing	-14.17%	-3.58%	0.00%	-10.95%	-2.49%	0.00%
% of Properties with No Price Change	--	40.24%	--	--	42.02%	--
% of Price Changes that are Increases	--	6.99%	--	--	8.63%	--
Square Feet	1180	1499	1974	1150	1478	1983
Year Built	1950	1957	1977	1948	1956	1975
Sales Price	370000	515000	720000	--	--	--

<sup>1</sup>Conditional on at least one list price change.

Note: All of the differences in means between the two groups are statistically significant. Expected Price is a predicted sales price based on a repeat sales regression.

FIGURE 3.9: Summary Statistics

Weeks Since Listing	% Adjusting List Price
1	4.52%
2	6.97%
3	8.60%
4	9.40%
5	9.68%
6	8.93%
7	8.69%
8	8.48%
9	8.36%
10	8.24%
11	7.99%
12	7.91%
13	8.10%
14	7.96%
15	7.56%
16	7.27%
17	7.23%
18	7.29%
19	6.96%
20	6.83%
21	6.96%
22	6.54%
23	6.54%
24	6.84%
>=25	7.23%

FIGURE 3.10: Percent of Sellers that Adjust List Price by Weeks Since Listing



Variable	Description	Estimate	Std. error
$h_0$	Intercept of st. dev. of initial prior.	0.0608	0.0043
$h_1$	How st. dev. of initial prior varies with the heterogeneity of surrounding homes.	0.1174	0.0067
$\sigma_{\eta}$	St. dev. of buyer valuations.	0.2224	0.0035
$c$	Weekly holding cost.	-0.0020	1.6E-04
$\kappa$	Buyer inspection cost.	-0.0090	0.0035
$\sigma_v$	Standard deviation of buyer uncertainty.	0.1998	0.0220
$\lambda$	Prob. that competition drives the sales price above the list price.	0.5427	0.0421
$\alpha$	Determines how seller beliefs depend on prices of previous comparables.	0.6466	0.0234
$\mu_{wL}$	Mean of unobserved utility of withdrawing (low type).	-0.0443	0.0217
$\mu_{wH}$	Mean of unobserved utility of withdrawing (high type).	0.5755	0.0050
$\sigma_w$	Standard deviation of unobserved utility of withdrawing.	0.1089	0.0093
$\gamma$	Probability seller is low type.	0.2726	0.0161
$\sigma_z$	St. dev. of signal about weekly decline in mean valuations.	0.0230	0.0078
$\beta$	Weekly Discount Factor.	0.9990	--
$\mu_p$	Mean of Belief about weekly decline in mean valuations.	-0.0030	--
$\sigma_p$	St. dev. of Belief about weekly decline in mean valuations.	0.0080	--

N=107591

Note: The final three parameters are fixed in estimation.

FIGURE 3.11: Parameter Estimates from the Structural Model

Note: All prices are in logs and pricehat is the log expected sales price. Weight is the diagonal element of the weighting matrix associated with each moment calculated from a first stage.

	Moment	Simulated Moment	Weight
1	Mean price-pricehat	-1.06%	1408
2	Std price-pricehat	20.02%	993
3	Mean of price-pricehat homes sold at week 5	-0.14%	1406
4	Mean of price-pricehat homes sold at week 10	0.55%	1449
5	Mean of price-pricehat homes sold at week 20	1.09%	720
6	Std of price-pricehat homes sold at week 5	19.93%	2102
7	Std of price-pricehat homes sold at week 10	20.67%	2166
8	Std of price-pricehat homes sold at week 20	20.67%	1077
9	Mean listprice-price	4.57%	901185
10	Std listprice-price	10.02%	821
11	Mean listprice-price homes sold at week 5	3.30%	12181
12	Mean listprice-price homes sold at week 10	3.65%	12554
13	Mean listprice-price homes sold at week 20	5.29%	6240
14	Std listprice-price homes sold at week 5	8.75%	2721
15	Std listprice-price homes sold at week 10	7.72%	2805
16	Std listprice-price homes sold at week 20	9.45%	1394
17	Mean listprice-pricehat	12.91%	53204
18	Mean listprice-pricehat in week 5	11.23%	21400
19	Mean listprice-pricehat in week 15	12.81%	13719
20	Std listprice-pricehat	22.77%	43498
21	% of prices above listprice	17.86%	1060
22	% of prices equal listprice	15.01%	316
23	mean sale price - list price when price above	4.02%	1340
24	std sale price - list price when price above	3.55%	1519
25	Mean listprice change in week 5	-0.45%	1185798
26	Mean listprice change in week 15	-0.40%	760174
27	Std listprice change in week 5	2.20%	4671
28	Std listprice change in week 15	2.43%	2995
29	.05 percentile of listprice change distribution in week 5	-3.52%	2199
30	% of sold homes that sell in week 1	3.73%	5320
31	% of sold homes that sell in week 2	3.85%	5320
32	% of sold homes that sell in week 3	3.13%	5320
33	% of sold homes that sell in week 4	3.04%	5320
34	% of sold homes that sell in week 5	3.48%	5320
35	% of sold homes that sell in week 6	4.01%	5320
36	% of sold homes that sell in week 7	4.40%	5320
37	% of sold homes that sell in week 8	4.34%	5320
38	% of sold homes that sell in week 9	3.92%	5320
39	% of sold homes that sell in week 10	3.59%	5320
40	% of sold homes that sell in week 11	3.51%	5320
41	% of sold homes that sell in week 12	3.12%	5320
42	% of sold homes that sell in week 13	3.35%	5320
43	% of sold homes that sell in week 14	2.74%	5320
44	% of sold homes that sell in week 15	2.56%	5320
45	% of sold homes that sell in week 16	2.54%	5320
46	% of sold homes that sell in week 17	2.54%	5320
47	% of sold homes that sell in week 18	2.13%	5320
48	% of sold homes that sell in week 19	2.05%	5320

