

NORTHWESTERN UNIVERSITY



Guthrie Center for Real Estate Research

An Extrapolative Model of House Price Dynamics

Charles G. Nathanson
Kellogg School of Management
Northwestern University
nathanson@kellogg.northwestern.edu

Edward L. Glaeser
Department of Economics
Harvard University
eglaeser@harvard.edu

March 2015

10004

An Extrapolative Model of House Price Dynamics*

Edward L. Glaeser

Department of Economics

Harvard University

eglaeser@harvard.edu

Charles G. Nathanson

Kellogg School of Management

Northwestern University

nathanson@kellogg.northwestern.edu

March 2015

Abstract

A modest approximation by homebuyers leads house prices to display three features that are present in the data but usually missing from perfectly rational models: momentum at one-year horizons, mean reversion at five-year horizons, and excess longer-term volatility relative to fundamentals. Valuing a house involves forecasting the current and future demand to live in the surrounding area. Buyers forecast using past transaction prices. Approximating buyers do not adjust for the expectations of past buyers, and instead assume that past prices reflect only contemporaneous demand, as with a capitalization rate formula. Consistent with survey evidence, this approximation leads buyers to expect increases in the market value of their homes after recent house price increases, to fail to anticipate the price busts that follow booms, and to be overconfident in their assessments of the housing market.

*First draft: November 2014. We thank Ian Dew-Becker, David Levine, Giacomo Ponzetto and seminar participants at Kellogg, UPF/CREI, and EUI for helpful comments, and Nina Tobio and Aidan McLoughlin for excellent research assistance. Nathanson thanks the Guthrie Center for Real Estate Research for financial support.

Metropolitan area housing prices display significant momentum (Case and Shiller, 1989), mean reversion (Cutler, Poterba and Summers, 1991), and excess variance relative to fundamentals (Glaeser et al., 2014). These features were spectacularly on display during the great housing convulsion that rocked the U.S., and the world, between 1996 and 2010. Yet these three phenomena characterized house price dynamics even before this episode, and continue to do so afterward. Case and Shiller’s seminal work on house price momentum was published in 1989, and Glaeser et al. (2014) document mean reversion and excess volatility in house prices between 1980 and 2003. A successful model of house price dynamics must therefore predict momentum, mean reversion, and excess volatility in steady state, and not just during periods of extraordinary turbulence.

A large literature has tried, with incomplete success, to accomplish this task with rational models of the housing market. Rationality is difficult to square with the strong predictability of house prices, as rational homebuyers should foresee predictable price changes and alter their bids quickly to arbitrage predictability away. Search models such as Head, Lloyd-Ellis and Sun (2014) and Guren (2014), in which prices fail to be Walrasian, introduce sluggish adjustment into the housing market, which produces momentum in price changes. Yet these models typically have more trouble explaining either excess volatility or longer-term mean reversion. Glaeser et al. (2014) show that longer-term mean reversion in fundamentals can explain the mean reversion observed in prices, but the momentum in fundamentals cannot explain the momentum in prices.

In this paper, we present a simple model of housing price formation that can fit these facts in which buyers use an approximation rather than fully fathoming the beliefs of past buyers. Valuing a home involves forecasting future house prices, as buyers expect to resell at some future point. To forecast prices, buyers look at past prices. This extrapolation, on its own, is perfectly rational and reasonable. Demand to live in the city determines prices, so buyers may learn about the state of housing demand by studying past prices. Both academics and real estate practitioners commonly use house prices as a measure of demand for living in an area, and use past prices to forecast future movements.

This sequential inference is difficult to do correctly. An economy of rational homebuyers will filter demand perfectly out of the history of prices (Proposition 1). But this filtering depends on an immense amount of coordination among the buyers. They all must use the same unintuitive formula to map the price history to current demand. Under some parameters, this formula puts *exponentially increasing weights on past prices* (Proposition 3). This bizarre formula works as long

as all past buyers used it as well. It fails spectacularly when they did not. The cognitive difficulty in calculating this formula, as well as its lack of robustness, leads to the consideration of alternate inference rules.

We follow Eyster and Rabin (2010) and assume that individuals are imperfect at inferring the belief processes of others. Following their terminology, homebuyers are “naive,” meaning that while they rationally calculate the correct price given their beliefs about demand and demand growth, they do not rationally infer how past prices were formed. They neglect that past buyers, like themselves, also used prices to update their beliefs. Instead, naive buyers simply think that past prices provide direct estimates of housing demand. Our inference rule is also an application of the cognitive hierarchy model of Camerer, Ho and Chong (2004); naive homebuyers are “level-1 thinkers.” Eyster, Rabin and Vayanos (2013) model similarly naive investors in financial markets. In their case, traders ignore the information implied by current prices, and in our case, traders use an approximation that leads to a misunderstanding of past prices.¹

One interpretation of naive inference is that buyers just make a mistake, but our preferred interpretation is that they are making a convenient approximation. Buyers know that on average, prices equal fundamentals divided by the interest rate. Naive buyers apply this approximation to compute demand from past prices. The approximation is accurate for deducing the general level of demand from house prices, but causes errors when used sequentially by buyers. As in Gabaix (2014)’s sparse reasoning theory, buyers err by using an approximation well-suited to one environment in a different setting.

This approximation seems plausible to us, especially since like many housing economists, we have been guilty of it ourselves. Somewhat surprisingly, it radically shifts the model’s predicted house price dynamics, and generates extrapolation, as in Barberis et al. (2015). Naive homebuyers infer the path of fundamentals from past price changes. If housing prices grew by five percent annually over a five year period, they infer that this change represents the underlying growth rate in fundamentals. The mistake is that this five year growth also reflects changes in beliefs about fundamentals, which means that the actual level of demand is much smaller. As a result, a positive demand shock leads to a wave of sequential upward revisions in beliefs, causing momentum in prices and eventual overshooting. During a boom, naive homebuyers overestimate housing demand, and

¹Adam, Beutel and Marcet (2014) study stock price dynamics when traders have extrapolative expectations of future stock prices. These extrapolative expectations are similar to those that result from naive inference, but they are motivated in a different way.

during a bust, they underestimate it.

We incorporate naive inference into a continuous time model of house prices. In this model, growth rates mean revert and the current growth rate is not directly observable. Importantly, while individuals observe their own level of demand, they do not observe the current level of market demand directly and must also infer that through transaction prices. We demonstrate analytically that if buyers only have limited access to past transactions (two periods) then a sinusoidal relationship between price changes results (Proposition 5). Prices load positively on the most recent lag, as naive buyers infer the level of demand from this lag, and negatively on the second lag, as a lower second lag signals a higher growth rate. Precisely this time series relationship leads to short-term momentum and long-term mean reversion (Lemma 4).

We then calibrate the model to see if the quantitative predictions match the data. Many of our parameters come from existing literature, and in some cases, we estimate parameters ourselves. When buyers are all rational, even a strong persistence of demand growth rates generates little positive momentum in housing price changes. With naive homebuyers, the model closely matches the empirical value of one- and two-period auto-correlations in housing prices. At lower frequencies, our model also matches mean reversion when buyers are naive, but fails if buyers are rational. The naive model generates over-shooting in response to an impulse response, and variance in excess of fundamentals, especially over longer time periods. We conduct sensitivity analysis to understand which parameter values are needed to generate these results. Growth rate persistence is most important, and the necessary value can be calibrated from data on metropolitan area income and rents.

We ask whether the naive inference can explain the patterns in home buyer beliefs documented by Case, Shiller and Thompson (2012). Homeowners appear to forecast significant increases in market values after recent price growth. This type of extrapolation is at odds with full rationality, as an economy of rational homeowners should realize that demand shocks are fully reflected in house prices almost immediately. Naive buyers, in contrast, do extrapolate, as they fail to understand the evolving beliefs of future buyers. Significantly, this failure leads naive buyers to *under*-extrapolate future price increases from past ones, a result that Case, Shiller and Thompson (2012) document. Our simulations match the quantitative extent of extrapolation that appears in survey data.

We find that reasonable parameter estimates yield predictions about momentum, mean reversion and volatility of prices that match the data reasonably well. The model also shows how these

predictions change as we change different parameter values. One of the more interesting findings is that the bubble-like features of markets disappear when information is either too good or too bad. If buyers have highly accurate direct signals about the state of demand, then momentum, mean reversion and excess volatility disappear. But these features also disappear if buyers have access to relatively limited data on the number of past housing transactions. The most extreme fluctuations occur when buyers have relatively good data about past prices, but limited data on the underlying fundamentals.

Our approach builds on a number of non-rational theories of momentum and mean reversion in housing, such as Piazzesi and Schneider (2009). Their work follows a long established “noise trader” tradition in finance of perturbing rational models by assuming irrationality for a small number of agents and then examining how those small number of agents shape prices (De Long et al., 1990). Barberis et al. (2015), for example, continue this tradition by examining the impact of modest numbers of extrapolators in formal asset markets. Hong and Stein (1999) present a model in which momentum caused by inattentive investors leads to mean reversion resulting from arbitrageurs.

Yet we think that it is easier to understand housing markets as being driven by small irrationality from the many rather than major irrationality from the few. It seems incorrect to view housing markets in 2004-2006 as being dominated by a small number of highly irrational investors. Millions of Americans bought homes during that time period. Polls of homebuyers during booms (Case, Shiller and Thompson, 2012) suggest that beliefs about perpetually high rates of future price appreciation were quite widespread. Furthermore, arbitrage by outside investors is notoriously difficult in the housing market (Glaeser and Gyourko, 2009), limiting the applicability of standard finance models to this context.

A voluminous literature in finance has argued that movements in discount rates can explain the excess volatility of asset prices relative to fundamentals. This literature is surveyed by Cochrane (2011). Discount rate variation provides a fully rational explanation of why high price-dividend ratios predict low returns.² This paper explains how homeowners may consistently overestimate returns when prices are high and underestimate them when prices are low. Consistent errors seem

²In our model, housing prices are driven by beliefs about fundamentals, rather than an independent (near) rational bubble. We made this choice because with endogenous supply, rational bubbles predict a positive probability of extreme long-run prices, which would generate extreme long-run investment and city size. Giglio, Maggiori and Stroebl (2014) compare long-run leases and prices to provide significant evidence that the transversality condition really does hold for housing, making rational bubbles impossible. Yet Adam, Kuang and Marcet (2011) have shown how almost-rational bubbles can generate impressive swings in prices and strongly link prices and interest rates.

more likely in the housing market than in other asset markets where arbitrage is easier and trade is more frequent.

All work on semi-rationality is troubled by Tolstoy’s corollary: there is only one way to make correct inferences but an uncountable number of ways to get things wrong. Naive extrapolation is one possible model of semi-rationality in housing markets. Many other forms of irrationality may exist, and it may be possible to discover a rational model that can reconcile all the facts. Yet it is remarkable that this relatively modest deviation from rationality predicts outcomes so much closer to reality than the standard rational model.

1 A Model of House Price Determination

1.1 Housing Market Fundamentals

We consider the choice of an individual who is deciding whether or not to purchase a home. This person is matched with one house, and if she purchases the house, she receives a flow of utility improvement of $D_{i,t}$ relative to her next best alternative. This flow utility can be interpreted as the overall benefit of living in the city relative to a reservation locale, but in that case, we must also assume that the opportunity to buy in the city is a once-in-a-lifetime chance. Alternatively, the reservation utility could include the opportunity of buying again in the city, but this makes interpretation slightly more difficult. The supply of housing is fixed.³

This overall utility combines an idiosyncratic element a_i with a city-specific component D_t :

$$D_{i,t} = D_t + a_i. \tag{1}$$

These elements include both the labor market returns and utility-related benefits from living in the city. The idiosyncratic component is drawn independently for each individual from a normal

³The interaction between uncertainty and housing supply has been explored elsewhere. Glaeser, Gyourko and Saiz (2008) examine the link between belief-based bubbles and housing supply. Nathanson and Zwick (2014) go further and show that in land markets, which can be dominated by small numbers of professional buyers, bubbles can appear more readily than in housing markets, in which ownership is far more dispersed. Gao, Sockin and Xiong (2014) is similar to this paper as they also focus on the process of learning about housing demand from prices, but their model is fully rational and they focus on the interaction between housing supply and belief aggregation.

distribution with mean 0 and standard deviation σ_a . The city-specific component of utility follows

$$dD = gdt + \sigma_D dW^D, \quad (2)$$

where W^D is a standard Wiener process and g is a stochastic process we define shortly.

In this specification, changes to city-level demand persist over time. Such persistence has some empirical basis in the fact that Gibrat's law seems to hold for metropolitan areas (Glaeser, Scheinkman and Shleifer, 1995; Eaton and Eckstein, 1997; Gabaix, 1999). There is correlation between past success and the future population or employment growth of the city. Typically, there has been mean reversion of incomes at the city level, but even that fact has declined over time (Berry and Glaeser, 2005; Ganong and Shoag, 2013). The absence of city-level mean reversion is for convenience, and the model could easily encompass this feature while leaving the results unchanged. The critical assumption is that growth rates shift over time and mean revert, so that

$$dg = -\lambda gdt + \sigma_g dW^g, \quad (3)$$

where W^g is a standard Wiener process that is independent from W^D . If growth rates were constant, then they would eventually be known, and the learning about growth that is a crucial element in the model would disappear. If growth rates did not mean revert, but instead followed a random walk, then the price dynamics would become too explosive and yield none of the mean reversion that we see in the data.

The persistence of growth rates is empirically debatable, and depends on how demand at the city level is measured. Head, Lloyd-Ellis and Sun (2014) find a correlation of 0.27 between annual income growth and lagged income growth at the city level. In our empirical work below, we find a larger correlation of 0.7 using metropolitan area rents as the proxy for demand. This persistence is certainly a necessary feature of our model, and we will show what the model implies when the persistence of growth rates is quite small.

The market assumption is that each transaction involves exactly one buyer and one seller, and the buyer pays a price that makes her indifferent between owning the house or not. The seller's willingness to accept is irrelevant. We do not mean to suggest that this is a realistic model of housing markets, in which most homes have multiple prospective buyers and most buyers consider

a number of homes. The role that the bargaining process can play in shaping housing dynamics has been examined elsewhere (Anenberg and Bayer, 2013; Guren, 2014) and we are interested in particularly examining the role of non-standard beliefs. As such, we have chosen a particularly simple market structure in order to focus how the inference process can shape demand volatility and the persistence of price movements.

Individuals remain in their homes unless they receive a shock that forces them to move. These exogenous mobility shocks are Poisson and arrive at a rate μ . Mobility shocks are the easiest way to generate resale and an interest by buyers in future prices. We agree strongly that an endogenous resale model would be more realistic. Moreover, the fixed sale assumption is compatible with the earlier assumption that prices are determined by the willingness to pay of the buyer, not any aspect of the seller.

We define $p_{i,t}$ to be the price of a house transacted at t to buyer i , and p_t to be the average price of all houses sold at t . The discounted value of owning the house and the willingness to pay therefore equal

$$p_{i,t} = \mathbf{E}_{i,t} \left[\int_t^T e^{-r(\tau-t)} D_{i,\tau} d\tau + e^{-r(T-t)} p_T \mid T - t \sim \text{Poisson}(\mu) \right], \quad (4)$$

where r is the discount rate. Price formation depends on each buyer's expectations about future prices, and therefore on each buyer's expectations of future buyers' expectations. When this expectation forecasting takes a specific form, which is general enough to encompass both the rational and naive rules we specify later, prices follow a simple linear structure:

Lemma 1. *Suppose that for all $T \geq t$, $\mathbf{E}_{i,t} \mathbf{E}_T D_T = \mathbf{E}_{i,t} D_T$ and $\mathbf{E}_{i,t} \mathbf{E}_T g_T = \phi_g(T - t) \mathbf{E}_{i,t} g_t$ for some function ϕ_g , where \mathbf{E}_T denotes the average expectation among buyers at T . An equilibrium for the average price of the houses transacted at t is given by*

$$p_t = \frac{1}{r} \left(\frac{r}{r + \mu} D_t^a + \frac{\mu}{r + \mu} \widehat{D}_t \right) + A_g \widehat{g}_t, \quad (5)$$

where D_t^a equals the average flow utility $D_{i,t}$ among buyers at t , \widehat{D}_t and \widehat{g}_t are their average beliefs about the current values of city demand D and its growth rate g , and A_g is a constant.

