Price and Time to Sale Dynamics in the Housing Market: the Role of Incomplete Information

Gabriel Ehrlich†
Congressional Budget Office
November 8, 2012

Abstract

I propose a stylized model of the house-selling process in which sellers possess incomplete information regarding the state of the housing market. Consistent with the data, the model generates a negative correlation between house prices and time on market. This result is robust to the presence of real estate agents with complete information. I construct a measure of homeowner perceptions of house prices and compare them to market-based house price indices to form a ‘misperceptions index’ for house prices. An increase in homeowners’ perceptions relative to actual prices is associated with a decrease in sales volumes, with price misperceptions explaining over one-fifth of the within-state variation in sales volumes from 2000 to 2010.

1 Introduction

It is a stylized fact of the market for existing homes that there is a strong positive correlation between sales prices and sales volumes and a negative correlation between sales prices and the average

*The views expressed in this paper are the author’s and should not be interpreted as the views of the Congressional Budget Office.
†I would like to thank David Albouy, Bob Barsky, Ruediger Bachmann, Dennis Capozza, Chris House, and Stephan Lauermann for their helpful comments.
time houses take to sell. Figure 1 illustrates this pattern using monthly data for the United States from January 2000 to December 2010.\(^1\) The top panel shows the log of the CoreLogic House Price Index for single family detached homes, the middle panel illustrates log single family home sales, and the bottom panel displays months’ supply of single family homes for sale (hereafter months’ supply).\(^2\) The correlation coefficient between prices and sales volumes during this period was 0.80 while the correlation coefficient between prices and months’ supply was -0.19; both correlations are statistically significant at the 5-percent level.

Although direct data on the time it takes houses to sell (hereafter time on market or time to sale) is not available at a national level, studies on the city level find a similar pattern when considering time on market directly. For instance, Genesove and Mayer (2001) document that in the Boston condominium market, fewer than 30 percent of listed units sold within 180 days during the trough year of 1992, while in 1997, after the market had recovered, more than 60 percent of new listings sold within 180 days. Miller and Sklarz (1986) document similar trends in Hawaii and Salt Lake City. Regarding the correlation between prices sales volumes, Stein (1995) shows that between 1986 and 1992, a 10 percent drop in prices reduces sales volumes by approximately 1.6 million units in a time span with average volume of three to four million units. Ortalo-Magne and Rady (1998) show the same qualitative pattern holds in the U.K. The robustness of this pattern leads Genesove and Mayer (1997) to describe it as “one of the most distinctive and puzzling macro features of the market for existing homes.”

This paper offers a stylized model to explain this pattern with three key features: first, that sellers have incomplete information regarding conditions in the housing market; second, that sellers face idiosyncratic variation in the offers they receive; and third, that misalignment between sellers’

---

\(^1\)Sales volume and months’ supply data are from the National Association of Realtors and price data is from CoreLogic.

\(^2\)Months’ supply is the ratio of the number of homes listed as being for sale at the end of the month divided by the number of sales that month. It is used as a proxy for the average time on market because nationally representative data for time on market is unavailable. All series have had a linear time trend removed, and prices were deflated using the CPI.
and realtors’ incentives prevents realtors from conveying their knowledge of market conditions to sellers efficiently.

The first assumption seems natural given most households’ infrequent participation in the housing market. Anily et al. (1999) report that the expected total time for a household to reside in an owner-occupied unit is 13 years. Case and Shiller (2003) note the relative absence of professional traders in the housing market: “Buyers and sellers in the housing market are overwhelmingly amateurs, who have little experience with trading. High transactions costs, moral hazard problems, and government subsidization of owner-occupied homes have kept professional speculators out of the market.” Furthermore, the idiosyncratic nature of housing units creates thinness in the market, leading to difficulty in inferring the market value of an individual house from recent sales prices of other units.

Idiosyncratic variation in the offers a seller receives may arise from differences in the match quality between the potential buyer and the house, variations in buyers’ eagerness to transact quickly, or other factors. Although there is limited data concerning the distributions of offers sellers receive, Merlo and Ortalo-Magne (2004) document both that offer amounts vary within a negotiation between a buyer and a seller, and that later bidders for a house tend to make higher offers than earlier bidders. The existence of idiosyncratic variation in offers seems well-accepted in the theoretical literature. For instance, Haurin (1988) discusses the optimal decision rule for sellers facing particular offer distributions.

Finally, recent research by Levitt and Syverson (2002, 2008) has shown that because realtors’ typical compensation structure leads to misalignment between realtors’ and sellers’ incentives, realtors may not be able to convey their knowledge of market conditions to sellers efficiently. They show that empirically, realtors seem to encourage their clients to sell their homes more quickly than would be optimal for a fully informed seller. The model in this paper demonstrates that in such an environment, expected time to sale remains negatively correlated with house prices even in the
presence of fully informed realtors.

I construct a measure of homeowner perceptions of housing market conditions using data from the American Community Survey and compare it to indices of actual market prices to derive a measure of homeowner misperceptions of market conditions. I find that elevated misperceptions of house prices are associated with lower sales volumes at both the state and Metropolitan Statistical Area (MSA) levels. A simple fixed effects regression of sales volumes on the misperceptions index explains 22% of the time series variation in sales volumes within states and 19% of the variation within MSAs. Furthermore, variation in the misperceptions index is related to volatility and persistence in prices and a measure of the difficulty of inferring house prices from observable characteristics.

A substantial literature attempts to explain the time series correlations discussed in the paper. Magnus (2010) and Ehrlich (2012) show that the strong correlations between prices, sales, and time on market are difficult to replicate in a general equilibrium search and matching model of the housing market. Stein (1995) models the role of downpayment requirements and credit constraints in generating these correlations, which Genesove and Mayer (1997) document empirically. Genesove and Mayer (2001) argue that prospect theory can explain the correlations if sellers use the price they paid for their house as a reference point to evaluate offers. They provide evidence that this reference point influences seller behavior. Albrecht et al. (2007) model a process in which sellers become more anxious to sell the longer the sales process takes, and therefore accept lower offers when a house has sat on the market for a long time. Finally, Lazear (2010) argues that consistent with the theory of monopoly pricing, sellers find it optimal to accept a lower probability of sale when demand, and therefore prices, fall.

The model in the present paper departs from the previous literature in assuming that sellers have imperfect information regarding the state of the housing market. It demonstrates that relaxing the assumption of perfect information can generate the observed patterns in the data, even when some
agents in the model do have full information. I show that empirically, homeowners’ misperceptions regarding housing values are correlated with changes in sales volumes: a one percent increase in homeowner perceptions of prices relative to market prices reduces sales volumes 1.2 percent in my baseline specification.

2 The Model

The model I consider is a partial equilibrium model in the sense that the distribution of offers that sellers receive is exogenous, and sellers receive exactly one offer per period. There is no negotiation in the model, leaving sellers with a choice only about whether to accept or reject a given offer. Once a seller has rejected an offer, they cannot recall it. I consider the situation in which sellers have no outside assistance in making their decisions, and the situation in which they are assisted by a real estate agent. I assume the agent has complete information regarding the state of the housing market but has incentives that are not entirely aligned with the seller’s. In the body of the paper I consider a two-period model; in the appendix I consider extending the model to multiple periods. In all cases, I assume that in the last period, the seller must sell the house for the final offered price.

2.1 The Offer Distribution

Offers are the sum of an ‘aggregate demand’ component, \( z \), which is constant across periods, and an idiosyncratic component, \( x_t \), which is distributed i.i.d. across periods. \( z \) is distributed \( U[z_L, z_H] \), and \( x_t \) is distributed \( U[x_L, x_H] \). \( z \) and \( x_t \) are independently distributed. Denoting the period \( t \) offer as \( \psi_t \),

\[
\psi_t = z + x_t
\]
is distributed according to a ‘modified triangular distribution’.\(^3\)

There is limited empirical evidence on the distribution of offers that home sellers receive, although Merlo and Ortalo-Magne (2004), using data from England, document that there is economically meaningful variation in the bids a seller receives on a given property. Haurin (1988) considers both uniform and normal distributions in his analysis of optimal seller behavior. The distribution analyzed here is chosen primarily for illustrative rather than for maximum realism. The essential feature of the distribution is the presence both of aggregate and of idiosyncratic variation in the offers.

2.2 The Model with no Realtor

First I consider a two-period model with no realtor. As outlined above, sellers receive one offer each period, and if they do not accept the first offer they must accept the second. I assume sellers are risk-neutral, perfectly patient, and do not bear any flow cost of leaving their homes on the market. Therefore, a seller’s goal is simply to obtain the highest possible price for their home. Accordingly, a seller will accept the period 1 offer, \(\psi_1\), if and only if it is greater than or equal to the expectation of the period 2 offer, \(E[\psi_2]\). Sellers cannot observe the state of market demand \(z\) or the idiosyncratic component of the offer \(x_t\) directly, so they must infer \(z\) using Bayes’ Theorem:

\[
f_Z(z|\psi) = \frac{f_{Z,\psi}(z,\psi)}{f_\psi(\psi)} = \frac{f_\psi(\psi|z)f_Z(z)}{f_\psi(\psi)}
\]

Define \(\tilde{z}_{L,1} = \max(z_L, \psi_1 - x_H)\) and \(\tilde{z}_{H,1} = \min(z_H, \psi_1 - x_L)\). Then the seller’s posterior belief about the distribution of \(z\) conditional on \(\psi_1\) is \(z \sim U[\tilde{z}_{L,1}, \tilde{z}_{H,1}]\).\(^4\) Knowing the seller’s belief about \(z\) as a function of \(\psi_1\) allows us to calculate the seller’s expectation of \(\psi_2\). The seller’s conditional expectation function, \(E[\psi_2|\psi_1]\), is plotted in Figure 2 for two cases, the first in which

---

\(^3\)The p.d.f. follows directly from the sum of two independent uniformly distributed variables and is defined in the appendix.

\(^4\)I demonstrate this result along with proofs of the propositions in the appendix.
\( x_H - x_L \geq z_H - z_L \), and the second in which \( z_H - z_L > x_H - x_L \), and is defined explicitly in the appendix. As the figure indicates, the seller’s expectation of \( \psi_2 \) is a weakly increasing function of \( \psi_1 \). In the first case, \( E[\psi_2|\psi_1] \geq \psi_1 \) when \( \psi_1 \leq \bar{\psi}_1 \); in the second case, \( E[\psi_2|\psi_1] \geq \psi_1 \) when \( \psi_1 \leq x_H + z_L \).

This allows us to state the main result of the model with no realtor.

**Proposition 1**: When the seller follows the optimal policy, the expected time to sale in the model with no realtor is weakly decreasing in the state of aggregate demand, \( z \), while the expected sales price is strictly increasing in \( z \).

Figure 3 illustrates the expected time to sale and expected sales price in the model with no realtor for the two cases. If the variance of the idiosyncratic component of the offer is greater than the variance of aggregate demand, then the expected time to sale will be falling for all values of \( z \). If the variance of aggregate demand is greater, however, the expected time to sale will be constant for very low and very high values of aggregate demand, while for intermediate values expected time to sale will be falling in \( z \). Therefore, the model implies that expected time to sale will be negatively correlated with expected prices.

Although this model is quite simple, it illustrates the essential mechanism by which incomplete information generates a negative correlation between the strength of housing demand and the expected time to sale. When sellers are uncertain about the state of aggregate demand, they follow a reservation price strategy in period 1, whereas if they could observe \( z \), they would ignore the aggregate demand component of the offer, which is stable over time, and make their decision based solely on the idiosyncratic component, \( x_1 \). With incomplete information about aggregate demand, offers with middling values of \( x_1 \) will be above the seller’s reservation price when \( z \) is high and below the reservation price when \( z \) is low. Therefore, the probability that the seller accepts the
first period offer is an increasing function of $z$.