FIGURE 3.12: Moments Used in Estimation

% of sold homes that sell in week 18	2.11%	2.10%	34275
% of sold homes that sell in week 19	2.05%	2.04%	34275
% of sold homes that sell in week 20	1.82%	1.96%	34275
% of sold homes that sell in week 21	1.83%	1.88%	34275
% of sold homes that sell in week 22	1.76%	1.77%	34275
% of sold homes that sell in week 23	1.63%	1.67%	34275
% of sold homes that sell in week 24	1.54%	1.62%	34275
% of sold homes that sell in week 25	1.61%	1.55%	34275
% of sold homes that sell in week >25	30.16%	30.67%	34275
Effect of Stdyrbt on Prices	0.10	0.12	7042
Effect of Prev. Price Changes on Prices	0.17	0.16	4177
Effect of Stdyrbt on Withdraw	0.43	0.29	35
Effect of Prev. Price Changes on Withdraw	0.19	0.58	8
% of withdrawn homes that withdraw in week 1	2.59%	2.63%	25866
% of withdrawn homes that withdraw in week 2	2.16%	2.98%	25866
% of withdrawn homes that withdraw in week 3	2.14%	2.93%	25866
% of withdrawn homes that withdraw in week 4	2.27%	2.86%	25866
% of withdrawn homes that withdraw in week 5	2.36%	2.79%	25866
% of withdrawn homes that withdraw in week 6	2.55%	2.75%	25866
% of withdrawn homes that withdraw in week 7	2.45%	2.75%	25866
% of withdrawn homes that withdraw in week 8	2.70%	2.66%	25866
% of withdrawn homes that withdraw in week 9	2.79%	2.66%	25866
% of withdrawn homes that withdraw in week 10	2.57%	2.60%	25866
% of withdrawn homes that withdraw in week 11	2.56%	2.60%	25866
% of withdrawn homes that withdraw in week 12	2.96%	2.56%	25866
% of withdrawn homes that withdraw in week 13	4.12%	2.55%	25866
% of withdrawn homes that withdraw in week 14	2.89%	2.43%	25866
% of withdrawn homes that withdraw in week 15	2.39%	2.43%	25866
% of withdrawn homes that withdraw in week 16	2.33%	2.45%	25866
% of withdrawn homes that withdraw in week 17	2.60%	2.40%	25866
% of withdrawn homes that withdraw in week 18	2.58%	2.33%	25866
% of withdrawn homes that withdraw in week 19	1.96%	2.23%	25866
% of withdrawn homes that withdraw in week 20	1.96%	2.20%	25866
% of withdrawn homes that withdraw in week 21	2.17%	2.19%	25866
% of withdrawn homes that withdraw in week 22	2.39%	2.08%	25866
% of withdrawn homes that withdraw in week 23	1.89%	2.04%	25866
% of withdrawn homes that withdraw in week 24	1.98%	1.99%	25866
% of withdrawn homes that withdraw in week 25	2.68%	1.89%	25866
% of withdrawn homes that withdraw in week >25	37.96%	37.99%	25866
% withdraw	55.95%	56.39%	17069
List price at delisting - initial list price for sales	-10.44%	-10.96%	19788
List price at delisting - initial list price for withdrawals	-8.02%	-7.37%	21477

FIGURE 3.13: Moments Used in Estimation Continued

Sales Price	List Price in week 1	List Price at delisting	Number of Offers	% Reduction in uncertainty	Bias About $\mu$ at delisting	% of Total Potential Sellers	% Sell	Average $v^w$ Among Potential Sellers
<i>Unmotivated Seller</i>								
\$303,120	\$351,560	\$372,286	0.7879	7.92%	7.27%	53.85%	10.09%	\$264,943
<i>Motivated Seller</i>								
\$220,540	\$283,679	\$237,435	2.8696	26.18%	3.48%	46.15%	85.73%	\$8,467

Note: All dollar figures are normalized by the mean of the valuation distribution. For example, Sales Price = 100,000 implies that prices are \$100,000 higher than the average buyer valuation.