To determine the expectations in the pricing formula, we now discuss each buyer's information set.

1.2 Information Available to Buyers

At time t , the buyer knows the current flow utility $D_{i,t}$ she receives from the house to which she is matched. She also observes a history of transaction prices of houses in the city. In particular, she learns the average price $p_{t'}$ every δ units of time before her purchase, that is, for $t' = t - \delta, t - 2\delta$, and so forth. The number of sales in each average equals N . This transaction history corresponds to a price index in the case N is large, or to simply a list of all transactions when $N = 1$. We derive some analytic results in the limit as $N \rightarrow \infty$, but allow for finite N in the quantitative exercises. We denote the set of observed prices by $\Omega_t^p = \{p_{t-m\delta} \mid m \in \mathbb{N}\}$, where \mathbb{N} is the set of positive integers.

It is natural that the buyer would know her own flow utility $D_{i,t}$ at t . After all, this is the utility the buyer currently receives from living in the city. Knowing past prices is also reasonable, as house prices are readily available from a number of sources. Homebuyers frequently examine the sale prices of similar homes before making a purchase. This practice, called “comparable analysis” or “comps,” is the foundation for property appraisals.

The buyer cannot directly observe the city demand D_t or its growth rate g_t . City-wide demand D_t is an aggregation of private information $D_{i,t}$. Less obvious is why the buyer cannot observe g_t , given that this growth rate directly affects the flow utility the buyer will receive. The intuition here is that the buyer’s utility rises and falls with the quality of the city, either through amenities or the labor market. As these fluctuations involve city-wide forces beyond the buyer’s control, it is reasonable to suppose that the buyer possesses no private information about the current growth rate.

In principle, a buyer could obtain information on D_t from looking at various economic indicators about the city. To permit this possibility, we allow buyers to observe noisy signals D_t^s of demand, where $D_t^s = D_t + s_t$ and s_t is an independently drawn normal error with standard deviation σ_s . Buyers have access to this information at the same frequency with which they observe prices: the set of signals known at t equals $\Omega_t^s = \{D_{t-m\delta}^s \mid m \in \mathbb{N}\}$.

In addition to $\{D_{i,t}\}$, Ω_t^p , and Ω_t^s , the buyer at t also observes stochastically revealed direct observations of the true state of demand. These observations occur at a set of times \mathcal{T} that are realizations of a continuous-time Poisson process with parameter $\rho > 0$, where ρ is small. That is, given realizations on \mathcal{T} up to some time, the cumulative distribution function for the time Δt until

the next realization is $1 - e^{-\rho\Delta t}$. We denote $\mathbf{x}_t = (D_t, g_t)'$ to be the state of demand at t . The buyer at t observes $\Omega_t^{\mathbf{x}} = \{\mathbf{x}_{t'} \mid t' \in \mathcal{T} \text{ and } t' \leq t\}$. This rare revelation of the true state of the world has important consequences for the uniqueness of equilibrium when buyers are rational. As long as $\rho > 0$, with probability 1 there exists some time in the past when the state of demand was revealed. Rational buyers begin at that state and then infer all of the demand shocks since then. This process of rational inference occurs anytime history has a beginning.

The complete information set for the buyer is

$$\Omega_{i,t} = \{D_{i,t}\} \cup \Omega_t^p \cup \Omega_t^a \cup \Omega_t^{\mathbf{x}}. \quad (6)$$

1.3 Inference about Demand

The buyer's inference problem is to use the data in $\Omega_{i,t}$ to infer the value of market demand D_t and its growth rate g_t . The best a buyer can do is to extract all the data that has been directly observed by buyers before t . In addition to the signals in Ω_t^s , this information includes all individual flow utility $D_{i',t'}$. Due to the normality assumptions, a sufficient statistic for the distribution of buyer flow utility at t' is $D_{t'}^a$, the average flow utility across the N buyers at that time. We denote $\Omega_t^a = \{D_{t-m\delta}^a \mid m \in \mathbb{N}\}$.

When buyers are rational and the rationality of all buyers is common knowledge, observing the infinite history of prices allows the buyers to know the history Ω_t^a :

Proposition 1. *Suppose it is common knowledge among buyers at all times that information sets take the form given in (6). Then each buyer can perfectly deduce the history Ω_t^a of average buyer flow utility.*

Given this proposition, inference for the rational buyers involves a standard signal extraction problem where D_t and g_t are inferred from the series of past noisy observations of D in Ω_t^a and Ω_t^s , as well as the noisy observation of D_t given by private utility $D_{i,t}$.

The proof of Proposition 1 goes as follows. Let $t' < t$ be a time at which buyers at t observe house sales. Conditional on all house prices and demand signals before t' , the price at t' is a strictly increasing function of average flow utility $D_{t'}^a$, as higher flow utility directly increases the pricing equation (5) and also increases the posterior means $\widehat{D}_{t'}$ and $\widehat{g}_{t'}$. Because the buyer at t observes all prices and demand signals before t' , she exactly infers $D_{t'}^a$ from observing the transacted price $p_{t'}$.

As this proof makes clear, deducing Ω_t^a requires a fairly hefty cognitive load on the part of the buyers, as they need to infer everything about beliefs in the past. As such, we introduce a second possibility: buyers believe that past buyers used the pricing formula $p_{i,t'} = D_{i,t'}/r$, which makes inference quite straightforward. Simply multiplying the price $p_{t'}$ by the constant r yields $D_{t'}^a$, and then the inference proceeds as in the rational case. We call this procedure *naive inference*.

Naive inference can be motivated as an approximation that results from inattention. Consider the pricing formula delivered by Lemma 1 evaluated at some past time t' . The long-run average of g is 0, so if the expectation $\widehat{g}_{t'}$ is unbiased, it equals 0 on average. Similarly, if $\widehat{D}_{t'}$ is unbiased it equals $D_{t'}^a$ on average. Therefore, as long as the conditions of Lemma 1 hold, the past price can be written as $p_{t'} = D_{t'}^a/r + \xi_{t'}$, where $\xi_{t'}$ is mean 0 measurement error. The quantity $rp_{t'}$ is an unbiased estimate for $D_{t'}^a$, and this is the estimate naive buyers use.

The problem with using $rp_{t'}$ to estimate $D_{t'}^a$ is that the measurement error $\xi_{t'}$ is correlated across time. This serial correlation arises from the non-independence of demand forecasts. If buyers believe that the level or growth rate of demand is higher than its true value or long-run average today, they are likely to believe this tomorrow as well. Proper Bayesians would recognize the dependence across time in the measurement error, but naive buyers do not. Naive inference is quite good for estimating demand from a single observation of house prices, but fails when estimating demand using a series of prices. Ignoring the serial correlation in $\xi_{t'}$ is a form of inattention that results from a procedure that is rational in a different context. This sort of inattention is studied in Gabaix (2014).

A complementary interpretation for naive inference is as the result of a simplified problem for individuals who lack the cognitive ability to make inferences about inferences. Buyers avoid this recursion by replacing each prior buyer's expectation by its expected value, i.e. $\widehat{D}_{t'}$ with $D_{t'}^a$ and $\widehat{g}_{t'}$ with 0. In so doing, naive buyers assume that past buyers were "simple," and used just their private information $\{D_{i,t'}\}$ to draw inferences rather than the full information set $\Omega_{i,t'}$. The idea that economic agents might assign this type of simplicity to others with whom they interact has been extensively explored by Eyster and Rabin (2005, 2010, 2014), and naive inference can be seen as an application of their work to financial markets. Throughout the paper, we focus on the two cases of hyper-rational homebuyers and naive homebuyers. In theory, it may be desirable to create a mixture of the two, and consider markets in which there are both types of buyers.

We now solve for the posteriors on D_t and g_t of rational and naive buyers. The state of demand

for the city at t' can be summarized by a 2×1 vector $\mathbf{x}_{t'} = (D_{t'}, g_{t'})'$. Given the laws of motion in (2) and (3), this state vector evolves linearly with normal noise. Over δ , the discrete length of time in between sales, this state vector changes according to the rule $\mathbf{x}_{t'+\delta} = \mathbf{F}\mathbf{x}_{t'} + \mathbf{w}_{t'}$, where $\mathbf{F} = \begin{pmatrix} 1 & (1-e^{-\delta\lambda})/\lambda \\ 0 & e^{-\delta\lambda} \end{pmatrix}$ and $\mathbf{w}_{t'}$ is identically distributed mean zero normal noise with covariance matrix \mathbf{Q} that is independent across time and from $\mathbf{x}_{t'}$. The news at t' consists of $D_{t'}^a$ and $D_{t'}^s$. We write this news as $\mathbf{H}_0\mathbf{x}_{t'} + \mathbf{v}_{t'}$, where $\mathbf{H}_0 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $\mathbf{v}_{t'}$ is normal mean zero noise with covariance $\mathbf{R}_0 = \begin{pmatrix} \sigma_a^2/N & 0 \\ 0 & \sigma_s^2 \end{pmatrix}$. At t , the news consists of just $D_{i,t}$, which we write as $\mathbf{H}\mathbf{x}_t + \mathbf{v}_t$, where $\mathbf{H} = (1, 0)$. The variance of \mathbf{v}_t is $\mathbf{R} = \sigma_a^2$.

The linear evolution of state variables as well as the normal structure of all noise allows the buyers to use a Kalman filter to derive the optimal posterior $\mathbf{x}_t | \Omega_{i,t}$. The resulting average of the posteriors across buyers at t gives $\hat{\mathbf{x}}_t = (\widehat{D}_t, \widehat{g}_t)'$, the belief terms that appear in the pricing function in (5). Lemma 2 solves for these posterior averages, as well as the covariance of each buyer's posterior.

Lemma 2. *Let $\mathbf{x}_t = (D_t, g_t)'$ denote the state of housing demand at t . For both rational and naive buyers, the posterior $\mathbf{x}_t | \Omega_{i,t}$ is a multivariate normal with the same covariance. As $\rho \rightarrow 0$, the mean of this posterior for rational buyers converges almost surely to*

$$\hat{\mathbf{x}}_t = \mathbf{K}D_t^a + (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{F} \sum_{m=1}^{\infty} [(\mathbf{I} - \mathbf{K}_0\mathbf{H}_0)\mathbf{F}]^{m-1} \mathbf{K}_0 (D_{t-m\delta}^a, D_{t-m\delta}^s)',$$

and for naive buyers, the mean converges almost surely to

$$\hat{\mathbf{x}}_t = \mathbf{K}D_t^a + (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{F} \sum_{m=1}^{\infty} [(\mathbf{I} - \mathbf{K}_0\mathbf{H}_0)\mathbf{F}]^{m-1} \mathbf{K}_0 (rp_{t-m\delta}, D_{t-m\delta}^s)',$$

where \mathbf{K} and \mathbf{K}_0 are matrices that depend on \mathbf{F} , \mathbf{Q} , \mathbf{R} , \mathbf{R}_0 , \mathbf{H} , and \mathbf{H}_0 . The covariance of these posteriors converges almost surely to a time-independent matrix \mathbf{P} .

The two types of buyers use the same filters, but naive buyers use $rp_{t'}$ in place of the true value of $D_{t'}^a$ that rational buyers use. A corollary is that naive buyers are overconfident in their estimates of housing demand. The naive posterior $\mathbf{x}_t | \Omega_{i,t}$ is a multivariate normal with covariance \mathbf{P} . This matrix \mathbf{P} is the optimal covariance that results from the correct application of the Kalman filter. However, naive buyers do not apply the Kalman filter correctly. They use $rp_{t'}$ in place of $D_{t'}^a$, and as a result, they use some alternate linear filter without realizing it. Because the Kalman filter is

optimal among all linear filters, the resulting error covariance is larger than \mathbf{P} , and hence larger than what naive buyers think it is. This following proposition sums up this argument.

Proposition 2. *Naive homebuyers are overconfident in their estimates of housing demand. Let \mathbf{P}_n denote the covariance matrix of the naive forecast error $\mathbf{x}_t - \mathbf{E}(\mathbf{x}_t \mid \Omega_{i,t})$. Then $\mathbf{P}_n > \mathbf{P}$, where \mathbf{P} is the covariance matrix of the naive posterior $\mathbf{x}_t \mid \Omega_{i,t}$. The inequality means that $\mathbf{P}_n - \mathbf{P}$ is positive definite.*

The relationship $\mathbf{P}_n > \mathbf{P}$ means that naive buyers overestimate the precision of their estimates of the level of demand D_t and its growth rate g_t . Furthermore, naive buyers underestimate the forecast error of any linear combinations of these quantities. In particular, they are overconfident in the valuations of their homes, as the pricing equation (5) is linear in \widehat{D}_t and \widehat{g}_t .

Overconfidence limits naive buyers' attention to noisy signals about market demand. They do not appreciate the imprecision of their inferences from prices, so they demand less information, as the marginal value of information decreases with additional certainty. As a clear example, consider the inference problem when N , the number of house transactions underlying each price, goes to infinity. In this case, $D_{t'}^a \rightarrow D_{t'}$ by the law of large numbers, so the history of buyer flow utility Ω_t^a perfectly reveals the history of market demand. As a result, buyers ignore the noisy demand signals Ω_t^s . They believe all such information is already factored into house prices, which they directly observe. This response is optimal for the rational buyers, but it is a mistake for the naive buyers. Their overconfidence leads them to ignore valuable information.

Unlike naive buyers, rational buyers correctly understand the precision of their forecasts, and their posterior has the optimal covariance matrix \mathbf{P} . They are able to achieve this optimal forecast by inferring the true value of $D_{t'}$ from past prices. As described in the proof of Proposition 1, extracting Ω_t^a from Ω_t^p involves knowing exactly how previous buyers form their own expectations of market demand. One of our motivations for introducing naive inference was the complexity of this procedure. To illustrate this claim, we solve directly for the rational posterior on D_t as a function of past prices, which is what buyers directly observe. In general, this expression is quite complicated. The following proposition gives an intuitive form that holds in a special case.

Proposition 3. *Let t_0 be the last time demand was directly observed, and suppose $n \equiv \lfloor (t - t_0)/\delta \rfloor > 1$. When the number N of transactions observed each period goes to infinity, naive buyers ignore*

information about demand and extrapolate lagged demand from the most recent price:

$$\mathbf{E}(D_{t-\delta} | \Omega_t^p \cup \Omega_t^s \cup \Omega_t^x) = rp_{t-\delta}.$$

When $N \rightarrow \infty$ and growth rates do not persist ($\lambda \rightarrow \infty$), the rational buyer's posterior on the level of demand is a telescoping sum of past prices:

$$\mathbf{E}(D_{t-\delta} | \Omega_t^p \cup \Omega_t^s \cup \Omega_t^x) = \left(-\frac{\alpha}{1-\alpha}\right)^{n-1} \frac{rp_{t-n\delta} - \alpha_0 D_{t_0}}{1-\alpha_0} + \sum_{m=1}^{n-1} \left(-\frac{\alpha}{1-\alpha}\right)^{m-1} \frac{rp_{t-m\delta}}{1-\alpha},$$

where

$$\alpha = \frac{\mu}{r + \mu} \frac{\sigma_a^2}{\sigma_a^2 + \delta\sigma_D^2}$$

is a constant between 0 and 1, and $\alpha_0 = \mu[r + \mu]^{-1} \sigma_a^2 [\sigma_a^2 + (t - n\delta - t_0)\sigma_D^2]^{-1}$.

When $\lambda \rightarrow \infty$, the growth rate is irrelevant and buyers must only infer the level of demand from past prices. Even in this simple case, the rational filter takes a starkly unintuitive form. Every other past price counts negatively towards the rational estimate of current demand. Furthermore, when $\alpha > 1/2$ the weights on past prices grow exponentially. The parameter α captures the dependence of current prices on past prices. A larger $\mu/(r + \mu)$ leads buyers to care more about resale, and hence about \widehat{D}_t , and a larger $\sigma_a^2/(\sigma_a^2 + \delta\sigma_D^2)$ makes prices better signals about D_t than the buyer's idiosyncratic utility $D_{i,t}$. In contrast, naive buyers simply estimate demand using the most recent price, and in fact use this rule even when they are inferring the growth rate as well.