### 2.3 A Two-Period Model with a Realtor

I now introduce a realtor to the two-period model. I assume the realtor can observe aggregate housing demand $z$ directly. However, in the model the realtor’s and the seller’s incentives are only partially aligned, so the realtor will not always be able to signal his knowledge of $z$ to the seller in a credible way. In the model, the only possible contract between sellers and realtors is one in which the realtor receives a fixed fraction $\alpha$ of the sales price. Furthermore, I assume the seller must employ a realtor when selling the house. If there is no sale in period 1, the realtor must pay flow cost $c_R$ at the beginning of period 2 in order to market the house.\footnote{Implicitly, one could imagine that $c_R$ is a flow cost the realtor must pay at the beginning of each period to market the house, but the cost paid at the beginning of period 1 is sunk and does not affect the realtor’s maximization problem.}

The realtor can only communicate with the seller by advising him on whether to accept an offer after it has been received. This assumption is less restrictive than it may seem. As Levitt and Syverson (2002) argue, if the realtor is constrained to advising the seller only after an offer has been received, there is no way for the realtor to report the intensity of his preferences credibly. Therefore it is not overly restrictive to limit the realtor to recommending either ‘accept’ or ‘reject’ after each offer is received.

To fix terminology, denote the realtor’s recommendation about the time $t$ offer as $\xi_t$, with $\xi_t = 0$ if the realtor recommends ‘reject’ and $\xi_t = 1$ if the realtor recommends ‘accept’. Let $\tilde{f}(\psi_2|\psi_1, \xi_1)$ be the seller’s posterior belief about the distribution of $\psi_2$ conditional on the first period offer $\psi_1$ and the realtor’s recommendation $\xi_1$. Finally, call the seller’s period 1 policy function $\gamma_1(\psi_1, \tilde{f}(\psi_2|\psi_1, \xi_1))$, with $\gamma_1 = 0$ indicating that the seller rejects the period 1 offer and $\gamma_1 = 1$ indicating that the seller accepts the period 1 offer. Then the seller’s period 1 and period 2
value functions can be written as:

\[ V_S(\psi_1, \xi_1, 1) = \max \{ \gamma_1(1 - \alpha)\{\psi_1, E[\hat{f}(\psi_2|\psi_1, \xi_1)]\} \} \]

\[ V_S(\psi_2, \xi_2, 2) = (1 - \alpha)\psi_2 \]

The realtor’s value functions can be written:

\[ V_R(\psi_1, z, 1) = \max_{\xi_1} \{ \gamma_1(\psi_1, \hat{f}(\psi_2|\psi_1, 0)\alpha\psi_1 + (1 - \gamma_1(\psi_1, \hat{f}(\psi_2|\psi_1, 0)))(E[V_R(\psi_2, z, 2)]) \} \]

\[ \gamma_1(\psi_1, \hat{f}(\psi_2|\psi_1, 1)\alpha\psi_1 + (1 - \gamma_1(\psi_1, \hat{f}(\psi_2|\psi_1, 1)))(E[V_R(\psi_2, z, 2)]) \} \]

\[ V_R(\psi_2, z, 2) = -c_R + \alpha\psi_2 \]

It is then natural to define a Bayesian Nash equilibrium of the game between realtors and sellers as a policy function \( \xi_1(\psi_1, z) \) for the realtor, a policy function \( \gamma_1(\psi_1, \hat{f}(\psi_2|\psi_1, \xi_1)) \) for the seller, and a belief updating strategy \( \hat{f}(\psi_2|\psi_1, \xi_1) \) for the seller such that:

1. \( \xi_1(\psi_1, z) \) maximizes the realtor’s value function for all \((\psi_1, z)\), taking \( \gamma_1(\psi_1, \hat{f}(\psi_2|\psi_1, \xi_1)) \) as given;

2. \( \gamma_1(\psi_1, \hat{f}(\psi_2|\psi_1, \xi_1)) \) maximizes the seller’s value function for all \(\psi_1, \xi_1\) taking \(\xi_1(\psi_1, z)\) as given; and

3. \( \hat{f}(\psi_2|\psi_1, \xi_1) \) is consistent with \(\xi_1(\psi_1, z)\).

Consider what the realtor would choose if he could decide which offers the seller would accept and which he would reject. After some algebra, the realtor’s value function in period 1 can be re-written

\[ V_R(\psi_1, z, 1) = \alpha z + \max\{\alpha x_1, \alpha \bar{x} - c_r\} \]
Therefore, in period 1 the realtor would prefer that the seller accept any offer \( \psi_1 \) such that \( \alpha x_1 \geq \alpha \bar{x} - c_r \), or equivalently, \( x_1 \geq \bar{x} - \frac{c_r}{\alpha} \). Denote \( \bar{x} - \frac{c_r}{\alpha} \) as \( \hat{x}_1 \), which represents the realtor’s cutoff value of \( x_1 \), i.e. the realtor would like the seller to accept all offers with an idiosyncratic component \( x_1 \) above \( \hat{x}_1 \) and reject all others.\(^6\) Note that \( \frac{c_r}{\alpha} \) measures the degree of misalignment between the seller’s and the realtor’s incentives. If \( \frac{c_r}{\alpha} \) were zero, the realtor’s and seller’s incentives would be perfectly aligned.

The following is a Bayesian Nash equilibrium of the game between the realtor and the seller. The realtor reports ‘accept’ (\( \xi_1 = 1 \)) for any offer such that \( x_1 \geq \hat{x}_1 \) and ‘reject’ (\( \xi_1 = 0 \)) for any offer such that \( x_1 < \hat{x}_1 \). Define \( \tilde{x}_{L,1} \) as \( x_L \) if the realtor recommends reject and \( \hat{x}_1 \) if the realtor recommends accept, and define \( \tilde{x}_{H,1} \) as \( \hat{x}_1 \) if the realtor recommends reject and \( x_H \) if the realtor recommends accept. Further, define \( \tilde{z}_{L,1} = \max(z_L, \psi_1 - \tilde{x}_{H,1}) \) and \( \tilde{z}_{H,1} = \min(z_H, \psi_1 - \tilde{x}_{L,1}) \). Then the seller’s posterior belief is that \( z \sim U[\tilde{z}_{L,1}, \tilde{z}_{H,1}] \).\(^7\)

Figure 4 displays the seller’s expectation of \( \psi_2 \) as a function of \( \psi_1 \) and the realtor’s recommendation \( \xi_1 \). The red lines show the expectation when the realtor recommends reject and the green lines show the expectation when the realtor recommends accept; the dashed portions of the red and green lines show off-equilibrium path recommendations. As before, I show two cases, differentiated by whether \( z \) or \( x \) has the higher variance conditional on the realtor’s recommendation. The figure illustrates that the seller’s expectation of \( \psi_2 \) is always greater than \( \psi_1 \) when the realtor recommends ‘reject’. Therefore, the seller will always reject the first offer when the realtor recommends it. However, the opposite is not true: the seller will sometimes reject an offer when the realtor recommends ‘accept’. These cases arise when the seller’s expectation of \( \psi_2 \) is greater than \( \psi_1 \) despite the realtor’s recommendation to accept the offer. They are illustrated in the figure as values of \( \psi_1 \) for which the green line is above the dashed blue line.

\(^6\)I assume that \( \alpha \) and \( c_R \) are such that \( \hat{x}_1 \geq x_L \).

\(^7\)I do not currently offer a proof but it would follow the proof of the seller’s belief updating strategy with no realtor very closely.
To verify that reporting his own preference truthfully is a best response for the realtor, note that in equilibrium, the realtor’s recommendation weakly increases the chance that the seller will take the realtor’s preferred action. For very high and very low values of $\psi_1$, the realtor’s recommendation will not affect the seller’s decision, so any policy the realtor follows is a best response. However, for medium values of $\psi_1$, the expected value of $\psi_2$ when the realtor recommends ‘accept’ is below $\psi_1$, but the expected value of $\psi_2$ when the realtor recommends ‘reject’ is above $\psi_1$. For these offers, the realtor’s recommendation will be decisive - the seller will follow the realtor’s recommendation. Therefore, for this range of offers it must also be a best response for the realtor to report his own preference truthfully. This allows us to state Proposition 2, which shows that the main result from the model with no realtor continues to hold when the realtor is added to the model.

**Proposition 2:** In the equilibrium I consider, the expected time to sale in the model with a realtor is weakly decreasing in the state of aggregate demand, $z$, while the expected sales price is strictly increasing in $z$.

Figure 5 illustrates the expected time to sale and sales price as a function $z$. The expected sales price is a strictly increasing function of $z$, but the expected time on market is weakly decreasing with $z$. Specifically, when the variance of $z$ is large, there will be a range in which the expected time to sale does not depend on $z$.

3 **Empirical Work**

A key prediction of the model presented above is that the correlations between prices, time on market, and sales volumes are driven by seller confusion over the true state of the housing market. To test this prediction, I construct a measure of homeowner perceptions of housing values using data from the American Community Survey. I then compare this measure to several market indices.
of home prices to construct a measure of homeowner misperceptions of housing values at the state and MSA levels. Finally, I regress sales volumes on the misperceptions index to test whether homeowner misperceptions of the state of the housing market influence sales levels. I conclude by exploring the determinants of homeowner misperceptions about the housing market.

3.1 Constructing the Misperceptions Indices

I begin by constructing time series indices of perceived house values at the state and MSA levels. I use data from the American Community Survey one-percent national samples, in which homeowners were asked to report a rich set of physical characteristics of their homes, and also to answer the question:

About how much do you think this house and lot, apartment, or mobile home (and lot, if owned) would sell for if it were for sale?

Amount - Dollars

$ \quad \quad \quad 0.00$

The data is available yearly from 2000 to 2010 at the state level, but MSA level identifiers are only available from 2005-2010. At the state level, there are approximately 6.5 million home value observations in the data set, just over 100,000 in year 2000, between 300,000 and 400,000 per year between 2001 and 2004, and about 840,000 per year thereafter. Approximately 680,000 of these observations per year are located in MSAs from 2005 to 2010.

Within each geographical area $i$ I run regressions of the form:

$$Price_{ijt} = \alpha_i + \beta X_{ijt} + \delta_{it} + \epsilon_{ijt}$$  \hspace{1cm} (2)

where $Price_{ijt}$ is the log self-reported value of housing unit $j$ in area $i$ and year $t$, $X_{ijt}$ is a vector of housing characteristics, and $\delta_{it}$ is a set of year dummies. The $X_{ijt}$ comprise 9 indicators of
building size, 9 indicators for number of rooms, 5 indicators for number of bedrooms, the number of rooms interacted with the number of bedrooms, 2 indicators for lot size, 13 indicators for when the structure was built, 2 indicators for complete plumbing and kitchen facilities, an indicator for commercial use, and an indicator for condominium status. I take the estimated coefficients on the year dummies, \( \hat{\delta}_t \), to be my perceived housing value time series indices for each jurisdiction.\(^8\) Because I omit the dummy for the first year in the data set (2000 in the state-level regressions, 2005 in the MSA-level regressions), the house price perception indices can be interpreted as the percent change in perceived prices from the base year in an area.

I define the house price misperceptions indices as the difference between the house price perception indices and market-based house price indices, also expressed as log changes from the relevant base year. At the state level, the primary house price index I use is the Federal Housing Finance Administration’s (FHFA) purchase only index. Because the geographical coverage of the FHFA’s MSA-level purchase only house price index is limited, I employ the FHFA all-transactions index, which includes data from appraisals used in re-financings, as my base MSA-level house price index. Figures 9A and 9B show the perceived and market-based price indices for the 50 states,\(^9\) while figure 11 shows the indices for 19 of the 20 cities in the Case-Shiller 20-city house price index.\(^10\) The misperceptions index is the vertical distance between the solid black lines, representing perceived price changes, and the dashed red lines, representing actual price changes. A positive value for the misperceptions index implies that homeowners are “too bullish” regarding price changes in an area since the base year, while a negative value implies homeowners are “too bearish”.

