House Price Dynamics with Dispersed Information*

Giovanni Favara† and Zheng (Michael) Song‡

September 2011
(first version September 2008)

Abstract

We use a user-cost model to study how dispersed information among housing market participants affects the equilibrium house price. In the model, agents are disparately informed about local economic conditions, consume housing services, and speculate on price changes. Information dispersion leads agents to have heterogeneous expectations about housing demand and prices. Optimists, who expect high house price growth, buy in anticipation of capital gains; pessimists, who expect capital losses, prefer to rent. Because of short-selling constraints on housing, pessimistic expectations are not incorporated in the price of owned houses and the equilibrium price is higher and more volatile relative to the benchmark case of common information. We present evidence supporting the model’s predictions in a panel of US cities.

JEL Classification Numbers: R21, R23, G10

Keywords: Housing prices; Information dispersion; Income dispersion

---

*We are grateful to Alessandro Beber, Darrell Duffie, Bernard Dumas, Simon Gilchrist, John Hassler, Christian Hellwig, Ethan Kaplan, Per Krusell, Rafael Lalíve, John Leahy, Prakash Loungani, Torsten Persson, Andrea Prat, David Romer, Martin Schneider, Pascal St. Amour, Gianluca Violante, Mirko Wiederholt, Alexandre Ziegler, Fabrizio Zilibotti, and seminar participants at HEC Lausanne (DEEP and IBF), the University of Zurich (IEW and ISB), Shanghai JiaoTong University, the Sverige Riksbank, Bocconi University, Luiss, the Einaudi Institute for Economics and Finance, IMF, ECB, George Washington University, St. Louis FED, the 2009 Society for Economic Dynamics Meeting, the 2009 LSE-FMG conference on Housing and the Macroeconomy, the 2008 North American Summer Meeting of the Econometric Society, the 2008 Royal Economic Society Meeting, for helpful discussions and comments. Financial support by the National Centre of Competence in Research "Financial Valuation and Risk Management" (NCCR FINRISK, IPA2) is gratefully acknowledged.

†International Monetary Fund and Swiss Finance Institute. Email: gfavara@imf.org
‡Booth School of Business, University of Chicago. E-mail: zheng.song@chicagobooth.edu
1 Introduction

The U.S. housing market has experienced substantial price fluctuations in the last two decades. Figure 1 gives an example of such fluctuations for the aggregate U.S. economy and a representative sample of U.S. cities. As shown, housing prices not only have different trends in different cities, but also display heterogeneous short-run dynamics. In the opinion of many housing-market observers (see, e.g., Glaeser and Gyourko, 2006, 2007) these dynamics are difficult to explain through the lens of a user cost model in which house prices are determined by an indifference condition between owning and renting. The reason is that in such a model (Poterba, 1984; Henderson and Ioannides, 1982), the cost of owning depends on variables that either do not vary much over time (e.g., property taxes) or are constant across markets (e.g., interest rates).

The goal of this paper is to propose an extension of the standard user cost model to rationalize the heterogeneous behavior of housing prices in the U.S. In our model agents have dispersed information about local economic conditions and thus hold heterogeneous expectations about house prices. Since the cost of owning is inversely related to the expected resale value of houses, optimists prefer to buy and pessimists prefer to rent. As a result, house prices, reflecting only the opinion of optimists, will be higher and more volatile the larger the difference in expectations. To the extent house price expectations depend on local economic conditions, and economic conditions vary across markets and time, our model provides a novel interpretation behind the price fluctuations displayed in Figure 1.

Our analysis is based on four assumptions: 1) income is the main determinant of housing demand; 2) agents hold heterogeneous expectations about house prices dynamics, and buy houses for speculative reasons; 3) housing supply is inelastic, and 4) it is impossible to short sell houses. These assumptions are motivated by several aspects of the US market. First, there is evidence that income affects the demand for housing either because richer agents can afford to spend more on houses (Poterba, 1991; Englund and Ioannides, 1997) or because higher income relaxes credit constraints (Ortalo-Magné and Rady, 2006; Almeida, Campiello

---

1In some cities, such as Los Angeles, housing prices have moved in tandem with the overall national index, though they have moved much more. In other cities, prices movements have been quite heterogeneous. In Miami, for example, the house price index has been steady for almost two decades before experiencing an exponential increase beginning in 2000; in San Antonio, it has declined since the 1980s; in Rochester, it has displayed an inverse “U-shaped” history; in Memphis, the same index has gone through periodic cycles.

2While there is consensus that differences in state level property taxes cannot explain the house price behavior across markets, the debate concerning the relationship between interest rates and house prices is less conclusive. McCarthy and Peach (2004) and Himmelberg, Mayer and Sinai (2005) argue that the recent house price boom in the U.S. was largely brought about by low interest rates. In contrast, Shiller (2005, 2006) documents a non-significant relationship between house prices and interest rates over a longer period of time.
and Liuet, 2006). Second, surveys of housing market participants (Case and Shiller, 1988, 2003; Piazzesi and Schneider, 2009) reveal that agents’ desire to buy is strongly influenced by their expectations to resell houses at higher prices. These surveys also document that home buyers disagree about the causes of house price movements, and expectations are largely influenced by past and current economic conditions (see, e.g., Case, Quigley and Shiller, 2003). Third, housing supply adjusts slowly to local demand shocks because of regulations, zoning laws or geographical constraints (see, e.g., Glaeser and Gyourko, 2003; Glaeser, Gyourko and Saks 2005, Saiz, 2010). Finally, the impossibility of selling housing short is a very natural assumption for the housing market, relative to almost any other asset markets.

Taken together, these three ingredients suggest a specific mechanism through which changes in income may generate more than proportional changes in house prices: if income not only influences housing demand, but also shapes expectations of future house prices, an income shock may initiate a dynamic process that, through heterogeneous expectations, the short-selling constraint, and the inelastic housing supply, runs from expected prices to house demand and back to house prices.

To formalize this mechanism, we propose a model of housing prices in which agents speculate on future price changes and consume housing services by either buying or renting. In our model, the supply of housing is inelastic and the demand fluctuates stochastically because information about local economic conditions is imperfect. To estimate the unknown state of the economy, agents rely on public and private signals, including their own income shocks. As a result, idiosyncratic income shocks translate into heterogeneous expectations of aggregate housing demand, and — given the fixed supply — into heterogeneous expectations of house prices.

As in the standard user-cost model of housing prices, the equilibrium price is pinned down by an indifference condition between owning and renting. The key departure from the standard model is that expectations are heterogeneous. Hence, the equilibrium price no longer reflects the indifference condition of the average market participant, but it is determined by the expectations of the most optimistic agents in the market. This is so because pessimists, who expect future capital losses, perceive the user cost to be higher than the cost of renting. Since these agents derive utility from housing services and cannot short sale houses, they move out of the market of homes for sale and rent from the optimists who, for speculative reasons, buy units in excess of their demand for housing services.

The direct implication is that the price of owned houses is higher and more volatile relative to a benchmark scenario where information is not dispersed. The price is higher because it reflects only the opinion of the optimists. The price is also more volatile because
the housing demand of the optimists is not only affected by fundamental shocks but also by noisy information. Were the rental market absent and short sales allowed, the equilibrium price would only reflect the average opinion, rather than the most optimistic opinion in the market.

This result is reminiscent of the Miller’s (1977) intuition that when agents have heterogeneous beliefs and short selling is not possible, asset prices may be above their fundamental value, since it is only the opinion of the most optimistic investors that is embedded in the equilibrium price. Because our set-up is more akin to a noisy rational expectations model than to a model with heterogeneous priors, we can show that house prices may exceed their fundamental value even if agents use the equilibrium price to update their inference about the state of the economy — provided the price is not fully revealing.

In our model credit frictions play no role even though mortgage credit is an important feature of the housing market. We abstract from credit frictions because the main result of our model would not change in a setting with borrowing and lending, provided short selling of houses is not allowed and there is a rental market. The reason is that optimists would continue to be the marginal buyers even if they were credit constrained. Of course, the pricing equation would be different, reflecting among other things the collateral value of houses, if houses are pledged as collateral (see e.g., Fostel and Geanakoplos, 2008 and Geanakoplos, 2009). However, our main result that the equilibrium price is higher the larger the difference in expectations would still hold true.

Central to the result that house prices are higher and more volatile the higher the dispersion of income is the mapping from income shocks to information dispersion. If income shocks did not affect the information set of market participants’ income dispersion would not influence the equilibrium price. In fact, when expectations are homogeneous everyone is indifferent between owning and renting. Thus, even if high income agents would demand more housing services, low income agents would demand less, leaving the equilibrium price unchanged.

An empirical evaluation of our model is difficult because there is no data on the dispersion of information about local market conditions. To overcome this problem, we follow the logic of the model and use the dispersion of city income shocks as a proxy for information dispersion about city income. In our model local house prices depend on expectations about local economic conditions. Income shocks not only influence housing demand, but also shape expectations of future house prices. Thus, if city residents are employed in different industries and are imperfectly informed about the city income, within-city industry income shock may be easily seen as a source of information about current local economic conditions. Using a large panel of US cities, we find, in line with the model’s predictions, that house prices are
higher and more volatile in cities where our proxy of information dispersion is higher.

The rest of the paper proceeds as follows. Section 2, relates our model to the relevant literature. Section 3, presents the baseline model and derives the main determinants of the equilibrium house price. Section 4, studies the benchmark case in which agents hold imperfect but common information about local economic conditions. Section 5, derives the main model’s predictions when information is imperfect and dispersed, and agents use the equilibrium price to infer the unknown state of the economy. Section 6, discusses our proxy for information dispersion and our empirical findings. Section 7 concludes, and all proofs are in the Appendix.

2 Related Literature

Methodologically, our paper follows the user-cost approach of Poterba (1984) and Henderson and Ioannides (1982), in which a prospective buyer is indifferent between renting and owning, and the cost of owning depends on, among other variables, property taxes, the opportunity cost of capital and the expected capital gains on the housing unit. While some papers have studied the house prices effects of changes in taxes (Poterba, 1991) and interest rates (Himmelberg, Mayer and Sinai, 2006; McCarthy and Peach, 2004), the role played by heterogeneity in the expected rate of price changes has remained so far unexplored. This is so because differences in expectations cannot arise in a standard user-cost model with homogeneous information. We complement this literature by showing that information dispersion across markets, and within markets over time, helps to rationalize part of the house price changes documented in Figure 1 — more than changes in property taxes, which are fairly constant over time, or interest rates, which are constant across markets.

The theme of our paper that changes in income may have more-than-proportional effects on house prices is similar in spirit to the work of Stein (1995) and Ortalo-Magné and Rady (2006). In these papers, agents buy houses by borrowing, and the ability to borrow is directly tied to the value of houses. Therefore, a positive income shock that increases the housing demand and price relaxes the borrowing constraint, which further increases the demand for houses. Our paper differs from Stein, and Ortalo-Magne and Rady, in three important ways. First, in our model agents do not borrow to buy houses and so the amplification mechanism runs only from changes in expected prices, via household income, to current prices, via changes in the speculative demand. Second, in our model, agents do not need to own houses to consume housing services; they can also use the rental market. Third, it is not only the level, but also the dispersion of income that affects house prices.