Eyster and Rabin (2014) study settings in which people extract information from observing the sequential actions of others. The correct action is positively correlated with the state of the world, which is only partially known. They show that hyper-rationality commonly leads to “anti-imitation.” The optimal action depends negatively on the actions of some previous people, even though all people have the same objectives. Proposition 3 provides an example of this phenomenon in the housing market. In this case, the action is the price paid for the house, and the information is the level and growth rate of housing demand.

Their work, and Proposition 3, call into question the robustness of rational updating in sequential settings. The divergent nature of the rational filter suggests that it will not work very well if previous prices were not formed by rational filters. If the rational filter indeed lacks this robustness property,

then a rational person should not use these rules if there is even a small chance that previous actors are not hyper-rational as well. Evaluating the robustness of rational filters is difficult, as all possible inference rules used by others must be considered. This paper stakes a small step in this direction by investigating the performance of the rational filter when previous buyers are actually naive. We perform this exercise in Section 3.

1.4 Price Change Forecasts

To close the model, we specify what buyers believe about the expectations of future buyers. The weight A_g on growth expectations in the pricing formula in Lemma 1 is determined by equation (4), and in turn by the buyer's expectation $\mathbf{E}_{i,t}p_T$ of future prices. This expectation, in turn, depends on the forecasts of forecasts $\mathbf{E}_{i,t}\widehat{D}_T$ and $\mathbf{E}_{i,t}\widehat{g}_T$.

A natural way to resolve these forecasts is to impose the law of iterated expectations, so that the first equals $\mathbf{E}_{i,t}D_T$ and the second $\mathbf{E}_{i,t}g_T$. Iterated expectations are consistent with the hyper-rationality we attributed to rational buyers in the inference problem. If common knowledge of rationality continues into the future, then future buyers, who have at least as much information as current buyers, should hold beliefs consistent on average with that of present buyers.

Iterated expectations are less consistent with naive inference. Naive buyers believe that past prices are given by D_t^a/r , and they reached this conclusion by assuming that other buyers' expectations equal their ex ante averages. A consistent forecast rule would assume that future prices are also given by $p_T = D_T^a/r$, which is equivalent to setting $\mathbf{E}_{i,t}\widehat{D}_T = \mathbf{E}_{i,t}D_T$ and $\mathbf{E}_{i,t}\widehat{g}_T = 0$.

We model buyer forecasts to allow for consistency with both types of inference. A buyer believes that with probability $1 - \phi$, future buyers are as sophisticated as herself, leading the law of iterated expectations to hold. With probability ϕ , future buyers are simple and base their expectations solely on private demand. This latter case results from buyers thinking that future buyers will be less sophisticated, or from our preferred explanation that current buyers choose a simple model of the behavior of others in order to lessen the cognitive load on themselves. The current buyer's forecasts hence equal $\mathbf{E}_{i,t}\widehat{D}_T = \mathbf{E}_{i,t}D_t$ and $\mathbf{E}_{i,t}\widehat{g}_T = (1 - \phi)\mathbf{E}_{i,t}g_T = (1 - \phi)e^{-\lambda(T-t)}\mathbf{E}_{i,t}g_t$. This forecasting rule satisfies the conditions of Lemma 1. Applying it and solving for A_g yields the following lemma.

Lemma 3. *The weight A_g on the growth rate expectation in the pricing formula in Lemma 1 is given by $A_g = [r(r + \lambda + \phi\mu)]^{-1}$, where ϕ denotes the perceived probability that future buyers are*

simple and do not use prices to draw inference, and $1 - \phi$ is the probability that future buyers are sophisticated enough for the law of iterated expectations to hold.

One goal of this paper is to make sense of survey evidence concerning expectations about the housing market. Much of the survey evidence on this topic is framed in terms of expectations of price growth. For instance, the Michigan Survey of Consumers asks whether the present is a “good time to buy [housing] for investment” (Burnside, Eichenbaum and Rebelo, 2014). A significant portion of respondents, around 30%, explicitly mention house prices to justify their view (Piazzesi and Schneider, 2009). Case, Shiller and Thompson (2012) elicit quantitative estimates of house price growth. They ask “How much of a change do you expect there to be in the value of your home over the next 12 months?” and “On average over the next ten years how much do you expect the value of your property to change each year?”

To match these surveys, we calculate buyers’ expectations of the *market value* of their house at current and future dates. The market value V_T of a house at time T is the expected price from selling the house at that time to a randomly selected buyer. Using the pricing function in Lemma 1, we calculate the expected market value as

$$\mathbf{E}_t V_T = \frac{\mathbf{E}_t D_T}{r} + \frac{(1 - \phi)\mathbf{E}_t g_T}{r(r + \lambda + \phi\mu)}. \quad (7)$$

Larger values of ϕ , the naivety of buyer forecasts, lead to stronger expectations about house value growth. When the law of iterated expectations holds ($\phi = 0$), buyers today believe their information is immediately incorporated into market values. As a result, they forecast very little changes in market values, even when they believe the growth rate is high. In contrast, buyers for whom $\phi = 1$ believe that growth rate news is never anticipated in market values. They therefore predict continued increases in the market value of their homes when the growth rate is high. The following proposition makes these results clear by solving for the expected change in market values as a function of the growth rate belief.

Proposition 4. *For a given belief about the growth rate, buyers expect greater increases in the market values of their home when they are more naive. The expected growth in the market value of a house equals*

$$\mathbf{E}_t(V_T - V_t) = \frac{r + \phi\lambda + \phi\mu}{r + \lambda + \phi\mu} \frac{1 - e^{-\lambda(T-t)}}{\lambda} \frac{\hat{g}_t}{r}.$$

Holding \hat{g}_t constant, this expression increases in ϕ , the perceived probability of selling to a simple buyer in the future.

Rational buyers are much more conservative than naive buyers in forecasting the growth in the market value of their homes. Indeed, the expected change when $\phi = 0$ equals the change when $\phi = 1$ times $r/(r + \lambda)$, and this fraction is always less than 1. The empirical value of this fraction falls *significantly* below 1. For instance, if the annual persistence of growth shocks is 0.3 (the value of income growth persistence at the metro-area level), then $\lambda = 1.2$; a value of $r = 0.04$ then leads to $r/(r + \lambda) = 0.03$. Under the parameters we use in Section 3, which assume more persistence in demand growth, the fraction rises to 0.07. This ratio falls well below 1 unless demand growth is very persistent.

Rational buyers believe public information gets priced into housing immediately. They therefore do not expect much growth in the market value of their homes. In contrast, naive buyers mistakenly believe that information about the growth rate never affects the market value of homes. This belief allows naive buyers to expect significant increases in market values when they perceive the growth rate of fundamentals to be high. According to the survey evidence mentioned above, expectations of changes to the market value of one's home are large and significant. A model in which all buyers use rational filtering is at odds with this empirical fact.

2 Autocorrelations of Naive House Price Changes

2.1 Requisite Time Series Properties

The goal of this paper is to explain the predictable booms and busts in house prices. These dynamics are described by the autocorrelation pattern of house price changes. In the data, annual house price changes are positively correlated at short lags (1 to 2 years), and negatively correlated at longer lags, with the autocorrelations decreasing over time. This pattern is fit by a dampened sinusoid: the autocorrelations gradually oscillate between positive and negative as their amplitude diminishes. We now show that house prices exhibit this feature when current prices depend positively on recent lags of prices, but negatively on further lags.

A time series displays sinusoidal autocorrelations when its characteristic polynomial has complex roots with absolute value exceeding 1. In this case, one of the factors of its characteristic polynomial

must be of the form $I - \beta_1 L + \beta_2 L^2$, where $(\beta_1/2)^2 < \beta_2 < 1$. If Δp_t has this property, we may write $b(L)(I - \beta_1 L + \beta_2 L^2)\Delta p_t = \mathbf{c}(L)\mathbf{z}_t$, where \mathbf{z}_t is independent and identically distributed across time (and possibly vectorial). Annual price changes then obey the equation

$$\Delta p_t = \beta_1 \Delta p_{t-\delta} - \beta_2 \Delta p_{t-2\delta} + \zeta_t, \quad (8)$$

where $\zeta_t = b(L)^{-1}\mathbf{c}(L)\mathbf{z}_t$.

If the behavior of ζ_t is unrestricted, then we cannot say much about the autocorrelations of Δp_t . However, when ζ_t is an AR(1) plus noise, β_1 measures the short-run autocorrelations of Δp_t while β_2 captures the cyclicity of price changes. The innovation ζ_t follows this form when it can be written $\zeta_t = \gamma_t + \epsilon_t$, with $\gamma_t = e^{-\delta\lambda\gamma}\gamma_{t-\delta} + \eta_t$, with each of ϵ_t and η_t independent and identically distributed over time and $\text{Cov}(\epsilon_t, \eta_t) \geq 0$. This case is of particular interest. As we show below, naive prices obey (8) with ζ_t an AR(1) plus noise. The following lemma, which is proved in the Appendix, describes the autocorrelations of Δp_t in this case.

Lemma 4. *Suppose house price changes are stationary and follow (8) with ζ_t an AR(1) plus noise. Then the correlation of price changes on once-lagged changes, given by $\text{Corr}(\Delta p_t, \Delta p_{t-\delta})$, is positive if $\beta_1 > 0$ and is strictly increasing in β_1 . If $(\beta_1/2)^2 < \beta_2 < 1$, then autocorrelations are given by*

$$\text{Corr}(\Delta p_t, \Delta p_{t-m\delta}) = A_\zeta e^{-m\delta\lambda\zeta} + A_\beta \beta_2^{m/2} \cos(m\theta + \omega),$$

where A_ζ , A_β , and ω are constants, and θ satisfies $\cos(\theta) = \beta_1/(2\sqrt{\beta_2})$.

Momentum, as measured by the autocorrelation of once-lagged price changes, is positive if $\beta_1 > 0$ and also increases with β_1 . Larger values of β_1 decrease the periodicity of further autocorrelations by lowering θ . In all three senses, β_1 captures the momentum in price changes. In contrast, β_2 measures the cyclical nature of house prices. When β_2 is high, specifically higher than $(\beta_1/2)^2$, the autocorrelations cycle, obeying the sinusoidal pattern given in the lemma. Furthermore, the higher is β_2 , the greater the amplitude of these cycles and the longer they last. The correlations dampen at the rate $\beta_2^{m/2}$, so a larger β_2 amplifies the sinusoidal nature of the autocorrelations.

2.2 Autoregressive Structure of Naive Prices

We show naive inference leads prices to obey (8) in a specific case of the model. The particular case we study imposes a number of simplifications. First, the number of home sales is large enough that prices aggregate information completely ($\sigma_a/\sqrt{N} \rightarrow 0$). Second, the individual utility noise is so large that buyers ignore their own utility and use only prices when inferring demand ($\sigma_a \rightarrow \infty$). Finally, buyers have access only to the two most recent lags of prices. This specification represents a scenario in which buyers learn about demand entirely from recent observations of a housing price index. The restricted information set in this special case is denoted $\Omega'_{i,t} = \{D_{i,t}\} \cup \{p_{t-\delta}\} \cup \{p_{t-2\delta}\}$. A naive buyer's posterior on the lagged level of demand is $\mathbf{E}(D_{t-\delta} \mid \Omega'_{i,t}) = rp_{t-\delta}$, and her estimate of the lagged growth rate equals $\mathbf{E}(g_{t-\delta} \mid \Omega'_{i,t}) = r(p_{t-\delta} - p_{t-2\delta})\lambda e^{-\delta\lambda}/(1 - e^{-\delta\lambda})$. The naive buyer simply extrapolates the level of demand from the level of prices, and the growth rate from the change in prices.

These simple formulas lead current prices to depend positively on the first lag of prices and negatively on the second lag. Higher values of $p_{t-\delta}$ increase the buyer's estimate of the level and growth rate of demand, and both of these estimates increase today's price p_t . Conversely, a higher value of $p_{t-2\delta}$ lowers the estimate of the growth rate, thereby negatively impacting p_t . As shown in Lemma 4, these relationships lead house price changes to exhibit sinusoidal autocorrelations, as long as certain additional technical conditions hold. The following proposition writes price changes in the form given by (8), and the Appendix proves that for certain parameters, the conditions in Lemma 4 are met.

Proposition 5. *Suppose naive buyers observe only the two most recent house prices. Then there exist parameters such that the autocorrelations of price changes obey the sinusoidal formula in Lemma 4. When $\sigma_a \rightarrow \infty$ and $\sigma_a/\sqrt{N} \rightarrow 0$, one-period house price changes obey the autoregressive equation*

$$\Delta p_t = \left(\frac{(1 + e^{-\delta\lambda})\mu}{r + \mu} + \frac{\lambda e^{-2\delta\lambda} r A_g}{1 - e^{-\delta\lambda}} \right) \Delta p_{t-\delta} - \left(\frac{e^{-\delta\lambda}\mu}{r + \mu} + \frac{\lambda e^{-2\delta\lambda} r A_g}{1 - e^{-\delta\lambda}} \right) \Delta p_{t-2\delta} + \frac{(1 - e^{-\delta\lambda})g_{t-\delta}}{\lambda(r + \mu)} + \frac{w_{t-\delta}^D}{r + \mu}.$$

The g and w terms constitute an AR(1) plus noise.

Proposition 5 shows that house prices fit the structure analyzed in Lemma 4 when buyers are naive. Therefore, house prices display momentum and a cyclical autocorrelation structure.

The momentum comes from the coefficient on $\Delta p_{t-\delta}$. Buyers' estimates of D_t and g_t positively influence their valuations of their homes, and both of these estimates depend positively on $p_{t-\delta}$. The cyclicalness comes from the coefficient on $\Delta p_{t-2\delta}$. This coefficient is negative because buyers' estimate of g_t depends *negatively* on $p_{t-2\delta}$. The lower this price, the higher buyers' estimate of the growth rate.

Naive inference succeeds at conceptually explaining the autocorrelation structure of house prices. To investigate whether naive updating can match these autocorrelations quantitatively, we simulate the richer model explicated in Section 1 using parameters calibrated from housing data.

3 Dynamics of Price Changes: Quantitative Results

3.1 Parameter Choices

This section calibrates the model using reasonable values of the parameters estimated from housing market data. Substantial uncertainty exists about the true values of the parameters, and they likely vary across space as well. Our approach, therefore, is to show that the model matches empirical house price dynamics using parameters within the range offered by the data. We perform sensitivity analysis with respect to the parameters of which we are most uncertain in Section 3.7.

Table 1 lists the parameters we estimate, which fall into two groups. The first are identified from data on city-wide demand. The second are identified from data on individual housing transactions. At no point do we use data on the time series of house price changes, which are the data we are trying to explain with our model.

3.1.1 Demand Parameters

The evolution of city-wide demand D is described by (2) and (3) and is governed by three parameters: the persistence λ of growth shocks, their volatility σ_g , and the volatility σ_D of non-growth demand shocks. These parameters are uniquely determined by the first three autocovariances of annual changes in D , which we denote $\gamma_0 = \text{Var}(\Delta D_t)$, $\gamma_1 = \text{Cov}(\Delta D_t, \Delta D_{t-1})$, and $\gamma_2 = \text{Cov}(\Delta D_t, \Delta D_{t-2})$; Δ denotes the difference over one year. As we show in the Appendix, the ratio γ_2/γ_1 uniquely determines λ . The autocorrelation γ_1/γ_0 then determines the ratio σ_g/σ_D , and finally γ_0 determines the level of these volatilities.

The literature has used two empirical proxies for housing fundamentals: rents (e.g. Campbell et al., 2009) and local incomes (e.g. Head, Lloyd-Ellis and Sun, 2014). Unfortunately, the autocovariances of these series differ significantly from each other, leaving us in the position of choosing between them. We choose intermediate values of these estimates and show that they work fairly well in allowing the model to capture house price dynamics.