FIGURE 3.14: Additional Results from Model. All Results are Averages Across Sellers.

Description of How Prior on $\mu$ is Formed	Seller Prior on $\mu$		Welfare Gain Relative to (1)				Mean Gain as Percent of Average Sales Price <sup>1</sup>
	Bias	Std Dev	Mean	25th Pctile	50th Pctile	75th Pctile	
<b>Variations on Estimated Prior (Welfare Figures Assume 6 Percent Realtor Commission)</b>							
(1) Estimated Prior at Average $S$	0.0362	0.0741	--	--	--	--	--
(2) Estimated Prior, No Bias at Average $S$	0.0000	0.0741	\$1,994	\$557	\$1,404	\$3,803	0.45%
(3) Estimated Prior at 10th Pctile of $S$	0.0362	0.0657	\$964	\$516	\$1,079	\$1,412	0.22%
(4) Estimated Prior at 90th Pctile of $S$	0.0362	0.0848	-\$990	-\$1,432	-\$1,093	-\$556	-0.22%
(5) Full Information	0.0000	0.0000	\$8,658	\$3,493	\$7,691	\$12,333	1.95%
<b>Alternative Demand Predictors (Welfare Figures Assume 3 Percent Realtor Commission)</b>							
(6) Zillow.com ( <i>Zestimates</i> )	0.0000	0.0971	-\$4,135	-\$11,361	-\$7,500	\$872	-0.93%
(7) Case-Shiller Data	0.0147	0.1271	-\$14,256	-\$24,641	-\$18,641	-\$5,620	-3.21%
(8) FHFA Data	0.0149	0.1680	-\$22,482	-\$32,887	-\$27,523	-\$15,930	-5.06%
(9) Hedonic Analysis; County, Yr, Sqft. Controls	0.0466	0.3629	-\$45,333	-\$48,928	-\$46,527	-\$41,041	-10.21%

<sup>1</sup>The average sales price when sellers have the prior in (1) is \$444,000 in this simulation.

Note: The labels for the alternative demand predictors refer to the data source used to form demand expectations. See appendix for a description of the assumptions used to calculate the parameters of each prior. The bias from zillow.com is conservatively set to zero because I do not have information to calculate biases in *Zestimates*.

FIGURE 3.15: Seller Welfare Under Alternative Priors Relative to Welfare at Estimated Prior

Dependent Variable: Annual Price Changes (Prices in Logs)

	Prices Not Aggregated		Prices Aggregated at Quarterly Level			
<b>OLS Estimates of AR(1) Coefficient</b>	-0.033	0.105	0.204	0.010	0.128	0.235
<b>Assumptions</b>						
Uncertainty Over Weekly Price Changes		x	x		x	x
Initial Beliefs Depend on Previous Comparable Sales			x			x

Note: The mean of the valuation distribution is assumed to follow a random walk at the weekly level with a drift of -.003. Sellers are assumed to have rational expectations about price changes during the selling horizon. All parameters are fixed at their estimated values. Simulated data is for 48,000 weeks and 20,000 sellers at each week.

FIGURE 3.16: Sales Price Regressions from Simulated Model

San Francisco and Los Angeles Samples Combined  
t-statistics in italics

Regressors	Dependent Variable			
	(1) List Price	(2) Sales Price	(3) TOM	(4) Withdraw
Deciles of Price Change Distribution ( $\Delta_6$ )				
d <sub>2</sub>	<b>0.006</b> <i>2.08</i>	<b>-0.012</b> <i>-3.93</i>	3.163 <i>12.43</i>	0.062 <i>12.71</i>
d <sub>3</sub>	0.006 <i>1.91</i>	-0.023 <i>-7.18</i>	4.651 <i>18.15</i>	0.075 <i>15.28</i>
d <sub>4</sub>	<b>0.011</b> <i>3.43</i>	<b>-0.030</b> <i>-9.31</i>	5.791 <i>22.37</i>	0.092 <i>18.81</i>
d <sub>5</sub>	0.026 <i>8.05</i>	-0.016 <i>-4.66</i>	7.001 <i>25.85</i>	0.101 <i>19.99</i>
d <sub>6</sub>	0.026 <i>7.60</i>	-0.024 <i>-7.03</i>	6.897 <i>24.17</i>	0.105 <i>19.85</i>
d <sub>7</sub>	0.025 <i>7.00</i>	-0.019 <i>-5.23</i>	6.775 <i>22.64</i>	0.099 <i>17.99</i>
d <sub>8</sub>	0.021 <i>5.81</i>	-0.031 <i>-8.76</i>	6.227 <i>20.42</i>	0.118 <i>21.18</i>
d <sub>9</sub>	0.019 <i>5.29</i>	-0.040 <i>-11.15</i>	6.269 <i>20.07</i>	0.127 <i>22.21</i>
d <sub>10</sub>	0.043 <i>11.88</i>	-0.010 <i>-2.94</i>	1.371 <i>4.26</i>	0.074 <i>12.34</i>
Controls for Expected Sales Price			X	X
Controls for Change in Expected Price Over Selling Horizon				X
# of Observations	82277	82277	82277	169866

