House Price Dynamics with Dispersed Information*

Giovanni Favara† and Zheng (Michael) Song‡

September 2012

Abstract

We use a user-cost model to study how dispersed information affects the equilibrium house price. In the model, agents are disparately informed about local economic conditions, consume housing services, and speculate on price changes. 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 an anonymous referee, Alessandro Beber, Darrell Duffie, Bernard Dumas, Simon Gilchrist, John Hassler, Christian Hellwig, Ethan Kaplan, Per Krusell, Rafael Lalive, 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, the University of Zurich, Shanghai JiaoTong University, the Sverige Riksbank, Bocconi University, Luiss, the Einaudi Institute for Economics and Finance, IMF, ECB, George Washington University, St. Louis FED, Philadelphia 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, 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.

†Federal Reserve Board. Email: giovanni.favara@frb.gov

‡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 both over time and across regions. 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.[1] 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).[2]

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

[1] In some cities, such as Los Angeles, housing prices have moved in tandem with the overall national index, though they have moved much less. In other cities, prices movements have been quite heterogeneous. In Miami, for example, the house price index has declined sharply for almost a decade and then increased exponentially by the end of the sample; in San Antonio, it has declined since the 1980s; in Rochester, it has displayed an inverse "U-shaped" history; in Memphis, it has gone through periodic cycles. Figure 1 plots the time series of these indices until 2000 because the empirical analysis in Section 6 focuses only on the sample period between 1980 and 2000. The same heterogeneity in trends and dynamics persists, however, in more recent years, including the housing boom and bust between 2005 and 2010.

[2] While 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.

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 and Liuet, 2006). Second, surveys of housing market participants (Case and Shiller, 1988, 2003, 2012; 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 also 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 four 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 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 housing 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 to isolate the role of heterogenous expectations and short sale constraints in the determination of the equilibrium house price. However, the main predictions 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 limited ability to borrow of the optimists as well as the collateral value of houses, if houses are pledged as collateral (see e.g., 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 Weill (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 Weill use a similar mechanism to explain both the level and the dispersion of house prices in the U.S., though in their model agents move across cities for productivity shocks rather than preference reasons. Our empirical findings that income dispersion correlates with the level and dispersion of house prices are thus similar to those in Van Nieuwerburgh and Weill. However, while they use a spatial equilibrium model of the housing market with agents indifferent between different locations, given local wages and amenities, the predictions of our model arise in a standard user cost model with no-arbitrage condition between owning and renting. In our framework, income shocks do not cause agents to move across areas, but 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.

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 search frictions and heterogeneous beliefs (as opposed to heterogeneous expectations) 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_t^j, at which labor is supplied inelastically, is equal to

$$W_t^j = \exp\left(\theta_t + \varepsilon_t^j\right), \tag{1}$$

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

$$\theta_t = \rho\theta_{t-1} + \eta_t, \quad \text{with} \quad \rho \in (0, 1], \tag{2}$$

where η_t is independently and normally distributed with zero mean and variance σ_η^2. When agents cannot observe the realization of θ_t, ε_t^j becomes a source of information heterogeneity. In other words, the individual wage W_t^j is the agent $j's$ noisy private signal about the unobservable aggregate shock, θ_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 θ_t:

Assumption 1: $\sum_j \varepsilon_t^j = 0$.

Throughout the paper we maintain the assumption that agents observe their idiosyncratic wages but do not observe the aggregate wage. This is akin to assume that agents take optimal decisions before news about the aggregate wage is released, as in the standard signal extraction model of Lucas (1972) in which only local, but not aggregate, variables are observable.[3]

[3] Alternatively, we may assume that agents have access to public information about θ but this information is plagued with noise due to, for example, measurement errors. In this modified setting, even if the precision of the public information is high, agents may remain uninformed about θ. As shown in Amador and Weill (2010), for example, increasing the precision of exogenous public information has the direct effect of providing

3.2 Preferences

Agents have logarithmic preferences over housing services, V_t^j, and second-period consumption, C_{t+1}^j,

$$\mathcal{U}_t^j = A_t^j \log V_t^j + E_t^j \log C_{t+1}^j, \tag{3}$$

where E_t^j denotes the expectation operator based on household j's information set at time t (to be specified later), and A_t^j is a preference shock,

$$A_t^j = \exp\left(a_t + \nu_t^j\right),$$

which consists of an aggregate taste shock, a_t, and an idiosyncratic noise ν_t^j. We assume that a_t and ν_t^j are independent and normally distributed with zero mean and variance σ_a^2 and σ_ν^2. We also consider the limiting case where the variance of ν_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 5.2 when we allow agents to use the equilibrium price to update their beliefs about θ.

Our specification of preferences makes important assumptions. First, it assumes away any intertemporal consumption-saving decision. This has, however, inessential consequences for our analysis given that the main focus is on the rental-owning margin. Second, it posits that agents do not have preferences for housing when old. This implies that agents make owning-renting decision only in the first period of life, as hypothetical first-time buyers would do. While this simplifying assumption has the virtue of making the model tractable, it also prevents the model from shedding lights on other important aspects of the housing market, such as agents' decision to retrade or to transit from ownership to renting. Lastly, in the model, housing units are homogeneous and provide the same quality of housing services. This assumption is standard in a user-cost model but it neglects the fact that richer agents with a preference for a minimum quality of houses may not have alternative to owning.[4]

new information, but may also crowd out private information, reducing the importance of private signals and thus the endogenous information efficiency of the price system. In some cases, this crowding out may increase rather than decrease aggregate uncertainty.

[4] See Landvoigt, Piazzesi and Schneider (2012) for a more elaborate user-cost model in which housing differ by quality. Their model, however, treats house price expectations parametrically, while the focus of this paper is on how agents form price expectations based on their limited information about the state of the economy.

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

$$H_t^j \geq 0. \tag{5}$$

The non negativity constraint (5) will play a crucial role in the analysis. It amounts to saying that houses cannot be sold short. When agents hold heterogeneous expectations this short sale constraint implies that the natural buyers are those with relatively more optimistic expectations about next period house prices.

3.4 Optimal housing 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:

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

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

where

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

denotes the (per-unit) user cost of housing, which decreases with next-period house price, P_{t+1}/R.[5]

[5] 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, $P_t H_t^j$. 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

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

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

$$v_t^j = w_t^j + a_t^j - q_t, \tag{9}$$

and

$$E_t^j u_t \geq q_t, \quad \text{and} \quad H_t^j \geqslant 0 \tag{10}$$

where

$$u_t \equiv \frac{(1+r)p_t - p_{t+1}}{r}, \quad r \equiv R - 1 > 0, \tag{11}$$

is the linearized user cost, and $a_t^j \equiv \left(a_t + \nu_t^j\right)/2$ 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.[6] With the convention that agents in group $j = 1$ are relatively more optimistic about the next-period house price, i.e., $E_t^1 p_{t+1} > E_t^0 p_{t+1}$, we can rewrite equation (10) as follows:

$$E_t^0 u_t \;\; > \;\; q_t \quad \text{and} \quad H_t^0 = 0 \tag{12}$$

$$E_t^1 u_t \;\; = \;\; q_t \quad \text{and} \quad H_t^1 > 0, \tag{13}$$

suggesting that with heterogeneous expectations pessimists choose to own no housing units,

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.