The data on rents come from the Bureau of Labor Statistics (BLS), which compiles rental price indices for 23 metropolitan areas. Campbell et al. (2009) describe these data further, and provide the dataset we use. Using these data, we compute the standard deviation of annual changes to be $\sqrt{\gamma_0} = \$250$, and the first and second autocorrelations to be $\gamma_1/\gamma_0 = 0.73$ and $\gamma_2/\gamma_0 = 0.44$.⁴

The BEA provides income data at the metropolitan area level. An alternate source of income data comes from HMDA, which gives the median income of new homebuyers in a metropolitan area. To the extent that flow utility D corresponds to that of the marginal homebuyer, HMDA may be more appropriate. Glaeser et al. (2014) describe both datasets. In the BEA data, $\sqrt{\gamma_0} = \$1900$, with autocorrelations $\gamma_1/\gamma_0 = 0.30$ and $\gamma_2/\gamma_0 = 0.11$.⁵ The HMDA data provide a higher standard deviation of annual changes at $\sqrt{\gamma_0} = \$2700$, with autocorrelations $\gamma_1/\gamma_0 = 0.29$ and $\gamma_2/\gamma_0 = 0.09$.

The rent data display much more persistence than the income data. The ratio γ_2/γ_1 , which determines the growth persistence λ , equals 0.60 in rents but only about 0.34 in income. Furthermore, γ_1/γ_0 , which determines the relative importance of growth shocks, is much higher in the rent data. We combine features from both datasets by setting $\gamma_1/\gamma_0 = 0.3$ and $\gamma_2/\gamma_1 = 0.6$ as our baseline figures. Although this selection is somewhat arbitrary, it falls within the numbers suggested by the data, and allows the model to match the dynamics of price changes quite well.

We adopt the value $\sqrt{\gamma_0} = \$325$, which is much closer to the volatility implied by rents. This value is largely unimportant for the results, as it simply scales the variances in the model and does not affect the autocorrelations of price changes. It allows the model to match the volatility of price changes, but as we discuss shortly, the model can compare the predicted volatility of prices and fundamentals, and this comparison is independent of the assumed value of $\sqrt{\gamma_0}$.

Finally, we set the discount rate $r = 0.04$, following Glaeser et al. (2014).

⁴To arrive at γ_0 , we convert the rent index provided by the BLS to levels. The standard deviation of annual changes in the index equals 3.2, and the mean of the index is 140. Therefore, the rent index change standard deviation equals 2.3% of the mean level. Median annual rent in the US is \$10,884, so the standard deviation of annual changes equals \$250.

⁵These figures use BEA income data adjusted for state taxes. Without the tax adjustment, the numbers are $\sqrt{\gamma_0} = \$2100$, $\gamma_1/\gamma_0 = 0.14$, and $\gamma_2/\gamma_0 = 0.08$.

3.1.2 Transaction Parameters

The remaining parameters are the flow probability μ of moving, the number N of observed sales each period, the standard deviation σ_a of individual flow utility around the city-wide average, and the standard deviation σ_s of direct signals about demand. We also must determine the length δ of each period.

We identify μ using data on the probability that an owner-occupant sells a house in a given year. The Census, accessed at <https://www.census.gov/prod/2000pubs/p23-200.pdf>, reports that the 5-year mobility rate for owners is 31.2%. Therefore $1 - e^{-5\mu} = 31.2\%$, and $\mu = 7.5\%$. This figure corresponds to an expected tenancy of 13 years.

To compute σ_a , we use the standard deviation of rents, controlling for location and housing characteristic fixed effects. Rent data at the housing unit level come from the 2000 Census; the rent data provide a snapshot at a given time (2000). In the model, all houses are identical, whereas in the data they possess different characteristics. We therefore augment (1) with unit characteristics to arrive at the estimating equation

$$D_i^c = D^c + \mathbf{h}_i\boldsymbol{\beta} + a_i,$$

where c denotes the location, and \mathbf{h}_i is a vector of unit characteristics (rooms, bedrooms, plumbing, kitchen, age of the building, number of units in the building, and an indicator for whether the building sits on more than 10 acres of land). We observe D_i^c and \mathbf{h}_i , so we estimate this equation as a fixed effects regression, and identify σ_a^2 as the variance of the residual. This procedure assumes that actual buyers also observe \mathbf{h}_i and know $\boldsymbol{\beta}$. The location identifiers we use are public use microdata areas (PUMAs), the standard location entity used by the Census. Each PUMA contains at least 100,000 people. The value of σ_a we estimate equals \$3,120.

We set the length of each period at half of a year ($\delta = 0.5$). This time represents the frequency at which buyers observe home sales and news about demand. As many house price series, such as Case-Shiller and FHFA, are published at quarterly frequencies, and because this information may take some time to disseminate, $\delta = 0.5$ seems like a natural starting point. To compute the number of observed sales each period, we take the number of owner-occupied homes in the average PUMA, which is 30,800, and multiply it times $1 - e^{-\delta\mu}$, the probability of sale within a unit δ of time. The result is $N = 1,130$.

The final parameter to choose is σ_s , the noise in the news about fundamentals. We set this value to $\sigma_s = \$1,000$. As the median annual rent in the United States is $\$10,884$, this noise equals about 10% of the level. Although this error seems low, it is high enough to make the fundamental news irrelevant in the simulation. Buyers believe they observe D_t^a , the average flow utility of buyers in a given period. This average is also a noisy signal of fundamentals, with standard deviation σ_a/\sqrt{N} . Given our parameter choices, $\sigma_a/\sqrt{N} = \$93$. Buyers ignore the news as it is an order of magnitude more noisy than prices. For news to be relevant, it must have an error rate on the order of 1%. We explore this possibility in Section 3.7.

3.2 Simulation Methodology

We simulate the model in the limit as $\rho \rightarrow 0$ and measure various statistics about the resulting prices. Each simulation begins with a choice of the initial state vector $(D_0, g_0)'$ and the mean of the priors. Because D is non-stationary, without loss of generality we set the initial value to $\$10,000$. We pick the initial value of $g_0 \sim \mathcal{N}(0, \sigma_g^2/(2\lambda))$, its stationary distribution. As we showed in Lemma 2, in the $\rho \rightarrow 0$ limit the covariance of the buyer posterior after observing any price and news history does not depend on t ; we denote it \mathbf{P}_0 . Motivated by this stationarity, we draw the mean of the prior from a multivariate normal with covariance \mathbf{P}_0 and mean $(D_0, g_0)'$, and set the covariance of the prior to \mathbf{P}_0 .⁶ After seeding the initial values, we iteratively update the prices, states, and beliefs using the formulas in Section 1.3. For the same evolution of fundamentals, we separately keep track of the markets in which all buyers are naive and in which they are all rational. We produce 1,000 simulations and analyze the pooled results. In addition to the naive and rational prices, we calculate the prices that would hold were city-wide demand directly observable. In this “Observable” specification, prices are given by (5) with D_t and g_t replacing \widehat{D}_t and \widehat{g}_t .

We compare the simulated prices to empirical house price data. Our dataset is comprised of the annual FHFA house price indices for a panel of the largest 115 metropolitan areas in the United States between 1980 and 2011. To convert the indices into levels, we multiply each city’s index by the median house price in the 2000 Census, following Glaeser et al. (2014). We run each simulation for 31 years to align the time horizon in the data and the simulation.

⁶We give the naive buyers the same initial mean as the rational buyers. As shown in Proposition 2, the covariance of the naive forecast error exceeds the covariance of their stationary posterior. To account for this fact, we experimented with “burning in” the simulations by discarding the first five years. Doing so did not materially affect the results.

TABLE 1
Calibrated Parameter Values

Demand Parameters		
λ	0.51	Demand growth reversion
σ_g	\$180	Volatility of growth shocks
σ_D	\$190	Volatility of demand shocks
r	0.04	Discount rate
Transaction Parameters		
σ_a	\$3,120	Volatility of idiosyncratic utility
μ	0.075	Probability of forced sale
δ	0.5	Length of period (years)
N	1,130	Sales observed per period
σ_s	\$1,000	Noise in observations of demand

Notes: We estimate these parameters from data on house prices and rents. As the Appendix shows, the first three autocovariances of annual city-level demand changes uniquely identify λ , σ_g , and σ_D . We take these autocovariances from time series on rents and incomes at the metropolitan area level. The discount rate r comes from Glaeser et al. (2014). The resale probability μ comes from Census data on mobility of owner-occupants. The volatility σ_a is identified from the residual of a hedonic regression of rents on property characteristics and location (PUMA) fixed effects. We choose δ to roughly capture the frequency at which housing price and other economic data is released, and σ_s to describe the accuracy of local economic indicators (about 10%). Finally, N is determined using the number of owner-occupied houses in a PUMA, together with μ and δ .

3.3 Price Autocorrelations

We calculate the autocorrelation of annual price changes in the resulting house prices. Let C denote the number of simulations (or cities) and let T denote the number of years of data. We focus on the sample autocorrelations of annual price changes, which we calculate as

$$\text{Corr}(\Delta p_t, \Delta p_{t+k}) = \frac{\sum_{c=1}^C \sum_{\tau=1/\delta}^{(T-k)/\delta} (\Delta p_{c,\delta\tau} - \nu_0)(\Delta p_{c,\delta\tau+k} - \nu_k)}{\sqrt{\sum_{c=1}^C \sum_{\tau=1/\delta}^{(T-k)/\delta} (\Delta p_{c,\delta\tau} - \nu_0)^2} \sqrt{\sum_{c=1}^C \sum_{\tau=1/\delta}^{(T-k)/\delta} (\Delta p_{c,\delta\tau+k} - \nu_k)^2}},$$

where $\nu_k = [C(T-k)/\delta]^{-1} \sum_{c=1}^C \sum_{\tau=1/\delta}^{(T-k)/\delta} \Delta p_{c,\delta\tau+k}$ is the sample mean of each annual price change. We use $\delta = 0.5$ for the simulated data, and $\delta = 1$ for the empirical data which exist at an annual frequency. These autocorrelations summarize the serial correlation of price changes over time, and have been the focus on the literature on the predictability of house prices, starting with Case and Shiller (1989) and appearing most recently in Glaeser et al. (2014) and Head, Lloyd-Ellis and Sun (2014).

Table 2 reports the autocorrelations, both in the FHFA housing data and in the simulations. The Naive specification of the model matches the data quite well, whereas the Rational and Observable specifications do not even come close. Empirically, house prices display strong momentum at 1- and 2-year horizons, followed by mean reversion at longer horizons. These results appear in Table 2 under the “Data” column. The Rational and Observable specifications fail to capture these dynamics. Information is observed with a one-period lag by rational buyers, leading to the modest autocorrelation at a one-year horizon of 0.11. This figure is substantially below the empirical momentum of 0.67, and the remainder of the empirical autocorrelation structure fails to appear in any way in the Observable and Rational specifications.

In contrast, the “Naive” specification well matches the general dynamics in empirical housing prices. It predicts strong momentum over a 1-year horizon, with a value of 0.75. At longer horizons it predicts mean reversion. Mean reversion begins around 3 years, similar to in the data. The magnitude of the mean reversion is similar to the data, although it is higher in the Naive model and ends faster.

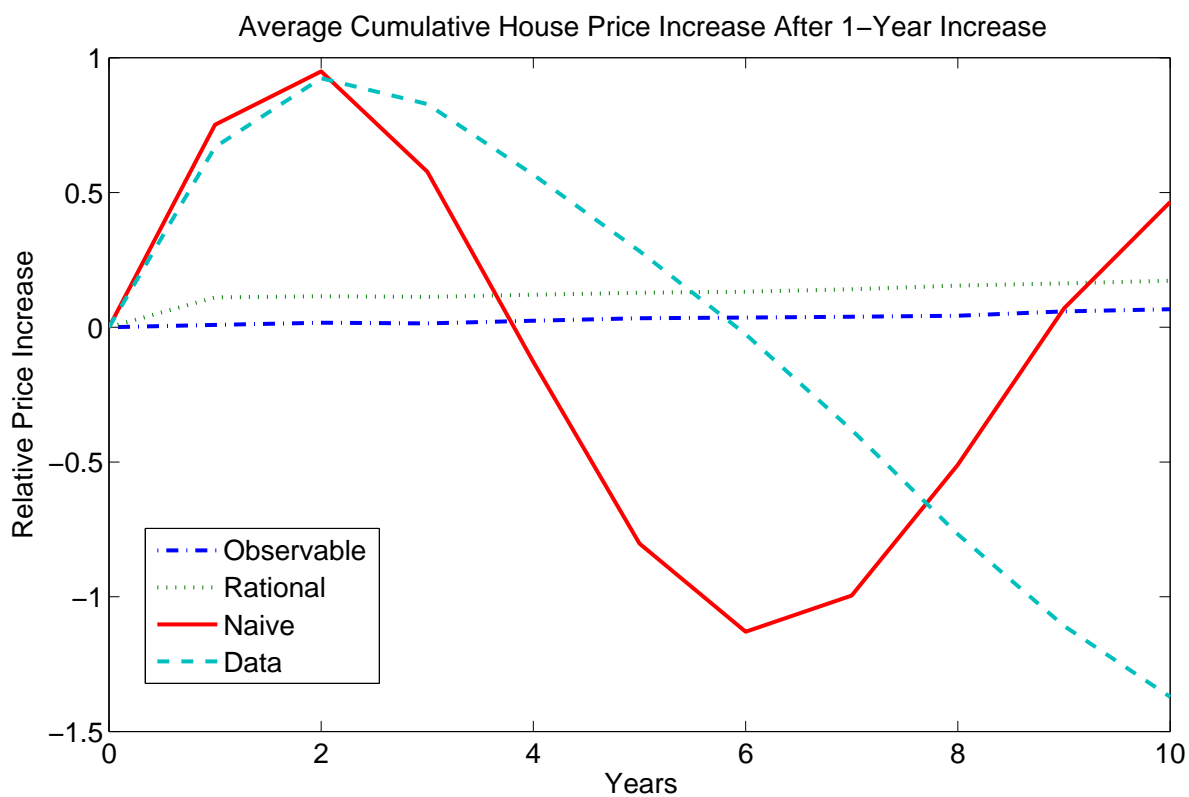
In the Observable and Rational models, prices are close to a random walk. This result is unsurprising in the Observable model, as news gets incorporated into prices immediately. More noteworthy is that this result holds in the Rational model as well. This similarity between the two

TABLE 2
Autocorrelations in House Price Changes, Annual Frequencies

	Model			
	Data	<i>Naive</i>	<i>Rational</i>	<i>Observable</i>
Corr($\Delta p_t, \Delta p_{t+1}$)	0.67	0.75	0.11	0.01
Corr($\Delta p_t, \Delta p_{t+2}$)	0.26	0.20	0.00	0.01
Corr($\Delta p_t, \Delta p_{t+3}$)	-0.10	-0.37	-0.00	-0.00
Corr($\Delta p_t, \Delta p_{t+4}$)	-0.26	-0.71	0.01	0.01
Corr($\Delta p_t, \Delta p_{t+5}$)	-0.28	-0.67	0.01	0.01
Corr($\Delta p_t, \Delta p_{t+6}$)	-0.31	-0.33	0.00	0.00
Corr($\Delta p_t, \Delta p_{t+7}$)	-0.35	0.13	0.01	0.00
Corr($\Delta p_t, \Delta p_{t+8}$)	-0.39	0.59	0.01	0.00
Corr($\Delta p_t, \Delta p_{t+9}$)	-0.34	0.58	0.01	0.02
Corr($\Delta p_t, \Delta p_{t+10}$)	-0.26	0.39	0.01	0.01

Notes: The Δ denotes an annual difference, so that $\Delta p_{t+k} = p_{t+k} - p_{t+k-1}$. “Observable” denotes the model in which buyers can observe the current state of demand, “Rational” denotes the model in which demand is unobservable but the buyers apply a rational filter, and “Naive” denotes the model in which buyers apply a naive filter. The correlations are estimated in both the data and the model by computing the correlation of all pairwise realizations of each pair of price changes in a panel of 100 cities over 30 years. Data come from the FHFA house price indices.