Bold coefficients indicate cases where the change in list price from moving up 1 decile is positive and significant and the change in sales price is negative and significant. All prices are measured in logs. All prices in the dependent variable are normalized by the expected sales price. TOM is measured in weeks. d1 is the excluded group. All specifications include quarter dummies and controls for whether the property is real estate owned (REO).

FIGURE 3.17: OLS Equations Illustrating Bias. t-statistics in italics.

San Francisco and Los Angeles Samples Combined  
t-statistics in italics

Regressors	Dependent Variable			
	(1) List Price	(2) Sales Price	(3) TOM	(4) Withdraw
S	0.100 <i>10.39</i>	0.006 <i>0.56</i>	-10.537 <i>-13.16</i>	0.189 <i>13.12</i>
Controls for Expected Sales Price			X	X
Controls for Change in Expected Price Over Selling Horizon				X
# of Observations	82277	82277	82277	169866

All prices are measured in logs. All prices in the dependent variable are normalized by the expected sales price. TOM is measured in weeks. All specifications include quarter dummies and controls for whether the property is real estate owned (REO). *S* is the standard deviation of the year built for all homes that sold within a .5 mile radius of house *i* in the 6 years prior to month *t*.

FIGURE 3.18: OLS Equations Illustrating Effects of Uncertainty. t-statistics in italics.



# Appendix A

## Appendix: Chapter 2

This appendix describes the estimation procedure used to produce the non-parametric plots.

In the first stage, I use the entire sample of repeat sales to estimate

$$p_{it} = v_i + \delta_t + w_{it} \tag{A.1}$$

where I assume that  $w_{it}$  is uncorrelated with the regressors. As described above, this assumption is a source of bias if  $LTV^*$  and  $LOSS^*$  affect prices because both variables are correlated with the right-hand-side variables.

With the estimates of  $\delta$  from the first stage, I calculate LTV as:

$$LTV_{it} = \frac{l_{it}}{\exp(p_{is} + \delta_t - \delta_s)} = \frac{l_{it}}{\exp(v_i + \delta_t + w_{is})} \tag{A.2}$$

where the second equality comes from substituting equation (A.1) for the actual selling price,  $p_{is}$ . Note that this definition of LTV differs from the true definition described above by  $w_{is}$  in the denominator.

Finally, I write down the second stage specification which states that the difference in the log selling price and the expected selling price is a linear function of

tenure controls, and some unknown function of LTV,  $f(LTV_{it})$ . I estimate  $f$  using partially linear regression. Intuitively, the idea is to partial out the non-parametric effect using 4th-differencing, get a consistent estimate of the linear parameters on tenure, calculate the residuals using the data with the non-parametric effect differenced out, and then use the residuals to non-parametrically estimate  $f$ . Yantchew (2000) provides the exact details. I use the optimal differencing weights reported in Yantchew (2000).

Now I summarize the steps used to run the simulation summarized in Table 3. First I describe the data generating process. By assuming that  $\epsilon_{it}$  is a normally distributed random variable that is iid over time and across houses, we can infer its mean and variance,  $\sigma_\epsilon^2$ , from the empirical distribution of the residuals from the first stage of the estimation routine<sup>1</sup>. I calculate that  $\sigma_\epsilon = .099$ .

I assume that the population distribution of house fixed effects is normal. To find its variance, I use the fact that  $\hat{v}_i = v_i + \frac{\sum_{t=1}^T \epsilon_{it}}{T}$  where  $T$  denotes the number of sales for house  $i$ . Using the variance of the empirical distribution of the estimated house fixed effects for houses that sold exactly twice and the estimate of  $\sigma_\epsilon^2$ , I calculate that  $std(v_i) = .445^2$ . I use these distributions to generate the data for the simulation. The other variables appearing in equation (4) are the loan amount and the quarters of sale dummies. To get the original loan amount, I sample directly from the empirical distribution of initial LTV ratios and multiply the LTV ratio by the purchase price. The outstanding loan amount is calculated by amortizing the mortgage as described above. During the years 1990-1997, I assume that a house sells in any given quarter with probability equal to 1/85. In the remaining years, I assume that the probability

---

<sup>1</sup> I use the empirical distribution for houses that sell exactly twice. When  $N$  is large and  $T = 2$ , the residual for the first sale for house  $i$  is  $\frac{1}{2}(\epsilon_{i1} - \epsilon_{i2})$ . Thus, I set  $\sigma_\epsilon^2$  equal to twice the variance of the estimated residuals from the first sale of houses that sell twice.