As the figures illustrate, there is substantial variation in the accuracy of homeowners’ perceptions of price movements across areas and time. The standard deviation of the state-level

---

\(^8\)I drop the most and least expensive 5% of observations in each jurisdiction-year to control for outliers.  
\(^9\)Washington, D.C. is omitted from the figures to facilitate presentation.  
\(^10\)Charlotte, NC is omitted because of a lack of sales data.
Misperceptions index is 6.8 percent. The average misperceptions index ranges from a minimum of -16 percent in Hawaii to a maximum of 8.9 percent in Michigan, and a minimum of -2.6 percent in 2005 to a maximum of 5.7 percent in 2010. The time series pattern of the misperceptions index suggests that homeowners underestimated both the house price appreciation in the early part of the 2000s and the declines in the later part of the decade.

I employ alternative house price indices to calculate the misperceptions index in order to gauge the robustness of my results. At the state level, I use the FHFA all-transactions index, FHFA median price index, and the Zillow Home Value Index (ZHVI). At the MSA level, I use the FHFA purchase only index, the Case-Shiller 20-city repeat sales index, and the MSA level ZHVI.

3.2 Misperceptions and Home Sales

Figures 10A, 10B, and 12 show the misperceptions indices as black solid lines, alongside the log change in sales volumes from the base year in dashed red lines. To test the model’s implication that homeowner misperceptions of housing market conditions should be negatively correlated with sales volumes, I regress sales volumes on house price misperceptions at the state and MSA levels. At the state level, the sales volume data is from the National Association of Realtors, while at the MSA level it is from Zillow.com.

The baseline specification at both levels is a fixed effects regression of the form

$$Sales_{it} = \alpha + \beta_1 Misperceptions_{it} + \beta_2 Price_{it} + \gamma_i + \delta_t + \epsilon_{it}$$

where $Sales_{it}$ is the log number of single family homes sold in area $i$ during year $t$, $Misperceptions_{it}$ is the

\footnote{To capture the disparate sizes of different geographical areas, I weight the summary statistics by the number of housing units used in calculating the house price perception indices.}

\footnote{I calculate a median house price perceptions index from the ACS data using quantile regression to use in calculating the misperceptions index with the FHFA median price index and the ZHVI, which is also a median price index.}

\footnote{Although I present the log change in sales volumes from the base year in the figures for presentational purposes, I use the log level of sales volumes in the regressions.}
is the value of the misperceptions index, $Price_{it}$ is the value of the relevant house price index (expressed as the log change from the base year), $\gamma_i$ is a set of location fixed effects, and $\delta_t$ is a set of year fixed effects. In simpler specifications, I omit the house price index variable and the year fixed effects. I also estimate the model in first differences as

$$\Delta Sales_{it} = \alpha + \beta_1 \Delta Misperceptions_{it} + \beta_2 \Delta Price_{it} + \gamma_i + \delta_t + \epsilon_{it}$$ (4)

In my baseline specifications, I weight the regressions by the number of housing units in the ACS price perceptions regressions to account for the varying sizes of different geographic areas.

Table 1 shows fixed effects regressions at the state level over the years 2000 to 2010. Column 1 shows results from the simplest specification, without prices or year fixed effects. The coefficient on the misperceptions index is -1.4, implying that a one percent increase in homeowner perceptions of prices relative to actual prices decreases home sales by 1.4 percent. The $R^2$-within of this regression is 0.22, suggesting that homeowner misperceptions of market conditions alone can account for more than one-fifth of the time series variation in sales volumes within a state. In column 2, I add the house price index to the regression, as many theories of sales volumes, such as those emphasizing loss aversion or equity constraints, suggest house prices should influence sales volumes independently from the homeowner perceptions channel. Adding the house price index does not have a major effect on the results, as the estimated coefficient is not significantly different from zero, and the estimated coefficient on the misperceptions index is essentially unchanged. Column 3 shows the results of the baseline specification described in equation (3), with year fixed effects added to the specification in column 2. The estimated coefficient on the misperceptions index is slightly smaller at -1.2, but remains highly statistically significant. Meanwhile, the estimated coefficient on the house price index becomes significantly negative, consistent with theories of loss aversion and downpayment constraints. I take this specification as my baseline specification for the sensitivity analyses that follow. In columns 4 through 6, I estimate the three models described above in first difference
form, as in equation (4). The estimated coefficients on the misperceptions index are smaller in these specifications, but remain highly statistically significant. In column 6, which corresponds to the baseline specification estimated in first differences, the elasticity of sales volumes to house price misperceptions is estimated to be -0.64.

Table 2 presents results using different house price indices and weights to assess the robustness of the results in table 1. All columns use the same specification as column 3 of table 1, described in equation (3). Column 1 of table 2 re-displays the results from column 3 of table 1 for ease of comparison. Column 2 shows results using the FHFA’s all transaction house price index, used both to calculate misperceptions and as the house price index in the regression. Using the all transactions index does not change the results appreciably. Column 3 uses the FHFA median house price index. Because this index is published on an ad hoc basis only for the period 2000q1-2010q2, I take the average of the first and second quarters as the whole year average for 2010. Again, the results in column 3 are broadly similar to the results in the baseline specification of column 1. Column 4 presents results from the Zillow Home Value Index, another median value index. The ZHVI has more limited geographical coverage than the FHFA indices with only 35 states included in the regression, but the ZHVI includes all non-distressed sales, whereas that FHFA indices include only transactions with conforming mortgages. The coefficient on the misperceptions index shrinks to -0.95 in this specification but remains highly statistically significant. Finally, column 5 presents results using the same specification in column 1, but without weighting the states by the number of housing units. The estimated coefficient on the misperceptions index is slightly smaller than in the baseline specification but the difference is not statistically significant. Overall, the results in table 2 suggest that the association between home sales and house price misperceptions reported in table 1 is not particular to a particular house price index or weighting scheme.

Table 3 shows the same regressions as table 1 at the MSA level rather than the state level. The regressions cover the period 2005 to 2010 in 105 MSAs and PMSAs. The coefficient on
the misperceptions index is negative and statistically significant in columns 1 through 3. The coefficients on the misperceptions index in columns 1 and 2 are larger in absolute value than the corresponding coefficient in table 1, but the coefficient in the baseline specification, column 3, is of roughly the same magnitude. The $R^2$-within of the regression in column 1 is 0.19, again suggesting that approximately one-fifth of the time series variation in sales volumes in a geographical area can be explained by homeowner misperceptions of house prices. Columns 4 through 6 show the model estimated in first differences. Here, the results are substantially weaker than in table 1, as the coefficient on the misperceptions index is statistically insignificant in all three specifications. This difference seems to be at least partly due to the use of the FHFA all-transactions index, which, as illustrated in appendix figure A, shows much smaller price declines over the period 2005-2010 than other indices, such as the ZHVI and the Case-Shiller 20-city index. Unfortunately, both the Case-Shiller 20-city index and the metropolitan level FHFA purchase only index have very limited geographic coverage. If the ZHVI is used as the house price index, the coefficients on house price misperceptions in the first difference specifications of the model (not shown in the tables) are highly statistically significant and larger in magnitude than the corresponding coefficients in table 1.

Table 4 shows results of the baseline specification at the MSA level using alternative house price indices. Again, column 1 reproduces the baseline specification in column 3 of table 3 for comparison. Column 2 presents results from the FHFA’s purchase only index, which is available in a limited set of metro areas. The FHFA publishes this index at the CBSA level, while the perceptions indices are constructed at the MSA and PMSA level; I construct MSA-level house price indices from the FHFA data using population-weighted allocation factors. As the table illustrates, the coefficient on house price misperceptions is larger in magnitude using the purchase only index, although consistent with the limited geographical coverage, the standard error is larger as well. Column 3 shows results using the Case-Shiller 20-city house price index; the coefficient is of roughly the same magnitude as in column 1, but again the standard error is larger. Column 4 shows the results using the ZHVI as
the house price index. The geographical sample in column 4 is the same as the sample in column 1 because the sales volume data at the MSA level is from Zillow. The magnitude of the coefficient on misperceptions is more than twice as large as in column 1, a difference which a Chow test finds to be highly statistically significant. Finally, column 5 presents results from an unweighted regression. As in the state-level results, the coefficient on the misperceptions index is slightly smaller in magnitude in the unweighted regression. Overall, however, the MSA-level sensitivity analysis suggests that the results of the baseline specification may understate the effect of house price misperceptions on sales volumes at the MSA level.

### 3.3 Determinants of Misperceptions

In table 5, I briefly explore the empirical determinants of homeowner misperceptions of housing market conditions. The dependent variable in every column is the standard deviation within a geographical area of the misperceptions index, which I take to be an indicator of how much perceptions in that area tend to deviate from actual market prices. I use this measure rather than alternatives such as the mean absolute value of the misperceptions index because of the potential ambiguity created by normalizing the misperceptions index to zero in the base year. For instance, in Hawaii the misperceptions index is negative every year from 2001 to 2010. This could reflect either that Hawaii homeowners were unaware of the extent of Hawaiian price appreciation after the year 2000, or that they had unduly low perceptions of prices in the year 2000 and more accurate perceptions in later years.14

Columns 1 through 3 explore misperceptions at the state level, while columns 4 and 5 examine the MSA level. In column 1 the standard deviation of misperceptions is regressed on the standard deviation of house prices: a one point increase in the standard deviation of prices is associated with a 0.29 point increase in the standard deviation of misperceptions, with an $R^2$ of 0.57. Column

---

14Because the sales volumes regressions in the previous section all include state- or MSA-level fixed effects, this potential ambiguity should not affect the results in those regressions.
2 includes the persistence of house price changes as measured by a simple autoregression of the change in prices on their own lag; more persistent prices are associated with slightly more accurate perceptions. Column 3 includes the median percentage point error of Zillow’s home value estimate for homes in a state; a higher value for this number is suggestive of greater difficulty in inferring home values from observable house characteristics. A one percentage point increase in the median error is associated with a 0.55 point increase in the standard deviation of house price misperceptions, implying that a reduced ability for Zillow to estimate home values accurately is associated with less homeowner awareness in price variation over time. Columns 4 and 5 report results from the same specifications as columns 1 and 2, respectively, at the MSA level rather than the state level. As in the state-level regressions, greater variability in house prices is associated with greater variability in misperceptions, while greater persistence of house price changes is associated with less variability in misperceptions, although the coefficients on these variables are smaller in magnitude at the MSA level, and explain less of the variation in the variability of misperceptions.\footnote{I do not consider the median error of the ZHVI at the MSA level because Zillow only publishes it for a very small number of large MSAs.}

Overall, the sources of variation in house price misperceptions remain somewhat puzzling, especially at the MSA level. For instance, the results of regressions including variables such as whether a state requires public disclosure of the price paid for a house, and measures of the heterogeneity of the housing stock such as the standard deviation of the number of rooms or bedrooms in a house or the age of the housing stock, have variable and inconsistent coefficients in regressions similar to the ones in table 5 (not shown). Therefore, my cautious assessment is that while the variability and persistence of house prices seem reliably and intuitively related to house price misperceptions, it remains difficult to assess what other aspects of the housing market determine house price misperceptions.
4 Conclusion

This paper demonstrates that a stylized model of the house-selling process in which sellers have incomplete information regarding the state of the housing market can generate the correlations between prices, sales volumes, and time on the market observed in the data. Importantly, this effect can persist if sellers employ realtors with complete information about the state of housing demand, as long as the incentives of realtors and sellers are not perfectly aligned. Empirically, an increase in homeowners’ perceptions of house prices relative to true market conditions predicts a decrease in sales volumes. Homeowner misperceptions of housing market conditions appear to account for more than one-fifth of the time series variation in within-state sales volumes from 2000 to 2010.