FIGURE 1
 Cumulative Autocorrelations in Annual Price Changes



Notes: This figure plots the cumulative autocorrelations for each column in Table 2. “Observable” denotes the model in which buyers can observe the current state of demand, “Rational” denotes the model in which demand is unobservable but the buyers apply a rational filter, and “Naive” denotes the model in which buyers apply a naive filter. The correlations are estimated in both the data and the model by computing the correlation of all pairwise realizations of each pair of price changes over 31 years. Data come from the FHFA house price indices.

models suggests that the rational buyers are extremely good at quickly filtering underlying demand from past prices, leading Rational prices to behave similarly to Observable prices. We explore this hypothesis below.

Figure 1 plots the cumulative autocorrelations for each specification in Table 2. The result equals the average movement in house prices after they initially rise for 1 year, relative to the initial increase. Only the Naive specification is able to capture the boom and bust profile of house prices that appears in the data.

3.4 Belief Dynamics

To explore the boom and bust profile documented in Figure 1, we study the evolution of prices after an exogenous demand shock. We decompose the resulting impulse response into three components: the idiosyncratic utility of the buyer, the buyer’s belief about the level of city demand, and the buyer’s belief about the growth rate. Explicitly,

$$p_t = \underbrace{\frac{1}{r + \mu} D_t^a}_{\text{idiosyncratic utility}} + \underbrace{\frac{\mu}{r(r + \mu)} \widehat{D}_t}_{\text{level belief}} + \underbrace{A_g \widehat{g}_t}_{\text{growth belief}}. \quad (9)$$

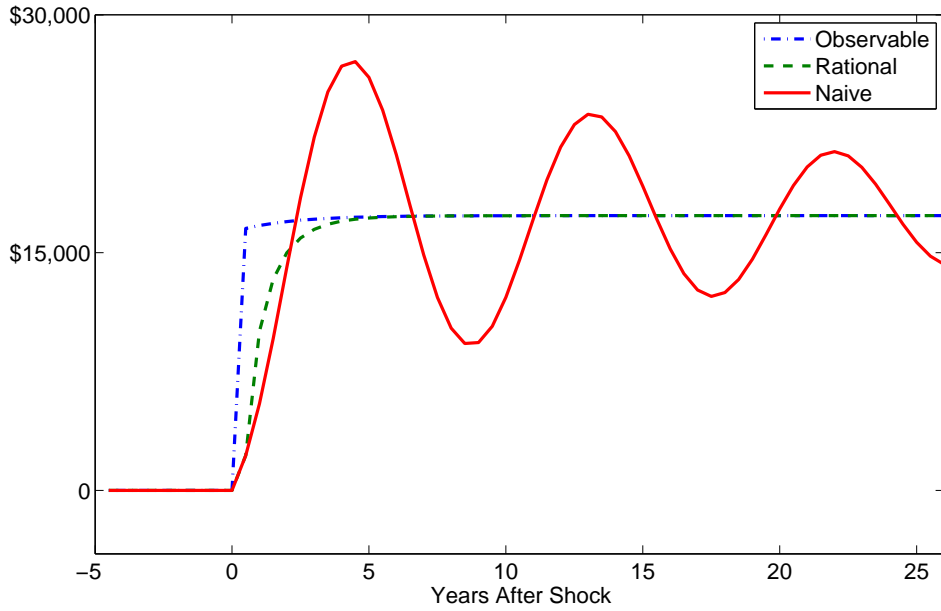
To calculate the impulse response, we simulate the model with and without a one-time, one standard deviation shock to the demand increments dW^D and dW^g . We report the average difference between the impulsed and non-impulsed simulations.

Figure 2 plots the impulse responses for prices, as well as for the “level belief” and “growth belief” components in (9). The axes are the same for each subfigure so that the relative importance of the belief components can be easily compared. Relative to Observable and Rational prices, the Naive prices substantially overshoot after a demand shock. Almost all of this overshooting comes from buyers overestimating the level of demand after the shock. They overestimate the demand level because they neglect the “growth belief” component of prices. Naive buyers erroneously believe that \widehat{g}_t , the growth belief of past buyers, never moves around. As they filter demand from past prices, naive buyers overestimate the level of demand when growth rates are high.

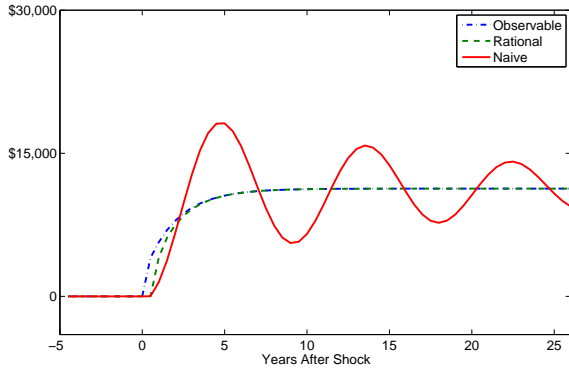
The profile of prices in Figure 2 captures the common accounts of “bubbles” found in a number of sources, such as Kindleberger and Aliber (2005), Shiller (2005), Pástor and Veronesi (2009), and Glaeser (2013). In this narrative, some fundamentally good shock, such as the discovery of

FIGURE 2
Evolution of Prices and Beliefs After a Demand Shock

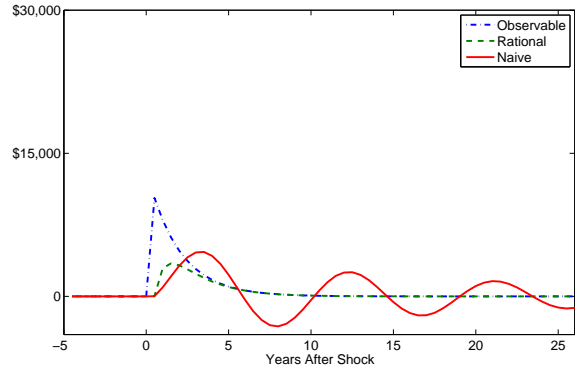
a) Prices



b) Level Beliefs



c) Growth Beliefs



Notes: We plot impulse responses from a one standard deviation shock to demand. The figures plot the average difference between the simulated model with and without the shock. Panel (a) displays transaction prices, and (b) is the component of prices related to beliefs about the level of demand D_t and (c) is the price component corresponding to beliefs about the growth rate of demand g_t . This decomposition appears in (9). “Observable” denotes the model in which buyers can observe the current state of demand, “Rational” denotes the model in which demand is unobservable but the buyers apply a rational filter, and “Naive” denotes the model in which buyers apply a naive filter.

a new technology, leads to increases in asset prices. Then, for some reason, this boom in asset values goes beyond what is justified by fundamentals, leading to an eventual bust. What causes the overshooting is a matter of debate, for which the sources above, as well as many other papers, offer competing explanations. Our explanation of this phenomenon is that buyers think the initial asset price boom conveys better information than it actually does, because the buyers neglect that part of this boom involves revisions to other buyers’ beliefs about the growth rate. This filtering error leads to overestimates of fundamentals, which cause an overshooting of prices and an eventual “bust” as prices return to fundamentals.

3.5 Expected Price Changes

As discussed in Section 1, homebuyers empirically extrapolate expected increases in the market value of their homes from past price increases. To explore this phenomenon in our model, we regress the expected annual gain in market value, as given by Proposition 4, on lagged annual price changes. Using our simulated data, we estimate the coefficients β_k in the equation

$$\mathbf{E}_t(V_{t+1} - V_t) = \sum_{k=0}^9 \beta_k \Delta p_{t-k} + \xi_t$$

using ordinary least squares for each of the three model specifications.

Table 3 displays the results. The Naive buyers strongly extrapolate future increases in the market values of their houses from past price increases. The Rational buyers, and those who can observe demand directly, do not. We can put these numbers in perspective using empirical survey evidence of homebuyer expectations. Case, Shiller and Thompson (2012) regress the reported expected one-year change in home prices on one-year lagged price changes, and find a coefficient of 0.23. They have 40 observations from four metropolitan areas between 2003 and 2012. Their simple regression has an R^2 of 0.73.

The naive buyers in our model match this empirical behavior quantitatively. When we regress the expected market value change on only the most recent annual price change, the coefficient equals 0.20. Naive buyers think news about growth rates does not get incorporated into prices. Therefore, when they see recent price increases, they infer a high growth rate, which leads them to expect increases in flow utility and hence prices in the future. According to Table 3, they draw

TABLE 3
Impact of Past Price Changes on Expected Appreciation

	Model		
	<i>Naive</i>	<i>Rational</i>	<i>Observable</i>
	$\mathbf{E}_t(V_{t+1} - V_t)$		
Δp_t	0.09	0.01	0.02
Δp_{t-1}	0.15	0.01	0.01
Δp_{t-2}	0.04	0.00	0.01
Δp_{t-3}	0.01	0.00	0.00
Δp_{t-4}	0.00	0.00	0.00
Δp_{t-5}	0.00	0.00	0.00
Δp_{t-6}	-0.00	0.00	0.00
Δp_{t-7}	0.00	0.00	0.00
Δp_{t-8}	-0.00	0.00	0.00
Δp_{t-9}	-0.00	0.00	0.00

Notes: This table reports the result from the regression $\mathbf{E}_t(V_{t+1} - V_t) = \sum_{k=0}^9 \beta_k \Delta p_{t-k} + \xi_t$ estimated using OLS with the simulated data from each model. The Δ denotes an annual difference, so that $\Delta p_{t+k} = p_{t+k} - p_{t+k-1}$. “Observable” denotes the model in which buyers can observe the current state of demand, “Rational” denotes the model in which demand is unobservable but the buyers apply a rational filter, and “Naive” denotes the model in which buyers apply a naive filter.

these inferences from price changes using several years of data.

In contrast, the rational buyers hardly extrapolate any increases in the market values of their homes. They (rightly) believe that all available information about growth rates already appears in the current market value of their home. Past price increases fail to convince them of any future appreciation in their house values. Although this behavior is perfectly rational, it is strongly at odds with the survey evidence on expectations.

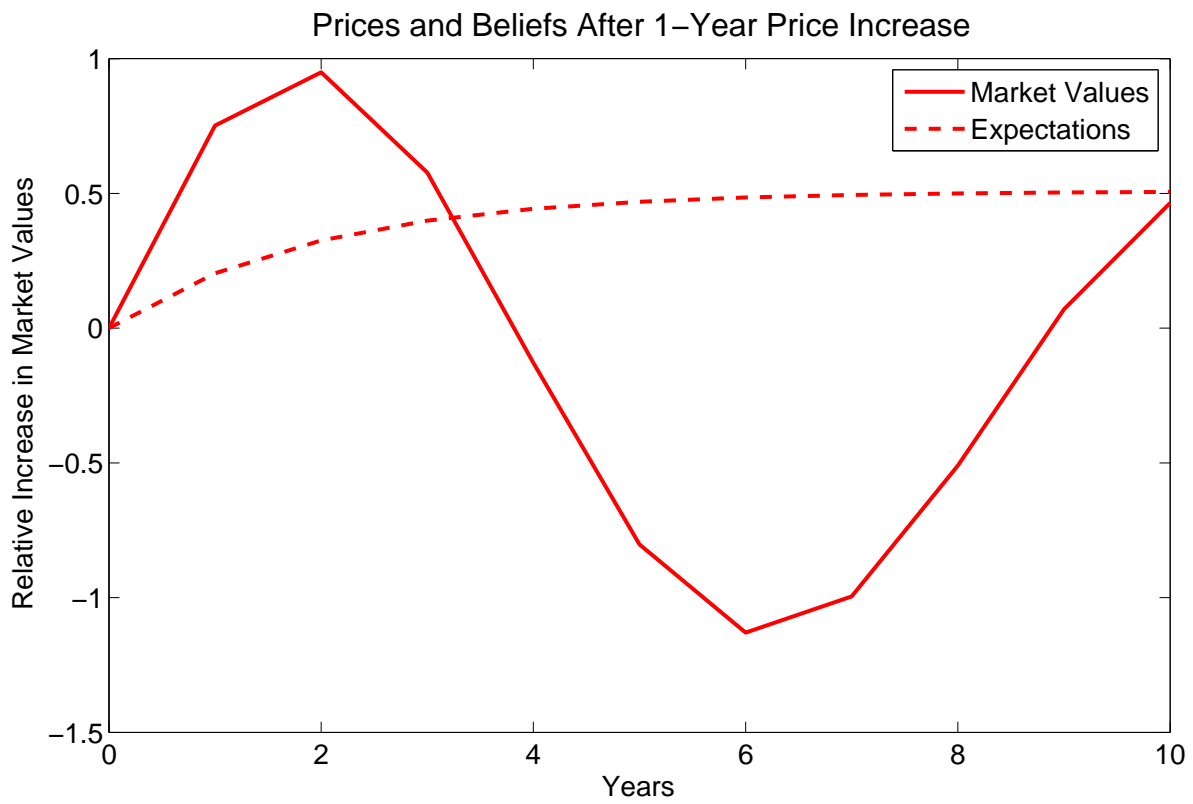
Rational beliefs about future prices must be correct on average, by definition. Naive beliefs do not have this restriction. To explore the forecasting ability of naive buyers, we plot the expected change in market values and the empirical change after a 1-year price increase. The realized change in market values exactly equals the price response in Figure 1. To compute the expected change, we use Proposition 4 to extend the expected one-year change conditional on a lagged one-year change (which is 0.20) to further years.

Figure 3 plots the results. After a 1-year increase in house prices, naive buyers *underestimate* the subsequent increase in their home values over the short-run. This result may at first seem surprising. Naive buyers extrapolate from prices much more strongly than rational buyers, and rational buyers extrapolate perfectly. One might guess, therefore, that naive buyers over-extrapolate. The reason they do not is that they fail to anticipate that future buyers will also revise their beliefs upwards after a price increase. This result—that naive buyers under-extrapolate in the short-run—is essential to the workings of this model. Momentum can exist in price increases only if buyers are continually being surprised by the extent of price increases. If naive buyers fully anticipated price increases, then these anticipations would become priced immediately, negating their realization.

Over longer horizons, naive buyers do over-extrapolate. As Figure 3 shows, these buyers completely fail to anticipate the eventual mean reversion in prices. Naive buyers believe market values follow the path of D_t , the city-wide demand. This demand exhibits no mean reversion, as it is a random walk with persistent drift. Actual prices, however, do exhibit mean reversion because naive buyers overestimate fundamentals after recent price increases.

Our model therefore microfounds the result that homebuyers fail to forecast busts. This phenomenon has recently been explored in a number of papers on “Natural Expectations” (Fuster, Laibson and Mendel, 2010; Fuster, Hebert and Laibson, 2010, 2011). This line of research studies consumers who, due to cognitive limitations, forecast macroeconomic variables as an AR(1) instead of an AR(2). This forecast restriction prohibits consumers from forecasting mean reversion in vari-

FIGURE 3
Price and Belief Evolution for Naive Buyers



Notes: This figure plots the naive price changes from Figure 1 along with naive buyers' expected change in market values after a one-year price increase. This expectation is calculated using Proposition 4 and a regression of \hat{g}_t on a one-year price increase.

ables such as house prices where it exists, while allowing them to forecast momentum. In contrast to consumers with natural expectations, naive buyers perfectly understand the underlying process for fundamentals. Their naive view that house prices reflect just fundamentals and not beliefs prohibits them from forecasting busts.

3.6 Volatility

Empirically, house prices exhibit excess volatility relative to the movements of underlying fundamentals (Glaeser et al., 2014; Head, Lloyd-Ellis and Sun, 2014). We compare the volatility of price changes to that of movements in flow utility in the three specifications of our model. The fundamental in our model is D_t , the city-wide flow utility at time t . We must scale D_t so that it is comparable to house prices p . On average, $g_t = 0$, so house prices on average are $p_t = D_t/r$. This result leads us to use D_t/r as our measure of fundamentals at t .

We estimate the volatility of a price change over a horizon of k years using the sample standard deviation of k -year price changes over all simulations and k -year price intervals. Specifically,

$$\text{Vol}(\Delta_k p_t) = \sqrt{[C(T - k)/\delta]^{-1} \sum_{c=1}^C \sum_{\tau=k/\delta}^{T/\delta} (\Delta_k p_{c,\delta\tau} - \nu_k)^2},$$

where $\nu_k = [C(T - k)/\delta]^{-1} \sum_{c=1}^C \sum_{\tau=1/\delta}^{k/\delta} \Delta_k p_{c,\delta\tau}$ and Δ_k is the k -year difference operator.