For this reason, our paper is also related to Gyourko, Mayer and Sinai (2006) and Van
Nieuwerburgh and Weil (2010). Gyourko et al. argue that the interaction between an inelastic supply of houses and the skewing of the income distribution generates significant price appreciations in superstar cities (i.e., cities with unique characteristics preferred by the majority of the population) because wealthy agents are willing to pay a financial premium to live in these areas, bidding up prices in the face of a relatively inelastic supply of houses. Van Nieuwerburgh and Weil use a similar mechanism to explain the correlation between the dispersion of U.S. house prices and the cross-sectional dispersion of U.S. wages, though in their model, agents move across cities for productive rather than preference reasons. Our paper differs from these contributions because it highlights a different channel through which income dispersion affects house prices. In our framework, income shocks affect agents’ perception of local economic conditions, leading to heterogeneous expectations about current and future economic fundamentals. As a consequence, differences in expectations are more pronounced when, ceteris paribus, income is more dispersed. Another important difference is methodological. In our model, prices are determined by a no-arbitrage condition between buying and renting, while in Gyourko et al. and van Nieuwerburgh and Weil, prices are determined by a spatial no-arbitrage condition with owners indifferent between different locations, given local wages and amenities. The spatial equilibrium approach is, however, more suitable for studying the long-run distribution of housing prices as opposed to high frequency price variations, which is the main focus of our analysis.

Our paper is also related to a large literature in macroeconomics and finance that studies the role of imperfect information among decision makers. In fact, our model can be seen as an application of the Phelps-Lucas hypothesis to the housing market, in the sense that imperfect information about the nature of disturbances to the economy makes agents react differently to changes in market conditions. Part of our work also shares many features with the literature on the pricing of financial assets in the presence of heterogeneous beliefs and short-sale constraints (e.g., Miller, 1977; Harrison and Kreps, 1979; Hong and Stein, 1999 and Sheinkman and Xiong, 2003). In this literature, if agents have heterogeneous beliefs about asset fundamentals and face short-sale constraints, the equilibrium asset price reflects the opinion of the most optimistic investors. We adapt the same idea to the housing market. In our model, pessimists would short their houses if they could. By consuming housing services through the rental market, they do not participate in the market of houses for sale and the price of owned houses ends up reflecting only the most optimistic opinion in the market, rather than the average opinion. In this sense, our model is related to the recent work of Piazzesi and Schneider (2009) and Burnside, Eichenbaum and Rebelo (2011). These papers, however, use heterogeneous beliefs (as opposed to heterogeneous expectations) and search frictions to explain why house prices fluctuate much more than fundamental shocks.
3 The Model

3.1 Information

The economy is populated by an infinite sequence of agents with unit mass that lives for two periods. In the first period, agents supply labor and make savings and housing decisions; in the second period, they consume the return on savings and housing. The wage $W^j_t$, at which labor is supplied inelastically, is equal to

$$W^j_t = \exp \left( \theta_t + \varepsilon^j_t \right),$$

where $\theta_t$ is the economy income and $\varepsilon^j_t$ an individual-specific wage shock. The individual-specific shocks, $\varepsilon^j_t$, which are the only source of income heterogeneity, are serially independent and have normal distribution with zero mean and variance $\sigma^2\varepsilon$. We make the assumption that $\theta_t$ follows an AR(1) process,

$$\theta_t = \rho \theta_{t-1} + \eta_t, \text{ with } \rho \in (0, 1],$$

where $\eta_t$ is independently and normally distributed with zero mean and variance $\sigma^2\eta$. When agents cannot observe the realization of $\theta_t$, $\varepsilon^j_t$ becomes a source of information heterogeneity. In other words, the individual wage $W^j_t$ is the agent $j$'s noisy private signal about the unobservable aggregate shock, $\theta_t$.

To make the analysis simple, we consider only two groups of agents, $j = 0, 1$, each with equal mass. We also make the standard assumption that idiosyncratic shocks cancel out in the aggregate or, equivalently, the average private signal is an unbiased estimate of $\theta_t$:

Assumption 1: $\sum_j \varepsilon^j_t = 0$.

3.2 Preferences

Agents have logarithmic preferences over housing services, $V^j_t$, and second-period consumption, $C^j_{t+1}$,

$$U^j_t = A^j_t \log V^j_t + E^j_t \log C^j_{t+1},$$

where $E^j_t$ denotes the expectation operator based on household $j$’s information set at time $t$ (to be specified later), and the parameter $A^j_t$ is a preference shock,

$$A^j_t = \exp \left( a_t + \nu^j_t \right),$$
which consists of an aggregate taste shock, \( a_t \), and an idiosyncratic noise \( \nu_t^j \). We assume that \( a_t \) and \( \nu_t^j \) are independent and normally distributed with zero mean and variance \( \sigma_a^2 \) and \( \sigma_{\nu}^2 \). We also consider the limiting case where the variance of \( \nu_t^j \) is arbitrarily large, so that knowing one’s own individual taste provides no information about the aggregate taste. Finally, the preference shock \( A_t^j \) is introduced to have another source of noise in the demand of housing. Preference shocks ensure that house prices are not fully revealing, a feature we exploit in Section 5 when we allow agents to use the equilibrium price to update their beliefs about \( \theta \).\(^3\)

### 3.3 Budget constraint

In the first period, after the realization of the idiosyncratic income, agents decide how many housing units to buy, \( H_t^j \geq 0 \), at the unit price, \( P_t \). They also choose the quantity of housing services to consume, \( V_t^j \), and the units to rent out, \( H_t^j - V_t^j \), at the rental price \( Q_t \). The stock of houses owned at time \( t \) is sold to agents entering the economy at \( t + 1 \). At the end of period \( t \), the residual income is saved at the gross interest rate, \( R \).

For type-\( j \) agents, the resource constraint is thus:

\[
C_{t+1}^j = R \left( W_t^j - P_t H_t^j + Q_t \left( H_t^j - V_t^j \right) \right) + P_{t+1} H_t^j, \tag{4}
\]

with

\[
H_t^j \geq 0. \tag{5}
\]

### 3.4 Optimal house demand

Agents’ intertemporal decisions consist of choosing \( H_t^j \) and \( V_t^j \) to maximize (3) subject to (4) and (5). It is immediate to establish that the optimal demand for \( V_t^j \) and \( H_t^j \) satisfy the following first-order conditions:

\[
\frac{A_t^j}{V_t^j} = E_t^j \left[ \frac{RQ_t}{C_{t+1}^j} \right], \tag{6}
\]

\[
E_t^j \left[ \frac{R(U_t - Q_t)}{C_{t+1}^j} \right] \geq 0, \quad \text{and} \quad H_t^j \geq 0 \tag{7}
\]

where

\[
U_t \equiv P_t - \frac{P_{t+1}}{R}, \tag{8}
\]

\(^3\)Our specification of preferences assumes away any intertemporal consumption-saving decision. This has, however, inessential consequences for the main focus of our analysis: the rental-owning margin.
denotes the (per-unit) user cost of housing, which decreases with next-period house price, $P_{t+1}/R$.

According to equation (6), agents consume housing services until the marginal benefit (the LHS) equals the marginal cost, defined in terms of next-period consumption (the RHS). The optimal demand for owned houses is implicit in equation (7), which relates the cost of owning, $U_t$, to the cost of renting housing services, $Q_t$.

### 3.5 The linearized optimality conditions

To deliver explicit solutions, we log-linearize equations (6) and (7) around the “certainty” equilibrium: i.e., the equilibrium prevailing when both aggregate and idiosyncratic shocks are zero. Using lower-case letters to denote variables in percentage deviations from the equilibrium with certainty, Appendix I shows that a log-linear approximation of (6), (7) and (8) leads to

\[ v^j_t = w^j_t + a^j_t - q_t, \quad (9) \]

and

\[ E^j_t u_t \geq q_t, \quad \text{and} \quad H^j_t \geq 0, \quad (10) \]

where

\[ u_t \equiv \frac{(1 + r)P_t - P_{t+1}}{r}, \quad r \equiv R - 1 > 0, \quad (11) \]

is the linearized user cost, and $a^j_t \equiv (a_t + v^j_t) / 2$ denotes the average preference shock in group $j$.

According to equations (9) and (10) the demand for housing services depends on current period variables (income, preferences and rental prices), while the decision to own houses depends on the expected cost of owning relative to renting.\(^5\) With the convention that agents in group $j = 1$ are relatively more optimistic about the next-period house price, i.e.,

\(^4\)Our specification of the user cost is deliberately simple. We could have assumed that for each unit owned, agents also incur a cost equal to a fraction $M_t$ of the nominal value of housing, $P_t H^j_t$. $M_t$ can be thought of as including maintenance and depreciation costs, property taxes, interest payments on mortgages, etc. Under this alternative specification, the user cost of housing would be

\[ U_t = P_t (1 + M_t) - \frac{P_{t+1}}{R}. \]

As long as housing-market participants are homogeneously informed about $M_t$, none of the results presented below is affected, though the algebra would be more cumbersome.

\(^5\)Notice that because a log-linearization of (7) does not involve $H^j_t$ the demand for owned houses is pinned down by (10) and the market clearing condition (see the next subsection). Notice also that in an equilibrium with homogeneous expectation the demand for owned houses is indeterminate since every agent would be indifferent between renting and owning: equation (10) would hold with equality for any $j$. 

\[ \]
$E^1_t p_{t+1} > E^0_t p_{t+1}$, we can rewrite equation (10) as follows:

$$E^0_t u_t > q_t \quad \text{and} \quad H^0_t = 0$$ (12)

$$E^1_t u_t = q_t \quad \text{and} \quad H^1_t > 0,$$ (13)

suggesting that with heterogeneous expectations pessimists choose to own no housing units, $H^0_t = 0$ (as they perceive the cost of ownership to be higher than the cost of renting) and optimists choose to own (as they expect higher prices in the future). As a result, in equilibrium, optimists own all the housing units, consume housing services, $V^1_t$, out of the units owned, $H^1_t$, and rent out the difference, $H^1_t - V^1_t$, to the pessimists.

### 3.6 The equilibrium rental and house price

Assuming a fixed housing supply, $S$, the rental price is pinned down by the market clearing condition for housing services:

$$S = \frac{V^1_t + V^0_t}{2}.$$  

Since $V^j_t = (1 + v^j_t) V$, and $V = S$ in the certainty equilibrium, the market clearing condition can be rewritten as $\sum_j v^j_t = 0$. Together with (9), it yields

$$q_t = \theta_t + a_t,$$ (14)

where

$$\theta_t = \frac{w^1_t + w^0_t}{2} \quad \text{and} \quad a_t = \frac{a^1_t + a^2_t}{2},$$

denote the average income and the average preference shock for housing services.