<sup>2</sup> When  $T$  is large, the empirical distribution of the estimated fixed effects exactly corresponds to the distribution of  $v_i$ .

of sale is 42 percent higher in each quarter based on the differences in sales volumes across hot and cold markets that I observe in the data. For a sample of  $N=1,150,497$  houses<sup>3</sup>, these probabilities of sale yield sample sizes in the second stage of estimation that are comparable to the number of observations reported in Table 4. It also leads to estimates of the house fixed effects based on a small number of sales. As discussed above, this is the potential source of bias that motivates the simulation exercise.

For the simulation, I assume that the coefficients on the quarter dummy variables are equal to the price indices reported by OFHEO for San Francisco, except I subtract 5 percent from the prices from 1990-1997 based on the results presented in section 7.5. Then, I simulate prices according to equation (5) for sales between 1988-1989 and 1998-2005, and I simulate prices according to equation (4) for the remainder of the sales. I run the estimation procedure described in Section 6 on the simulated data and report the results in Table 3.

---

<sup>3</sup> This corresponds to the number of observations that I observe in my sample as reported in Table 1.

# Appendix B

## Appendix: Chapter 3

### *B.0.1 Detail on Calculation of Expected Prices*

$\hat{p}_{it}$  is the log expected sales price for house  $i$  in month  $t$ .<sup>1</sup> 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

$$p_{ijt}^* = v_i + \delta_{jt} + \epsilon_{ijt} \quad (\text{B.1})$$

where  $v$  is a house fixed effect,  $\delta_{jt}$  is a neighborhood specific time dummy, and  $\epsilon_{ijt}$  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

---

<sup>1</sup> In the model, the time period is a week. I calculate weekly expected prices by distributing monthly price changes equally across weeks within a month.

for zip code  $i$  using

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

where  $\phi$  is the standard normal pdf,  $dist$  is the distance between the centroids of the zip codes  $i$  and  $j$ , and  $h$  is a bandwidth.<sup>2</sup> I use this weighting scheme because sometimes the number of sales in a particular zip code in a particular month is not large.

The second expected price,  $\hat{p}_{it}^L$ , is the log expected list price in the initial week of listing. If we had a longer time series of list prices, we could construct this price exactly as explained above, except substitute list prices in the initial week of listing for sales prices. Since I do not observe the previous log list price for each home listed during my sample, I proxy for it with  $p_0^* + \hat{\delta}_0$ , where  $p_0^*$  is the previous log sales price and  $\hat{\delta}_0$  is the sales price index (calculated from the repeat sales analysis described above) during the month of previous sale. Then, for each listing in my sample, I regress  $p^L - (p_0^* + \hat{\delta}_0)$  on month dummies, where  $p^L$  is the list price in the initial week of listing. I also use weighted regression here, although the small numbers problem is not as problematic because the number of sales is a subset of the number of listings.

### B.0.2 Proof of Theorem 1

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

$$(1 - \lambda) \int_{p^L}^{\infty} (v - p^L) \frac{1}{\sigma_{\hat{v}}} \phi\left(\frac{v - \hat{v}}{\sigma_{\hat{v}}}\right) dv \geq -\kappa \quad (\text{B.3})$$

---

<sup>2</sup> 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. I also include a dummy variable for foreclosures to capture the forced sales discount (see Campbell, Giglio, and Pathak (2009)).

where  $\phi$  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 (B.3) is increasing in  $\hat{v}$ . Using properties of the truncated normal distribution, we rewrite the integral as

$$(\hat{v} - p^L)(1 - \Phi(\frac{p^L - \hat{v}}{\sigma_{\hat{v}}})) + \sigma_{\hat{v}}\phi(\frac{p^L - \hat{v}}{\sigma_{\hat{v}}}) \quad (\text{B.4})$$

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

$$(1 - \Phi(\frac{p^L - \hat{v}}{\sigma_{\hat{v}}})) + (\hat{v} - p^L)\phi(\frac{p^L - \hat{v}}{\sigma_{\hat{v}}}) + (p^L - \hat{v})\phi(\frac{p^L - \hat{v}}{\sigma_{\hat{v}}}) = 1 - \Phi(\frac{p^L - \hat{v}}{\sigma_{\hat{v}}}) > 0. \quad (\text{B.5})$$