Two broad patterns emerge from the results, which appear in Table 4. First, all three simulated price paths display excess volatility relative to fundamentals. This feature has a simple explanation. Because growth is persistent, a rise in fundamentals today conveys news about future fundamentals. Because prices are forward-looking, they move more strongly than current fundamentals, and hence exhibit more volatility.

The second pattern is that naive prices are more volatile than rational and observable prices. This relationship holds because naive buyers over-extrapolate fundamentals from demand shocks. This added volatility in naive prices explains the majority of excess volatility at longer horizons. Over 5 years, the volatility in rational price changes is only 17% higher than fundamentals, whereas the naive volatility is 2.1 times fundamental volatility. Head, Lloyd-Ellis and Sun (2014) calculate that empirical house prices display volatility 2.2 times higher than fundamentals.

TABLE 4
Volatility of Price Changes at Different Horizons

	Horizon		
	1 Year	3 Years	5 Years
Data	\$16,000	\$40,000	\$50,000
<i>Simulated Prices</i>			
Naive	\$16,000	\$41,000	\$51,000
Rational	\$11,000	\$21,000	\$28,000
Observable	\$12,000	\$22,000	\$28,000
<i>Simulated Fundamentals</i>			
D_t/r	\$8,000	\$17,000	\$24,000

Notes: We calculate volatility as the sample standard deviation of changes over the given time horizon, for all realizations over this horizon and over all simulations we run. “Observable” denotes the model in which buyers can observe the current state of demand, “Rational” denotes the model in which demand is unobservable but the buyers apply a rational filter, and “Naive” denotes the model in which buyers apply a naive filter. The fundamental D_t/r denotes the present value of the current city-wide level of flow utility. Empirical volatility is computed using FHFA house price data for 115 metropolitan areas from 1980-2011.

3.7 Predictions for Alternate Parameters

Table 5 explores how the key house price moments change under different parameter values. We focus on three moments: the autocorrelation of annual price changes over a one-year horizon, the autocorrelation over a five-year horizon, and the standard deviation of five-year price changes. These moments encompass the three stylized facts of house prices that motivate this paper—momentum, mean reversion, and excess volatility. To explore the sensitivity of these moments to the parameters, we change a single parameter at a time, holding the ones used in the main analysis constant.⁷

We first adjust λ , the persistence of demand growth. As mentioned earlier, the autocovariance ratio γ_2/γ_1 determines this parameter. The empirical value of this ratio depends on how one measures demand at the city level. Using income yields a ratio of 0.3, smaller than the 0.6 we estimate from BLS rents. In Table 5, we report the ratio γ_2/γ_1 corresponding to the values of λ used; we adjust σ_D for each λ to keep the volatility of demand changes ($\sqrt{\gamma_0}$) constant. According to Table 5, smaller values for the persistence significantly attenuate the mean reversion and excess volatility of naive prices, yet leave the momentum largely unchanged. Overshooting occurs in the model when buyers incorrectly attribute price growth from revisions about \hat{g} to increases in fundamentals. When growth persistence is small, growth shocks are not of much quantitative importance, and growth rate expectations enter into prices only slightly. In terms of (5), A_g is small. Momentum persists because shocks to dW^D slowly incorporate into prices when buyers are naive.

The next parameter we consider is tenancy length. This input determines μ , as the expected tenancy equals $1/\mu$. When tenancy length is shorter, buyers expect to resell their houses more quickly and therefore care more about market demand. This change enhances momentum in naive prices, but has a non-monotonic affect on mean reversion and volatility.

More precise information about demand moves naive prices closer to rational ones. This result is unsurprising, as with perfect news, demand is known and buyers no longer rely on house prices for inference, rendering the two types of buyers identical. News must be extremely precise to dampen the effects of naive inference. Momentum persists even when the error is on the order of \$100. The reason is that prices already aggregate information about demand quite well due to the central limit theorem and the large number of observed sales, so news must be extremely precise to make

⁷In some cases, extreme values of the parameters lead to non-stationarity of naive price changes, causing explosive behavior of the price paths. We do not analyze any such cases in Table 5.

TABLE 5
Simulated Moments for a Range of Inputs

	1-Year Momentum		5-Year Reversion		5-Year Volatility	
	<i>Naive</i>	<i>Rational</i>	<i>Naive</i>	<i>Rational</i>	<i>Naive</i>	<i>Rational</i>
Annual Growth Persistence						
0	0.50	0.17	0.03	0.00	\$16,000	\$18,000
0.3	0.56	0.15	0.00	0.00	\$18,000	\$19,000
0.6	0.75	0.11	-0.67	0.01	\$51,000	\$28,000
Expected Tenancy (Years)						
2	0.94	0.04	-0.07	-0.00	\$46,000	\$28,000
10	0.80	0.11	-0.80	0.00	\$62,000	\$28,000
30	0.60	0.16	-0.14	-0.01	\$35,000	\$28,000
News Noise (σ_s)						
\$10	0.12	0.10	0.01	0.00	\$28,000	\$28,000
\$100	0.44	0.11	0.00	0.00	\$29,000	\$28,000
\$1,000	0.75	0.11	-0.67	0.01	\$51,000	\$28,000
\$10,000	0.76	0.12	-0.71	0.00	\$54,000	\$28,000
Houses in the Area						
1,000	0.21	0.12	-0.03	0.00	\$28,000	\$28,000
10,000	0.76	0.14	-0.47	0.00	\$39,000	\$28,000
50,000	0.75	0.11	-0.72	-0.01	\$59,000	\$28,000
Utility Noise (σ_a)						
\$100	0.05	0.10	0.01	0.00	\$27,000	\$28,000
\$1,000	0.69	0.11	-0.51	0.01	\$59,000	\$28,000
\$10,000	0.61	0.16	-0.17	-0.00	\$31,000	\$28,000

Notes: “Annual Growth Persistence” is the ratio of the second and first autocorrelations of annual changes in city-wide demand D . This input uniquely determines λ ; as we change λ we alter σ_D to maintain the volatility of annual demand changes. “Expected Tenancy” equals the average time spent in a house before a forced sale; its inverse equals the moving probability μ . “News Noise” gives the standard deviation of error in direct signals about demand. “Houses in the Area” determines the number N of sales observed each period. “Utility Noise” is the standard deviation of idiosyncratic utility for each buyer. In each row, we simulate the model with naive and rational homebuyers using the parameters of our main analysis (shown in Table 1) with the exception of the parameter of interest studied in that row. “1-Year Momentum” is the correlation of annual price changes on lagged changes, “5-Year Reversion” is the correlation on a 5-year lag, and “5-Year Volatility” is the standard deviation of 5-year house price changes.

a difference.

By the same token, naive inference produces stronger mean reversion and volatility when there are more houses in the area. The number of houses in the area determines how much weight naive buyers place on past housing prices. When the number of houses is very low, they do not weight the past much and the momentum attenuates. In a sense, this result suggests a non-monotonic relationship between information flows and housing fluctuations. With no information, momentum and mean reversion disappear. With good information about fundamentals, the same effect occurs. It is only when there is good information about past market behavior but not about fundamentals that momentum, mean reversion and volatility become most pronounced.

Finally, the standard deviation of idiosyncratic utility has a non-monotonic affect on the moments of interest. Low values of σ_a mean that buyers know much about the city-wide demand from observing their own demand. Hence in this case they do not rely on market prices for inference very much, dampening the volatility, momentum, and mean reversion of naive prices. At high values of σ_a , house prices are again not of much use. House prices average away the noise in idiosyncratic utility, but for a given N , this average is less precise when σ_a is larger. However, momentum and mean reversion exist to some degree even at $\sigma_a = \$10,000$, three times higher than the baseline value in Table 1 that is implied by rent data.

3.8 Forecast Accuracy

This section evaluates the relative accuracy of the naive and rational filters. Each buyer's goal is to infer D_t and g_t to minimize the error in valuing her home. The error in this forecast is the difference between the price in (5) under \widehat{D}_t and \widehat{g}_t and that under the true values D_t and g_t :

$$\text{error} = \frac{\mu}{r + \mu} \frac{\widehat{D}_t - D_t}{r} + A_g(\widehat{g}_t - g_t). \quad (10)$$

The accuracy of the filter is the standard deviation of this error. We calculate this standard deviation as the square-root of the pooled variance across all simulations and time periods of the above error.

We additionally use this error to evaluate the robustness of the rational and naive filters. As Proposition 3 suggests, the rational filter is quite sensitive. It performs poorly when applied to price paths produced by buyers using non-rational filters. We document this phenomenon quantitatively here. We calculate the estimates of \widehat{D}_t and \widehat{g}_t obtained using the rational filter applied to *naive*

TABLE 6
Forecast Accuracy

		Market	
		<i>Rational</i>	<i>Naive</i>
Individual	<i>Rational</i>	\$11,000	\$600,000,000,000,000,000,000,000,000
	<i>Naive</i>	\$12,000	\$22,000

Notes: This table reports the standard deviation of the forecast error in (10), calculated as the square-root of the variance pooled over all simulations and time periods. The column denotes the filtering used by buyers to determine the price path, and the row specifies the filtering used by an individual, the error of whose forecasts is evaluated. The parameters used in the simulations are given by Table 1.

prices.⁸ We also compute the forecast accuracy for the naive filter applied to rational prices.

Table 6 presents the results. The row chooses the filter used by the buyer we are studying, and the column specifies the filter used to obtain market prices. The rational equilibrium is twice as accurate as the naive one. When all buyers are rational, the forecast error has standard deviation \$11,000; it rises to \$22,000 when all buyers are naive. When the market uses the rational filter, the marginal cost of using the naive filter is very small in terms of lost accuracy. The standard deviation of the error rises only to \$12,000 from \$11,000. The largest error comes from employing the rational filter when other buyers are in fact naive. In this case, the standard deviation of the error equals essentially an infinite number ($(\$6)10^{32}$).

These results suggest that the naive filter is much more robust to uncertainty about previous buyers. The equilibrium in which all buyers are rational is extraordinarily fragile. If there is even a small probability all other buyers are naive, an individual buyer is much better off being naive as well. Furthermore, the cost of naivety is quite small when others are rational, as the naive buyer essentially free-rides off the information aggregation provided by others.

4 Conclusion

Many salient features of house prices—excess volatility, momentum, and mean reversion—can be explained by a model in which homebuyers make a small error in filtering information out of past prices. These naive buyers expect the market value of their home to rise after recent house price increases, and they fail to forecast busts after booms. They are overconfident in their assessments of the housing market.

Several promising avenues for future research arise from this paper’s results. One is to investigate robust filtering rules buyers could use to reduce house price volatility. The rational filter in this paper has two flaws: it is very complicated, and it works only if *all* other buyers use it. Perhaps there exists a simple filtering rule whose use would reduce house price volatility and whose accuracy is independent of the actions of other homebuyers. It is also possible that the naive filter already is optimal for reducing volatility among simple filters.

A second direction for future research is to explore the implications of naive filtering for con-

⁸The rational buyer begins with a prior on \mathbf{x}_0 which we seed randomly, as explained above. At each step, the rational buyer at t extracts $D_{t'}^a$ out of $p_{t'}$ by assuming that $\mathbf{E}(\mathbf{x}_{t'} \mid \Omega_{t'}^p \cup \Omega_{t'}^a \cup \Omega_{t'}^x)$ is the same for buyers at t' as it is for herself. Then, using this extracted value for $D_{t'}^a$, the buyer updates her posterior using Lemma 2.

sumption. Several recent papers have documented the explosion of consumption financed by home equity during the 2000-2006 boom. Because naive buyers do not forecast the bust that follows booms, they may over-consume out of house price increases relative to rational buyers. This over-consumption may be important for understanding the extent of leverage homeowners took on during the boom.

Finally, incorporating construction into the model is of first-order importance. Supply responses have the potential to temper price increases caused by naive homebuyers. However, if the homebuilders are also using naive filters, it is unclear whether housing supply would attenuate or even amplify the effects documented in this paper.

References

- Adam, Klaus, Johannes Beutel, and Albert Marcet.** 2014. “Stock Price Booms and Expected Capital Gains.” *Working Paper*.
- Adam, Klaus, Pei Kuang, and Albert Marcet.** 2011. “House Price Booms and the Current Account.” In *NBER Macroeconomics Annual 2011, Volume 26*. 77–122. University of Chicago Press.
- Anenberg, Elliot, and Patrick J. Bayer.** 2013. “Endogenous Sources of Volatility in Housing Markets: The Joint Buyer-Seller Problem.” *Working paper*.
- Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer.** 2015. “X-CAPM: An Extrapolative Capital Asset Pricing Model.” *Journal of Financial Economics*, 115(1): 1–24.
- Berry, Christopher R., and Edward L. Glaeser.** 2005. “The Divergence of Human Capital Levels Across Cities.” *Papers in Regional Science*, 84(3): 407–444.
- Burnside, Craig, Martin Eichenbaum, and Sergio Rebelo.** 2014. “Understanding Booms and Busts in Housing Markets.” *Working Paper*.
- Camerer, Colin F., Teck-Hua Ho, and Juin-Kuan Chong.** 2004. “A Cognitive Hierarchy Model of Games.” *Quarterly Journal of Economics*, 861–898.
- Campbell, Sean D., Morris A. Davis, Joshua Gallin, and Robert F. Martin.** 2009. “What Moves Housing Markets: A Variance Decomposition of the Rent–Price Ratio.” *Journal of Urban Economics*, 66(2): 90–102.
- Case, Karl E., and Robert J. Shiller.** 1989. “The Efficiency of the Market for Single-Family Homes.” *American Economic Review*, 125–137.
- Case, Karl E., Robert J. Shiller, and Anne K. Thompson.** 2012. “What Have They Been Thinking? Homebuyer Behavior in Hot and Cold Markets.” *Brookings Papers on Economic Activity*, 265–315.
- Cochrane, John H.** 2011. “Presidential Address: Discount Rates.” *Journal of Finance*, 66(4): 1047–1108.
- Cutler, David M., James M. Poterba, and Lawrence H. Summers.** 1991. “Speculative Dynamics.” *Review of Economic Studies*, 58(3): 529–546.
- De Long, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann.** 1990. “Noise Trader Risk in Financial Markets.” *Journal of Political Economy*, 703–738.
- Eaton, Jonathan, and Zvi Eckstein.** 1997. “Cities and Growth: Theory and Evidence from France and Japan.” *Regional Science and Urban Economics*, 27(4): 443–474.
- Eyster, Erik, and Matthew Rabin.** 2005. “Cursed Equilibrium.” *Econometrica*, 73(5): 1623–1672.