REFERENCES


### TABLE 1: HOUSE PRICE MISPERCEPTIONS AND HOME SALES AT THE STATE LEVEL

<table>
<thead>
<tr>
<th>Specification</th>
<th>Log Homes Sold (1)</th>
<th>Log Homes Sold (2)</th>
<th>Log Homes Sold (3)</th>
<th>Log Homes Sold (4)</th>
<th>Log Homes Sold (5)</th>
<th>Log Homes Sold (6)</th>
<th>Log Homes Sold (7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Misperceptions Index</td>
<td>-1.396 (0.118)</td>
<td>-1.394 (0.119)</td>
<td>-1.193 (0.096)</td>
<td>-0.878 (0.143)</td>
<td>-0.807 (0.154)</td>
<td>-0.639 (0.116)</td>
<td></td>
</tr>
<tr>
<td>FHFA House Price Index</td>
<td>0.024 (0.037)</td>
<td>-0.730 (0.047)</td>
<td>0.094 (0.074)</td>
<td>-0.873 (0.080)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year Fixed Effects?</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Number of States</td>
<td>51</td>
<td>51</td>
<td>51</td>
<td>51</td>
<td>51</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td>Number of Observations</td>
<td>556</td>
<td>556</td>
<td>556</td>
<td>503</td>
<td>503</td>
<td>503</td>
<td></td>
</tr>
<tr>
<td>R-squared within</td>
<td>0.216</td>
<td>0.217</td>
<td>0.716</td>
<td>0.077</td>
<td>0.080</td>
<td>0.593</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Fixed Effects estimates using yearly state-level data from 2000-2010. All specifications include a constant term. Homes sold data from the National Association of Realtors. The misperceptions index is the log change in housing values constructed from American Community Survey data minus the log change in the FHFA purchase-only house price index. All regressions weighted by the number of housing units in ACS price regressions.
<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Log Homes Sold</th>
<th>Log Homes Sold</th>
<th>Log Homes Sold</th>
<th>Log Homes Sold</th>
<th>Log Homes Sold</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FHFA Purchase Only Index (1)</td>
<td>FHFA All Transactions Index (2)</td>
<td>FHFA Median Index (3)</td>
<td>Zillow Home Value Index (4)</td>
<td>FHFA Purchase Only Index (5)</td>
</tr>
<tr>
<td>Misperceptions Index</td>
<td>-1.193 (0.096)</td>
<td>-1.268 (0.131)</td>
<td>-1.163 (0.081)</td>
<td>-0.949 (0.088)</td>
<td>-1.055 (0.105)</td>
</tr>
<tr>
<td>House Price Index</td>
<td>-0.730 (0.047)</td>
<td>-0.694 (0.049)</td>
<td>-0.611 (0.043)</td>
<td>-0.643 (0.051)</td>
<td>-0.609 (0.053)</td>
</tr>
</tbody>
</table>

Year Fixed Effects? | Yes | Yes | Yes | Yes | Yes
Weighted by Housing Units? | Yes | Yes | Yes | Yes | No
Number of States | 51 | 51 | 51 | 35 | 51
Number of Observations | 556 | 556 | 556 | 381 | 556
R-squared within | 0.716 | 0.711 | 0.725 | 0.725 | 0.697

Notes: Fixed Effects estimates using yearly state-level data from 2000-2010. All specifications include a constant term. Homes sold data from the National Association of Realtors.
**TABLE 3: HOUSE PRICE MISPERCEPTIONS AND HOME SALES AT THE MSA LEVEL**

<table>
<thead>
<tr>
<th>Specification</th>
<th>Log Homes Sold</th>
<th>Log Homes Sold</th>
<th>Log Homes Sold</th>
<th>Log Homes Sold</th>
<th>Log Homes Sold</th>
<th>Log Homes Sold</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Levels (1)</td>
<td>Levels (2)</td>
<td>Levels (3)</td>
<td>First Differences (4)</td>
<td>First Differences (5)</td>
<td>First Differences (6)</td>
</tr>
<tr>
<td>Misperceptions Index</td>
<td>-2.627</td>
<td>-2.868</td>
<td>-1.093</td>
<td>-0.113</td>
<td>-0.075</td>
<td>-0.122</td>
</tr>
<tr>
<td></td>
<td>(0.234)</td>
<td>(0.247)</td>
<td>(0.169)</td>
<td>(0.217)</td>
<td>(0.187)</td>
<td>(0.178)</td>
</tr>
<tr>
<td>FHFA House Price Index</td>
<td>0.024</td>
<td>-0.730</td>
<td>-1.021</td>
<td>-1.737</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.037)</td>
<td>(0.047)</td>
<td>(0.085)</td>
<td>(0.117)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year Fixed Effects?</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Number of MSAs</td>
<td>105</td>
<td>105</td>
<td>105</td>
<td>105</td>
<td>105</td>
<td>105</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>630</td>
<td>630</td>
<td>630</td>
<td>525</td>
<td>525</td>
<td>525</td>
</tr>
<tr>
<td>R-squared within</td>
<td>0.193</td>
<td>0.207</td>
<td>0.734</td>
<td>0.001</td>
<td>0.258</td>
<td>0.474</td>
</tr>
</tbody>
</table>

Notes: Fixed Effects estimates using yearly MSA-level data from 2005-2010. All specifications include a constant term. Homes sold data from Zillow.com. The misperceptions index is the log change in housing values constructed from American Community Survey data minus the log change in the FHFA all-transactions house price index. All regressions weighted by the number of housing units in ACS price regressions.
<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Log Homes Sold FHFA All Transactions Index (1)</th>
<th>Log Homes Sold FHFA Purchase Only Index (2)</th>
<th>Log Homes Sold Case-Shiller Index (3)</th>
<th>Log Homes Sold Zillow Home Value Index (4)</th>
<th>Log Homes Sold FHFA All Transactions Index (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Misperceptions Index</td>
<td>-1.093</td>
<td>-1.638</td>
<td>-1.032</td>
<td>-2.255</td>
<td>-0.917</td>
</tr>
<tr>
<td>House Price Index</td>
<td>(0.169)</td>
<td>(0.366)</td>
<td>(0.323)</td>
<td>(0.165)</td>
<td>(0.166)</td>
</tr>
<tr>
<td>Misperceptions Index</td>
<td>-1.134</td>
<td>-1.539</td>
<td>-1.437</td>
<td>-1.264</td>
<td>-0.994</td>
</tr>
<tr>
<td>House Price Index</td>
<td>(0.068)</td>
<td>(0.171)</td>
<td>(0.184)</td>
<td>(0.067)</td>
<td>(0.063)</td>
</tr>
<tr>
<td>Year Fixed Effects?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Weighted by Housing Units?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Number of MSAs</td>
<td>105</td>
<td>20</td>
<td>19</td>
<td>105</td>
<td>105</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>630</td>
<td>120</td>
<td>114</td>
<td>630</td>
<td>630</td>
</tr>
<tr>
<td>R-squared within</td>
<td>0.734</td>
<td>0.778</td>
<td>0.756</td>
<td>0.758</td>
<td>0.695</td>
</tr>
</tbody>
</table>

Notes: Fixed Effects estimates using yearly MSA-level data from 2005-2010. All specifications include a constant term. Homes sold data from Zillow.com.
### TABLE 5: CAUSES OF HOUSE PRICE MISPERCEPTIONS

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>States (1)</td>
<td>States (2)</td>
<td>States (3)</td>
<td>MSAs (4)</td>
<td>MSAs (5)</td>
</tr>
<tr>
<td>Std. Dev. Of House Price Index</td>
<td>0.288 (0.035)</td>
<td>0.276 (0.035)</td>
<td>0.224 (0.036)</td>
<td>0.146 (0.025)</td>
<td>0.120 (0.029)</td>
</tr>
<tr>
<td>House Price Persistence</td>
<td>-0.040 (0.022)</td>
<td>-0.090 (0.024)</td>
<td>-0.011 (0.007)</td>
<td>-0.011 (0.007)</td>
<td>-0.011 (0.007)</td>
</tr>
<tr>
<td>Zillow Estimate Median Error</td>
<td>0.554 (0.220)</td>
<td>0.554 (0.220)</td>
<td>0.554 (0.220)</td>
<td>0.554 (0.220)</td>
<td>0.554 (0.220)</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>51</td>
<td>51</td>
<td>36</td>
<td>105</td>
<td>105</td>
</tr>
<tr>
<td>Adjusted R-squared</td>
<td>0.571</td>
<td>0.590</td>
<td>0.697</td>
<td>0.240</td>
<td>0.252</td>
</tr>
</tbody>
</table>

**Notes:** All specifications include a constant term and are weighted by the number of housing units in ACS price regressions. House price persistence is the autoregressive parameter from regressing the change in house prices on the lagged change within an area. Zillow estimate median error is the median error of Zillow’s house price estimate within an area.
Figure 1: Housing Market Time Series January 2000 - December 2010

Note: All series have had a linear time trend removed.
Figure 2: Expected Values of $\psi_2$ as a function of $\psi_1$

Case 1: $x_H - x_L \geq z_H - z_L$

Case 2: $z_H - z_L \geq x_H - x_L$
Figure 3: Expected Time to Sale and Sales Price without Realtor
Figure 4: Expected Values of $\psi_2$ as a function of $\psi_1$
Figure 5: Expected Time to Sale and Sales Price with Realtor

Case 1: $x_H - \check{x}_1 \geq z_R - x_L$ and $\check{x}_1 + z_H \leq \underline{\pi} + \overline{\pi}$

Case 2: $z_H - z_L \geq x_H - \check{x}_1$ and $\check{x}_1 + z_H \geq \underline{\pi} + \overline{\pi}$
Figure 6: Realtor’s Optimal Cutoff Rule
Figure 7: Seller’s Expected Value of waiting for Period 2 in 3-period Model
Figure 8: Expected Time to Sale and Sales Price in 3-period Model
Figure 9A: Log Price Change from 2000: Perceived and Actual (States)

Graphs by (mean) statefip
Figure 9B: Log Price Change from 2000: Perceived and Actual (States)

Census year

- **American Community Survey**
- **FHFA HPI**

Graphs by (mean) statefip
Figure 10A: Misperceptions and Home Sales (States)

Graphs by (mean) state pops

Misperceptions Index

Home Sales

Census year

Alabama - Alaska - Arizona - Arkansas - California

Colorado - Connecticut - Delaware - Florida - Georgia

Hawaii - Idaho - Illinois - Indiana - Iowa

Kansas - Kentucky - Louisiana - Maine - Maryland

Massachusetts - Michigan - Minnesota - Mississippi - Missouri
Figure 10B: Misperceptions and Home Sales (States)

Misperceptions Index (log change from 2000)

Home Sales (log change from 2000)

Census year

- Misperceptions Index
- Home Sales

Graphs by (mean) statefip
Figure 11: Percent Price change from 2005: Perceived and Actual (MSAs)

Census year

- American Community Survey
- FHFA

Graphs by (mean) pmsa
Figure 12: Misperceptions Index and Home Sales (MSAs)

Home Sales (log change from 2005)

<table>
<thead>
<tr>
<th>City</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dallas</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Miami</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>San Diego</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cleveland</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Los Angeles</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Portland</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Washington, D.C.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chicago</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Las Vegas</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phoenix</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tampa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boston</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Detroit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New York</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seattle</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atlanta</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denver</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minneapolis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>San Francisco</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Misperceptions Index (log change from 2005)
Appendix

A Proofs of Propositions

A.1 The p.d.f. of $\psi$

Let $\psi_L = x_L + z_L$ and $\psi_H = x_H + z_L$. Then the p.d.f. of $\psi$ is:

$$f(\psi) = \begin{cases} \frac{\psi - \psi_L}{(x_H - x_L)(z_H - z_L)} & \text{if } \psi_L \leq \psi \leq \psi_L + \min(x_H - x_L, z_H - z_L) \\ \frac{1}{\max(x_H - x_L, z_H - z_L)} & \text{if } \psi_L + \min(x_H - x_L, z_H - z_L) \leq \psi \\ \frac{\psi_H - \psi}{(x_H - x_L)(z_H - z_L)} & \text{if } \psi_L + \max(x_H - x_L, z_H - z_L) \leq \psi \leq \psi_H \end{cases}$$

A.2 Seller’s posterior belief about $z$

First note

$$f_\Psi(\psi|z) = \begin{cases} \frac{1}{(x_H - x_L)} & \text{if } z + x_L \leq \psi \leq z + x_H \\ 0 & \text{otherwise} \end{cases}$$

There are multiple cases to consider to verify that the seller’s posterior distribution for $z$ is $z \sim U[\hat{z}_L, \hat{z}_H]$.