[6] Notice that because a log-linearization of (7) does not involve H_t^j 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

$H_t^0 = 0$ (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_t^1, out of the units owned, H_t^1, and rent out the difference, $H_t^1 - V_t^1$, 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:

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

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

$$q_t = \theta_t + a_t, \tag{14}$$

where

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

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:

$$p_t = \frac{r}{1+r} q_t + \frac{1}{1+r} E_t^1 p_{t+1}, \tag{15}$$

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

$$p_t = \frac{r}{1+r} f_t + \frac{1}{1+r} \overline{E}_t p_{t+1} + \frac{1}{1+r} \widetilde{E}_t p_{t+1}, \tag{16}$$

where

$$f_t \equiv \theta_t + a_t \tag{17}$$

summarizes average fundamental variables, and

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

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, $\widetilde{E}_t p_{t+1}$, 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 $\overline{E}_t p_{t+1}$ and $\widetilde{E}_t p_{t+1}$ 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, θ_t, and thus rely only on public information, θ_{t-1}, to infer θ_t. In other words, agents share a common information set. In this case individual expectations coincide with the average expectation, i.e., $E_t^j p_{t+1} = \overline{E}_t p_{t+1}$ and the difference in expectations is zero, $\widetilde{E}_t p_{t+1} = 0$.

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

$$\overline{E}_t p_{t+1} = \phi \rho \theta_{t-1}, \tag{18}$$

with

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

The average expectation depends on past fundamentals, θ_{t-1}, because θ_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 $\widetilde{E}_t p_{t+1} = 0$, we have

Proposition 1 *The equilibrium house price with homogeneous information, p^*, is equal to*

$$p_t^* = f_t + \Lambda_t, \tag{19}$$

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 is, by Assumption 1, 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_t^j, as well as the public signal, θ_{t-1}, to make an optimal inference about θ_t. Agent j's information set at t is,[7]

$$\Omega_t^j = \left\{w_t^j, \theta_{t-1}\right\} \quad j = 0, 1.$$

It is important to notice that the equilibrium house price is not included in Ω_t^j. 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 the results.[8]

With signals w_t^j and θ_{t-1}, the ability of agent j to estimate θ_t depends on the relative magnitude of σ_ε^2 and σ_η^2. Given the assumption of independently and normally distributed errors, the projection theorem implies

$$E_t^j\theta_t = (1-\lambda)\rho\theta_{t-1} + \lambda w_t^j, \tag{20}$$

where the weight $\lambda \equiv \sigma_\eta^2 / \left(\sigma_\eta^2 + \sigma_\varepsilon^2\right)$ 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 $E_t^j\theta_t$ depends on w_t^j, 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,

$$\begin{aligned} \widetilde{E}_t p_{t+1} &= \phi\lambda i_t, &(21)\\ \overline{E}_t p_{t+1} &= \phi\rho\theta_{t-1} + \frac{\phi\lambda}{r}I + \phi\lambda\left(\theta_t - \rho\theta_{t-1}\right), &(22)\end{aligned}$$

[7] It is superfluous to know the entire history of aggregate shocks since θ_t follows an $AR(1)$ process. Similarly, knowing the past realization of agents' private signals is irrelevant, given the *iid* assumption for ε_t^j.

[8] A way to think of this assumption is to consider the special case where the variance of the aggregate preference shock, σ_a^2, is arbitrarily large. In such a case, the house price (16) becomes uninformative about θ_t and housing-market participants do not learn much upon observing p_t. In excluding p_t from agents' information set, we make our analysis akin to models where agents do not condition on the equilibrium price because they do not know how to use prices correctly (e.g., they display bounded rationality, as in Hong and Stein, 1999) or because they exhibit behavioral biases (e.g., they are overconfident, as in Scheinkman and Xiong, 2003).

where

$$i_t \equiv \varepsilon_t^1 - \varepsilon_t^0,$$

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

$$I \equiv \int_0^{\infty} x d\Gamma(x)$$

measures the average degree of information heterogeneity in the economy (with Γ 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 θ_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, θ_{t-1}, to make the optimal inference about θ_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*

$$p_t = p_t^* + \lambda\Upsilon_t, \tag{23}$$

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

$$\Upsilon_t \equiv \phi\frac{(\theta_t - \rho\theta_{t-1})}{1+r} + \phi\frac{I}{r(1+r)} + \phi\frac{i_t}{1+r} \tag{24}$$

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 Υ_t is positive, implying that information dispersion leads to a higher equilibrium house price. This is the case because optimists estimate a higher θ_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 ε_t^1 increases relative to ε_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, σ_η^2, and the larger the variance of information dispersion, σ_i^2:

$$V(p_t) - V(p_t^*) = \left(\frac{\lambda\phi}{1+r}\right)^2 \left(\sigma_\eta^2 + \sigma_i^2\right) > 0. \tag{25}$$

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 also arise in a setting that abstracts from heterogeneous information but features credit frictions. It turns out that the implications of our model do not hinge on the assumption that the demand for housing is independent of credit conditions. To see why, notice that if agents have homogeneous information (i.e., $E_t^j u_t = E_t u_t$) equation (10) implies that either $E_t u_t = q_t$ or $E_t u_t < q_t$. When $E_t u_t < q_t$ all agents prefer owning to renting and so everyone must be constrained. If they were not, the optimal demand for housing h_t^j would increase until the borrowing constraint is binding for any one, irrespective of their wages. Conversely, when $E_t u_t = q_t$ all agents are indifferent about the number of housing units to own, which is equivalent to say that no one will be constrained: in equilibrium those with lower income will demand fewer housing units, and those with higher income will demand more. In both cases, the price of housing will depend on the average expectation in the market, or equivalently (in the model) the average income, irrespective of the credit constraint. Accordingly, in a setting with borrowing constraints and common information the price of housing cannot be higher the larger the dispersion in income — it will be higher only if the average price expectation (or average income) is higher.

Our model's predictions would also continue to hold if agents had heterogeneous expectations (as in our model) *and* faced credit constraints. The reason is that with heterogeneous expectations the short sale constraint implies that the optimists are the marginal buyers, even if they are credit constrained. Of course, the pricing equation would be different, possibly reflecting the collateral value of houses (if these assets are pledged as collateral as e.g., in Geanakoplos, 2009) and the fact that optimists' demand for housing is limited by their ability to borrow. However, our main intuition that the equilibrium price is higher the larger the difference in expectations would continue to hold.

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., $i_t = \left|\varepsilon_t^i - \varepsilon_t^j\right|$. But, since i_t is not normally distributed, p_t has a non-Gaussian distribution, and standard linear filtering methods cannot be used.[9]

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 $\mathcal{A}$, with zero mean and variance σ_a^2. Moreover, a_t is such that $a_t + i_t \equiv \delta_t \sim N$ ($\bar{\imath}$, σ_δ^2) where $\bar{\imath}$ denotes the unconditional mean of i_t and σ_δ^2 the variance of $a_t + i_t$.[10]

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*

$$p_t = p_t^* + \pi_2 \Upsilon_t + \pi_3 \Phi_t, \tag{26}$$

[9]See Appendix IV for a derivation of the exact distribution of i_t.