The equilibrium house price is pinned down by the indifference conditions of the optimists (13), which can be written as:

$$p_t = \frac{r}{1 + r} q_t + \frac{1}{1 + r} E^1_t p_{t+1},$$ (15)

or, using (14) to substitute out $q_t$, as:

$$p_t = \frac{r}{1 + r} f_t + \frac{1}{1 + r} E^1_t p_{t+1} + \frac{1}{1 + r} \tilde{E}^1_t p_{t+1},$$ (16)

where

$$f_t \equiv \theta_t + a_t$$ (17)
summarizes average fundamental variables, and

$$\bar{E}_t p_{t+1} \equiv \frac{E_t^1 p_{t+1} + E_t^0 p_{t+1}}{2}, \quad \tilde{E}_t p_{t+1} \equiv \frac{E_t^1 p_{t+1} - E_t^0 p_{t+1}}{2},$$

denotes, respectively, the average expectation and the difference in expectations about tomorrow's price.

In equation (16), as in a standard house pricing equation, $p_t$ depends on fundamentals, $f_t$, and the average expectation on the future house price. The extra term, $\tilde{E}_t p_{t+1}$, is non-standard and arises because agents may hold heterogeneous expectations. In the next two sections, we make different assumptions about agents' information sets in order to evaluate how $\bar{E}_t p_{t+1}$ and $\tilde{E}_t p_{t+1}$ influence the determination of the equilibrium house price.

4 Homogeneous Information

We start with the benchmark case in which agents are homogeneously informed about the state of the economy, $\theta_t$, and thus rely only on public information, $\theta_{t-1}$, to infer $\theta_t$. In other words, agents share a common information set. In this case individual expectations coincide with the average expectation, i.e., $E_t^j p_{t+1} = \bar{E}_t p_{t+1}$ and the difference in expectations is zero, $\tilde{E}_t p_{t+1} = 0$.

Iterating equation (16) forward and imposing a stationary condition on prices, Appendix II shows that the average expectation of tomorrow's price can be written as

$$\bar{E}_t p_{t+1} = \phi \rho \theta_{t-1},$$

with

$$\phi \equiv \frac{r \rho}{1 + r - \rho}.$$  

The average expectation depends on past fundamentals, $\theta_{t-1}$, because $\theta_t$, which is unobservable, follows an $AR(1)$, but does not depend on the preference shock, $a_t$, because by assumption it has zero mean. Inserting (18) into (16), and recalling that $\tilde{E}_t p_{t+1} = 0$, we have

**Proposition 1** The equilibrium house price with homogeneous information, $p^*$, can be written as

$$p_t^* = f_t + \Lambda_t,$$

where $f_t$ is given in (17) and

$$\Lambda_t \equiv \frac{\phi \rho \theta_{t-1} - \theta_t - a_t}{1 + r}$$
is an expectation error.

We interpret $p^*_t$ as the “fundamental” price of owned houses, because it reflects the average opinion in the market which, by Assumption 1, is an unbiased estimate of the unknown fundamental.

5 Heterogeneous Information

We now consider a setting where agents use the current realization of their income, $w^j_t$, as well as the public signal, $\theta_{t-1}$, to make an optimal inference about $\theta_t$. Agent $j$’s information set at $t$ is,

$$\Omega^j_t = \{w^j_t, \theta_{t-1}\} \quad j = 0, 1.$$  

It is important to notice that the equilibrium house price is not included in $\Omega^j_t$. This assumption is made only to simplify the characterization of the channels through which information dispersion affects the equilibrium price. As we will discuss in Section 5.1, this assumption is not essential for our results.

With signals $w^j_t$ and $\theta_{t-1}$, the ability of agent $j$ to estimate $\theta_t$ depends on the relative magnitude of $\sigma^2_\varepsilon$ and $\sigma^2_\eta$. Because of our assumption of independently and normally distributed errors, the projection theorem implies

$$E^j_t \theta_t = (1 - \lambda) \rho \theta_{t-1} + \lambda w^j_t, \quad (20)$$

where the weight $\lambda \equiv \sigma^2_\eta / (\sigma^2_\eta + \sigma^2_\varepsilon)$ reflects the relative precision of the two signals. With $\lambda > 0$, expectations among agents are heterogeneous, and both average expectations and differences in expectations become important determinants of the equilibrium price. Moreover, since expectations depend on $w^j_t$, the optimists (pessimists) are those with higher (lower) realization of the idiosyncratic shock.

Iterating equations (16) and (20) forward and excluding explosive price paths, Appendix III shows that difference in expectations, and the average expectation of the future price are,
respectively,

\[
\begin{align*}
\tilde{E}_t p_{t+1} &= \phi \lambda i_t, \\
E_t p_{t+1} &= \phi \rho \theta_t \theta_{t-1} + \frac{\phi \lambda}{r} I + \phi \lambda (\theta_t - \rho \theta_{t-1}),
\end{align*}
\]

(21) (22)

where

\[ i_t \equiv \varepsilon_t - \varepsilon_t^0, \]

denotes the dispersion of information between the two groups of agents, and

\[ I \equiv \int_0^{\infty} xd\Gamma(x) \]

measures the average degree of information heterogeneity in the economy, with \( \Gamma \) denoting the distribution of \( i_t \).

Equation (21), stems from the fact that agents are disparately informed and assign a positive weight to their private signal in estimating \( \theta_t \). Differences in expectations are, therefore, proportional to the dispersion in private signals. Equation (22) is the equivalent of equation (18). It differs from (18) because dispersed information introduces two additional terms, each proportional to the weight agents assign to their private signals. The first term, \( \phi \lambda I/r \), arises because prices are forward-looking: it is not only the current dispersion of information that influences the price of housing, but also the dispersion of future information. The second term, \( \phi \lambda (\theta_t - \rho \theta_{t-1}) \), capturing the average misperception in the economy, arises because agents use only part of the information contained in the public signal, \( \theta_{t-1} \), to make the optimal inference about \( \theta_t \). The slow reaction to changes in fundamentals has the effect of introducing inertia in the way average expectations are formed, which accords well with the idea that housing market expectations tend to be extrapolative (see Case and Shiller, 1988, 2003). Plugging these expressions in (16), we have

**Proposition 2** The equilibrium house price with heterogeneous information is

\[
p_t = p_t^* + \lambda \Upsilon_t,
\]

(23)

where, \( p_t^* \), is the fundamental price given in (19), and

\[
\Upsilon_t \equiv \phi \left( \frac{\theta_t - \rho \theta_{t-1}}{1 + r} \right) + \phi \frac{I}{r(1 + r)} + \phi \frac{i_t}{1 + r}
\]

(24)

summarizes the role of information dispersion.

With heterogeneous information (i.e., \( \lambda > 0 \)), \( p_t \) is higher than \( p_t^* \) for two reasons. First,
the unconditional mean of $\Upsilon_t$ is positive, implying that information dispersion leads to a higher equilibrium house price. This is the case because optimists estimate a higher $\theta_t$ (see equation (20)) and, thus, expect higher future prices (see equation (22)); conversely, pessimists expect capital losses. As discussed in Section 3, this implies that pessimists prefer to consume housing services through the rental market and so move out of the market of homes for sale. Hence, the equilibrium house reflecting only the opinion of optimists stays above its fundamental value. Second, the price misalignment becomes more pronounced the larger the information dispersion, $i_t$: when $\varepsilon_t^1$ increases relative to $\varepsilon_t^0$, optimistic agents demand more houses for speculative reasons, while pessimists continue to demand no housing units. Overall, these two effects lead to the prediction that housing prices unambiguously increase with information dispersion.

Another testable prediction arises in comparing (23) and (19). It is straightforward to see that relative to the benchmark case of homogeneous information, the volatility of house prices is higher the larger the average misperception in the economy, $\sigma_t^2$, and the larger the variance of information dispersion, $\sigma_i^2$:

$$V(p_t) - V(p_t^*) = \left( \frac{\lambda \phi}{1 + r} \right)^2 (\sigma_t^2 + \sigma_i^2) > 0.$$  

(25)

The extra source of price volatility arises because the equilibrium price with dispersed information is influenced not only by fundamental shocks but also by noise shocks.

## 5.1 Credit Constraint

Before proceeding, it is worth discussing whether our model’s predictions may also arise in a setting that abstracts from heterogeneous information but features financial frictions. It turns out that the implications of our model do not depend on the assumption that the demand for housing is independent of credit conditions. What is crucial for our findings is that information is heterogeneous and it is not possible to short sell houses. To see why, consider a model where information is not dispersed, and the ability of the two groups of agents to borrow is proportional to their own income. In such a setting, even if the demand for housing would depend on the agents ability to borrow, housing prices would reflect the average expectation in the market. As a result positive and negative idiosyncratic income shocks would cancel out in aggregate. A model with only credit frictions would thus generate different predictions relative to a setting in which information is dispersed and expectations on the state of the economy depend on idiosyncratic income shocks. Our model’s predictions would also continue to hold if agents had dispersed information and faced credit constraints,
provided short selling is not allowed. The reason is that with dispersed information and short sale constraints the optimists would continue to be the marginal buyers even if they were credit constrained. Of course, the pricing equation would be different, possibly reflecting the collateral value of houses — if these assets are pledged as collateral (see Fostel and Geanakoplos, 2008). However, our main intuition that the equilibrium price is higher the larger the difference in expectations would continue to be correct.

5.2 Learning from the equilibrium price

We now relax the assumption that agents disregard the equilibrium price to infer the unknown state of the economy. This extension is desirable because house prices, like any other financial prices, summarize most of the dispersed information in the economy. In extending our analysis to a setup where households learn from the equilibrium price we run, however, into a non-trivial problem. As discussed in the previous section, if households receive symmetrically dispersed signals and have the option to consume housing services by either buying or renting, the housing market is segmented, and the equilibrium price depends on the dispersion of information, i.e., \(i_t = |\varepsilon_t^e - \varepsilon_t^f|\). But, since \(i_t\) is not normally distributed, \(p_t\) has a non-Gaussian distribution, and standard linear filtering methods cannot be used.\(^8\)

To circumvent this problem we make the assumption that \(a_t\) — the aggregate preference shock — is an independent and identically distributed random variable, drawn from a distribution \(A\), with zero mean and variance \(\sigma_a^2\). Moreover, \(A\) is such that \(a_t + i_t \equiv \delta_t \sim N(\bar{\delta}, \sigma_\delta^2)\) where \(\bar{\delta}\) denotes the unconditional mean of \(i_t\) and \(\sigma_\delta^2\) the variance of \(a_t + i_t\).\(^9\)

Although ad-hoc, this assumption enables us to use standard methods to characterize the filtering problem since it ensures that the equilibrium price is Gaussian. In addition, as in a typical noisy rational expectation model à la Grossman and Stiglitz (1976) and Hellwig (1980), this assumption guarantees that the equilibrium price is not fully revealing. Specifically, households cannot tell whether prices are high because aggregate economic conditions improve or because unobservable taste shocks drive housing demand.