Case 1: $x_H - x_L \geq z_H - z_L$

Case 1a: $\psi \leq x_L + z_H$. In this range, $f(\psi) = \frac{\psi - \psi_L}{(x_H - x_L)(z_H - z_L)}$. If $z < z_L$ or $z > z_H$, $f(z) = 0$. If $z < \psi - x_H$ or $z > \psi - x_L$, $f(\psi|z) = 0$. Therefore, $f(z|\psi) = 0$ if $z \leq \max(z_H, \psi - x_H) \equiv \hat{z}_L$ or if $z \geq \min(z_H, \psi - x_L) \equiv \hat{z}_H$. In the range $[\hat{z}_L, \hat{z}_H]$, $f_Z(z|\Psi = \psi) = f_\Psi(\psi|z)f_Z(z) = \frac{1}{x_H - x_L} \cdot \frac{1}{z_H - z_L} = \frac{1}{\psi - \psi_L}$.

To see that $f_Z(z|\Psi = \psi)$ is a proper density, note that in this case $\hat{z}_L = z_L$ and $\hat{z}_H = \psi - x_L$, so that $\hat{z}_H - \hat{z}_L = \psi - x_L - z_L = \psi - \psi_L$.

Case 1b: $x_L + z_H \leq \psi \leq x_H + z_L$. In this case, $f(\psi) = \frac{1}{x_H - x_L}$, $f(z) = \frac{1}{z_H - z_L}$, $f(\psi|z) = \frac{1}{x_H - x_L}$, $\hat{z}_L = z_L$, and $\hat{z}_H = z_H$. Then in the range $[\hat{z}_L, \hat{z}_H]$, $f_Z(z|\Psi = \psi) = f_\Psi(\psi|z)f_Z(z) = \frac{1}{x_H - x_L} \cdot \frac{1}{z_H - z_L} = \frac{1}{z_H - z_L}$ and $f_Z(z|\Psi = \psi) = 0$ elsewhere.

Case 1c: $x_H + z_L < \psi$. In this case $f(\psi) = \frac{\psi_H - \psi}{(x_H - x_L)(z_H - z_L)}$, $f(\psi|z) = \frac{1}{x_H - x_L}$, $f(z) = \frac{1}{z_H - z_L}$.
\[ \tilde{z}_L = \psi - x_H, \text{ and } \tilde{z}_H = z_H. \text{ Then in the range } [\tilde{z}_L, \tilde{z}_H], \]

\[ f_Z(z|\Psi = \psi) = \frac{f_{\Psi}(\psi|z)f_Z(z)}{f_{\Psi}(\psi)} = \frac{1}{x_H-x_L} \cdot \frac{1}{\psi_H - \psi} = \frac{1}{\psi_H - \psi} \]

and \( f_Z(z|\Psi = \psi) = 0 \) elsewhere. Because \( \tilde{z}_H - \tilde{z}_L = x_H + z_H - \psi = \psi_H - \psi \), the posterior distribution is a proper density.

Case 2: \( z_H - z_L > x_H - x_L \).

Case 2a: \( \psi < x_H + z_L \). In this case the proof is the same as in case 1a.

Case 2b: \( x_H + z_L \leq \psi \leq x_L + z_H \). In this case \( f(\psi) = \frac{1}{z_H - z_L} \), \( f(z) = \frac{1}{x_H - x_L} \), \( f(\psi|z) = \frac{1}{x_H - x_L} \), \( \tilde{z}_L = \psi - x_H \), and \( \tilde{z}_H = \psi - x_L \). Then in the range \([\tilde{z}_L, \tilde{z}_H]\),

\[ f_Z(z|\Psi = \psi) = \frac{f_{\Psi}(\psi|z)f_Z(z)}{f_{\Psi}(\psi)} = \frac{1}{x_H-x_L} \cdot \frac{1}{\psi_H - \psi} = \frac{1}{\psi_H - \psi} \]

and \( f_Z(z|\Psi = \psi) = 0 \) elsewhere. Because \( \tilde{z}_H - \tilde{z}_L = \psi - x_L - (\psi - x_H) = x_H - x_L \), the posterior distribution is a proper density.

Case 2c: \( x_L + z_H < \psi \). In this case the proof is the same as in case 1c.

A.3 The seller’s expectation of \( \psi_2 \) conditional on \( \psi_1 \)

Let \( \overline{\psi} = \frac{x_L + x_H}{2} \) and \( \tilde{\psi}_1 = \frac{z_L + z_H}{2} \). Then \( E[\psi_2|\psi_1] = \overline{\psi} + \tilde{\psi}_1 \). If we further define \( \overline{\psi} = \frac{x_L + z_H}{2} \), we can write the unconditional expectation of \( \psi \) as \( E[^{\psi}] = \overline{\psi} = \overline{\psi} + \overline{\psi} \). If \( x_H - x_L \geq z_H - z_L \), we can write:

\[ E[\psi_2|\psi_1] = \begin{cases} \frac{\psi_1 + x_H + z_L}{\psi} & \text{if } \psi_1 \leq \psi \leq x_H + z_H \\ \frac{\psi_1 + x_L + z_H}{\psi} & \text{if } x_L + z_L \leq \psi \leq x_H + z_L \\ \frac{\psi_1 + x_L + z_H}{\psi} & \text{if } x_H + z_L \leq \psi \leq \psi_H \end{cases} \]

Then for all \( \psi_1 < \overline{\psi}, \psi_1 < E[\psi_2|\psi_1] \), while for all \( \psi_1 \geq \overline{\psi}, \psi_1 \geq E[\psi_2|\psi_1] \).

If \( z_H - z_L > x_H - x_L \):

\[ E[\psi_2|\psi_1] = \begin{cases} \frac{\psi_1 + x_H + z_L}{\psi} & \text{if } \psi_1 \leq \psi \leq x_H + z_L \\ \frac{\psi_1 + x_L + z_H}{\psi} & \text{if } x_L + z_L \leq \psi \leq x_H + z_H \\ \frac{\psi_1 + x_L + z_H}{\psi} & \text{if } x_L + z_H \leq \psi \leq \psi_H \end{cases} \]

Then for all \( \psi_1 < x_H + z_L, \psi_1 < E[\psi_2|\psi_1] \), while for all \( \psi_1 \geq x_H + z_L, \psi_1 \geq E[\psi_2|\psi_1] \).
A.4 Proof of Proposition 1

We assume the seller will accept any offer \( \psi_1 \geq E[\psi_2|\psi_1] \). Let \( \bar{t}(z) \) denote the expected number of periods the seller leaves his house on the market. Then if \( x_H - x_L \geq z_H - z_L \):

\[
\bar{t}(z) = Pr(\psi_1 \geq \bar{\psi}) + 2Pr(\psi_1 < \bar{\psi}) \\
= Pr(x_1 \geq \bar{\psi} - z) + 2Pr(x_1 < \bar{\psi} - z) \\
= 1 - \frac{\bar{\psi} - x_L - z}{x_H - x_L} + 2 \frac{\bar{\psi} - x_L - z}{x_H - x_L} \\
= 1 + \frac{\bar{\psi} - x_L - z}{x_H - x_L}
\]

Then

\[
\frac{\partial \bar{t}}{\partial z} = \frac{-1}{x_H - x_L} < 0
\]

Now consider the case in which \( z_H - z_L > x_H - x_L \):

\[
\bar{t}(z) = Pr(\psi_1 \geq x_H + z_L) + 2Pr(\psi_1 < x_H + z_L) \\
= Pr(x_1 \geq x_H + z_L - z) + 2Pr(x_1 < x_H + z_L - z)
\]

If \( z > x_H - x_L + z_L \), \( Pr(x_1 \geq x_H + z_L - z) = 1 \), so \( \bar{t} = 1 + 0 = 1 \). If \( z \leq x_H - x_L + z_L \),

\[
\bar{t}(z) = 1 - \frac{x_H - x_L + z_L - z}{x_H - x_L} + 2 \frac{x_H - x_L + z_L - z}{x_H - x_L} \\
= 1 + \frac{x_H - x_L + z_L - z}{x_H - x_L} \\
= 2 - \frac{z - z_L}{x_H - x_L}
\]

Then when \( z_H - z_L > x_H - x_L \),

\[
\bar{t}(z) = \begin{cases} 
2 - \frac{z - z_L}{x_H - x_L} & \text{if } z_L \leq z \leq x_H - x_L + z_L \\
1 & \text{if } z > x_H - x_L + z_L
\end{cases}
\]

and

\[
\frac{\partial \bar{t}}{\partial z} = \begin{cases} 
\frac{-1}{x_H - x_L} & \text{if } z_L \leq z \leq x_H - x_L + z_L \\
0 & \text{if } x_H - x_L + z_L \leq z \leq z_H
\end{cases}
\]
Therefore, \( \frac{\partial \bar{p}}{\partial z} \) must always be weakly negative.

Let \( \bar{p}(z) \) denote the expected sales price for the house. When \( x_H - x_L \geq z_H - z_L \),
\[
\bar{p}(z) = Pr(\psi_1 \geq \bar{\pi} + \bar{z})E[\psi_1 | \psi_1 \geq \bar{\pi} + \bar{z}] + Pr(\psi_1 < \bar{\pi} + \bar{z})E[\psi_2]
\]
\[
= Pr(x_1 \geq \bar{\pi} + \bar{z} - z)E[\psi_1 | \psi_1 \geq \bar{\pi} + \bar{z}] + Pr(x_1 < \bar{\pi} + \bar{z} - z)E[\psi_2]
\]
\[
= (1 - \frac{\bar{\pi} - x_L + \bar{z} - z}{x_H - x_L})(\frac{\bar{\pi} + x_H + \bar{z} + z}{2}) + (\frac{\bar{\pi} - x_L + \bar{z} - z}{x_H - x_L})(\bar{\pi} + z)
\]
\[
= \frac{\bar{\pi} + x_H + \bar{z} + z}{2} + (\frac{\bar{\pi} - x_L + \bar{z} - z}{x_H - x_L})(\bar{\pi} - x_H - \bar{z} + z)
\]

This implies \( \frac{\partial \bar{p}}{\partial z} = 1 + \frac{\bar{z} - \bar{\pi}}{x_H - x_L} \). To see that this must always be positive, consider the case \( z = z_H \).