[10]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

$$\Phi_t \equiv \frac{\phi\left(\theta_t - \rho\theta_{t-1}\right)}{1+r} + \frac{\phi r}{\left(1+r\right)\left(r+\phi\pi_2\right)} a_t + \frac{\phi^2 \pi_2}{\left(1+r\right)\left(r+\phi\pi_2\right)} i_t$$

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 (η_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 Υ_t, for the reasons discussed in the previous section, and an indirect effect, via Φ_t, because of the magnification of shocks induced by the process of learning.

A comparison of (26) and (23) also reveals that the difference in the equilibrium price with and without learning depends on $\left(\pi_2 - \lambda\right)\Upsilon_t$ and $\pi_3\Phi_t$. Since $\pi_2 \leq \lambda$ and $\pi_3 \geq 0$, it follows that learning weakens the direct effect of i_t via Υ_t (i.e., $\left(\pi_2 - \lambda\right)\Upsilon_t < 0$) but it exacerbates the indirect effect of i_t via Φ_t (i.e., $\pi_3\Phi_t > 0$).[11] As shown in Appendix IV, however, the direct effect of information heterogeneity via Υ_t always prevails over its indirect effect via Φ_t. Accordingly,

Corollary 1 *The equilibrium house price with learning has a higher mean than that without learning.*

All in all, the effect of information dispersion on the equilibrium housing price survives in a more general setting with learning, even though the magnitude of such an effect is muted.[12]

[11] Appendix IV shows that $\pi_2 < \lambda$ and $\pi_3 > 0$ but $\pi_2 \to \lambda$ and $\pi_3 \to 0$ as the noise in the preference for housing services increases, $\sigma_a^2 \to \infty$. In this latter case the equilibrium price with learning (26) and the one without learning (23) are identical.

[12] With learning, the volatility of the equilibrium house price also 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.

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 information dispersion; 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.[13] To overcome this limit, our strategy is to 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

[13] Case and Shiller (1988, 2003, 2012) provide survey data on house price expectations in 1988 and for each year between 2003 and 2012 for four U.S. metropolitan areas. These surveys can be used to measure local house price expectations, but cannot be used to test the prediction of our model that the price of housing increases with disagreement among housing market participants. The reason is that these surveys collect the opinion of people that have actually bought a house, but neglect home seekers and those that prefer renting to buying. An alternative to the Case and Shiller survey is the Michigan Consumers Survey used, for example, in Piazzesi and Schneider (2009). The Michigan survey, however, has also important limitations for the purpose of testing our theory: it does not report the location of the survey respondent, and provides only an average measurement of house price expectations across U.S. cities.

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.[14] 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.[15]

In a second step, we run ten regressions, one for each industry, in which we pool the growth rate of industry earnings, $\Delta w_{k,t}^j$, for the full sample of MSAs:

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

where γ_t is a time fixed effect. This specification is based on equation (27) in the model, and adds lags of $w_{k,t}^j$ and $\theta_{k,t}$ to account for a minimum of industry and city income dynamics. As a result, the residuals $\varepsilon_{k,t}^j$ record shocks to industry-j's earnings in city-k, controlling for nationwide effects, γ_t, industry-MSA specific earning dynamics, $\Delta w_{k,t-1}^j$, and exogenous MSA income dynamics, $\Delta\theta_{k,t}$ and $\Delta\theta_{k,t-1}$.[16]

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,

$$i_{k,t} = \sum_{j=1}^{10} \omega_{k,t}^j \left|\varepsilon_{k,t}^j\right|, \tag{30}$$

where the weights $\omega_{k,t}^j$ denote the share of MSA workers employed in industry j, to control for the size of each industry.[17] Accordingly, this variable captures the dispersion of local

[14] 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.

[15] 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.

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

[17] 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

income shocks that are orthogonal to changes in local income, via (28), and to changes in aggregate income, via the period fixed effects, γ_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 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.[18] 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

coefficients in the house price regressions reported below can easily be interpreted.

[18] 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.

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:

$$\Delta p_{k,t} = \gamma_t + \gamma_k + X_{k,t}\beta + \delta_1 i_{k,t} + \delta_2(i_{k,t} \times \eta_k^S) + \epsilon_{k,t}, \tag{31}$$

where $\Delta p_{k,t}$ is the log change of the real house price index in MSA k in year t, γ_t is a year effect common to all markets, γ_k, is a time-invarying 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.[19]

The parameters of interest are δ_1 and δ_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 δ_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 η_k^S. 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

[19] 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.

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 δ_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

Almeida, H., M. Campello and C. Liu (2006), "The Financial Accelerator: Evidence From International Housing Markets," *Review of Finance*, 10, 1-32.

Amador M. and P. Weill (2010), "Learning from Prices: Public Communication and Welfare," *Journal of Political Economy*, 118, 866-907.

Bartik, T.J., (1991), "Who Benefits from State and Local Economic Development Policies?," W.E. Upjohn Institute for Employment Research: Kalamazoo, Mich.

Blanchard, O.J. and L.F. Katz (1992), "Regional evolutions," *Brookings Papers on Economic Activity*, 1-75.

Burnside C., M., Eichenbaum, and S. Rebelo (2011), "Understanding Booms and Busts in Housing Markets," *mimeo.*

Case, K. E., and R. J. Shiller (1988), "The Behavior of Home Buyers in Boom and Post-Boom Markets," *New England Economic Review* (November/December) 29-46.

Case, K. E., and R. J. Shiller (1989), "The Efficiency of the Market for Single Family Homes," *American Economic Review*, 79(1), 125-137.

Case, K. E., and R. J. Shiller (2003) "Is There a Bubble in the Housing Market," *Brooking Papers on Economic Activity*, 2, 299-342.

Case, K. E., and R. J. Shiller (2012) "What Have They Been Thinking? Home Buyer Behavior in Hot and Cold Markets," NBER WP 18400.

Case, K. E., J. M., Quigley and R. J., Shiller (2003), "Home-buyers, Housing and the Macroeconomy," in Anthony Richards and Tim Robinson, eds., Asset Prices and Monetary Policy, Canberra: Reserve Bank of Australia, 149-188.

Davis S. J., P. Loungani and R. Mahidhara (1997), "Regional Labor Fluctuations: Oil Shocks, Military Spending, and Other Driving Forces", Board of Governors of the Federal Reserve System WP. 578

Englund, P. and Y.M. Ioannides (1997), "House Price Dynamics: An International Empirical Perspective," *Journal of Housing Economics*, 6, 119-136.

Geanakoplos, J. (2009), "The Leverage Cycle", In D. Acemoglu, K. Rogoff and M. Woodford, eds., *NBER Macroeconomic Annual* 24, 1-65, University of Chicago Press

Glaeser, E. and J. Gyourko (2003), "The Impact of Zoning on Housing Affordability," Federal Reserve Bank of New York, Economic Policy Review, 9(2), 21-39.

Glaeser, E. and J. Gyourko (2006), "Housing Dynamics," NBER W.P. 12787

Glaeser, E. and J. Gyourko (2007), "Arbitrage in Housing Markets," NBER W.P. 13704

Glaeser, E., J. Gyourko and A. Saiz (2008), "Housing Supply and Housing Bubbles," NBER W.P. 14193

Glaeser, E., J. Gyourko and R. Saks (2005a), "Why Have Housing Prices Gone Up?" *American Economic Review Papers and Proceedings*, 95(2), 329–333.

Grossman, S. J. and J. E. Stiglitz (1976), "Information and Competitive Price Systems." *American Economic Review,* 66, 246-253.