Using a linear solution method, Appendix IV proves that

**Proposition 3** The equilibrium house price with heterogeneous expectations and learning is

\[p_t = p_t^* + \pi_2 Y_t + \pi_3 \Phi_t,\]

\(^8\)See Appendix IV for a derivation of the exact distribution of \(i_t\).

\(^9\)Hellwig, Mukherji and Tsyvinski (2006), follow the same strategy to solve a noisy rational expectation model with non-Gaussian disturbances.
where $\pi_2 > 0$ and $\pi_3 > 0$ are the weights on the private and the endogenous public signal (the price), respectively, and

$$\Phi_t \equiv \phi \frac{\eta_t}{1 + r} + \frac{r\phi}{(1 + r)(r + \phi\pi_2)} a_t + \frac{\phi^2\pi_2}{(1 + r)(r + \phi\pi_2)} i_t$$

summarizes the degree of magnification of shocks induced by the process of learning from the price.

Intuitively, in the presence of unobservable shocks, households who observe a change in house prices do not understand whether this change is driven by changes in aggregate income ($\eta_t$), preferences ($a_t$), or private signals ($i_t$). Thus, with $\pi_3 > 0$, each of these shocks will have an amplified effect on equilibrium prices, since households respond to whatever is the source of movement in the house prices.

A key observation to make in comparing equations (26) and (23) is that $i_t$ — our measure of information dispersion — continues to shift the equilibrium price above its fundamental value, $p_t^*$. More specifically, $i_t$ exerts a direct effect on $p_t$, via $\Upsilon_t$, for the reasons discussed in the previous section, and an indirect effect, via $\Phi_t$, because of the magnification of shocks induced by the process of learning.

The relative importance of $\Upsilon_t$ and $\Phi_t$ depends, however, on $\pi_2$ and $\pi_3$. As shown in Appendix IV, when the variance of the preference shock is finite, then $\pi_2 < \lambda$ and $\pi_3 > 0$. However, as the noise in the preference for housing services increases, $\sigma_a^2 \to \infty$, the equilibrium price (26) becomes non-informative. In this case $\pi_2 \to \lambda$ and $\pi_3 \to 0$, and the equilibrium price is identical to the one prevailing in the absence of learning (23).

6 Empirical evidence

In this section we present some empirical evidence supporting our model’s main predictions: (1) the deviation of house prices from their fundamental value increases with the dispersion of information; and (2) the volatility of house prices is higher the larger the volatility of information dispersion.

6.1 The proxy of information dispersion

The obvious challenge in testing our model is to measure information dispersion. There is no data available and there is no natural candidate for a proxy. For this reason our strategy is to

\[10\] With learning, the volatility of the equilibrium house price remains higher than in the benchmark scenario of imperfect but homogeneous information. By comparing (26) with (19), it is straightforward to see that (25) still holds.
construct a proxy of information dispersion following the logic of the model. Our model can be interpreted as describing the house prices dynamics in a typical city where the speculative demand for housing depends on expectations about local economic conditions. If one assumes that city residents are employed in different industries and they are imperfectly informed about city’s income, it is then natural to think that industry-specific income shocks may convey useful information to estimate the average city income — as in the signal extraction problem discussed in the previous sections.

With this interpretation of the model, equations (1) and (2) can be rewritten as follows,

\[ w_{k,t}^j = \theta_{k,t} + \varepsilon_{k,t}^j \quad \text{and} \quad \theta_{k,t} = \rho \theta_{k,t-1} + \eta_{k,t} \tag{27} \]

where \( w_{k,t}^j \) is the time-\( t \) earning of residents in city \( k \) employed in industry \( j \), \( \theta_{k,t} \) the time-\( t \) average income in city \( k \), and \( \varepsilon_{k,t}^j \) the time-\( t \) industry-\( j \) shock in city \( k \). A proxy of information dispersion about \( \theta_{k,t} \) can then be computed using the dispersion of city-industry-earnings shocks \( \varepsilon_{k,t}^j \). For this purpose, we consider a large sample of U.S. Metropolitan Areas (MSA) and infer the time series properties of local income shocks based on annual earnings data for 10 one-digit industries.

With this data, we compute the dispersion of city earnings shocks in three steps. First, we use variation in national earnings by industry, and variation in the industry mix by cities, to compute exogenous changes in local income. Specifically, for each MSA and year, the change in income, \( \Delta \theta_{k,t} \), is computed as a weighted average (over the 10 one-digit industries) of the growth rate of national industry earnings, \( w_t^j \), with weights \( \omega_{k,t}^j \) given by the fraction of MSA people employed in each industry:

\[ \Delta \theta_{k,t} = \sum_{j=1}^{10} \omega_{k,t}^j \Delta w_t^j. \tag{28} \]

This variable measures the predicted change in city income \( \theta_{k,t} \), had each sector in city \( k \) grown at the national growth rate.\(^{11}\) This approach of imputing exogenous income shocks for local economies follows the literature on local business shocks and cycles (see e.g., Neumann and Topel, 1991, Bartik, 1991, Blanchard and Katz, 1992, Davis, Loungani and Mahidhara, 1997, among others) and rests on the plausible assumption that national industry earnings growth is uncorrelated with local labor supply shocks.\(^{12}\)

\(^{11}\)A regression of per capita income changes in city \( k \) on the predicted income changes based on (28), with MSA and year fixed effects, yields a coefficient of 1.47 (s.e. 0.173) and an overall R-squared of 0.342. Thus, the predicted MSA income predicts well actual MSA income.

\(^{12}\)In a recent paper, Guerrieri, Hartley and Hurst (2010) use also the same methodology in their study of the effects of neighborhood income shocks on the price of housing in a sample of 20 U.S. cities.
In a second step, we run ten regressions, one for each industry, in which we pool the growth rate of industry earnings, $\Delta w_{j,k,t}$, for the full sample of MSAs:

$$\Delta w_{j,k,t} = \alpha_{0j} + \alpha_{1j}\Delta w_{j,k,t-1} + \alpha_{2j}\Delta \theta_{k,t} + \alpha_{3j}\Delta \theta_{k,t-1} + \gamma_t + \varepsilon_{j,k,t}^j \quad \text{for} \quad j = 1, 2, \ldots, 10, \quad (29)$$

where $\gamma_t$ is a time fixed effect. This specification is based on equation (27) in the model, and adds lags of $w_{j,k,t}$ and $\theta_{k,t}$ to account for a minimum of industry and city income dynamics. As a result, the residuals $\varepsilon_{j,k,t}^j$ record shocks to industry-$j$’s earnings in city-$k$, controlling for nationwide effects, $\gamma_t$, industry-MSA specific earning dynamics, $\Delta w_{j,k,t-1}$, and MSA exogenous income dynamics, $\Delta \theta_{k,t}$ and $\Delta \theta_{k,t-1}$.

In a third and final step, we measure the dispersion of earnings shocks across the $j$ industries within each MSA as the weighted average of the absolute value of industry-MSA shocks,

$$i_{k,t} = \sum_{j=1}^{10} \omega_{k,j}^j |\varepsilon_{j,k,t}^j|, \quad (30)$$

where the weights $\omega_{k,j}^j$ denote the share of MSA workers employed in industry $j$, to control for the size of each industry. Accordingly, this variable captures the dispersion of local income shocks that are orthogonal to changes in local income, via (28), and to changes in aggregate income, via the period fixed effects, $\gamma_t$.

### 6.2 Data description and summary statistics

We use MSA and national industry data from the BEA, and construct our proxy of information dispersion with annual earnings data for the following industries: (1) Farm, (2) Mining, (3) Construction, (4) Manufacturing, (5) Transportation and public utilities, (6) Wholesale trade, (7) Retail trade, (8) Finance, insurance, and real estate, (9) Services, and (10) Government and government enterprises. We collect this data from 1980, the first year in which the FHFA house price index is available, until 2000, the year in which the Standard Industrial Classification (SIC) system had been replaced by the North American Industry Classification System (NAICS). Unfortunately, the different system for classifying economic activity makes it impossible to extend our data beyond 2000. Since available data based on the NAICS system covers only the period 2001 to 2008, we use the SIC classification codes

---

13 We have also experimented with specifications that does not include lags of $\Delta w_{j,k,t}$ All the results reported below are robust to such changes.

14 None of the results presented below change if we use squared deviations rather than absolute deviations. We prefer to use absolute deviations to keep the same units as the change in industry earnings, so that the coefficients in the house price regressions reported below can easily be interpreted.
to exploit the longer time series dimension of the data.

MSA level house price indices come from the Federal Housing Finance Agency (the formerly OFHEO indices). These are repeat sale indices for single-family, detached properties bought using conventional conforming loans.\footnote{A prominent alternative is to use the Case-Shiller-Weiss index, which also measures changes in housing market prices given a constant level of quality. The advantage of the Case-Shiller-Weiss index is that it is not limited to properties purchased with conventional mortgages. The disadvantage is that it has a limited geographical coverage, 20 MSA as opposed to 340 for the FHFA indices. For this reason our MSA analysis uses only the FHFA index.} Local economic and demographic conditions are proxied by MSA income per capita and MSA population, both obtained from the BEA. These variables will be used in our regressions to hold constant conventional determinants of housing demand. In addition, our regressions will also control for observable MSA heterogeneity in the supply of housing, with the index of supply elasticity compiled by Saiz (2010). The noteworthy feature of this index is that it does not depend on local market conditions but only on geographical and topographical constraints on house construction. All nominal variables in our data are converted into real dollars using the national CPI index from the Bureau of Labor Statistics.

Table 1 lists the variables contained in our dataset, along with their definitions and data sources. Table 2 reports some summary statistics. Over the full period 1980-2000, Dispersion — our proxy of information dispersion — has a mean value of 2.5\% and a standard deviation of 1.2\%. Most of its variation is within MSAs, but there is also a considerable variation across MSAs. It is less than 1.2\% in Atlanta, Dallas, Minneapolis, New Orleans, and greater than 4\% in Boston, Miami, New York, San Diego, to mention a few MSAs. Over the same period, real house prices increased at an average annual rate of 0.4\%, about one-third of the average MSA real per capita income and population growth. As for our proxy of information dispersion, the observed variation in house prices comes mostly from time variation. The same is true for per capita income growth. Finally, the predicted MSA personal income based on national industry earnings has a mean of 6.6\%, very similar to the average MSA personal income.