Then \( \frac{\partial \bar{p}}{\partial z} = 1 + \frac{\bar{z} - \bar{\pi}}{x_H - x_L} = \frac{x_H + z_H - \bar{\pi} - z_L}{x_H - x_L} \). \( \bar{\pi} \geq \bar{\pi} \), so \( z_H - \bar{\pi} < z_H - z_L \). By assumption, \( x_H - x_L \geq z_H - z_L \). Therefore the numerator of this expression is positive, and \( \frac{\partial \bar{p}}{\partial z} > 0 \) when \( x_H - x_L \geq z_H - z_L \).

When \( z_H - z_L > x_H - x_L \),
\[
\bar{p}(z) = Pr(\psi_1 \geq x_H + z_L)E[\psi_1 | \psi_1 \geq x_H + z_L] + Pr(\psi_1 < x_H + z_L)E[\psi_2]
\]
\[
= Pr(x_1 \geq x_H + z_L - z)E[\psi_1 | \psi_1 \geq x_H + z_L] + Pr(x_1 < x_H + z_L - z)E[\psi_2]
\]
If \( z \geq x_H - x_L + z_L \), \( Pr(\psi_1 \geq x_H + z_L) = 1 \), and \( E[\psi_1 | \psi_1 \geq x_H + z_L] = E[\psi_1] \). Then \( \bar{p} = \bar{\pi} + z \), so \( \frac{\partial \bar{p}}{\partial z} = 1 \). If \( z < x_H - x_L + z_L \), \( Pr(\psi_1 \geq x_H + z_L) = \frac{z - z_L}{x_H - x_L} \). Then
\[
\bar{p} = \frac{z - z_L}{x_H - x_L}(x_H + \frac{z_L + z}{2}) + (1 - \frac{z - z_L}{x_H - x_L})(\bar{\pi} + z)
\]
\[
= \bar{\pi} + z + \frac{z - z_L}{x_H - x_L}(\frac{x_H - x_L + z_L - z}{2})
\]

implying
\[
\frac{\partial \bar{p}}{\partial z} = 1 + \frac{x_H - x_L + z_L - 2z + z_L}{2(x_H - x_L)}
\]
\[
= \frac{3}{2} - \frac{z - z_L}{x_H - x_L}
\]

Therefore when \( z_H - z_L > x_H - x_L \),
\[
\bar{p}(z) = \begin{cases} 
\bar{\pi} + z + \frac{z - z_L}{x_H - x_L}(\frac{x_H - x_L + z_L - z}{2}) & \text{if } z_L \leq z \leq x_H - x_L + z_L \\
\bar{\pi} + z & \text{if } z > x_H - x_L + z_L
\end{cases}
\]
and
\[
\frac{\partial \bar{p}}{\partial z} = \begin{cases} 
\frac{3}{2} - \frac{z - z_L}{x_H - x_L} & \text{if } z_L \leq z \leq x_H - x_L + z_L \\
1 & \text{if } x_H - x_L + z_L \leq z \leq z_H
\end{cases}
\]
By assumption \( z < x_H - x_L + z_L \), so \( \frac{z - z_L}{x_H - x_L} < 1 \). Therefore \( \frac{\partial f}{\partial z} > 0 \) when \( z_H - z_L > x_H - x_L \) and \( z < x_H - x_L + z_L \), implying that \( \frac{\partial f}{\partial z} > 0 \) in all cases. \( \blacksquare \)

A.5 Proof of Proposition 2

Let \( \tilde{\psi}_L = \tilde{z}_{L,1} + x_L \) and \( \tilde{\psi}_L = \tilde{z}_{L,1} + x_L \). The seller’s posterior belief about the distribution of \( \psi_2 \) is then

\[
\tilde{f}(\psi_2|\psi_1, \xi_1) = \begin{cases} 
\frac{\psi - \psi_L}{(x_H - x_L)(\tilde{z}_{H,1} - \tilde{z}_{L,1})} & \text{if } \tilde{\psi}_L \leq \psi \leq \psi_L + \min(x_H - x_L, \tilde{z}_{H,1} - \tilde{z}_{L,1}) \\
\frac{1}{\max(x_H - x_L, \tilde{z}_{H,1} - \tilde{z}_{L,1})} & \text{if } \tilde{\psi}_L + \min(x_H - x_L, \tilde{z}_{H,1} - \tilde{z}_{L,1}) \leq \psi \leq \psi_L + \max(x_H - x_L, \tilde{z}_{H,1} - \tilde{z}_{L,1}) \\
\frac{\psi_L - \psi}{(x_H - x_L)(\tilde{z}_{H,1} - \tilde{z}_{L,1})} & \text{if } \tilde{\psi}_L + \max(x_H - x_L, \tilde{z}_{H,1} - \tilde{z}_{L,1}) \leq \psi \leq \tilde{\psi}_L 
\end{cases}
\]

Consider \( E[\psi_2|\psi_1, \xi_1 = 1] \). If \( x_H - \hat{x}_1 \geq z_H - z_L \) (call this case 1),

\[
E[\psi_2|\tilde{f}(\psi_2|\psi_1, 1)] = \begin{cases} 
\bar{\pi} + \frac{\psi_1 - \hat{x}_1 + z_L}{2} & \text{if } \hat{x}_1 + z_L \leq \psi_1 \leq \hat{x}_1 + z_H \\
\bar{\pi} + \frac{\psi_1 - x_H - z_L}{2} & \text{if } x_H + z_L \leq \psi_1 \leq \psi_H 
\end{cases}
\]

If \( z_H - z_L > x_H - \hat{x}_1 \) (case 2),

\[
E[\psi_2|\tilde{f}(\psi_2|\psi_1, 1)] = \begin{cases} 
\bar{\pi} + \frac{\psi_1 - \hat{x}_1 + z_L}{2} & \text{if } x_H + z_L \leq \psi_1 \leq \hat{x}_1 + z_H \\
\bar{\pi} + \frac{\psi_1 - \hat{x}_1 + z_L}{2} & \text{if } x_H + z_L \leq \psi_1 \leq \hat{x}_1 + z_H 
\end{cases}
\]

In case 1, if \( \hat{x}_1 + z_H \leq \bar{\pi} + \bar{\pi} \) (call this case 1a), \( \psi_1 \geq E[\psi_2|\tilde{f}(\psi_2|\psi_1, 1)] \) when \( \psi_1 \geq \bar{\pi} + \bar{\pi} \) and \( \psi_1 < E[\psi_2|\tilde{f}(\psi_2|\psi_1, 1)] \) otherwise. If \( \hat{x}_1 + z_H > \bar{\pi} + \bar{\pi} \) (case 1b), \( \psi_1 \geq E[\psi_2|\tilde{f}(\psi_2|\psi_1, 1)] \) when \( \psi_1 \geq x_L + x_H - \hat{x}_1 + z_L \) and \( \psi_1 < E[\psi_2|\tilde{f}(\psi_2|\psi_1, 1)] \) otherwise. In case 2 \( \psi_1 \geq E[\psi_2|\tilde{f}(\psi_2|\psi_1, 1)] \) when \( \psi_1 \geq x_L + x_H - \hat{x}_1 + z_L \) and \( \psi_1 < E[\psi_2|\tilde{f}(\psi_2|\psi_1, 1)] \) otherwise.

Again, we assume the seller accepts any offer \( \psi_1 \geq E[\psi_2|\psi_1, \xi_1] \). Recall that the seller always rejects an offer when the realtor recommends ‘reject’. Then in case 1a \( (x_H - \hat{x}_1) \geq z_H - z_L \) and
\[ \hat{x}_1 + z_H \leq \bar{x} + \bar{z}, \]

\[ \bar{t}(z) = 2Pr(\xi_1 = 0) + Pr(\psi_1 \geq \bar{x} + \bar{z} \cap \xi_1 = 1) + 2Pr(\psi_1 < \bar{x} + \bar{z} \cap \xi_1 = 1) \]

\[ = 2(\hat{x}_1 - x_L) + (1 - \frac{\hat{x}_1 - x_L}{x_H - x_L})(Pr(x_1 \geq \bar{x} + \bar{z} - z|\xi_1 = 1) + 2Pr(x_1 < \bar{x} + \bar{z} - z|\xi_1 = 1)) \]

\[ = \frac{2(\hat{x}_1 - x_L)}{x_H - x_L} + (1 - \frac{\hat{x}_1 - x_L}{x_H - x_L})(\frac{\bar{x} - \hat{x}_1 + \bar{z} - z}{x_H - \hat{x}_1} + 2(\frac{\bar{x} - \hat{x}_1 + \bar{z} - z}{x_H - \hat{x}_1})) \]

Therefore in case 1a,

\[ \frac{\partial \bar{t}}{\partial z} = \frac{-1}{x_H - x_L} < 0 \]

In cases 1b and 2 \((\hat{x}_1 + z_H > \bar{x} + \bar{z})\),

\[ \bar{t}(z) = 2Pr(\xi_1 = 0) + Pr(\psi_1 \geq x_L + x_H - \hat{x}_1 + z_L \cap \xi_1 = 1) + 2Pr(\psi_1 < x_L + x_H - \hat{x}_1 + z_L \cap \xi_1 = 1) \]

\[ = \frac{2(\hat{x}_1 - x_L)}{x_H - x_L} + (1 - \frac{\hat{x}_1 - x_L}{x_H - x_L})(Pr(x_1 \geq x_L + x_H - \hat{x}_1 + z_L - z|\xi_1 = 1) \]

\[ + 2Pr(x_1 < x_L + x_H - \hat{x}_1 + z_L - z|\xi_1 = 1)) \]

If \( z > x_L + x_H - 2\hat{x}_1 + z_L \), \( Pr(x_1 \geq x_L + x_H - \hat{x}_1 + z_L - z|\xi_1 = 1) = 1 \). In that case \( \bar{t} = 2Pr(\xi_1 = 0) + Pr(\xi_1 = 1) = \frac{2(\hat{x}_1 - x_L)}{x_H - x_L} + (1 - \frac{\hat{x}_1 - x_L}{x_H - x_L}) = 1 + \frac{\hat{x}_1 - x_L}{x_H - x_L} \). If \( z < x_L + x_H - 2\hat{x}_1 + z_L \),

\[ \bar{t}(z) = \frac{2(\hat{x}_1 - x_L)}{x_H - x_L} + (1 - \frac{\hat{x}_1 - x_L}{x_H - x_L})(1 - \frac{\hat{x}_1 - x_L}{x_H - x_L}) + 2(\frac{x_L + x_H - 2\hat{x}_1 + z_L - z}{x_H - \hat{x}_1}) \]

\[ = \frac{2(\hat{x}_1 - x_L)}{x_H - x_L} + (1 - \frac{\hat{x}_1 - x_L}{x_H - x_L})(1 + \frac{x_L + x_H - 2\hat{x}_1 + z_L - z}{x_H - \hat{x}_1}) \]

\[ = 1 + \frac{\hat{x}_1 - x_L}{x_H - x_L} + (\frac{x_H + x_H - 2\hat{x}_1 + z_L - z}{x_H - \hat{x}_1}) \]

Then in cases 1b and 2,

\[ \bar{t}(z) = \begin{cases} 
1 + \frac{x_H - \hat{x}_1 + z_L - z}{x_H - x_L} & \text{if } z_L \leq z < x_L + x_H - 2\hat{x}_1 + z_L \\
1 + \frac{\hat{x}_1 - x_L}{x_H - x_L} & \text{if } x_L + x_H - 2\hat{x}_1 + z_L \leq z \leq z_H 
\end{cases} \]
Therefore, in case 2 \((z_H - z_L > x_H - \hat{x}_1)\),

\[
\tilde{t}(z) = 2Pr(\xi_1 = 0) + Pr(\psi_1 \geq x_L + x_H - \hat{x}_1 + z_L \cap \xi_1 = 1) + 2Pr(\psi_1 < x_L + x_H - \hat{x}_1 + z_L \cap \xi_1 = 1)
\]

\[
= 2(\hat{x}_1 - x_L) + (1 - \hat{x}_1 - x_L)(Pr(x_1 \geq x_L + x_H - \hat{x}_1 + z_L - z|\xi_1 = 1)
+ 2Pr(x_1 < x_L + x_H - \hat{x}_1 + z_L - z|\xi_1 = 1))
\]