Guerrieri V., D. Hartley and E. Hurst (2010), "Endogenous Gentrification and Housing Price Dynamics", NBER W.P. 16237

Gyourko, J., C. Mayer and T. Sinai (2006), "Superstar Cities," NBER W.P. 12355.

Harrison, M., and D. Kreps (1978), "Speculative investor behavior in a stock market with heterogeneous expectations", *Quarterly Journal of Economics,* 92, 323-336.

Hellwig, C., Mukherji, A., and Tsyvinski A., (2006), "Self-Fulfilling Currency Crises: The Role of Interest Rates." *American Economic Review,* 96(5), 1769-1787.

Hellwig, M. F., (1980), "On the Aggregation of Information in Competitive Markets." *Journal of Economic Theory,* 22(3), 477-498.

Henderson, V., and Y. Ioannides (1983), "A Model of Housing Tenure Choice," *American Economic Review,* 73(1), 93-113.

Himmelberg, C., C. Mayer and T. Sinai (2005), "Assessing High House Prices: Bubbles, Fundamentals and Misperceptions," *Journal of Economic Perspective*, 89, 323-336.

Hong, H., and J. Stein (1999),"A Unified Theory of Underreaction, Momentum Trading and Overreaction in the Asset Market", *Journal of Finance*, 45, 2143-2184.

Lamont, O. and J. Stein (1999), “Leverage and House-Price Dynamics in U.S. Cities,” *Rand Journal of Economics*, 30, 498-514.

Landvoigt, T., M. Piazzesi and M. Schneider (2011), “The Housing Market(s) of San Diego”, NBER W.P. 17723.

Lucas, R. (1972), “Expectations and the Neutrality of Money,” *Journal of Economic Theory*, 4, 103-124.

McCarthy, J. and R. Peach, (2004), “Are Home Prices the Next Bubble?”, *Economic Policy Review,* 10(3), 1-17.

Miller, E. M. (1977), “Risk, uncertainty, and divergence of opinion”, *Journal of Finance,* 32, 1151-68.

Neumann G., and R. Topel (1991) “Employment Risk, Diversification, and Unemployment”, *Quarterly Journal of Economics*, 106(4), 1341-1365

Ortalo-Magné, F. and S. Rady (2006), “Housing Market Dynamics, on the Contributions of Income Shocks and Credit Constraints,” *Review of Economic Studies,* 73, 459-485.

Piazzesi, M. and M. Schneider (2009), “Momentum Traders in the Housing Market: Survey Evidence and a Search Model,” *American Economic Review Papers and Proceedings,* 99(2), 406-411.

Poterba, J. (1984), “Tax Subsidies to Owner-Occupied Housing: An Asset Market Approach”, *Quarterly Journal of Economics*, 99, 729-52.

Poterba, J. (1991), “House Price Dynamics: the Role of Tax Policy and Demography”, *Brooking Papers on Economic Activity*, 2, 143-203.

Saiz, A., (2010), “On Local Housing Supply Elasticity” forthcoming *Quarterly Journal of Economics*

Scheinkman, J. and W. Xiong (2003), "Overconfidence and Speculative Bubbles", *Journal of Political Economy,* 111, 1183-1219.

Stein, J. C. (1995), “Prices and Trading Volume in the Housing Market: A model with Downpayment Constraints,” *Quarterly Journal of Economics*, 110(2), 379-406.

Van Nieuwerburgh, S. and P.O. Weil (2010), “Why Has House Price Dispersion Gone Up?” *Review of Economic Studies,* 77(4) 1567–1606

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

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

$$\begin{aligned} V^j &= V > 0 \Longrightarrow V = \frac{C}{R}\frac{1}{Q}, && (32) \\ Q &= U. && (33) \end{aligned}$$

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

$$\frac{C}{R} = W - VQ. \tag{34}$$

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,

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

Rearranging,

$$\begin{aligned} E_t^j \left[\frac{RP}{C} p_t - \frac{RQ}{C} q_t - \frac{P}{C} p_{t+1} - c_{t+1}^j \left(\frac{RP}{C} - \frac{RQ}{C} - \frac{P}{C} \right) \right] &\geq 0 \Rightarrow \\ E_t^j \left[RP p_t - RQ q_t - P p_{t+1} \right] &\geq 0 \Rightarrow \\ E_t^j \left[p_t - \frac{Q}{P} q_t - \frac{1}{R} p_{t+1} \right] &\geq 0, \end{aligned}$$

we obtain

$$p_t \geq \frac{r}{1+r} q_t + \frac{1}{1+r} E_t^j p_{t+1}, \tag{35}$$

where

$$r = R - 1.$$

Notice, also, that a linearization (8) gives

$$\begin{aligned} u_t &= \frac{P}{U} p_t - \frac{P}{RU} p_{t+1} \\ &= \left(\frac{1+r}{r}\right) p_t - \frac{1}{r} p_{t+1}. \end{aligned}$$

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

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

$$\frac{(1+r)p_t - E_t^j p_{t+1}}{r} \geq q_t. \tag{36}$$

Since $E_t^1 p_{t+1} > E_t^0 p_{t+1}$, equation (36) holds with strict inequality for $j = 0$ and so pessimists choose to own no housing units, $H_t^0 = 0$.

Linearization of (6)

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

$$\begin{aligned} E_t^j \frac{RQ}{C} \left(q_t - c_{t+1}^j\right) &= \frac{A}{V}(2a_t^j - v_t^j), \\ E_t^j \frac{1}{S} \left(q_t - c_{t+1}^j\right) &= \frac{1}{V}(2a_t^j - v_t^j), \end{aligned}$$

which defines the optimal demand of housing services

$$v_t^j = 2a_t^j - q_t + E_t^j c_{t+1}^j. \tag{37}$$

The term $E_t^j c_{t+1}^j$ in (37) is obtained by linearizing the flow of budget constraint (4), that for the two groups of agents reads as follows:

$$C_{t+1}^1 = R\left(W_t^1 - P_t H_t^1 + Q_t\left(H_t^1 - V_t^1\right)\right) + P_{t+1} H_t^1, \tag{38}$$

$$C_{t+1}^0 = R\left(W_t^0 - Q_t V_t^0\right), \tag{39}$$

where the second equation uses the fact that $H_t^0 = 0$. A bit of algebra establishes[20]

$$E_t^1 c_{t+1}^1 = 2w_t^1 - v_t^1 - \left(\frac{r+1}{r}\right) p_t + \frac{1}{r} E_t^1 p_{t+1} = 2w_t^1 - v_t^1 - E_t^1 u_t, \tag{40}$$

[20] Linearizing (38) yields

$$\begin{aligned} E_t^1 c_{t+1}^1 &= \frac{RW}{C} w_t^1 - \frac{RPH}{C}(p_t + h_t^1) + \frac{RQH}{C}(q_t + h_t^1) - \frac{RQV}{C}(q_t + v_t^1) \\ &\quad + \frac{PH}{C}(E_t^1 p_{t+1} + h_t^1) \\ &= 2w_t^1 - \frac{P}{U}(p_t + h_t^1) + (q_t + h_t^1) - (q_t + v_t^1) + \frac{P}{RU}(E_t^1 p_{t+1} + h_t^1) \end{aligned}$$

Rearranging this equation gives (40). Proceeding in a similar way, one obtains (41).

$$E_t^0 c_{t+1}^0 = 2w_t^0 - v_t^0 - q_t. \tag{41}$$

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

$$\begin{aligned} v_t^1 &= w_t^1 + a_t^1 - \frac{1}{2}\left(q_t + E_t^1 u_t\right) = w_t^1 + a_t^1 - q_t, \\ v_t^0 &= w_t^0 + a_t^0 - q_t. \end{aligned}$$

These establish equation (9).