### 6.3 House price changes and information dispersion

To evaluate the empirical prediction that information dispersion leads to higher house prices, we estimate regressions of the following form:

$$\Delta p_{k,t} = \gamma_t + \gamma_k + X_{k,t} \beta + \delta_1 i_{k,t} + \delta_2 (i_{k,t} \times \eta_k^S) + \epsilon_{k,t},$$

(31)
where $\Delta p_{k,t}$ is the log change of the real house price index in MSA $k$ in year $t$, $\gamma_t$ is a year effect common to all markets, $\gamma_k$, is a time-invariant MSA effect, and $X_{k,t}$ is a vector of observable factors that are likely to influence local house prices. This vector includes current and past changes in income per capita, population and house prices. Time and MSA fixed effects are included to hold constant aggregate and local unobservable determinants of house prices.\footnote{Regressions are performed on first-differenced variables to put non-stationarity concerns to rest, and to follow the standard approach in the literature. Himmelberg, Mayer and Sinai (2005), for example, suggest using log differences in the FHFA house price index because this index is not standardized to the same representative house across markets. Thus, price levels cannot be compared across MSAs, but they can be used to calculate growth rates.}

The parameters of interest are $\delta_1$ and $\delta_2$. The first parameter traces the direct effect on real estate prices of a change in $i_{k,t}$, our proxy of information dispersion. In light of our theoretical model, we expect a positive estimate of $\delta_1$. The second parameter measures the differential impact of $i_{k,t}$ across MSAs, depending on the elasticity of local housing supply. Because our model’s prediction rests on the assumption that the stock of housing is fixed, we want to hold constant the supply of houses. We do so using the Saiz (2010) index of housing supply elasticity, denoted $\eta^S_k$. We expect $\delta_2 < 0$, that is house prices should respond less to an increase in information dispersion in MSAs with less supply restrictions.

Table 3 presents the OLS estimates of (31) with standard errors clustered at the MSA level to allow for within-MSA autocorrelation in the errors. Column 1 reports the results with current and lagged changes in MSA income per capita as the only controls. These two controls are suggested by the price equation (23) derived in Section 5. As shown, the prediction that information dispersion is associated with higher house prices is strongly supported by the data. The estimated effect is not only statistically significant but also sizeable: a 1% increase in $i_{k,t}$ results in a 0.25% increase in the growth rate of house prices. This means that an exogenous increase in $i_{k,t}$, from the 10th percentile value (which is approximately 1.2%) to the 90th percentile value (which is approximately 4%), implies a 0.7% annual acceleration in the growth rate of house price, which is large considering that the average annual growth rate of real house prices during the 1980-2000 period is 0.4%.

The estimates in column 2 show that $\delta_1$ is significant not only unconditionally, but also when we control for the elasticity of housing supply. This result assures us that movements in $i_{k,t}$ engender changes in housing demand, which have more pronounced effects on house prices the more inelastic the supply of housing. The estimates in column 2 indicate that a 1% increase in $i_{k,t}$ is associated with a 0.6% increase in the growth rate of house prices in “highly inelastic” MSAs, i.e., those that fall in the bottom 10% of the distribution of the Saiz index.
The results obtained so far, although based on the price equation derived in our model, do not control for other important determinants of house price dynamics. Thus, in column 3 and 4 we add three lags of the dependent variable and control also for changes in MSA population. We include lagged changes in house prices because it is well known that house prices exhibit momentum and mean reversion over time (Case and Shiller, 1989). Population growth is included to control for the possibility that the demand for housing is also affected by demographic factors. Despite the larger set of controls, our core findings are unaffected: our proxy of information dispersion significantly explains changes in house prices, and the estimated effect is stronger in MSA with a topography that makes new house construction difficult.

Table 4 explores the robustness of our findings to an alternative empirical specification suggested by the work of Lamont and Stein (1999). In their study of house price dynamics in U.S. cities, Lamont and Stein find that house prices (a) exhibit short-run movements, (b) respond to contemporaneous income shocks, and (c) display a long-run tendency to fundamental reversion. Accordingly, in the vector of controls, \( X_{k,t} \), we include the lagged change in house prices, current change in per capita income, and the lagged ratio of house prices to per-capita income. As shown in columns 1 and 2, these variables have the expected signs and our proxy of information dispersion continues to be related significantly to house price changes: the growth rate of house prices is higher in cities where local income shocks are more dispersed, and the effect is muted in MSA with high supply elasticity. These results are confirmed in column 3 and 4, where population growth is included as additional control.

### 6.4 House price and information dispersion volatility

We now turn to the second prediction of the model that the volatility of house prices increases with the variance in the dispersion of information. To examine the strength of this prediction, we compute the volatility of house prices by running a pooled regression for the change in house prices, controlling for year effects, and then by taking the standard deviation of the residuals in each MSA. We follow the same procedure to compute the volatility of our proxy of information dispersion. This gives us a measure of the volatility of house prices and information dispersion within a metropolitan area, controlling for aggregate effects. Next, with one observation for MSA, we exploit the cross-sectional variation of house price volatility and regress our measure of house price volatility on the volatility of information dispersion in each MSA.

The OLS estimates are in Table 5 and illustrated in Figure 2, which graphs the volatility of house price against the fitted values from the regression. As can be seen, MSAs with
large dispersion of information also have more volatile house prices. Interestingly, this result continues to hold even if we control for the standard deviation of aggregate MSA income, as shown in the second column of Table 5.

7 Conclusion

We have used a user-cost model of the housing market to study how information dispersion about local economic conditions affects the equilibrium price of housing. The equilibrium housing price is higher the larger the difference in expectations about future house prices. The reason is that all agents face a short-sale constraint in housing and derive utility from consuming housing services. Therefore, those who hold pessimistic expectations about future prices decide to rent to avoid capital losses, while those who have optimistic expectations decide to buy in anticipation of future price increases. The result is that the equilibrium price of owner-occupied houses reflects only the expectations of optimists and is, thus, higher and more volatile relative to an environment of homogeneous information.

We provide empirical evidence supporting the model’s predictions in a panel of U.S. cities, using the dispersion in industry income shocks as a proxy for the dispersion in information about local economic conditions. This proxy is motivated by our model’s assumption that different realizations of individual income lead agents to form different views of the economy.

To keep our model simple we have abstracted from a number of issues. For example, we have abstracted from the general equilibrium effects of the interest rate. Changes in \( R \), however, may affect our analysis since the return on the safe asset influences agents’ choice of renting and owning, for a given level of house price expectations. We have also prevented agents from re-trading. An extension of the model that allows for re-trading, as in Stein (1995) or Ortalo-Magné and Rady (2006), may shed new light on whether information dispersion induces a positive correlation between house prices and housing transactions. These extensions are left for future research.

References


Appendix I: Linearization of equation (6), (7) and (8).

We linearize equations (6) and (7) around the equilibrium with “certainty,” i.e., when $\varepsilon^j_t = 0$, $\eta_t = 0$, $a_t = 0$ and $\nu^j_t = 0 \ \forall t$. Denoting with $X$ any variable $X_t$ in the “certainty” equilibrium, the first-order conditions (6) and (7) can be written as

$$ V^j = V > 0 \implies V = \frac{C}{RQ}, \quad (32) $$

$$ Q = U. \quad (33) $$

Moreover, using equations (4), (8) and (33), we have

$$ \frac{C}{R} = W - VQ. \quad (34) $$

Thus, combining (34) and (32) one obtains

$$ V = \frac{W}{2Q}. $$

Under the assumption of fixed housing supply, $S$, the market clearing condition is

$$ V = S, $$

which implies that the following relationships must hold in a certainty equilibrium:

$$ U = Q, \quad Q = \frac{W}{2S}, \quad C = \frac{RW}{2}. $$
Linearization of (7) and (8)

Denoting with lower-case letters variables in percent deviation from the equilibrium with certainty, and recalling our definition of user cost, (8), a linearization of (7) around the certainty equilibrium yields,

\[ E^j_i \left[ \frac{RP}{C} (1 + p_t - c_{t+1}^j) - \frac{RQ}{C} (1 + q_t - c_{t+1}^j) - \frac{P}{C} (1 + p_{t+1} - c_{t+1}^j) \right] \geq 0. \]

Rearranging,

\[ E^j_i \left[ \frac{RP}{C} p_t - \frac{RQ}{C} q_t - \frac{P}{C} c_{t+1}^j - c_{t+1}^j \right] \geq 0 \Rightarrow \]

\[ E^j_i [RP_p - RQ_q - P_{pt+1}] \geq 0 \Rightarrow \]

\[ E^j_i \left[ p_t - \frac{Q}{P} q_t - \frac{1}{R} P_{pt+1} \right] \geq 0, \]

we obtain

\[ p_t \geq \frac{r}{1 + r} q_t + \frac{1}{1 + r} E^j_i p_{t+1}, \quad (35) \]

where

\[ r = R - 1. \]

Notice, also, that a linearization (8) gives

\[ u_t = \frac{P}{U} p_t - \frac{P}{RU} p_{t+1} \]

\[ = \left( \frac{1 + r}{r} \right) p_t - \frac{1}{r} p_{t+1}. \]

Therefore, (35) can be rewritten as (10).

Moreover, using (11), equation (10) can be written as:

\[ \frac{(1 + r)p_t - E^j_i p_{t+1}}{r} \geq q_t. \quad (36) \]

Since \( E^j_i p_{t+1} > E^j_0 p_{t+1} \), equation (36) holds with strict inequality for \( j = 0 \) and so pessimists choose to own no housing units, \( H^0_t = 0. \)

Linearization of (6)

A linearization of equation (6), around the certainty equilibrium, gives

\[ E^j_i \frac{RQ}{C} (q_t - c_{t+1}^j) = \frac{A}{V} (2a^j_i - v^j_i), \]

\[ E^j_i \frac{1}{S} (q_t - c_{t+1}^j) = \frac{1}{V} (2a^j_i - v^j_i), \]

which defines the optimal demand of housing services

\[ v^j_i = 2a^j_i - q_t + E^j_i c_{t+1}^j. \quad (37) \]
The term $E^j_t c^j_{t+1}$ in (37) is obtained by linearizing the flow of budget constraint (4), that for the two groups of agents reads as follows:

\[
\begin{align*}
C^1_{t+1} &= R(W^1_t - P_t H^1_t + Q_t (H^1_t - V^1_t)) + P_{t+1} H^1_t, \\
C^0_{t+1} &= R(W^0_t - Q_t V^0_t),
\end{align*}
\]

where the second equation uses the fact that $H^0_t = 0$. A bit of algebra establishes\(^\text{17}\)

\[
E^1_t c^1_{t+1} = 2w^1_t - v^1_t - \left(\frac{r + 1}{r}\right) p_t + \frac{1}{r} E^1_t p_{t+1} = 2w^1_t - v^1_t - E^1_t u_t,
\]

\[
E^0_t c^0_{t+1} = 2w^0_t - v^0_t - q_t.
\]

Plugging these expressions into (37) and using equation (10) for $j = 1$, it follows that

\[
v^1_t = w^1_t + a^1_t - \frac{1}{2} (q_t + E^1_t u_t) = w^1_t + a^1_t - q_t,
\]

\[
v^0_t = w^0_t + a^0_t - q_t.
\]

These establish equation (9).