To show the particular form of  $\bar{v}$ , using properties of the truncated normal distribution, we rewrite equation (B.3) for  $\hat{v} = \bar{v}$  as

$$(1 - \lambda) \left[ (\bar{v} - p^L)(1 - \Phi(\frac{p^L - \bar{v}}{\sigma_{\bar{v}}})) + \sigma_{\bar{v}}\phi(\frac{p^L - \bar{v}}{\sigma_{\bar{v}}}) \right] + \kappa = 0. \quad (\text{B.6})$$

Let  $z$  be the left hand side of (B.6). It is clear from (B.6) 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 (B.6), we get

$$(1 - \lambda) \left[ (T^*)(1 - \Phi(\frac{-T^*}{\sigma_{\hat{v}}})) + \sigma_{\hat{v}}\phi(\frac{-T^*}{\sigma_{\hat{v}}}) \right] + \kappa = 0. \quad (\text{B.7})$$

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

### B.0.3 Detail on Alternative Demand Predictors

In this section, I describe how I generate the priors used in the simulations discussed in Section 7.2. I first describe the ‘back-of-the-envelope’ calculation used to

estimate the amount of uncertainty that arises from using zillow.com. I use data from zillow.com on how *Zestimates* compare to sales prices for homes sold in Los Angeles from July-September 2010. Zillow.com reports the percent of estimates that fall within 5 percent, 10 percent, and 20 percent of the sales price. Assuming that the distribution of the percentage difference between sales price and *Zestimate* is  $N(0, \sigma^2)$ , I estimate  $\sigma$  using method of moments with the weighting matrix set to the identity matrix and the three statistics reported from zillow.com as moments. This gives an estimate of 0.1414. However, this will overstate the amount of uncertainty for a given house because some amount of price dispersion arises from heterogeneity in holding costs, randomness in the search process, etc. Fortunately, the structural model provides an estimate of this dispersion. Then the amount of uncertainty is calculated as the residual between  $.144^2$  and the amount of price dispersion conditional on quality predicted by the model. I adjust the results slightly for the fact that *Zestimates* are for condo and single family home sales whereas my parameters are estimated off of a sample of single family home sales using the estimate of  $h_1$ .

I now discuss how I calculate the priors in specifications (7)-(9). In order to calculate  $\mu$ , sellers need to know  $p_0 + \delta_{jt} - \delta_{j0} + e_{it}$ , where  $p_0$  is the log purchase price,  $j$  indexes the seller's neighborhood, and  $t$  indexes the month.  $\delta_{jt} - \delta_{j0}$  measures the amount of appreciation or depreciation since the time of purchase. Appendix A.1 describes how I estimate these price indexes using a local linear repeat sales estimator.

There will be a prediction error associated with each of the alternative demand predictors in Table 6 because all of or part of the true  $\delta_{jt} - \delta_{j0}$  is unavailable. For example, the Case-Shiller price index is reported with a 2-month lag, so  $\delta_{jt}$  will be unavailable. Furthermore, the Case-Shiller index measures price changes at the metropolitan (not the zip-code) level.

I calculate the prediction error as

$$error = (\delta_{jt} - \delta_{j0}) - (\hat{\delta}_{jt} - \hat{\delta}_{j0}) \quad (\text{B.8})$$

for each zip-code in each month where  $\hat{\delta}$  indicates the time effect estimated under each alternative data source.<sup>3</sup> Since I estimate the model using a sample where home prices are falling, I calculate the prediction error for the months during the years 1990-1997 and 2007-2009 for Los Angeles and San Francisco. During these years, home prices were falling. I set the mean and standard deviation of the prediction error equal to the mean and standard deviation of the prior on  $\mu$ . However, there is also uncertainty over  $e_{it}$ . I assume that this is equal to the uncertainty that arises from using zillow.com. This is reasonable because zillow.com has the data to calculate localized price indexes, but it cannot in general account for renovations or other time-varying unobserved characteristics that would be captured in  $e_{it}$ .

#### *B.0.4 Expectation Bias: Additional Results*

Although correlation between  $\Delta_6$  and unobserved heterogeneity in motivation to sell or house quality cannot explain the price results in Table 8, correlation between  $\Delta_6$  and the amount of uncertainty about home values can generate similar pricing patterns. If the list price is viewed as a commitment to sell at that price as in Chen and Rosenthal (1996), then the theoretical explanation is that uncertain sellers should set high list prices in order to test the market before committing to sell at a price that may be too low. Uncertainty should have a cost, however, and this may be reflected in lower sales prices on average.

In Appendix Table 1, I test for these effects using  $S_{it}$  as a proxy for the amount of uncertainty facing each seller. The mean of  $S_{it}$  is 0.13 and the standard deviation

---

<sup>3</sup> I assume that the length since time of purchase is 5 years, but the results are not sensitive to this assumption.



is 0.07. Uncertainty slightly (but significantly) increases list prices and the propensity to withdraw; it has no effect on sales prices; and it slightly (but significantly) *decreases* TOM. The results are similar if the standard deviation of square feet is used instead of  $S$ . Assuming that any unobserved uncertainty that is correlated with  $\Delta_6$  affects selling behavior similarly to  $S$ , then the TOM and sales price results presented here suggest that the results in Table 8 cannot be explained by unobserved heterogeneity in the amount of uncertainty.