Appendix II: Proof for Proposition 1

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

$$p_{t+1} = \frac{r}{1+r}\left(\theta_{t+1} + a_{t+1}\right) + \frac{1}{1+r}\bar{E}_{t+1} p_{t+2}.$$

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

$$\overline{E}_t p_{t+1} = \frac{r}{1+r}\sum_{\tau=0}^{\infty}\left(\frac{1}{1+r}\right)^{\tau}\overline{E}_t\left(\theta_{t+1+\tau} + a_{t+1+\tau}\right),$$

Since θ_t and a_t are unobservable at time t and

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

we have

$$\overline{E}_t\left[\theta_{t+1} + a_{t+1}\right] = \rho^2\theta_{t-1}.$$

It is, therefore, immediate to obtain

$$\overline{E}_t p_{t+1} = \overline{E}_t f_t = \phi\rho\theta_{t-1}, \tag{42}$$

where $\phi \equiv \frac{r\rho}{1+r-\rho}$. Plugging (42) back into (16) and recalling that $\tilde{E}_t p_{t+1} = 0$, the equilibrium price under common information can then be written as

$$p_t^* = (\theta_t + a_t) + \frac{1}{1+r}\left((\phi\rho\theta_{t-1} - \theta_t) - a_t\right).$$

Appendix III: Proof for Proposition 2

In the presence of heterogeneous expectations, $E_t^j p_{t+1} \neq \overline{E}_t p_{t+1}$ and $\widetilde{E}_t p_{t+1} \neq 0$. Shifting equation **(16)** one period forward

$$p_{t+1} = \frac{r}{1+r}(\theta_{t+1} + a_{t+1}) + \frac{1}{1+r}\overline{E}_{t+1}p_{t+2} + \frac{1}{1+r}\widetilde{E}_{t+1}p_{t+2}$$

denoting,

$$i_t = \left|\varepsilon_t^j - \varepsilon_t^i\right| \quad \text{for} \quad i \neq j.$$

and guessing that $\widetilde{E}_t\left[p_{t+1}\right] = \phi\lambda i_t$, we have

$$\begin{aligned}
E_t^j p_{t+1} &= \frac{r}{1+r} E_t^j(\theta_{t+1} + a_{t+1}) + \frac{1}{1+r} E_t^j \overline{E}_{t+1} p_{t+2} + \frac{\phi\lambda}{1+r} I \\
\overline{E}_t p_{t+1} &= \frac{r}{1+r} \bar{E}_t(\theta_{t+1} + a_{t+1}) + \frac{1}{1+r} \overline{E}_t \overline{E}_{t+1} p_{t+2} + \frac{\phi\lambda}{1+r} I, \\
\widetilde{E}_t p_{t+1} &= \frac{r}{1+r} \widetilde{E}_t \theta_{t+1} + \frac{1}{1+r} \widetilde{E}_t \overline{E}_{t+1} p_{t+2},
\end{aligned}$$

where the last equality holds because agents hold heterogeneous expectations with respect to θ_{t+1} but not with respect to a_{t+1}. In the expressions above,

$$I \equiv \int_0^\infty x d\Gamma\left(x\right),$$

is the average degree of information heterogeneity where Γ is the density of i_t.

Iterating these expressions forward and excluding explosive price paths, we obtain:

$$\begin{aligned}
E_t^j p_{t+1} &= \frac{r}{1+r-\rho} E_t^j \theta_{t+1} + \frac{\phi\lambda}{r} I, \\
\bar{E}_t p_{t+1} &= \frac{r}{1+r-\rho} \bar{E}_t \theta_{t+1} + \frac{\phi\lambda}{r} I, \\
\tilde{E}_t p_{t+1} &= \frac{r}{1+r-\rho} \tilde{E}_t \theta_{t+1}.
\end{aligned}$$

Moreover, using equation equation (20), it is easy to see that:

$$E_t^j \theta_{t+1} = \rho E_t^j \theta_t = \rho\left[(1-\lambda)\rho\theta_{t-1} + \lambda w_t^j\right],$$

and, thus,

$$\begin{aligned}
\bar{E}_t p_{t+1} &= \phi\left(\rho(1-\lambda)\theta_{t-1} + \lambda\theta_t\right) + \frac{\phi\lambda}{r} I, \\
&= \left(\phi\rho\theta_{t-1}\right) + \phi\lambda\left(\theta_t - \rho\theta_{t-1}\right) + \frac{\phi\lambda}{r} I, \\
\widetilde{E}_t p_{t+1} &= \phi\lambda i_t.
\end{aligned}$$

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

$$\begin{aligned}
p_t &= \left(\theta_t + a_t\right) + \frac{1}{1+r}\left(\left(\phi\rho\theta_{t-1} - \theta_t\right) - a_t\right) \\
&\quad + \frac{\phi\lambda}{1+r}\left(\theta_t - \rho\theta_{t-1}\right) + \frac{\phi\lambda}{r(1+r)} I + \frac{\phi\lambda}{1+r} i_t. \\
&= p_t^* + \lambda\Upsilon_t
\end{aligned}$$

where

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

Appendix IV: Learning from the Equilibrium House 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, θ_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 μ_t, is such that sum of i_t and μ_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 μ_t that makes the equilibrium price normally distributed. Finally, using a method of undetermined coefficients, we characterize the inference problem for θ_t and the resulting equilibrium price.

The distribution of $i = \left|\varepsilon^i - \varepsilon^j\right|$ for $i \neq j$

Consider two independent random variables, ε^i and ε^j, distributed normally with zero mean and equal variance σ_ε^2. Define,

$$\tilde{\varepsilon} = \varepsilon^j - \varepsilon^i \sim \mathcal{N}(0, 2\sigma_\varepsilon^2).$$

The cumulative distribution function of $i = |\tilde{\varepsilon}|$ is

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

and the associated density,

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

Denote with $\bar{\imath}$, the mean of i,

$$\bar{\imath} = \int_0^\infty y f_i\left(y\right) dy.$$

The distribution of the aggregate preference shock, a.