\section*{Appendix II: Proof for Proposition 1}

When information is imperfect but homogeneous, $E^j_t p_{t+1} = \hat{E}^j_t p_{t+1}$ and $\tilde{E}^j_t p_{t+1} = 0$. Therefore, equation (16), shifted one period forward, gives

\[
p_{t+1} = \frac{r}{1 + r} (\theta_{t+1} + a_{t+1}) + \frac{1}{1 + r} E^0_{t+1} p_{t+2}.
\]

Taking expectations on both sides conditional on time $t$ information, and excluding explosive price paths, a forward iteration of the expression above gives

\[
E_t p_{t+1} = \frac{r}{1 + r} \sum_{\tau=0}^{\infty} \left(\frac{1}{1 + r}\right)^\tau E_t (\theta_{t+1+\tau} + a_{t+1+\tau}),
\]

Since $\theta_t$ and $a_t$ are unobservable at time $t$ and

\[
\theta_t = \rho \theta_{t-1} + \eta_t, \quad \text{with} \quad \rho \in (0, 1],
\]

\(^\text{17}\)Linearizing (38) yields

\[
E^1_t c^1_{t+1} = \frac{R}{C} w^1_t - \frac{RP}{C} (p_t + h^1_t) + \frac{RQH}{C} (q_t + h^1_t) - \frac{RQV}{C} (q_t + v^1_t) + \frac{PH}{C} (E^1_t p_{t+1} + h^1_t)
\]

\[
= 2w^1_t - \frac{P}{U} (p_t + h^1_t) + (q_t + h^1_t) - (q_t + v^1_t) + \frac{P}{RU} (E^1_t p_{t+1} + h^1_t)
\]

Rearranging this equation gives (40). Proceeding in a similar way, one obtains (41).
we have
\[ E_t[\theta_{t+1} + a_{t+1}] = \rho^2 \theta_{t-1}. \]

It is, therefore, immediate to obtain
\[ \bar{E}_{t+1}p_t = \bar{E}_t f_t = \phi \rho \theta_{t-1}, \tag{42} \]
where \( \phi \equiv \frac{r}{1 + r - \rho} \). Plugging (42) back into (16) and recalling that \( \bar{E}_{t+1}p_t = 0 \), the equilibrium price under common information can then be written as
\[ p^*_t = (\theta_t + a_t) + \frac{1}{1 + r} ((\phi \rho \theta_{t-1} - \theta_t) - a_t). \]

Appendix III: Proof for Proposition 2

In the presence of heterogeneous expectations, \( E_t^j p_t \neq \bar{E}_t p_t \) and \( \bar{E}_t p_t \neq 0 \). Shifting equation (16) one period forward
\[ p_{t+1} = \frac{r}{1 + r} (\theta_{t+1} + a_{t+1}) + \frac{1}{1 + r} \bar{E}_{t+1}p_{t+2} + \frac{1}{1 + r} \bar{E}_{t+1} + I, \]
denoting,
\[ i_t = |\varepsilon_i - \varepsilon_j| \quad \text{for} \quad i \neq j, \]
and guessing that \( \bar{E}_t [p_{t+1}] = \phi \lambda i_t \), we have
\[
\begin{align*}
E_t^j p_{t+1} &= \frac{r}{1 + r} E_t^j (\theta_{t+1} + a_{t+1}) + \frac{1}{1 + r} E_t^j \bar{E}_{t+1}p_{t+2} + \frac{\phi \lambda}{1 + r} I, \\
\bar{E}_t p_{t+1} &= \frac{r}{1 + r} \bar{E}_t (\theta_{t+1} + a_{t+1}) + \frac{1}{1 + r} \bar{E}_t \bar{E}_{t+1}p_{t+2} + \frac{\phi \lambda}{1 + r} I, \\
\bar{E}_t \bar{E}_t p_{t+1} &= \frac{r}{1 + r} \bar{E}_t \theta_{t+1} + \frac{1}{1 + r} \bar{E}_t \bar{E}_t \bar{E}_{t+1}p_{t+2},
\end{align*}
\]
where the last equality holds because agents hold heterogeneous expectations with respect to \( \theta_{t+1} \) but not with respect to \( a_{t+1} \). In the expressions above,
\[ I \equiv \int_0^\infty x d\Gamma(x), \]
is the average degree of information heterogeneity where \( \Gamma \) is the density of \( i_t \).

Iterating these expressions forward and excluding explosive price paths, we obtain:
\[
\begin{align*}
E_t^j p_{t+1} &= \frac{r}{1 + r - \rho} E_t^j \theta_{t+1} + \frac{\phi \lambda}{r} I, \\
\bar{E}_t p_{t+1} &= \frac{r}{1 + r - \rho} \bar{E}_t \theta_{t+1} + \frac{\phi \lambda}{r} I, \\
\bar{E}_t \bar{E}_t p_{t+1} &= \frac{r}{1 + r - \rho} \bar{E}_t \theta_{t+1}.
\end{align*}
\]
Moreover, using equation (20), it is easy to see that:

\[ E_j^t \theta_{t+1} = \rho E_j^t \theta_t = \rho \left[ (1 - \lambda) \rho \theta_{t-1} + \lambda w_j^t \right], \]

and, thus,

\[ \tilde{E}_t p_{t+1} = \phi \left( \rho (1 - \lambda) \theta_{t-1} + \lambda \theta_t \right) + \frac{\phi \lambda}{r} I, \]
\[ = (\phi \rho \theta_{t-1}) + \phi \lambda (\theta_t - \rho \theta_{t-1}) + \frac{\phi \lambda}{r} I, \]
\[ \tilde{E}_t p_{t+1} = \phi \lambda i_t. \]

so that \( \tilde{E}_t p_{t+1} = \phi \lambda i_t \) as claimed. Plugging \( \tilde{E}_t p_{t+1} \) into (16), the equilibrium house prices can be written as

\[ p_t = (\theta_t + \alpha_t) + \frac{1}{1 + r} \left( \left( \phi \rho \theta_{t-1} - \theta_t \right) - \alpha_t \right) + \frac{\phi \lambda}{1 + r} (\theta_t - \rho \theta_{t-1}) + \frac{\phi \lambda}{r(1 + r)} I + \frac{\phi \lambda}{1 + r} i_t. \]

where

\[ Y_t = \frac{\phi (\theta_t - \rho \theta_{t-1})}{1 + r} + \frac{\phi I}{r(1 + r)} + \frac{\phi i_t}{1 + r}. \]

**Appendix IV: Learning from the equilibrium rental price**

In this appendix, we provide a solution to the signal extraction problem when agents condition on the house price to learn the unknown fundamental, \( \theta_t \). As explained in Section 5.1, the inference problem is involved since the equilibrium price in the presence of heterogeneous information is not normally distributed. To characterize this non-standard signal extraction problem, we assume that the distribution of the preference shock \( \mu_t \) is such that sum of \( i_t \) and \( \mu_t \) follows a normal distribution. This assumption enables us to recover a Gaussian distribution for the equilibrium price and allows us to apply standard linear filtering techniques.

We proceed in three steps. First, we define the exact distribution for \( i_t \). Next, we determine the form of the distribution of \( \mu_t \) that makes the equilibrium price normally distributed. Finally, using a method of undetermined coefficients, we characterize the inference problem for \( \theta_t \) and the resulting equilibrium price.

**The distribution of** \( i = |\varepsilon^i - \varepsilon^j| \) **for** \( i \neq j \)

Consider two independent random variables, \( \varepsilon^i \) and \( \varepsilon^j \), distributed normally with zero mean and equal variance \( \sigma^2 \). Define,

\[ \tilde{\varepsilon} = \varepsilon^j - \varepsilon^i \sim N(0, 2\sigma^2). \]
The cumulative distribution function of \( i = |\tilde{z}| \) is

\[
F_i(y) = \Pr (i = |\tilde{z}| \leq y) = 2 \int_0^y \frac{1}{\sqrt{2\pi} \sqrt{2\sigma_\varepsilon}} \exp \left(-\frac{1}{2} \frac{z^2}{2\sigma_\varepsilon^2}\right) dz,
\]
and the associated density,

\[
f_i(y) = \begin{cases} 
\frac{\partial F_i(y)}{\partial y} = \frac{2}{\sqrt{2\pi} \sqrt{2\sigma_\varepsilon}} \exp \left(-\frac{1}{2} \frac{y^2}{2\sigma_\varepsilon^2}\right) & \text{if } y \geq 0 \\
0 & \text{otherwise}
\end{cases}.
\]  
(43)

Denote with \( \bar{i} \), the mean of \( i \),

\[
\bar{i} = \int_0^{\infty} y f_i(y) dy.
\]

The distribution of the aggregate preference shock, \( a \).

We wish to find the distribution of a random variable, \( a \), with zero mean and variance \( \sigma_a^2 \), such that

\[ a + i \sim N(\bar{i}, \sigma_a^2 + \sigma_i^2). \]

The cumulative function of \( a + i \) is

\[
F_{a+i}(y) = \Pr (a + i \leq y) = \int_{-\infty}^{\infty} \left( \int_{-\infty}^{y-a} f_i(i) \, di \right) f_a(a) \, da,
\]
where \( f_a \) is the density of \( a \) and \( f_i \) is defined in (43). Differentiating \( F_{a+i}(y) \) with respect to \( y \) yields the probability density of \( a + i \),

\[
f_{a+i}(y) = \int_{-\infty}^{\infty} f_i(y-a) f_a(a) \, da.
\]

Since, by assumption, \( a + i \) follows a normal distribution, it must be

\[
f_{a+i}(y) = \frac{1}{\sqrt{2\pi} \sqrt{\sigma_a^2 + \sigma_i^2}} \exp \left(-\frac{1}{2} \frac{(y - \bar{i})^2}{\sigma_a^2 + \sigma_i^2}\right).
\]

Therefore, the density \( f_a(a) \) is recovered by solving the following integral:

\[
\int_{-\infty}^{\infty} f_i(y-a) f_a(a) \, da = \frac{1}{\sqrt{2\pi} \sqrt{\sigma_a^2 + \sigma_i^2}} \exp \left(-\frac{1}{2} \frac{(y - \bar{i})^2}{\sigma_a^2 + \sigma_i^2}\right).
\]

Lemma 4

Lemma 4 The correlation coefficient between \( \varepsilon^j \) and \( i \equiv |\varepsilon^j - \varepsilon^i| \) is zero.
Proof.

\[
\begin{align*}
\text{Cov} (\varepsilon^j, |\varepsilon^j - \varepsilon^i|) &= \text{Cov} (\varepsilon^j, \varepsilon^j - \varepsilon^i) \Pr (\varepsilon^j > \varepsilon^i) + \text{Cov} (\varepsilon^j, -(\varepsilon^j - \varepsilon^i)) \Pr (\varepsilon^j < \varepsilon^i) \\
&= \text{Cov} (\varepsilon^j, \tilde{\varepsilon}) \Pr (\varepsilon^j > \varepsilon^i) - \text{Cov} (\varepsilon^j, \tilde{\varepsilon}) \Pr (\varepsilon^j < \varepsilon^i) \\
&= \text{Cov} (\varepsilon^j, \tilde{\varepsilon}) [\Pr (\varepsilon^j > \varepsilon^i) - \Pr (\varepsilon^j < \varepsilon^i)] = 0
\end{align*}
\]