# Bibliography

- Akerberg, D. A. (2003), “Advertising, Learning, and Consumer Choice in Experience Good Markets: An Empirical Examination,” *International Economic Review*, 44, 1007–1040.
- Albrecht, J., Anderson, A., Smith, E., and Vroman, S. (2007), “OPPORTUNISTIC MATCHING IN THE HOUSING MARKET,” *International Economic Review*, 48, 641–664.
- Albrecht, J., Gautier, P., and Vroman, S. (2008), “Directed Search in the Housing Market,” *mimeo*, *Georgetown University*.
- Anenberg, E. (2011), “Loss aversion, equity constraints and seller behavior in the real estate market,” *Regional Science and Urban Economics*, 41, 67 – 76.
- Arnold, M. (1999), “Search, Bargaining, and Optimal Asking Prices,” *Real Estate Economics*, 27, 453–482.
- Bayer, P., McMillan, R., Murphy, A., and Timmins, C. (2008), “A Dynamic Model of the Housing Market,” *mimeo*, *Duke University*.
- Benitez-Silva, H., Eren, S., Heiland, F., and Jimenez-Martin, S. (2011), “How Well do Individuals Predict the Selling Prices of their Homes,” *mimeo*.
- Berk, E., Ickler, and Levine, R. A. (2007), “Bayesian demand updating in the lost sales newsvendor problem: A two-moment approximation,” *European Journal of Operational Research*, 182, 256 – 281.
- Bernheim, D. and Meer, J. (2007), “How Much Value do Real Estate Brokers Add: A Case Study,” *mimeo*, *Stanford University*.
- Campbell, J., Giglio, S., and Pathak, P. (2009), “Forced Sales and House Prices,” .
- Carrillo, P. (2010), “An Empirical Stationary Equilibrium Search Model of the Housing Market,” *forthcoming International Economic Review*.
- Case, K. E. and Shiller, R. J. (1989), “The Efficiency of the Market for Single-Family Homes,” *The American Economic Review*, 79, 125–137.

- Chan, S. (1997), "Spatial Lock-in: Do Falling House Prices Constrain Residential Mobility?" *Journal of Urban Economics*, 49, 567–568.
- Chen, Y. and Rosenthal, R. W. (1996a), "Asking Prices as Commitment Devices," *International Economic Review*, 37, 129–155.
- Chen, Y. and Rosenthal, R. W. (1996b), "On the Use of Ceiling-Price Commitments by Monopolists," *The RAND Journal of Economics*, 27, pp. 207–220.
- Cho, M. (1996), "House Price Dynamics: A Survey of Theoretical and Empirical Issues," *Journal of Housing Research*, 7, 145–172.
- Crawford, G. S. and Shum, M. (2005), "Uncertainty and Learning in Pharmaceutical Demand," *Econometrica*, 73, 1137–1173.
- DeGroot, M. (1970) *Optimal Statistical Decisions*.
- Engelberg, J. and Parsons, C. (2010), "Learning from Prices: Evidence from Real Estate Transactions," *mimeo*, *University of North Carolina*.
- Erdem, T. and Keane, M. P. (1996), "Decision-Making under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets," *Marketing Science*, 15, pp. 1–20.
- Genesove, D. and Mayer, C. (2001), "Loss Aversion and Seller Behavior: Evidence from the Housing Market," *The Quarterly Journal of Economics*, 116, 1233–1260.
- Genesove, D. and Mayer, C. J. (1997), "Equity and Time to Sale in the Real Estate Market," *The American Economic Review*, 87, 255–269.
- Glaeser, E. and Gyourko, J. (2007), "Housing Dynamics," *mimeo*, *Harvard University*.
- Glower, M., Haurin, D., and Hendershott, P. (1998), "Selling Time and Selling Price: the Influence of Seller Motivation," *Real Estate Economics*, 26, 719–740.
- Goettler, R. L. and Clay, K. (July 2009), "Tariff Choice with Consumer Learning and Switching Costs," *mimeo*, *University of Chicago*.
- Han, L. and Hong, S.-H. (2010), "Testing Cost Inefficiency under Free Entry in the Real Estate Brokerage Industry," *mimeo*, *University of Toronto*.
- Haurin, D. (1988), "The Duration of Marketing Time of Residential Housing," *Real Estate Economics*, 16, 396–410.
- Hendel, I., Nevo, A., and Ortalo-Magne, F. (2009), "The Relative Performance of Real Estate Marketing Platforms: MLS versus FSBOMadison.com," *American Economic Review*, 99, 1878–98.