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

$$a + i \sim \mathcal{N}(\bar{\imath}, \quad \sigma_a^2 + \sigma_i^2).$$

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

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

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

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

Lemma 4

Lemma 4 *The correlation coefficient between* ε^j *and* $i \equiv \left|\varepsilon^j - \varepsilon^i\right|$ *is zero.*

Proof.

$$\begin{aligned} Cov\left(\varepsilon^j, \left|\varepsilon^j-\varepsilon^i\right|\right) &= Cov\left(\varepsilon^j, \varepsilon^j-\varepsilon^i\right)\Pr\left(\varepsilon^j>\varepsilon^i\right) + Cov\left(\varepsilon^j, -(\varepsilon^j-\varepsilon^i)\right)\Pr\left(\varepsilon^j<\varepsilon^i\right) \\ &= Cov\left(\varepsilon^j, \tilde{\varepsilon}\right)\Pr\left(\varepsilon^j>\varepsilon^i\right) - Cov\left(\varepsilon^j, \tilde{\varepsilon}\right)\Pr\left(\varepsilon^j<\varepsilon^i\right) \\ &= Cov\left(\varepsilon^j, \tilde{\varepsilon}\right)\left[\Pr\left(\varepsilon^j>\varepsilon^i\right) - \Pr\left(\varepsilon^j<\varepsilon^i\right)\right] = 0 \end{aligned}$$

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

The method of undetermined coefficients

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

$$p_t = b_0 + b_\theta \rho \theta_{t-1} + b_\eta \eta_t + b_a a_t + b_i i_t, \tag{44}$$

where b_0, b_θ, b_η, b_a and b_i are undetermined coefficients. It is convenient to rewrite equation (44) as

$$p_t = b_\eta \eta_t + b_a a_t + b_i i_t + X_t, \tag{45}$$

where

$$X_t \equiv b_0 + b_\theta \rho \theta_{t-1}$$

is non-stochastic. Defining

$$\widehat{p}_t \equiv \frac{p_t - X_t}{b_\eta},$$

equation (45) can be written as

$$\widehat{p}_t = \eta_t + \delta_t,$$

where,

$$\delta_t = \frac{b_a}{b_\eta} a_t + \frac{b_i}{b_\eta} i_t. \tag{46}$$

Under the assumption made on the distribution of a_t, δ_t is normally distributed,

$$\delta_t \sim \mathcal{N}\left(\frac{b_i}{b_\eta}\bar{\imath}, \quad \left(\frac{b_a}{b_\eta}\right)^2 \sigma_a^2 + \left(\frac{b_i}{b_\eta}\right)^2 \sigma_i^2\right)$$

and, as a consequence $\widehat{p}_t$, is also normally distributed,

$$\widehat{p}_t \sim \mathcal{N}\left(\frac{b_i}{b_\eta}\bar{\imath}, \quad \sigma_\eta^2 + \frac{b_a^2 \sigma_a^2 + b_i^2 \sigma_i^2}{b_\eta^2}\right). \tag{47}$$

The inference problem

Agent j estimates the unknown fundamental θ_t by solving a standard filtering problem, based on the normally distributed (a) private signal, w_t^j, (b) exogenous public signal, θ_{t-1}, and (c) endogenous public signal, $\widehat{p}_t$. Recalling that

$$\begin{aligned} \theta_t &= \rho \theta_{t-1} + \eta_t, \\ w_t^j &= \theta_t + \varepsilon_t^j, \\ \widehat{p}_t &= \eta_t + \delta_t, \end{aligned}$$

and using (47) and Lemma 4, the log-likelihood function can be written as

$$L = -\frac{1}{2\sigma_\eta^2}\left(\rho\theta_{t-1} - E_t^j \theta_t\right)^2 - \frac{1}{2\sigma_\varepsilon^2}\left(w_t^j - E_t^j \theta_t\right)^2 - \frac{1}{2\sigma_\delta^2}\left(\widehat{p}_t - E_t^j \eta_t\right)^2.$$

Thus, the optimal filtering solves the following first-order condition,

$$-\frac{1}{\sigma_\eta^2}\left(-E_t^j \eta_t\right) + \frac{1}{\sigma_\varepsilon^2}\left(w_t^j - \rho\theta_{t-1} - E_t^j \eta_t\right) + \frac{1}{\sigma_\delta^2}\left(\widehat{p}_t - E_t^j \eta_t\right) = 0,$$

or,

$$E_t^j \eta_t = \frac{\sigma_\eta^2 \sigma_\delta^2 \left(w_t^j - \rho\theta_{t-1}\right) + \sigma_\eta^2 \sigma_\varepsilon^2 \widehat{p}_t}{\sigma_\varepsilon^2 \sigma_\delta^2 + \sigma_\eta^2 \sigma_\delta^2 + \sigma_\eta^2 \sigma_\varepsilon^2}.$$

The best linear estimate of θ_t is, therefore,

$$E_t^j \theta_t = (\pi_1 + \pi_3)\, \rho\theta_{t-1} + \pi_2 w_t^j + \pi_3 \widehat{p}_t, \tag{48}$$

where

$$\pi_1 = \frac{\sigma_\varepsilon^2 \sigma_\delta^2}{\sigma_\varepsilon^2\sigma_\delta^2 + \sigma_\eta^2\sigma_\delta^2 + \sigma_\eta^2\sigma_\varepsilon^2} \tag{49}$$

$$\pi_2 = \frac{\sigma_\eta^2 \sigma_\delta^2}{\sigma_\varepsilon^2\sigma_\delta^2 + \sigma_\eta^2\sigma_\delta^2 + \sigma_\eta^2\sigma_\varepsilon^2}, \tag{50}$$

$$\pi_3 = \frac{\sigma_\eta^2 \sigma_\varepsilon^2}{\sigma_\varepsilon^2\sigma_\delta^2 + \sigma_\eta^2\sigma_\delta^2 + \sigma_\eta^2\sigma_\varepsilon^2}. \tag{51}$$

Notice that if $\sigma_\delta^2 \to \infty$ (for example, because $\sigma_a^2 \to \infty$, i.e., the preference shock has a very large variance), then

$$\pi_1 \to \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_\eta^2} = 1 - \lambda, \quad \pi_2 \to \frac{\sigma_\eta^2}{\sigma_\varepsilon^2 + \sigma_\eta^2} = \lambda \quad \text{and} \quad \pi_3 \to 0.$$

In other words, agents have nothing to learn from the equilibrium price, the weights used for inferring the unobservable aggregate fundamental are the same as in Section 5. Conversely, if $\sigma_\delta^2 \leq \infty$, then $\pi_2 < \lambda$, i.e., the equilibrium price conveys useful information and agents put less weight on their private signals.

The equilibrium price

To solve for the equilibrium price, we follow the same steps as in Appendix III. By guessing that $\widetilde{E}_t p_{t+1} = \phi\pi_2 i_t$, we have

$$\begin{aligned}
E_t^j p_{t+1} &= \frac{r}{1+r-\rho} E_t^j \theta_{t+1} + \frac{\phi\pi_2}{r} I, \\
\overline{E}_t p_{t+1} &= \frac{r}{1+r-\rho} \overline{E}_t \theta_{t+1} + \frac{\phi\pi_2 I}{r}, \\
\widetilde{E}_t p_{t+1} &= \frac{r}{1+r-\rho} \widetilde{E}_t \theta_{t+1}.
\end{aligned}$$

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

$$\begin{aligned}
\overline{E}_t p_{t+1} &= \phi\left(\rho\theta_{t-1} + \pi_2\eta_t + \pi_3\widehat{p}_t\right) + \frac{\phi\pi_2 I}{r}, \\
\widetilde{E}_t p_{t+1} &= \phi\pi_2 i_t.
\end{aligned}$$