The last equation holds because \( \varepsilon^j \) and \( \varepsilon^i \) are independent and identically distributed normal random variable with zero mean and equal variance, so that \( \Pr (\varepsilon^j > \varepsilon^i) - \Pr (\varepsilon^j < \varepsilon^i) = 0 \). ■

The method of undetermined coefficients

Starting from equation (23), we guess that the equilibrium price is a linear function of the past observable fundamental \( \theta_{t-1} \), the current unobservable fundamental \( \theta_t \), preference shock \( a_t \), and the difference in households’ private signals \( i_t \); i.e.,

\[
p_t = b_0 + b_y \rho \theta_{t-1} + b_\eta \eta_t + b_a a_t + b_i i_t,
\]

where \( b_0, b_y, b_\eta, b_a \) and \( b_i \) are undetermined coefficients. It is convenient to rewrite equation (44) as

\[
p_t = b_\eta \eta_t + b_a a_t + b_i i_t + X_t,
\]

where

\[
X_t \equiv b_0 + b_y \rho \theta_{t-1}
\]

is non-stochastic. Defining

\[
\hat{p}_t \equiv \frac{p_t - X_t}{b_\eta},
\]

equation (45) can be written as

\[
\hat{p}_t = \eta_t + \delta_t,
\]

where,

\[
\delta_t = \frac{b_a}{b_\eta} a_t + \frac{b_i}{b_\eta} i_t.
\]

Under the assumption made on the distribution of \( a_t, \delta_t \) is normally distributed,

\[
\delta_t \sim \mathcal{N} \left( \frac{b_i}{b_\eta}, \frac{b_a^2}{b_\eta^2} \sigma_a^2 + \left( \frac{b_i}{b_\eta} \right)^2 \sigma_i^2 \right)
\]

and, as a consequence \( \hat{p}_t \), is also normally distributed,

\[
\hat{p}_t \sim \mathcal{N} \left( \frac{b_i}{b_\eta}, \sigma_\eta^2 + \frac{b_a^2 \sigma_a^2 + b_i^2 \sigma_i^2}{b_\eta^2} \right).
\]

The inference problem

Agent \( j \) estimates the unknown fundamental \( \theta_t \) by solving a standard filtering problem, based on the normally distributed (a) private signal, \( w^j_t \), (b) exogenous public signal, \( \theta_{t-1} \), and (c) endogenous
public signal, \( \tilde{p}_t \). Recalling that
\[
\theta_t = \rho \theta_{t-1} + \eta_t, \\
w^i_t = \theta_t + \varepsilon^i_t, \\
\tilde{p}_t = \eta_t + \delta_t,
\]
and using (47) and Lemma 4, the log-likelihood function can be written as
\[
L = -\frac{1}{2\sigma^2_\eta} \left( \rho \theta_{t-1} - E^i_t \theta_t \right)^2 - \frac{1}{2\sigma^2_\varepsilon} \left( w^i_t - E^i_t \theta_t \right)^2 - \frac{1}{2\sigma^2_\delta} \left( \tilde{p}_t - E^i_t \eta_t \right)^2.
\]
Thus, the optimal filtering solves the following first-order condition,
\[
-\frac{1}{\sigma^2_\eta} \left( -E^i_t \eta_t \right) + \frac{1}{\sigma^2_\varepsilon} \left( w^i_t - \rho \theta_{t-1} - E^i_t \eta_t \right) + \frac{1}{\sigma^2_\delta} \left( \tilde{p}_t - E^i_t \eta_t \right) = 0,
\]
or,
\[
E^i_t \eta_t = \frac{\sigma^2_\eta \sigma^2_\delta \left( w^i_t - \rho \theta_{t-1} \right) + \sigma^2_\eta \sigma^2_\varepsilon \tilde{p}_t}{\sigma^2_\varepsilon \sigma^2_\delta + \sigma^2_\eta \sigma^2_\delta + \sigma^2_\eta \sigma^2_\varepsilon}.
\]
The best linear estimate of \( \theta_t \) is, therefore,
\[
E^i_t \theta_t = (\pi_1 + \pi_3) \rho \theta_{t-1} + \pi_2 w^i_t + \pi_3 \tilde{p}_t,
\]
where
\[
\pi_1 = \frac{\sigma^2_\eta \sigma^2_\delta}{\sigma^2_\varepsilon \sigma^2_\delta + \sigma^2_\eta \sigma^2_\delta + \sigma^2_\eta \sigma^2_\varepsilon}, \\
\pi_2 = \frac{\sigma^2_\eta \sigma^2_\delta}{\sigma^2_\varepsilon \sigma^2_\delta + \sigma^2_\eta \sigma^2_\delta + \sigma^2_\eta \sigma^2_\varepsilon}, \\
\pi_3 = \frac{\sigma^2_\eta \sigma^2_\varepsilon}{\sigma^2_\varepsilon \sigma^2_\delta + \sigma^2_\eta \sigma^2_\delta + \sigma^2_\eta \sigma^2_\varepsilon}.
\]
Notice that if \( \sigma^2_\delta \to \infty \) (for example, because \( \sigma^2_\eta \to \infty \), i.e., the preference shock has a very large variance), then
\[
\pi_1 \to \frac{\sigma^2_\eta}{\sigma^2_\varepsilon + \sigma^2_\eta} = 1 - \lambda, \quad \pi_2 \to \frac{\sigma^2_\eta}{\sigma^2_\varepsilon + \sigma^2_\eta} = \lambda \quad \text{and} \quad \pi_3 \to 0.
\]
In other words, agents have nothing to learn from the equilibrium price, and the weights used for inferring the unobservable aggregate fundamental are the same as in Section 5.

**The equilibrium price**

To solve for the equilibrium price, we follow the same steps as in Appendix III. By guessing that
\( \tilde{E}_t p_{t+1} = \phi \pi_{2t} \), we have

\[
E^j_t p_{t+1} = \frac{r}{1 + r - \rho} E^j_{t+1} + \frac{\phi \pi_{2t}}{r}
\]
\[
E_t p_{t+1} = \frac{r}{1 + r - \rho} E_t p_{t+1} + \frac{\phi \pi_{2t}}{r}
\]
\[
\tilde{E}_t p_{t+1} = \frac{r}{1 + r - \rho} \tilde{E}_t p_{t+1}
\]

Moreover, using (48), the last two equations can be written as:

\[
E_t p_{t+1} = \phi (\rho \theta_{t-1} + \pi_{2t} \eta_t + \pi_{3t} \tilde{p}_t) + \frac{\phi \pi_{2t}}{r}
\]
\[
\tilde{E}_t p_{t+1} = \phi \pi_{2t}.
\]

The second line confirms the claim that \( \tilde{E}_t p_{t+1} = \phi \pi_{2t} \). Inserting \( E_t p_{t+1} \) and \( \tilde{E}_t p_{t+1} \) in (16) now, the equilibrium price becomes

\[
p_t = \frac{r}{1 + r} \left( \rho \theta_{t-1} + \eta_t + a_t \right) + \frac{1}{1 + r} \left( \phi \rho \theta_{t-1} + \phi \pi_{2t} \eta_t + \phi \pi_{3t} \tilde{p}_t + \frac{\phi \pi_{2t}}{r} \right) + \frac{\phi \pi_{2t}}{1 + r}
\]

from which it follows,

\[
p_t = \frac{\phi}{1 + r} \left( \frac{\pi_{2t}}{r} - \frac{\pi_{3t}}{b_0} \right) + \frac{\phi + \rho \phi + \phi \pi_{3t}}{1 + r} \left( \rho \theta_{t-1} + \frac{\phi \pi_{2t}}{r} \eta_t + \frac{\phi \pi_{3t}}{1 + r} a_t + \frac{\phi \pi_{2t}}{1 + r} \right)
\]

The undetermined coefficients can, therefore, be written as

\[
b_0 = \frac{\phi}{r(1 + r)} I
\]
\[
b_\theta = \frac{\phi}{r + \phi}
\]
\[
b_\eta = \frac{\phi}{r + \phi} \left( \frac{\pi_{2t}}{r} + \frac{\pi_{3t}}{r + \phi \pi_{2t}} \right)
\]
\[
b_a = \frac{\phi}{r + \phi} \left( \frac{\pi_{2t}}{r + \phi \pi_{2t}} \right)
\]
\[
b_i = \frac{\phi}{1 + r} \left( \frac{\pi_{2t}}{r + \phi \pi_{2t}} \right)
\]

and the equilibrium price as,

\[
p_t = \frac{\phi \pi_{2t}}{r(1 + r)} + \frac{\phi}{1 + r} \rho \theta_{t-1} + \frac{\phi (\pi_{2t} + \pi_{3t})}{1 + r} \eta_t
\]
\[
+ \frac{\phi \pi_{2t}}{1 + r} \left( 1 + \frac{\phi \pi_{3t}}{r + \phi \pi_{2t}} \right) a_t + \frac{\phi \pi_{2t}}{1 + r} \left( 1 + \frac{\phi \pi_{3t}}{r + \phi \pi_{2t}} \right) i_t.
\]

or, after some manipulation, as

\[
p_t = p_t^* + \pi_{2t} \Upsilon_t + \pi_{3t} \Phi_t.
\]

As in Section 4 and 5, \( p_t^* \) denotes the fundamental price, and \( \Upsilon_t \) measures the degree of dispersion.
in beliefs. The new term,
\[
\Phi_t = \frac{\phi}{1 + r} (\theta_t - \rho \theta_{t-1}) + \frac{r \phi}{(1 + r) (r + \phi \pi_2)} a_t + \frac{\phi^2 \pi_2}{(1 + r) (r + \phi \pi_2)} i_t,
\]
captures, instead, the degree of magnification of shocks induced by the learning process of learning from price.