If \(z > x_L + x_H - 2\hat{x}_1 + z_L\), \(Pr(x_1 \geq x_L + x_H - \hat{x}_1 + z_L - z|\xi_1 = 1) = 1\). In that case \(\tilde{t} = 2Pr(\xi_1 = 0) + Pr(\xi_1 = 1) = 2(\hat{x}_1 - x_L)/(x_H - x_L) + (1 - \hat{x}_1 - x_L) = 1 + \hat{x}_1 - x_L/(x_H - x_L)\). Therefore \(\frac{\partial \tilde{t}}{\partial z} = 0\). If \(z < x_L + x_H - 2\hat{x}_1 + z_L\),

\[
\tilde{t}(z) = \frac{2(\hat{x}_1 - x_L)}{x_H - x_L} + (1 - \frac{\hat{x}_1 - x_L}{x_H - x_L})(1 - \frac{x_L + x_H - \hat{x}_1 + z_L - z}{x_H - \hat{x}_1} + 2(\frac{x_L + x_H - \hat{x}_1 + z_L - z}{x_H - \hat{x}_1}))
\]

In this case,

\[
\frac{\partial \tilde{t}}{\partial z} = (1 - \frac{\hat{x}_1 - x_L}{x_H - x_L}) \frac{-1}{x_H - \hat{x}_1} < 0
\]

Therefore, in case 2

\[
\frac{\partial \tilde{t}}{\partial z} = \begin{cases} 
(1 - \frac{\hat{x}_1 - x_L}{x_H - x_L}) \frac{-1}{x_H - \hat{x}_1} & \text{if } z_L \leq z < x_L + x_H - 2\hat{x}_1 + z_L \\
0 & \text{if } x_L + x_H - 2\hat{x}_1 + z_L \leq z \leq z_H 
\end{cases}
\]

Thus, \(\frac{\partial \tilde{t}}{\partial z}\) must always be weakly negative.

In case 1a,

\[
\overline{\psi}(z) = Pr(\xi_1 = 0)E[\psi_2] + Pr(\xi_1 = 1)[Pr(\psi_1 \geq \overline{\psi} + \overline{\xi}_1 = 1)E[\psi_1 | \psi_1 \geq \overline{\psi} + \overline{\xi}_1, \xi_1 = 1]
+ Pr(\psi_1 < \overline{\psi} + \overline{\xi}_1 | \xi_1 = 1)E[\psi_2]]]
\]

\[
= (\frac{\hat{x}_1 - x_L}{x_H - x_L})(\overline{\psi} + z) + (1 - \frac{\hat{x}_1 - x_L}{x_H - x_L}) \left[(1 - \frac{\overline{\psi} - \hat{x}_1 + z_L - z}{x_H - \hat{x}_1})(z + \frac{\overline{\psi} + x_H + z - z}{2})
+ (\frac{\overline{\psi} - \hat{x}_1 + z - z}{x_H - \hat{x}_1})(\overline{\psi} + z)\right]
\]

\[
= \frac{1}{2}(\overline{\psi}^2 + x_H^2 - \overline{\psi}^2 - z^2) - x_L\overline{\psi} + (x_H - x_L + \overline{\psi})z
\]

\(\frac{x_H - x_L}{x_H - x_L}\)
Therefore in case 1a, 
\[
\frac{\partial p}{\partial z} = \frac{x_H - x_L + \xi - z}{x_H - x_L}
\]

To see that this must be positive, consider the case in which \(z = z_H\). Then
\[
\frac{\partial p}{\partial z} = \frac{x_H - x_L + \xi - z_H}{x_H - x_L} = \frac{x_H - x_L - \frac{z_H - z_L}{2}}{x_H - x_L}
\]

By assumption \(x_H - x_L \geq x_H - \hat{x}_1 \geq z_H - z_L\), so \(\frac{\partial p}{\partial z} \geq 0\) in case 1a. In cases 1b and 2, if \(z > x_L + x_H - 2\hat{x}_1 + z_L\) (so that the seller will always accept the period 1 offer if the realtor recommends accept),
\[
p = Pr(\xi_1 = 0)E[\psi_2] + Pr(\xi_1 = 1)E[\psi_1|x_1 \geq \hat{x}_1]
\]
\[
= (\frac{\hat{x}_1 - x_L}{x_H - x_L})(\overline{x} + z) + (1 - \frac{\hat{x}_1 - x_L}{x_H - x_L})(z + \frac{\hat{x}_1 + x_H}{2})
\]
\[
= (\frac{\hat{x}_1 - x_L}{x_H - x_L})(x_L + x_H) + (x_H - \hat{x}_1)(\hat{x}_1 + x_H) + z
\]

If \(z < x_L + x_H - 2\hat{x}_1 + z_L\),
\[
p(z) = Pr(\xi_1 = 0)E[\psi_2] + Pr(\xi_1 = 1)(Pr(\psi_1 \geq x_L + x_H - \hat{x}_1 + z_L|x_1 \geq \hat{x}_1)
\]
\[
\times E[\psi_1|x_1 \geq x_L + x_H - \hat{x}_1 + z_L, x_1 \geq \hat{x}_1] + Pr(\psi_1 < x_L + x_H - \hat{x}_1 + z_L|x_1 \geq \hat{x}_1)E[\psi_2])
\]
\[
= (\frac{\hat{x}_1 - x_L}{x_H - x_L})(\overline{x} + z) + (1 - \frac{\hat{x}_1 - x_L}{x_H - x_L})\left(1 - \frac{x_L + x_H - 2\hat{x}_1 + z_L - z}{x_H - \hat{x}_1}\right)
\]
\[
\times (\frac{z + x_L + x_H - \hat{x}_1 + z_L - z + x_H}{x_H - \hat{x}_1} + (\frac{x_L + x_H - 2\hat{x}_1 + z_L - z}{x_H - \hat{x}_1})(\overline{x} + z))
\]
\[
= (\frac{1}{x_H - x_L})\left(-\frac{1}{2}(x_L^2 + \hat{x}_1^2 + z_L^2 + z^2) + (\overline{x} + z_L)\hat{x}_1 + (x_H + z_L)\overline{x}
\]
\[
- (x_H + z_L + z_H)x_L + (2x_H - \overline{x} - \hat{x}_1 - x_L + z_L)z\right)
\]

Then in case 1b,
\[
p(z) = \begin{cases} 
-\frac{1}{2}(x_L^2 + \hat{x}_1^2 + z_L^2 + z^2) + (\overline{x} + z_L)\hat{x}_1 + (x_H + z_L)x_L + (2x_H - \overline{x} - \hat{x}_1 - x_L + z_L)z & \text{if } z_L \leq z < x_L + x_H - 2\hat{x}_1 + z_L \\
(\hat{x}_1 - x_L)(x_L + x_H) + (x_H - \hat{x}_1)(x_L + x_H) + z & \text{if } x_L + x_H - 2\hat{x}_1 + z_L \leq z \leq z_H
\end{cases}
\]

and
\[
\frac{\partial p}{\partial z} = \begin{cases} 
\frac{2x_H - \overline{x} - \hat{x}_1 - x_L + z_L - z}{x_H - x_L} & \text{if } z_L \leq z < x_L + x_H - 2\hat{x}_1 + z_L \\
1 & \text{if } x_L + x_H - 2\hat{x}_1 + z_L \leq z \leq z_H
\end{cases}
\]
To see that $\frac{\partial V}{\partial z} > 0$ when $z \leq x_L + x_H - 2\hat{x}_1 + z_L$, note that $2x_H - \bar{x} - \hat{x}_1 + z_L \geq x_L + x_H - 2\hat{x}_1 + z_L$ is equivalent to $x_H - x_L \geq \bar{x} - \hat{x}_1$, which must be true because $x_H \geq \bar{x}$ and $\hat{x}_1 \geq x_L$. Therefore $\frac{\partial V}{\partial z} > 0$ in cases 1b and 2. ■

B Extending the Model to Multiple Periods

This section of the paper will relax one of the more restrictive assumptions in the previous sections, that the seller must accept the second period offer if he rejects the first period offer. Although the present paper extends the model only to allow for a third period, hopefully this extension will illustrate the key differences between a two period and multiple period model and present an equilibrium concept that is compatible with any finite number of periods. There are two main differences between the equilibria of the two period model and the three period model. First, in the first period of the three period model, the realtor’s cutoff rule for recommending ‘accept’ will be a function of $z$. Second, in equilibrium there will be ranges of first-period offers for which the realtor will “babble”: his recommendation will contain no information regarding the state of demand, and consequently the seller will ignore it when updating his beliefs. Without this feature of equilibrium, the realtor will sometimes have an incentive to misreport his own preference regarding the seller’s decision in period 1 in order to manipulate the seller’s beliefs regarding the state of demand.

The notation of the three-period model will follow the notation of the two-period model closely. The realtor’s recommendation in period $t$ shall be denoted $\xi_t$ and the seller’s decision in period $t$ shall be denoted $\gamma_t$, where in both cases a value of 1 indicates ‘accept’ and a value of 0 indicates ‘reject’. Let $\hat{x}_1(z)$ and $\hat{x}_2$ denote the realtor’s cutoff value for reporting ‘accept’ in periods 1 and 2, respectively. As discussed, in equilibrium there will be some values for $\psi_1$ such that the realtor will babble; for all other values of $\psi_1$ the realtor will employ a cutoff rule in $x_1$ to determine his recommendation, but the cutoff will be a function of $z$. Let $\tilde{z}_{L,t}$ and $\tilde{z}_{H,t}$ denote the seller’s beliefs about the lowest and highest possible values of $z$ after receiving the period $t$ offer and recommendation. Then the realtor’s value function can be written:

\[ V_R(\psi_1, z, 1) = \max \{ \gamma_1(\psi_1, \tilde{E}(\psi_1, 0)\alpha_{\psi_1} + (1 - \gamma_1(\psi_1, \tilde{E}(\psi_1, 0)))(E[V_R(\psi_2, z, \tilde{z}_{L,1}, \tilde{z}_{H,1}, 2)|\xi_1 = 0]), \gamma_2(\psi_1, \tilde{E}(\psi_1, 1)\alpha_{\psi_1} + (1 - \gamma_1(\psi_1, \tilde{E}(\psi_1, 1)))(E[V_R(\psi_2, z, \tilde{z}_{L,1}, \tilde{z}_{H,1}, 2)|\xi_1 = 1]) \} \]

\[ V_R(\psi_2, z, \tilde{z}_{L,1}, \tilde{z}_{H,1}, 2) = \max \{ \gamma_2(\psi_2, \tilde{E}(\psi_2, 0, \tilde{z}_{L,1}, \tilde{z}_{H,1})\alpha_{\psi_2} + (1 - \gamma_2(\psi_2, \tilde{E}(\psi_2, 0, \tilde{z}_{L,1}, \tilde{z}_{H,1})))(E[V_R(\psi_3, z, \tilde{z}_{L,2}, \tilde{z}_{H,2}, 3)|\xi_2 = 0]), \gamma_2(\psi_2, \tilde{E}(\psi_1, 1, \tilde{z}_{L,1}, \tilde{z}_{H,1})\alpha_{\psi_2} + (1 - \gamma_2(\psi_2, \tilde{E}(\psi_1, 1, \tilde{z}_{L,1}, \tilde{z}_{H,1})))(E[V_R(\psi_3, z, \tilde{z}_{L,2}, \tilde{z}_{H,2}, 3)|\xi_2 = 1]) \} \]
The seller’s value function can be written:

\[ V_R(\psi_3, z, \tilde{z}_{L,2}, \tilde{z}_{H,2}, 3) = -c_R + \alpha \psi_3 \]

Then we can update our equilibrium definition as follows.