The second line confirms the claim that $\widetilde{E}_t p_{t+1} = \phi\pi_2 i_t$. Inserting $\overline{E}_t p_{t+1}$ and $\widetilde{E}_t p_{t+1}$ in (16) now, the equilibrium price becomes

$$p_t = \frac{r}{1+r}\left(\rho\theta_{t-1} + \eta_t + a_t\right) + \frac{1}{1+r}\left(\phi\rho\theta_{t-1} + \phi\pi_2\eta_t + \phi\pi_3\widehat{p}_t + \frac{\phi\pi_2 I}{r}\right) + \frac{\phi\pi_2 i_t}{1+r}$$

from which it follows,

$$p_t = \frac{\frac{\phi}{1+r}\left(\frac{\pi_2 I}{r} - \frac{\pi_3 b_0}{b_\eta}\right) + \frac{r+\phi-\phi\pi_3\frac{b_\theta}{b_\eta}}{1+r}\rho\theta_{t-1} + \frac{r+\phi\pi_2}{1+r}\eta_t + \frac{r}{1+r}a_t + \frac{\phi\pi_2 i_t}{1+r}}{1-\frac{\phi\pi_3}{(1+r)b_\eta}}$$

The undetermined coefficients can, therefore, be written as

$$\begin{aligned}
b_0 &= \tfrac{\phi\pi_2}{r(1+r)}I \\
b_\theta &= \tfrac{r+\phi}{1+r} \\
b_\eta &= \tfrac{r+\phi(\pi_2+\pi_3)}{1+r} \\
b_a &= \tfrac{r}{1+r}\left(1+\tfrac{\phi\pi_3}{r+\phi\pi_2}\right) \\
b_i &= \tfrac{\phi\pi_2}{1+r}\left(1+\tfrac{\phi\pi_3}{r+\phi\pi_2}\right)
\end{aligned}$$

and the equilibrium price as,

$$\begin{aligned}
p_t \quad = \quad & \frac{\phi\pi_2}{r\,(1+r)}I + \frac{r+\phi}{1+r}\rho\theta_{t-1} + \frac{r+\phi\,(\pi_2+\pi_3)}{1+r}\eta_t \\
& + \frac{r}{1+r}\left(1+\frac{\phi\pi_3}{r+\phi\pi_2}\right)a_t + \frac{\phi\pi_2}{1+r}\left(1+\frac{\phi\pi_3}{r+\phi\pi_2}\right)i_t.
\end{aligned}$$

or, after some manipulation, as

$$p_t = p_t^* + \pi_2\Upsilon_t + \pi_3\Phi_t.$$

As in Section 4 and 5, p_t^* denotes the fundamental price, and Υ_t measures the degree of dispersion in beliefs. The new term,

$$\Phi_t \equiv \frac{\phi}{1+r}\left(\theta_t - \rho\theta_{t-1}\right) + \frac{r\phi}{(1+r)\,(r+\phi\pi_2)}a_t + \frac{\phi^2\pi_2}{(1+r)\,(r+\phi\pi_2)}i_t,$$

captures, instead, the degree of magnification of shocks induced by the the process of learning from price.

Finally, since

$$\sigma_\delta^2 = \left(\frac{b_a}{b_\eta}\right)^2\sigma_a^2 + \left(\frac{b_i}{b_\eta}\right)^2\sigma_i^2, \tag{52}$$

π_1, π_2 and π_3 are functions of σ_δ^2, which, in turn, depend on b_η, b_a and b_i. To pin down these undetermined coefficients, it is thus necessary to use equations (50), (51) and (52). This leads to

$$b_\eta = \frac{r}{1+r} + \frac{\phi}{1+r}\left(\frac{\sigma_\eta^2\left(\frac{b_a^2\sigma_a^2+b_i^2\sigma_i^2}{b_\eta^2}\right)+\sigma_\eta^2\sigma_\varepsilon^2}{\left(\sigma_\varepsilon^2+\sigma_\eta^2\right)\left(\frac{b_a^2\sigma_a^2+b_i^2\sigma_i^2}{b_\eta^2}\right)+\sigma_\eta^2\sigma_\varepsilon^2}\right),$$

$$\frac{b_i}{b_a} = \frac{\phi\pi_2}{r} = \frac{\phi}{r}\left(\frac{\sigma_\eta^2\left(\frac{b_a^2\sigma_a^2+b_i^2\sigma_i^2}{b_\eta^2}\right)}{\left(\sigma_\varepsilon^2+\sigma_\eta^2\right)\left(\frac{b_a^2\sigma_a^2+b_i^2\sigma_i^2}{b_\eta^2}\right)+\sigma_\eta^2\sigma_\varepsilon^2}\right),$$

and

$$b_\eta = b_a + b_i,$$

which define a system of three equations in the three unknowns, b_η, b_a and b_i. Unfortunately, this system of equations does not admit closed-form solutions. However, numerical values can easily be computed.

Proof of Corollary 1

The mean equilibrium price with learning (26) is strictly smaller than the one without learning (23) if

$$(\lambda - \pi_2)\, E\Upsilon_t > \pi_3 E\Phi_t.$$

Using the definitions of λ, π_2 and π_3, and the fact that $Ea_t = 0$ and $Ei_t = I$, this inequality can be written as

$$\frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\varepsilon^2} - \frac{\sigma_\eta^2 \sigma_\delta^2}{\sigma_\varepsilon^2 \sigma_\delta^2 + \sigma_\eta^2 \sigma_\delta^2 + \sigma_\eta^2 \sigma_\varepsilon^2} > \frac{\sigma_\eta^2 \sigma_\varepsilon^2}{\sigma_\varepsilon^2 \sigma_\delta^2 + \sigma_\eta^2 \sigma_\delta^2 + \sigma_\eta^2 \sigma_\varepsilon^2} \frac{r\phi\pi_2}{(1+r)\,(r+\phi\pi_2)},$$

or

$$\frac{\sigma_\eta^2 \sigma_\varepsilon^2 \sigma_\eta^2}{\sigma_\eta^2 + \sigma_\varepsilon^2} > \sigma_\eta^2 \sigma_\varepsilon^2 \frac{r\phi\pi_2}{(1+r)\,(r+\phi\pi_2)},$$