Finally, since
\[
\sigma^2 = \left( \frac{b_a}{b_\eta} \right)^2 \sigma^2_a + \left( \frac{b_i}{b_\eta} \right)^2 \sigma^2_i,
\]
(52)
\[ \pi_1, \pi_2 \text{ and } \pi_3 \text{ are functions of } \sigma^2_\delta, \text{ which, in turn, depend on } b_\eta, b_a \text{ and } b_i. \] To pin down these undetermined coefficients, it is thus necessary to use equations (50), (51) and (52). This leads to
\[
b_\eta = \frac{r}{1 + r} + \frac{\phi}{1 + r} \left( \frac{\sigma^2_\eta \left( \frac{b_a^2 \sigma^2_a + b_i^2 \sigma^2_i}{b_\eta^2} \right) + \sigma^2_\eta \sigma^2_\varepsilon}{(\sigma^2_\varepsilon + \sigma^2_\eta) \left( \frac{b_a^2 \sigma^2_a + b_i^2 \sigma^2_i}{b_\eta^2} \right) + \sigma^2_\eta \sigma^2_\varepsilon} \right),
\]
\[
b_i = \frac{\phi \pi_2}{r} = \frac{\phi}{r} \left( \frac{\sigma^2_\eta \left( \frac{b_a^2 \sigma^2_a + b_i^2 \sigma^2_i}{b_\eta^2} \right)}{(\sigma^2_\varepsilon + \sigma^2_\eta) \left( \frac{b_a^2 \sigma^2_a + b_i^2 \sigma^2_i}{b_\eta^2} \right) + \sigma^2_\eta \sigma^2_\varepsilon} \right),
\]
and
\[
b_\eta = b_a + b_i,
\]
which define a system of three equations in the three unknowns, \( b_\eta, b_a \) and \( b_i \). Unfortunately, this system of equations does not admit closed-form solutions. However, numerical values can easily be computed.
Figure 1: Real U.S. House Price Index (1980 = 100)

Source: FHFA and BLS
Figure 2: Housing price and income dispersion volatility in 331 MSAs
Table 1. Description of variables and data sources

<table>
<thead>
<tr>
<th>Variable name</th>
<th>Variable description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dispersion</td>
<td>Proxy of information dispersion within MSA, using the dispersion of MSA earnings in 10 one-digit industries, as explained in Section 6</td>
<td>BEA</td>
</tr>
<tr>
<td>House price</td>
<td>MSA repeat-sales price index of existing single-family houses</td>
<td>FHFA</td>
</tr>
<tr>
<td>Income per capita</td>
<td>MSA income per capita</td>
<td>BEA</td>
</tr>
<tr>
<td>Population</td>
<td>MSA population (in thousands)</td>
<td>BEA</td>
</tr>
<tr>
<td>Personal income</td>
<td>MSA personal income</td>
<td>BEA</td>
</tr>
<tr>
<td>Predicted personal income</td>
<td>Predicted MSA income growth based on national industry earnings growth and the MSA industry mix, as explained in Section 6</td>
<td>BEA</td>
</tr>
<tr>
<td>Index of housing supply elasticity</td>
<td>Land-topology based measure of housing supply elasticity</td>
<td>Saiz (2010)</td>
</tr>
</tbody>
</table>

Table 2  Summary statistics

Summary statistics of MSA-year pooled data. Except for the index of the Saiz index of housing supply elasticity, summary statistics refer to the annual log change of each variable during the period 1980-2000.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>Between SD</th>
<th>Within SD</th>
<th>10th pc</th>
<th>90th pc</th>
<th>Number of MSAs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dispersion</td>
<td>0.0254</td>
<td>0.0126</td>
<td>0.0076</td>
<td>0.0102</td>
<td>0.0127</td>
<td>0.0413</td>
<td>341</td>
</tr>
<tr>
<td>House price</td>
<td>0.0040</td>
<td>0.0446</td>
<td>0.0124</td>
<td>0.0428</td>
<td>-0.0439</td>
<td>0.0486</td>
<td>380</td>
</tr>
<tr>
<td>Income per capita</td>
<td>0.0154</td>
<td>0.0255</td>
<td>0.0061</td>
<td>0.0248</td>
<td>-0.0152</td>
<td>0.0445</td>
<td>381</td>
</tr>
<tr>
<td>Population</td>
<td>0.0121</td>
<td>0.0148</td>
<td>0.0114</td>
<td>0.0094</td>
<td>-0.0032</td>
<td>0.0295</td>
<td>381</td>
</tr>
<tr>
<td>Personal income</td>
<td>0.0640</td>
<td>0.0311</td>
<td>0.0131</td>
<td>0.0283</td>
<td>0.0300</td>
<td>0.1023</td>
<td>363</td>
</tr>
<tr>
<td>Predicted personal income</td>
<td>0.0661</td>
<td>0.0165</td>
<td>0.0030</td>
<td>0.0162</td>
<td>0.0461</td>
<td>0.0882</td>
<td>363</td>
</tr>
<tr>
<td>Index of housing supply elasticity</td>
<td>2.5397</td>
<td>1.4403</td>
<td>1.4403</td>
<td>0.0000</td>
<td>1.0592</td>
<td>4.3916</td>
<td>263</td>
</tr>
</tbody>
</table>
Tab 3   House Price and dispersion of MSA earnings

MSA panel regressions of the log change in the real FHFA house price index on Dispersion -- our proxy of MSA information dispersion. Controls include: current and lagged log change in MSA’s Income per capita, lagged log change in House Prices, current and lagged log change in Population, and the Saiz (2010) index of supply elasticity. All variables are defined in Table 1. The sample period is 1980-2000. All regressions include MSA and year fixed effects. Standard errors are clustered at the MSA level. Estimates followed by ***, **, and * are statistically different from zero with 0.01, 0.05 and 0.10 significance levels, respectively.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Income per capita</td>
<td>0.504***</td>
<td>0.528***</td>
<td>0.359***</td>
<td>0.332***</td>
</tr>
<tr>
<td>(0.056)</td>
<td>(0.063)</td>
<td>(0.035)</td>
<td>(0.041)</td>
<td></td>
</tr>
<tr>
<td>Lagged income per capita</td>
<td>0.729***</td>
<td>0.740***</td>
<td>0.176***</td>
<td>0.181***</td>
</tr>
<tr>
<td>(0.055)</td>
<td>(0.071)</td>
<td>(0.040)</td>
<td>(0.044)</td>
<td></td>
</tr>
<tr>
<td>One lag house price</td>
<td>0.421***</td>
<td>0.419***</td>
<td>(0.023)</td>
<td>(0.026)</td>
</tr>
<tr>
<td>(0.023)</td>
<td>(0.031)</td>
<td>(0.030)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two lags house prices</td>
<td>0.274***</td>
<td>0.301***</td>
<td>(0.031)</td>
<td>(0.030)</td>
</tr>
<tr>
<td>(0.014)</td>
<td>(0.017)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Three lags house prices</td>
<td>-0.109***</td>
<td>-0.109***</td>
<td>(0.014)</td>
<td>(0.017)</td>
</tr>
<tr>
<td>(0.014)</td>
<td>(0.017)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population</td>
<td>1.511***</td>
<td>1.404***</td>
<td>(0.164)</td>
<td>(0.196)</td>
</tr>
<tr>
<td>(0.164)</td>
<td>(0.196)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lagged Population</td>
<td>-0.324**</td>
<td>-0.275</td>
<td>(0.138)</td>
<td>(0.172)</td>
</tr>
<tr>
<td>(0.138)</td>
<td>(0.172)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dispersion</td>
<td>0.254***</td>
<td>0.997***</td>
<td>0.159***</td>
<td>0.521***</td>
</tr>
<tr>
<td>(0.100)</td>
<td>(0.238)</td>
<td>(0.061)</td>
<td>(0.175)</td>
<td></td>
</tr>
<tr>
<td>Dispersion × housing supply elasticity</td>
<td>-0.375***</td>
<td>-0.174**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.107)</td>
<td>(0.074)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>3454</td>
<td>2601</td>
<td>2760</td>
<td>2106</td>
</tr>
<tr>
<td>N. of MSAs</td>
<td>294</td>
<td>226</td>
<td>231</td>
<td>218</td>
</tr>
<tr>
<td>R2</td>
<td>0.260</td>
<td>0.273</td>
<td>0.570</td>
<td>0.571</td>
</tr>
</tbody>
</table>
Tab 4  House price and dispersion of MSA earnings (Lamont & Stein’s specification)

MSA panel regressions of the log change in the real FHFA house price index on Dispersion -- our proxy of MSA information dispersion. Controls include: lagged log changes in House Prices, log change in MSA’s Income per capita, log change in Population, the lagged price to income ratio, and Saiz (2010) index of supply elasticity. All variables are defined in Table 1. The sample period is 1980-2000. All regressions include MSA and year fixed effects. Standard errors are clustered at the MSA level. Estimates followed by ***, **, and * are statistically different from zero with 0.01, 0.05 and 0.10 significance levels, respectively.

<table>
<thead>
<tr>
<th>Dependent Variables</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>One lag house price</td>
<td>0.577***</td>
<td>0.594***</td>
<td>0.460***</td>
<td>0.476***</td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
<td>(0.019)</td>
<td>(0.028)</td>
<td>(0.029)</td>
</tr>
<tr>
<td>Income per capita</td>
<td>0.282***</td>
<td>0.325***</td>
<td>0.344***</td>
<td>0.359***</td>
</tr>
<tr>
<td></td>
<td>(0.037)</td>
<td>(0.045)</td>
<td>(0.035)</td>
<td>(0.043)</td>
</tr>
<tr>
<td>Lagged Price/Income</td>
<td>-0.164***</td>
<td>-0.159***</td>
<td>-0.141***</td>
<td>-0.135***</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.008)</td>
<td>(0.007)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>Population</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.192***</td>
<td>1.194***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.064)</td>
<td>(0.123)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dispersion</td>
<td>0.161**</td>
<td>0.600***</td>
<td>0.191***</td>
<td>0.522***</td>
</tr>
<tr>
<td></td>
<td>(0.067)</td>
<td>(0.159)</td>
<td>(0.064)</td>
<td>(0.156)</td>
</tr>
<tr>
<td>Dispersion × housing supply elasticity</td>
<td>-0.201***</td>
<td>-0.151***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.063)</td>
<td>(0.055)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>3295</td>
<td>2504</td>
<td>3295</td>
<td>2504</td>
</tr>
<tr>
<td>N, of MSAs</td>
<td>314</td>
<td>224</td>
<td>314</td>
<td>224</td>
</tr>
<tr>
<td>R2 within</td>
<td>.554</td>
<td>.578</td>
<td>.596</td>
<td>.616</td>
</tr>
</tbody>
</table>
Table 5  House price volatility and the volatility of earnings dispersion

MSA cross-sectional regressions of the volatility of house price on the volatility of MSA dispersion of industry earnings and the volatility of MSA income per capita. The MSA volatility of house prices (industry earning dispersion, and income per capita) is the MSA standard deviation of the residuals of a pooled regression of the log change in MSA house prices (industry earning dispersion, and income per capita) on year fixed effects. The sample period is 1980-2000. Estimation is by OLS. Standard errors are robust to heteroskedasticity. Estimates followed by ***, **, and * are statistically different from zero with 0.01, 0.05 and 0.10 significance levels, respectively.

<table>
<thead>
<tr>
<th>Dependent Variables</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volatility of house price</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Volatility of dispersion</td>
<td>1.219***</td>
<td>0.974***</td>
</tr>
<tr>
<td></td>
<td>(0.256)</td>
<td>(0.319)</td>
</tr>
<tr>
<td>Volatility of income</td>
<td>0.210</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.154)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>331</td>
<td>331</td>
</tr>
<tr>
<td>R2</td>
<td>0.08</td>
<td>0.09</td>
</tr>
</tbody>
</table>