**Definition:** a **Bayesian Nash equilibrium** of the game between realtors and sellers is a set of policy functions \( \xi_1(\psi_1, z) \) and \( \xi_2(\psi_2, z, \tilde{z}_{L,1}, \tilde{z}_{H,1}) \) for the realtor, a set of policy functions \( \gamma_1(\psi_1, \tilde{f}(\psi_2|\psi_1, \xi_1)) \) and \( \gamma_2(\psi_2, \tilde{f}(\psi_3|\psi_2, \xi_2, \tilde{z}_{L,1}, \tilde{z}_{H,1})) \) for the seller, and a set of belief updating strategies \( \tilde{f}(\psi_2|\psi_1, \xi_1) \) and \( \tilde{f}(\psi_3|\psi_2, \xi_2, \tilde{z}_{L,1}, \tilde{z}_{H,1}) \) for the seller such that:

1. \( \xi_1(\psi_1, z) \) and \( \xi_2(\psi_2, z, \tilde{z}_{L,1}, \tilde{z}_{H,1}) \) solve the realtor’s problem taking the seller’s policy functions and belief updating strategies as given;
2. \( \gamma_1(\psi_1, \tilde{f}(\psi_2|\psi_1, \xi_1)) \) and \( \gamma_2(\psi_2, \tilde{f}(\psi_3|\psi_2, \xi_2, \tilde{z}_{L,1}, \tilde{z}_{H,1})) \) solve the seller’s problem taking the realtor’s policy functions as given; and
3. \( \tilde{f}(\psi_2|\psi_1, \xi_1) \) and \( \tilde{f}(\psi_3|\psi_2, \xi_2, \tilde{z}_{L,1}, \tilde{z}_{H,1}) \) are consistent with the realtor’s policy functions.

As discussed, the first main difference between the equilibria of three-period model and the two-period model is that in the first period of the three-period model, the realtor’s optimal cutoff strategy is a function of \( z \). This is because when \( z \) is low, it is more likely that \( \psi_2 \) will be so low enough that the seller rejects the offer regardless of the realtor’s recommendation, and this loss of control lowers the realtor’s payoff. To determine the realtor’s optimal cutoff rule in period 1, consider the realtor’s payoff minus \( \alpha z \). If the seller accepts an offer in period \( t \), this value will be \( \alpha(x_t - c_R(t - 1)) \). Call this value the idiosyncratic component of the realtor’s payoff. The realtor will receive the \( \alpha z \) portion of his payoff no matter which offer the seller accepts, but the idiosyncratic portion of his payoff depends on the seller’s decisions and on the realizations of \( x_t \). If the seller accepts the first period offer, the idiosyncratic portion of the realtor’s payoff is \( \alpha x_1 \). Therefore the realtor would prefer that the seller accept any period 1 offers such that \( \alpha x_1 \geq E[\alpha(x_t - c_R(t - 1))] \) and to reject all others. Recall that the realtor’s preference in the penultimate period is for the seller to accept any offers such that \( x_2 \geq x_R - \frac{c_R}{\alpha} \) and reject all others. Call this value \( \bar{x}_2 \). Then the realtor’s optimal cutoff rule in period 1 is defined implicitly by the following mapping (where \( \bar{x}_1(z) \) is written \( \tilde{x}_1 \) on the right-hand side for simplicity):
As the value of $z$ is high) than if the realtor recommends ‘accept’ (indicating that
have a higher expectation of future offers if the realtor recommends ‘reject’ (thus indicating that
and will always accept offers
In this range. Similarly, when
ψ
1
< ψ
1
† as the fixed point of the seller’s expected value of waiting for
x
2
2
+(x
2
2
−x
H
−x
L
)(x
H
−x
2
−x
L
(x
H
−x
L
(2)) in the hypothetical equilibrium. No matter what the realtor recommends, the seller will reject the offer. However, because the seller expects the realtor to report his own preference regarding the offer truthfully, the seller will have a higher expectation of future offers if the realtor recommends ‘reject’ (thus indicating that $z$ is high) than if the realtor recommends ‘accept’ (indicating that $z$ is low). Intuitively, this state
\[ \tilde{x}_1(z) = \begin{cases} \frac{1}{2} (x_H^2 - x_H^2 - \tilde{x}_1(\tilde{x}_1)^2 - z^2) + (x_H - x_L + \tilde{x}_1(\tilde{x}_1)) (\tilde{x}_2 - x_H - \tilde{x}_1(\tilde{x}_1)) + x_L (\tilde{x}_2 - x_H - \tilde{x}_1(\tilde{x}_1)) \tilde{x}_2, & \text{if } \tilde{x}_2 + \tilde{z}_{H,1}(\tilde{x}_1) \leq \tilde{\tau} + \tilde{x}_1(\tilde{x}_1) \\ \frac{\tilde{x}_2 + x_H}{2} + \left( \frac{x_H - x_L}{x_H - x_L} \right) \tilde{\tau}_2, & \text{if } \tilde{x}_2 + \tilde{z}_{H,1}(\tilde{x}_1) \geq \tilde{\tau} + \tilde{x}_1(\tilde{x}_1) \end{cases} \]
\[ \tilde{z}_{L,1} = \begin{cases} \max(z_L, \tilde{z}), & \text{if } \gamma_1 = 0 \\ \max(z_L, \psi_1 - x_H), & \text{if } \gamma_1 = 1 \end{cases} \text{ and } \tilde{z}_{H,1} = \begin{cases} \min(z_H, \psi_1 - x_L), & \text{if } \gamma_1 = 0 \\ \min(\tilde{\tau}, \psi_1 - x_L), & \text{if } \gamma_1 = 1 \end{cases} \]
Because $\tilde{x}_1(z)$ is the fixed point of this functional equation, I have estimated it numerically. A description of the estimation algorithm is included at the end of this section. Figure 6 illustrates
the realtor’s optimal period 1 cutoff rule as a function of $z$. Given the realtor’s cutoff rule, the seller will update his beliefs concerning $z$ as follows. Define $\tilde{z}$ as the value of $z$ such that $\psi_1 - z = \tilde{x}_1(z)$. Further define
Then the seller’s posterior belief about the distribution of $z$ will again be that $z \sim U[\tilde{z}_{L,1}, \tilde{z}_{H,1}].$ In period 2, the seller’s belief updating strategy will be the same as the one outlined in the two-period model.
The second main difference between the equilibria of the two and three period models is the presence of ‘babbling regions’ in period 1 of the three period model. To see that these regions are a necessary feature of equilibrium, consider a hypothetical equilibrium in which the realtor always reports his own preference truthfully to the seller in period 1, and in which the realtor’s and seller’s behavior in the second period is the same as their behavior in the first period of the two-period model. Then the seller’s expected value of rejecting the first offer would look as in the top panel of figure 7. Define $\psi_1^*$ as the fixed point of the seller’s expected value of waiting for period 2 conditional on the realtor recommending ‘accept’, and $\psi_1^†$ as the fixed point of the seller’s expected value of waiting for period 2 conditional on the realtor recommending ‘reject’. Then for $\psi_1 < \psi_1^*$, the expected value of waiting for period 2 is higher than $\psi_1$ whether the the realtor recommends ‘accept’ or ‘reject’–the seller’s action will not depend on the realtor’s recommendation in this range. Similarly, when $\psi_1 > \psi_1^†$, the expected value of waiting for period 2 will be below $\psi_1$ no matter what the realtor recommends. Therefore, the seller will always reject offers $\psi_1 < \psi_1^*$ and will always accept offers $\psi_1 > \psi_1^†$. Consider the realtor’s best response when $\psi_1 < \psi_1^*$ and $x_1 > \tilde{x}_1(z)$ in the hypothetical equilibrium. No matter what the realtor recommends, the seller will reject the offer. However, because the seller expects the realtor to report his own preference regarding the offer truthfully, the seller will have a higher expectation of future offers if the realtor recommends ‘reject’ (thus indicating that $z$ is high) than if the realtor recommends ‘accept’ (indicating that $z$ is low). Intuitively, this state
\[ \psi_1^* = \psi_1^† = \frac{1}{2} \left( x_H + x_L \right). \]
Again I omit the proof, but it will follow the proof of Lemma 1 closely.
of affairs cannot be optimal for the realtor: we have shown in the two-period model that in the second-to-last period, the seller will always reject an offer that the realtor recommends rejecting. However, there are second-period offers the realtor recommends accepting that the seller will not accept. Therefore, the realtor will always prefer that the seller be more pessimistic (have a lower expectation of $\psi$) in the second period: a pessimistic seller is more likely to accept offers the realtor would like him to accept, but will always reject offers the realtor would like him to reject. Therefore, the realtor has a unilateral incentive to deviate from his proposed strategy in the hypothetical equilibrium. For any offer $\psi_1 < \psi_1^*$, the realtor should recommend ‘reject’.

Babbling regions solve the problem of the realtor’s incentive to misreport his preferences in the first period. As we have seen, if the realtor’s recommendation in period 1 changes the seller’s expectation of future offers but does not change the seller’s action, the realtor will always choose to send the message that will make the seller more pessimistic. Then in equilibrium, it cannot be the case that the realtor’s first period recommendation changes the seller’s beliefs when it does not change the seller’s first period action. Therefore, in the equilibrium I will consider, the realtor will babble in the first period when his recommendation will not change the seller’s action, and will report his own preference truthfully when his recommendation is decisive for the seller’s action. The bottom panel of figure 7 illustrates the period 1 equilibrium, assuming that in the second period the seller and the realtor play the same strategies as they did in the first period of the two-period model. The second period equilibrium will then look like it does in Figure 4.

Because $\hat{x}_1(z)$ must be estimated numerically, the expected sales price and expected time to sale must be simulated as well. Figure 8 shows the results from such a simulation. The general pattern from the two-period model remains intact: as aggregate demand rises, the expected sales price rises and the expected time on market mostly falls. However, there is a slight bump in the expected time to sale for high levels of $z$. This is due to the upward-sloping portion of $\hat{x}_1(z)$, which causes the realtor to recommend rejecting a higher percentage of offers when $z$ is high. Overall, however, the correlation between expected sales price and expected time to sale is negative, and the simulated results of the three period model are consistent with the stylized facts observed in the data.

I used the following algorithm for finding the realtor’s cutoff rule $\hat{x}_1(z)$ in the three period model:

1. Pick a candidate schedule for $\hat{x}_1(z)$. In practice I chose $\hat{x}_1(z) = \pi$ for all $z$.

2. On a fine grid of points for $z$:
   (a) Go through a fine grid of points for all values of $x$ to create a grid of all values of $\psi$ consistent with each value of $z$.
   (b) Calculate $\tilde{z}_{L,1}$ and $\tilde{z}_{H,1}$ for each value of $\psi$ conditional on the realtor recommending ‘reject’.
   (c) For each value of $\psi$, find the expected value of the idiosyncratic component of the realtor’s payoff if the seller rejects the first period offer. Denote this value $\pi_R$. For a fixed $z$, this gives $\pi_R$ as a function of $x$. 

xii
(d) Find the fixed point of $x_R(x)$; use this value as the new candidate for $\hat{x}_1(z)$.

3. Repeat this procedure using the new schedule for $\hat{x}_1(z)$ until the maximum distance between the old and new schedules is below a specified tolerance level.
Figure A: Percent Price change from 2005: FHFA, Case-Shiller, and Zillow

Graphs by (mean) pmsa