which is equivalent to

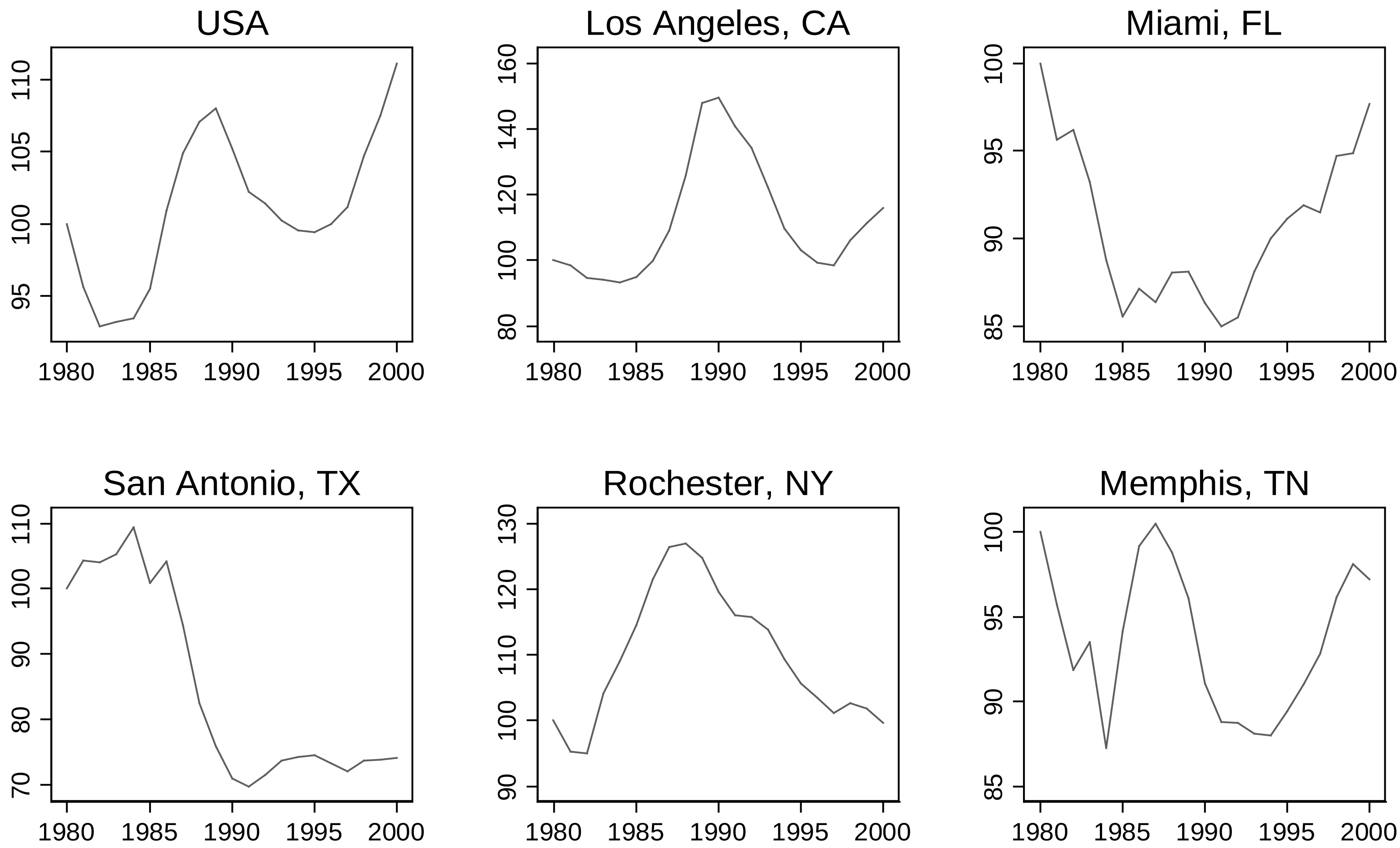
$$\frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\varepsilon^2} \equiv \lambda > \frac{r\phi\pi_2}{(1+r)\,(r+\phi\pi_2)} = \frac{r}{(1+r)\left(1 + \frac{r}{\phi\pi_2}\right)}.$$

Since the expression on the RHS of this inequality is maximized at $r^2 = \phi\pi_2$, it is sufficient to show that

$$\lambda > \frac{r^2}{(1+r)\left(r + \frac{r^2}{\phi\pi_2}\right)} = \frac{(\phi\pi_2)^2}{(1+r)^2},$$

which is always true since $\phi < 1$ and $\lambda > \pi_2$.

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

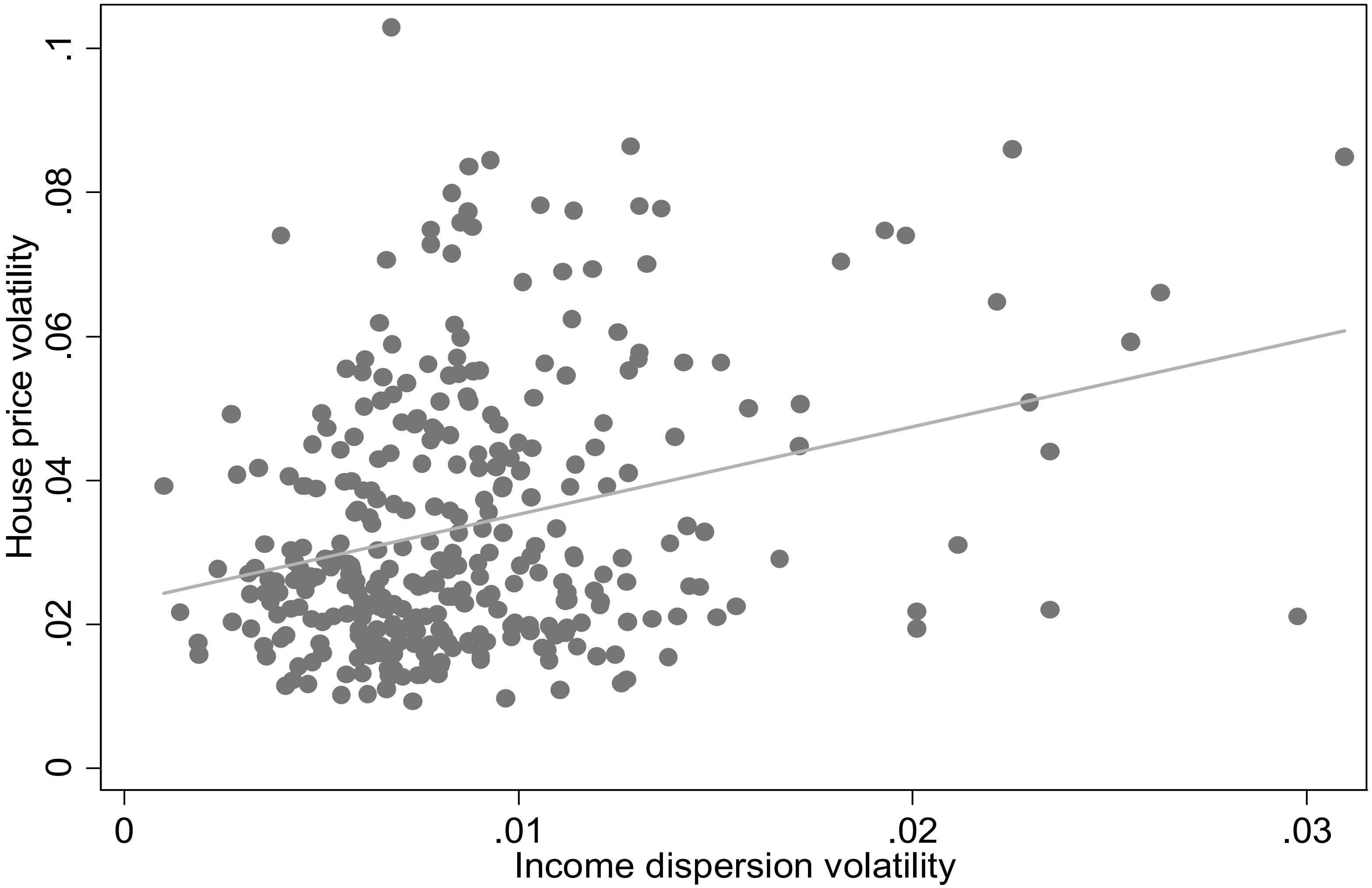


Table 1. Description of variables and data sources

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

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.

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

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.

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

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 incone 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.

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

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.

	Dependent Variables	
	Volatility of house price	
	(1)	(2)
Volatility of dispersion	1.219***	0.974***
	(0.256)	(0.319)
Volatility of income		0.210
		(0.154)
Observations	331	331
R2	0.08	0.09