- Eyster, Erik, and Matthew Rabin.** 2010. “Naive Herding in Rich-Information Settings.” *American Economic Journal: Microeconomics*, 2(4): 221–243.
- Eyster, Erik, and Matthew Rabin.** 2014. “Extensive Imitation is Irrational and Harmful.” *Quarterly Journal of Economics*.
- Eyster, Erik, Matthew Rabin, and Dimitri Vayanos.** 2013. “Financial Markets where Traders Neglect the Informational Content of Prices.” *Working Paper*.
- Fuster, Andreas, Benjamin Hebert, and David Laibson.** 2010. “Investment Dynamics with Natural Expectations.” *International Journal of Central Banking*, 8(81): 243.
- Fuster, Andreas, Benjamin Hebert, and David Laibson.** 2011. “Natural Expectations, Macroeconomic Dynamics, and Asset Pricing.” In *NBER Macroeconomics Annual 2011, Volume 26*. 1–48. University of Chicago Press.
- Fuster, Andreas, David Laibson, and Brock Mendel.** 2010. “Natural Expectations and Macroeconomic Fluctuations.” *Journal of Economic Perspectives*, 24(4): 67–84.
- Gabaix, Xavier.** 1999. “Zipf’s Law for Cities: An Explanation.” *Quarterly journal of Economics*, 739–767.
- Gabaix, Xavier.** 2014. “A Sparsity-Based Model of Bounded Rationality.” *The Quarterly Journal of Economics*, 129(4): 1661–1710.
- Ganong, Peter, and Daniel Shoag.** 2013. “Why Has Regional Income Convergence in the U.S. Declined?” *Working Paper*.
- Gao, Zhenyu, Michael Sockin, and Wei Xiong.** 2014. “Learning about the Neighborhood: The Role of Supply Elasticity for Housing Cycles.” *Working Paper*.
- Giglio, Stefano, Matteo Maggiori, and Johannes Stroebel.** 2014. “No-Bubble Condition: Model-Free Tests in Housing Markets.” *Working Paper*.
- Glaeser, Edward L.** 2013. “A Nation of Gamblers: Real Estate Speculation and American History.” *American Economic Review*, 103(3): 1–42.
- Glaeser, Edward L., and Joseph Gyourko.** 2009. “Arbitrage in Housing Markets.” In *Housing Markets and the Economy: Risk, Regulation, and Policy*. Chapter 5, 113–146. Lincoln Institute of Land Policy.
- Glaeser, Edward L., José A. Scheinkman, and Andrei Shleifer.** 1995. “Economic Growth in a Cross-Section of Cities.” *Journal of Monetary Economics*, 36(1): 117–143.
- Glaeser, Edward L., Joseph Gyourko, and Albert Saiz.** 2008. “Housing Supply and Housing Bubbles.” *Journal of Urban Economics*, 64(2): 198–217.
- Glaeser, Edward L., Joseph Gyourko, Eduardo Morales, and Charles G. Nathanson.** 2014. “Housing Dynamics: An Urban Approach.” *Journal of Urban Economics*, 81: 45–56.
- Guren, Adam.** 2014. “The Causes and Consequences of House Price Momentum.” *Working Paper*.

- Hamilton, James D.** 1994. *Time Series Analysis*. Princeton University Press.
- Head, Allen, Huw Lloyd-Ellis, and Hongfei Sun.** 2014. “Search, Liquidity, and the Dynamics of House Prices and Construction.” *American Economic Review*, 104(4): 1172–1210.
- Hong, Harrison, and Jeremy C. Stein.** 1999. “A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets.” *Journal of Finance*, 54(6): 2143–2184.
- Kindleberger, Charles P., and Robert Aliber.** 2005. “Manias, Panics, and Crashes: A History of Financial Crises.” *Wiley, 5th Edition*.
- Nathanson, Charles G., and Eric Zwick.** 2014. “Arrested Development: Theory and Evidence of Supply-Side Speculation in the Housing Market.” *Working Paper*.
- Pástor, Ľuboš, and Pietro Veronesi.** 2009. “Technological Revolutions and Stock Prices.” *American Economic Review*, 99(4): 1451–1483.
- Piazzesi, Monika, and Martin Schneider.** 2009. “Momentum Traders in the Housing Market: Survey Evidence and a Search Model.” *American Economic Review*, 406–411.
- Shiller, Robert J.** 2005. *Irrational Exuberance*. Princeton University Press.

Appendix

Proof of Lemma 1. We posit a pricing function of the form $p_t = A_0 D_t + A_c \widehat{D}_t + A_g \widehat{g}_t$ and show that this type of function is an equilibrium. If the random sale happens at $T > t$, then the realized value to the buyer equals

$$\int_t^T e^{-r(\tau-t)} D_{i,\tau} d\tau + e^{-r(T-t)} p_T = \int_t^T e^{-r(\tau-t)} \left(D_{i,t} + \int_t^\tau g_{\tau'} d\tau' \right) d\tau + e^{-r(T-t)} (A_0 D_T^a + A_c \widehat{D}_T + A_g \widehat{g}_T).$$

We first solve for the expected value $\mathbf{E}_{i,t}$ of the above quantity. It is a standard result from stochastic calculus that $\mathbf{E}_{i,t} g_\tau = e^{-\lambda(\tau-t)} \mathbf{E}_{i,t} g_t$ for all $\tau > t$, given the specification in (3). It follows that

$$\begin{aligned} \mathbf{E}_{i,t} \int_t^T \int_t^\tau e^{-r(\tau-t)} g_{\tau'} d\tau' d\tau &= \int_t^T \int_t^\tau e^{-r(\tau-t)} e^{-\lambda(\tau'-t)} \mathbf{E}_{i,t} g_t d\tau' d\tau \\ &= \int_t^T e^{-r(\tau-t)} \frac{1 - e^{-\lambda(\tau-t)}}{\lambda} d\tau \mathbf{E}_{i,t} g_t \\ &= \left(\frac{1 - e^{-r(T-t)}}{r\lambda} - \frac{1 - e^{-(r+\lambda)(T-t)}}{\lambda(r+\lambda)} \right) \mathbf{E}_{i,t} g_t. \end{aligned}$$

We have $\mathbf{E}_{i,t} \int_t^T e^{-r(\tau-t)} D_{i,t} d\tau = (1 - e^{-r(T-t)}) D_{i,t}/r$, and $\mathbf{E}_{i,t} D_T^a = \mathbf{E}_{i,t} D_T$ by (1). This expectation equals

$$\mathbf{E}_{i,t} D_T = \mathbf{E}_{i,t} D_t + \mathbf{E}_{i,t} \int_t^T g_\tau d\tau = \mathbf{E}_{i,t} D_t + \frac{1 - e^{-\lambda(T-t)}}{\lambda} \mathbf{E}_{i,t} g_t.$$

Next note that $\mathbf{E}_{i,t} \widehat{D}_T = \mathbf{E}_{i,t} D_T$ by assumption, and that $\mathbf{E}_{i,t} \widehat{g}_T = \phi_g(T-t) \mathbf{E}_{i,t} g_t$. The price $p_{i,t}$ is the expected value of the realized value to the buyer, given that $T-t$ is distributed with probability distribution function $\mu e^{-\mu(T-t)}$:

$$\begin{aligned} p_{i,t} &= \int_t^T \mu e^{-\mu(T-t)} \left[\frac{1 - e^{-r(T-t)}}{r} D_{i,t} + \left(\frac{1 - e^{-r(T-t)}}{r\lambda} - \frac{1 - e^{-(r+\lambda)(T-t)}}{\lambda(r+\lambda)} \right) \mathbf{E}_{i,t} g_t + \right. \\ &\quad \left. e^{-r(T-t)} (A_0 + A_c) \left(\mathbf{E}_{i,t} D_t + \frac{1 - e^{-\lambda(T-t)}}{\lambda} \mathbf{E}_{i,t} g_t \right) + e^{-r(T-t)} A_g \phi_g(T-t) \mathbf{E}_{i,t} g_t \right] dT \end{aligned}$$

Evaluating this integral and collecting terms yields

$$p_{i,t} = \frac{r}{r+\mu} \frac{D_{i,t}}{r} + \frac{\mu}{r+\mu} (A_0 + A_c) \mathbf{E}_{i,t} D_t + \left(\frac{1 + (A_0 + A_c)\mu}{(r+\mu)(r+\lambda+\mu)} + \phi^* A_g \right) \mathbf{E}_{i,t} g_t,$$

where $\phi^* = \int_t^\infty \mu e^{-(\mu+r)(T-t)} \phi_g(T-t) dT$. Taking the average over all buyers i at t and then using the method of undetermined coefficients yields $A_0 = (r/(r+\mu))/r$, $A_c = (\mu/(r+\mu))/r$, and $A_g = 1/(r(r+\lambda+\mu)(1-\phi^*))$. These formulas exactly match (5).

Proof of Proposition 1. For any two times τ_1 and τ_2 , we say a buyer at τ_1 observes τ_2 if $(\tau_1 - \tau_2)/\delta \in \mathbb{N}$, the set of positive integers. Let t' be any time that buyers at t observe. Let t_0

denote the maximal member of \mathcal{T} such that $t_0 \leq t'$. The time t_0 exists because arbitrarily negative values of \mathcal{T} exist, as $\rho > 0$. We denote \mathcal{T}' to be the set of times at least t_0 that the buyer at t' observes. This set is clearly finite.

We claim that the buyer at t' forms her posterior on $\mathbf{x}_{t'}$ conditional only on the finite set $\Omega'_{i,t'} = \{D_{i,t'}\} \cup \{p_\tau \mid \tau \in \mathcal{T}'\} \cup \{D_\tau^s \mid \tau \in \mathcal{T}'\} \cup \{\mathbf{x}_{t_0}\}$. It is clear from (2) and (3) that \mathbf{x} is a Markov process (this statement is proved formally in the proof of Lemma 2). Thus, a rational inference on \mathbf{x}_{t_1} will not use any information from times before t_0 , as \mathbf{x}_{t_0} is observed. The totality of $\Omega_{i,t'}$ that occurs no earlier than t_0 is $\Omega'_{i,t'}$.

The state \mathbf{x} evolves linearly with normal noise, and the elements of $\Omega'_{i,t'}$ are all observations of some linear function of a lag of \mathbf{x} with normal noise, or they are linear combinations of posteriors on lags of \mathbf{x} and such noisy observations (in the case of prices). Hence, standard Kalman filtering (which we make explicit in the proof of Lemma 2) leads the posterior $\mathbf{x}_{t'} \mid \Omega'_{i,t'}$ to be linear in the observations in $\Omega'_{i,t'}$. The linear weights are common knowledge to all buyers, as they depend only on the parameters governing the noise and evolution of the state. The average posterior $\widehat{\mathbf{x}}_{t'}$ across buyers at t' is thus a linear function of $D_{t'}^a$ and $\mathbf{x}_{t'} \mid \Omega'_{i,t'} \setminus \{D_{i,t'}\}$. As the price at t' follows (5) in Lemma 1, $p_{t'}$ is a linear combination of $D_{t'}^a$ and $\mathbf{x}_{t'} \mid \Omega'_{i,t'} \setminus \{D_{i,t'}\}$ whose weights are common knowledge. The buyer at t observes $\Omega'_{i,t'} \setminus \{D_{i,t'}\}$ and $p_{t'}$, and therefore can perfectly deduce $D_{t'}^a$.

Proof of Lemma 2. First, we prove that the state variable evolves as described in the text. Consider the evolution of D_t between times $t = 0$ and $t = \delta$. We will prove that we can write $D_\delta = D_0 + \beta g_0 + w^D$ and $g_\delta = e^{-\delta\lambda} g_0 + w^g$, where w^D and w^g are independent from D_0 and g_0 and have mean 0 conditional on data at $t = 0$. We also calculate the covariance matrix of $w = (w^D, w^g)'$. From (2) and (3), we have $D_\delta = D_0 + \int_0^\delta g_t dt + \sigma_D \int_0^\delta dW_t^D$ and $g_\delta = e^{-\delta\lambda} g_0 + \sigma_g \int_0^\delta e^{-\lambda(\delta-t)} dW_t^g$. We find β as the coefficient from regressing $D_\delta - D_0$ on g_0 . This coefficient equals

$$\beta = \frac{\text{Cov}\left(g_0, \int_0^\delta g_t dt + \sigma_D \int_0^\delta dW_t^D\right)}{\text{Var}(g_0)} = \frac{\text{Cov}\left(g_0, \int_0^\delta g_t dt\right)}{\text{Var}(g_0)} = \int_0^\delta e^{-\lambda t} dt = \frac{1 - e^{-\lambda\delta}}{\lambda},$$

where we have used the fact that for stochastic processes of the form specified in (3), $\text{Cov}(g_0, g_t) = e^{-\lambda t} \text{Var}(g_0)$ for all t . The variance of w^D equals the variance of this regression's forecast error, which is $\text{Var}\left(\int_0^\delta g_t dt + \sigma_D \int_0^\delta dW_t^D\right) - \beta^2 \text{Var}(g_0) = \text{Var}\left(\int_0^\delta g_t dt\right) + \delta\sigma_D^2 - \beta^2 \text{Var}(g_0)$. We solve for the first variance on the right as

$$\text{Cov}\left(\int_0^\delta g_t dt, \int_0^\delta g_s ds\right) = \int_0^\delta \int_0^\delta e^{-\lambda|t-s|} \text{Var}(g_0) dt ds = \frac{2(e^{-\delta\lambda} - 1 + \delta\lambda)}{\lambda^2} \text{Var}(g_0).$$

Another standard fact about the stochastic process specified in (3) is that $\text{Var}(g_0) = \sigma_g^2 / (2\lambda)$. We substitute this expression into the above equations to conclude that

$$\text{Var}(w^D) = \delta\sigma_D^2 + \frac{\sigma_g^2}{2\lambda^3} (-3 + 2\delta\lambda + 4e^{-\delta\lambda} - e^{-2\delta\lambda}).$$

We turn now to proving the equation $g_\delta = e^{-\lambda\delta} g_0 + w^g$. Another standard fact about the process in (3) is that for all t , we may write $g_t = e^{-\lambda t} g_0 + \int_0^t \sigma_g e^{-\lambda(t-\tau)} dW_\tau^g$. Substituting $\delta = t$ yields the

equation we desire. The variance of w^g equals $\sigma_g^2 \int_0^\delta e^{-2\lambda(\delta-t)} dt$, so that

$$\text{Var}(w^g) = \frac{\sigma_g^2}{2\lambda}(1 - e^{-2\lambda\delta}).$$

The last task is to calculate the covariance of w^D and w^g , which equals $\text{Cov}(g_\delta - e^{-\delta\lambda}g_0, \sigma_D \int_0^\delta dW_t^D + \int_0^\delta g_t dt - \beta g_0)$. The dW^D term drops out because W^D is independent from W^g . The remainder can be written as the sum of four covariances. The first is $\text{Cov}(g_\delta, \int_0^\delta g_t dt) = \frac{\sigma_g^2}{2\lambda} \int_0^\delta e^{-t\lambda} dt = \frac{\sigma_g^2}{2\lambda^2}(1 - e^{-\delta\lambda})$, the next one is $\text{Cov}(g_\delta, -\beta g_0) = -\frac{\sigma_g^2}{2\lambda^2}(e^{-\delta\lambda} - e^{-2\delta\lambda})$, the third covariance equals $\text{Cov}(-e^{-\delta\lambda}g_0, \int_0^\delta g_t dt) = -e^{-\delta\lambda} \frac{\sigma_g^2}{2\lambda} \int_0^\delta e^{-t\lambda} dt = -\frac{\sigma_g^2}{2\lambda^2}(e^{-\delta\lambda} - e^{-2\delta\lambda})$, and the fourth and final covariance is $\text{Cov}(-e^{-\delta\lambda}g_0, -\beta g_0) = \frac{\sigma_g^2}{2\lambda^2}(e^{-\delta\lambda} - e^{-2\delta\lambda})$. The total covariance equals the sum of these four terms:

$$\text{Cov}(w^g, w^D) = \frac{\sigma_g^2}{2\lambda^2} (1 - e^{-\delta\lambda})^2.$$

We conclude that $(D_\delta, g_\delta)' = \mathbf{F}(D_0^c, g_0)' + \mathbf{w}$, where \mathbf{F} is the matrix given in the text, and the covariance matrix of w equals

$$\mathbf{Q} = \begin{pmatrix} \delta\sigma_D^2 + \frac{\sigma_g^2}{2\lambda^3}(-3 + 2\delta\lambda + 4e^{-\delta\lambda} - e^{-2\delta\lambda}) & \frac{\sigma_g^2}{2\lambda^2}(1 - e^{-\delta\lambda})^2 \\ \frac{\sigma_g^2}{2\lambda^2}(1 - e^{-\delta\lambda})^2 & \frac{\sigma_g^2}{2\lambda}(1 - e^{-2\delta\lambda}) \end{pmatrix}.$$