- Hitsch, G. J. (2006), "An Empirical Model of Optimal Dynamic Product Launch and Exit Under Demand Uncertainty," *Marketing Science*, 25, 25–50.
- Horowitz, J. L. (1992), "The role of the list price in housing markets: Theory and an econometric model," *Journal of Applied Econometrics*, 7, 115–129.
- Hsieh, C.-T. and Moretti, E. (2003), "Can Free Entry Be Inefficient? Fixed Commissions and Social Waste in the Real Estate Industry," *The Journal of Political Economy*, 111, pp. 1076–1122.
- Jia, P. and Pathak, P. A. (2010), "The Impact of Commissions on Home Sales in Greater Boston," *American Economic Review*, 100, 475–79.
- Kahneman, D. and Tversky, A. (1979), "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47, 263–291.
- Keane, M. P. and Wolpin, K. I. (1994), "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence," *The Review of Economics and Statistics*, 76, 648–672.
- Knight, J. R. (1996), "Listing Price, Time on Market, and Ultimate Selling Price: Causes and Effects of Listing Price Changes," *Real Estate Economics*, 30, 213–237.
- Kuzmenko, T. (2010), "The Accuracy of Owner-Provided House Values: Evidence from the Census Data," *mimeo Duke University*.
- Lamont, O. and Stein, J. C. (1999), "Leverage and House-Price Dynamics in U.S. Cities," *The RAND Journal of Economics*, 30, 498–514.
- Lazear, E. P. (1986), "Retail Pricing and Clearance Sales," *The American Economic Review*, 76, 14–32.
- Lee, D. and Wolpin, K. I. (2010), "Accounting for wage and employment changes in the US from 1968-2000: A dynamic model of labor market equilibrium," *Journal of Econometrics*, 156, 68 – 85.
- Levitt, S. D. and Syverson, C. (2008), "Market Distortions When Agents Are Better Informed: The Value of Information in Real Estate Transactions," *Review of Economics and Statistics*, 90, 599–611.
- Lewis, M. (forthcoming), "Asymmetric Price Adjustment and Consumer Search: An Examination of the Retail Gasoline Market," *Journal of Economics and Management Strategy*.
- Linneman, P. and Wachter, S. (1989), "The Impacts of Borrowing Constraints on Homeownership," *Real Estate Economics*, 17, 389–402.
- Ljungqvist, L. and Sargent, T. (2004) *Recursive Macroeconomic Theory*.

- McCall, J. J. (1970), "Economics of Information and Job Search," *The Quarterly Journal of Economics*, 84, pp. 113–126.
- Merlo, A. and Ortalo-Magne, F. (2004), "Bargaining over residential real estate: evidence from England," *Journal of Urban Economics*, 56, 192–216.
- Murphy, A. (2008), "A Dynamic Model of Housing Supply," *mimeo*, Washington University in St. Louis.
- Narayanan, S., Chintagunta, P., and Miravete, E. (2007), "The role of self selection, usage uncertainty and learning in the demand for local telephone service," *Quantitative Marketing and Economics*, 5, 1–34.
- Ortalo-Magne, F. and Rady, S. (2006), "Housing Market Dynamics: On the Contribution of Income Shocks and Credit Constraints," *Review of Economic Studies*, 73, 459–485.
- Pagan, A. and Ullah, A. (1999), "Nonparametric Econometrics," *Cambridge University Press: Cambridge, UK*.
- Ross, S. L. (1989), "Mortgage Lending, Sample Selection and Default," *Real Estate Economics*, 28, 581–621.
- Shiller, R. J. (1991), "Arithmetic repeat sales price estimators," *Journal of Housing Economics*, 1, 110–126.
- Silverman, B. (1986), "Density Estimation for Statistics and Data Analysis," *Mono-graphs on Statistics and Applied Probability*.
- Taylor, C. R. (1999), "Time-on-the-Market as a Sign of Quality," *Review of Economic Studies*, 66, 555–78.
- Tversky, A. and Kahneman, D. (1991), "Loss Aversion in Riskless Choice: A Reference-Dependent Model," *The Quarterly Journal of Economics*, 106, 1039–1061.
- Yang, H. and Ye, L. (2008), "Search with learning: understanding asymmetric price adjustments," *The RAND Journal of Economics*, 39, pp. 547–564.
- Yatchew, A. (2000), "Scale Economies in Electricity Distribution: A Semiparametric Analysis," *Journal of Applied Econometrics*, 15, 187–210.

# Biography

Elliot Anenberg was born in Washington DC on June 11, 1982.

He holds a B.A. in Economics from Northwestern University, earned in 2004; a M.A. in Economics from Duke University, earned in 2007; and a Ph.D. in Economics from Duke University, earned in 2011.

After graduation, he will work as an Economist at the Federal Reserve Board of Governors.