The formulas in Lemma 2 result from applying a Kalman filter to the problem specified. For an exposition of Kalman filtering, see, for instance, Hamilton (1994). We solve for the posterior conditional on observing $\{D_{i,t}\} \cup \Omega_t^a \cup \Omega_t^s \cup \Omega_t^x$; the naive buyers substitute $rp_{t'}$ for $D_{t'}$. As in the Proof of Proposition 1, we let t_0 denote the time of the most recent observation of \mathbf{x} ; the buyer ignores all data that occurs before t_0 . Let t_1 be the first time not before t_0 that the buyer at t observes. The buyer forms a posterior on \mathbf{x}_{t_1} from observing just \mathbf{x}_{t_0} as well as $D_{t_1}^a$ and $D_{t_1}^s$; this posterior is a normal distribution with mean $\hat{\mathbf{x}}_1$ and covariance \mathbf{P}_1 . As we show below, we do not need to solve for this posterior directly. At each subsequent period, the buyer learns $D_{t_k}^a$ and $D_{t_k}^s$ and iteratively applies the Kalman filter using the formulas $\hat{\mathbf{x}}_k = \mathbf{K}_k(D_{t_k}^a, D_{t_k}^s)' + (\mathbf{I} - \mathbf{K}_k\mathbf{H}_0)\mathbf{F}\hat{\mathbf{x}}_{k-1}$ and $\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_0)(\mathbf{F}\mathbf{P}_{k-1}\mathbf{F}' + \mathbf{Q})$, where $\mathbf{K}_k = (\mathbf{F}\mathbf{P}_{k-1}\mathbf{F}' + \mathbf{Q})\mathbf{H}_0'(\mathbf{H}_0(\mathbf{F}\mathbf{P}_{k-1}\mathbf{F}' + \mathbf{Q})\mathbf{H}_0' + \mathbf{R}_0)^{-1}$. Let n be such that $t_n = t - \delta$. Then the posterior $\mathbf{x}_{t-\delta} | \Omega_{i,t} \setminus \{D_{i,t}\}$ is a normal with mean $\hat{\mathbf{x}}_n$ and covariance \mathbf{P}_n . The final posterior $\mathbf{x}_t | \Omega_{i,t}$ updates this posterior based on the information in $\{D_{i,t}\}$ and the evolution of time between $t - \delta$ and t . The mean of this posterior equals $\mathbf{E}(\mathbf{x}_t | \Omega_{i,t}) = \mathbf{K}D_{i,t} + (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{F}\mathbf{E}(\mathbf{x}_{t-\delta} | \Omega_{i,t} \setminus \{D_{i,t}\})$, where $\mathbf{K} = (\mathbf{F}\mathbf{P}_n\mathbf{F}' + \mathbf{Q})\mathbf{H}'(\mathbf{H}(\mathbf{F}\mathbf{P}_n\mathbf{F}' + \mathbf{Q})\mathbf{H}' + \mathbf{R})^{-1}$. The covariance of the posterior equals $\mathbf{P} = (\mathbf{I} - \mathbf{K}\mathbf{H})(\mathbf{F}\mathbf{P}_0\mathbf{F}' + \mathbf{Q})$. Averaging the posterior mean across i gives

$$\hat{\mathbf{x}}_t = \mathbf{K}D_t^a + (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{F} \left[\sum_{m=1}^{n-1} \left(\prod_{k=1}^{m-1} (\mathbf{I} - \mathbf{K}_{n-k+1}\mathbf{H}_0)\mathbf{F} \right) \mathbf{K}_{n-m+1}(D_{t-m\delta}^a, D_{t-m\delta}^s)' + \hat{\mathbf{x}}_1 \prod_{k=1}^{n-1} (\mathbf{I} - \mathbf{K}_{n-k+1}\mathbf{H}_0)\mathbf{F} \right].$$

We now show that as $\rho \rightarrow 0$, this expression converges almost surely to the formula in the

lemma. We demonstrate pointwise convergence, and hence show convergence of the coefficients on the dividend terms. According to Proposition 13.2 of Hamilton (1994), $\lim_{n \rightarrow \infty} \mathbf{P}_n = \mathbf{P}_0$, where \mathbf{P}_0 is the unique solution to

$$\mathbf{P}_0 = (\mathbf{I} - (\mathbf{F}\mathbf{P}_0\mathbf{F}' + \mathbf{Q})\mathbf{H}'_0 (\mathbf{H}_0(\mathbf{F}\mathbf{P}_0\mathbf{F}' + \mathbf{Q})\mathbf{H}'_0 + \mathbf{R}_0)^{-1} \mathbf{H}_0)(\mathbf{F}\mathbf{P}_0\mathbf{F}' + \mathbf{Q}).$$

We may apply this proposition because Q is strictly positive definite (as it is the covariance matrix variables that are not linear combinations of each other). As written in Hamilton (1994), the proposition requires the eigenvalues of \mathbf{F} to lie inside the unit circle (which they do not, as 1 is an eigenvalue), but the proof shows that the strict positive definiteness of \mathbf{Q} is sufficient. We define $\mathbf{K}_0 = (\mathbf{F}\mathbf{P}_0\mathbf{F}' + \mathbf{Q})\mathbf{H}'_0 (\mathbf{H}_0(\mathbf{F}\mathbf{P}_0\mathbf{F}' + \mathbf{Q})\mathbf{H}'_0 + \mathbf{R}_0)^{-1}$; note that $\lim_{n \rightarrow \infty} \mathbf{K}_n = \mathbf{K}_0$.

We also claim that $\lim_{n \rightarrow \infty} \prod_{k=1}^n (\mathbf{I} - \mathbf{K}_{n-k+1}\mathbf{H}_0)\mathbf{F} = 0$. A direct computation shows that \mathbf{F} has an eigenvalue less than 1 in magnitude and an eigenvalue of 1 with eigenvector $(1, 0)'$. Similarly, for any \mathbf{K}_k , $\mathbf{I} - \mathbf{K}_k\mathbf{H}_0$ has an eigenvalue less than 1 in magnitude (and that is independent of k) and an eigenvalue of 1 with eigenvector $(0, 1)'$. As the eigenvectors with eigenvalue 1 are not collinear, and the other eigenvalues are less than 1 in magnitude, the limit is 0 as claimed.

It follows that given an error tolerance ϵ for the coefficients, we can choose n large enough so that the coefficients in the equation for $\widehat{\mathbf{x}}_t$ are within ϵ of those in the formula in the lemma. We first choose n_1 large enough so that $|((\mathbf{I} - \mathbf{K}_0\mathbf{H}_0)\mathbf{F})^m - \prod_{k=1}^m (\mathbf{I} - \mathbf{K}_k\mathbf{H}_0)\mathbf{F}| < \epsilon$ for $m \geq n_1$ and any valid \mathbf{K}_k ; as these products both converge to 0, n_1 exists. Then we choose n_0 such that for $n \geq n_0$, $|((\mathbf{I} - \mathbf{K}_0\mathbf{H}_0)\mathbf{F})^m - \prod_{k=1}^m (\mathbf{I} - \mathbf{K}_{n-k+1}\mathbf{H}_0)\mathbf{F}| < \epsilon$ for $m \leq n - n_1$; n_0 exists because $\lim_{n \rightarrow \infty} \mathbf{K}_n = \mathbf{K}_0$. Thus convergence occurs as long as $n \geq n_0$, which happens with probability at most $1 - e^{-\rho(1+n_0)\delta}$. This probability converges to 0 as $\rho \rightarrow 0$.

Proof of Proposition 3. In the limit as $N \rightarrow \infty$, the noise in $D_{t'}^a$, which is σ_a^2/N , goes to 0. Naive buyers believe that $rp_{t-\delta} = D_{t-\delta}^a = D_{t-\delta}$, so they neglect all information in $\Omega_t^p \cup \Omega_t^s \cup \Omega_t^x$ before $t - \delta$, which is all information other than $p_{t-\delta}$ and $D_{t-\delta}^s$; they ignore $D_{t-\delta}^s$ as it provides a noisy signal of $D_{t-\delta}$. This argument proves the naive formula.

We now prove the rational formula. Note that from Lemma 1, $rp_{t'} = D_{t'}r/(r + \mu) + \widehat{D}_{t'}\mu/(r + \mu)$. At $t' = t - n\delta$, the expectation is formed using only \mathbf{x}_{t_0} and $D_{i,t-n\delta}$, as older data is obviated by \mathbf{x}_{t_0} . The noise in using D_{t_0} as a measure of $D_{t-n\delta}$ is $(t - n\delta - t_0)\sigma_D^2$, and the noise in using $D_{i,t-n\delta}$ is σ_a^2 . Therefore

$$\widehat{D}_{t-n\delta} = \frac{(t - n\delta - t_0)\sigma_D^2}{\sigma_a^2 + (t - n\delta - t_0)\sigma_D^2} D_{t-n\delta} + \frac{\sigma_a^2}{\sigma_a^2 + (t - n\delta - t_0)\sigma_D^2} D_{t_0}.$$

Prices therefore are given by $rp_{t-n\delta} = (1 - \alpha_0)D_{t-n\delta} + \alpha_0 D_{t_0}$, where α_0 is as defined in the text.

At all other times t' , from Proposition 1, we know that the rational buyer can infer all observed demand. Therefore $\mathbf{E}(D_{t'-\delta} | \Omega_{t'}^p \cup \Omega_{t'}^s \cup \Omega_{t'}^x) = D_{t'-\delta}$ and there is no noise in this estimate. It follows that the posterior on $D_{t'}$ combines this estimate and $D_{i,t'}$, with weights equal to the relative variance. The variance of the lagged demand estimate is $\delta\sigma_D^2$ and the variance of the idiosyncratic estimate is σ_a^2 . Therefore $rp_{t'} = (1 - \alpha)D_{t'} + \alpha\mathbf{E}(D_{t'-\delta} | \Omega_{t'}^p \cup \Omega_{t'}^s \cup \Omega_{t'}^x)$, where α is as defined in the proposition. Hence

$$\mathbf{E}(D_{t'-\delta} | \Omega_{t'}^p \cup \Omega_{t'}^s \cup \Omega_{t'}^x) = \frac{rp_{t'-\delta} - \alpha\mathbf{E}(D_{t'-2\delta} | \Omega_{t'-\delta}^p \cup \Omega_{t'-\delta}^s \cup \Omega_{t'-\delta}^x)}{1 - \alpha}$$

for $t' > t - (n - 1)\delta$. At $t' = t - (n - 1)\delta$,

$$\mathbf{E}(D_{t-n\delta} \mid \Omega_{t-(n-1)\delta}^p \cup \Omega_{t-(n-1)\delta}^s \cup \Omega_{t-(n-1)\delta}^x) = \frac{rp_{t-n\delta} - \alpha_0 D_{t_0}}{1 - \alpha_0}.$$

Iterating the first equation until the second is employed yields the formula in the proposition.

Proof of Lemma 3. Using the terminology from the proof of Lemma 1, $\phi^* = \int_0^\infty \mu e^{-(r+\mu)\tau} (1 - \phi) e^{-\lambda\tau} d\tau = (1 - \phi)\mu / (r + \mu + \lambda)$. We then use the formula from that same proof that $A_g = [r(r + \lambda + \mu)(1 - \phi^*)]^{-1}$ to arrive at the result.

Proof of Proposition 4. Using (7) and the forecasting rules in the text, we write $\mathbf{E}_t(V_T - V_t)$ as

$$\frac{\mathbf{E}_t(D_T - D_t)}{r} + \frac{(1 - \phi)(\mathbf{E}_t g_T - \hat{g}_t)}{r(r + \lambda + \phi\mu)} = \frac{1}{r} \int_t^T \mathbf{E}_t g_\tau d\tau - \frac{(1 - \phi)(1 - e^{-\lambda(T-t)})\hat{g}_t}{r(r + \lambda + \phi\mu)},$$

which reduces to the formula in the Proposition. This expression increases in ϕ . The derivative of the first fraction is positive when $(r + \lambda + \phi\mu)(\lambda + \mu) > (r + \phi\lambda + \phi\mu)\mu$, which is true because $r + \lambda + \phi\mu > r + \phi\lambda + \phi\mu$ and $\lambda + \mu > \mu$, as $\lambda > 0$.

Proof of Lemma 4. We write $\Delta p_t = \beta_1 \Delta p_{t-\delta} - \beta_2 \Delta p_{t-2\delta} + \gamma_t + \epsilon_t$. We let $\rho_m = \text{Corr}(\Delta p_t, \Delta p_{t-m\delta})$. By taking the covariance with respect to Δp_{t-1} of the above equation and simplifying, we obtain

$$\rho_1 = \frac{\beta_1}{1 + \beta_2} + \frac{\xi \text{Cov}(\Delta p_t, \gamma_t)}{(1 + \beta_2) \text{Var}(\Delta p_t)},$$

where $\xi = e^{-\delta\lambda}$ and we have used $\text{Cov}(\gamma_t, \Delta p_{t-\delta}) = \text{Cov}(\xi\gamma_{t-\delta} + \eta_t, \Delta p_{t-\delta}) = \xi \text{Cov}(\gamma_{t-\delta}, \Delta p_{t-\delta})$. Similarly, note that $\text{Cov}(\gamma_t, \Delta p_{t-2\delta}) = \xi^2 \text{Cov}(\gamma_t, \Delta p_t)$. Therefore, by taking the covariance of the first equation with respect to γ_t and simplifying, we obtain

$$\text{Cov}(\Delta p_t, \gamma_t) = \frac{\text{Var}(\gamma_t) + \text{Cov}(\eta_t, \epsilon_t)}{1 - \beta_1 \xi + \beta_2 \xi^2},$$

where we have used $\text{Cov}(\gamma_t, \epsilon_t) = \text{Cov}(\eta_t + \xi\gamma_{t-\delta}, \epsilon_t) = \text{Cov}(\eta_t, \epsilon_t)$. This expression is nonnegative: by assumption, $\text{Cov}(\eta_t, \epsilon_t) \geq 0$, and the polynomial $1 - \beta_1 L + \beta_2 L^2$ has no roots between 0 and 1, as it is a factor of the characteristic polynomial of the stationary time series Δp_t (stationarity forces the roots to have magnitude exceeding 1). Therefore $\rho_1 > 0$ if $\beta_1 > 0$, and both terms in the expression for ρ_1 are increasing in β_1 . We turn now to the general expression for the autocorrelations ρ_m . We may write $(I - \xi L)(I - \beta_1 L + \beta_2 L^2)\Delta p_t = \eta_t + \epsilon_t - \xi\epsilon_{t-1}$. By taking the covariance of this equation with respect to $\Delta p_{t-m\delta}$ for $m > 2$, and then dividing by the variance of Δp_t , we uncover that the autocorrelations satisfy the recursion $(I - \xi L)(I - \beta_1 L + \beta_2 L^2)\rho_m = 0$. Let r_1 and r_2 be such that $(I - r_1 L)(I - r_2 L) = I - \beta_1 L + \beta_2 L^2$. These roots satisfy $r_1 + r_2 = \beta_1$ and $r_1 r_2 = \beta_2$, and are given by $r_j = (\beta_1 \pm \sqrt{\beta_1^2 - 4\beta_2})/2$ for $j \in \{1, 2\}$. These roots are complex when $\beta_1^2 < 4\beta_2$. In this case, their absolute value is given by $\sqrt{(\beta_1^2 - (\beta_1^2 - 4\beta_2))/4} = \sqrt{\beta_2}$. The autocorrelations satisfy $(I - \xi L)(I - r_1 L)(I - r_2 L)\rho_m = 0$, so $\rho_m = A_\gamma \xi^m + A_1 r_1^m + A_2 r_2^m$ for some constants A_γ , A_1 , and A_2 . As long as $\beta_2 < 1$, then $|r_1|, |r_2| < 1$ and the autocorrelations are well-defined. When $\beta_1^2 < 4\beta_2$, r_1^m and r_2^m are complex conjugates. As ρ_m is real, A_1 and A_2 must be conjugates as well. We write

$A_1 = A_\beta e^{i\omega}/2$ and $A_2 = A_\beta e^{-i\omega}/2$ with A_β real, and $r = \sqrt{\beta_2} e^{\pm i\theta}$, where $\cos(\theta) = \beta_1/(2\sqrt{\beta_2})$. Then

$$\rho_m = A_\gamma \xi^m + A_\beta \beta_2^{m/2} e^{im\theta+i\omega}/2 + A_\beta \beta_2^{m/2} e^{-im\theta-i\omega}/2 = A_\gamma e^{-m\delta\lambda} + A_\beta \beta_2^{m/2} \cos(m\theta + \omega).$$

We proved this formula for $m \geq 3$, but it holds for $m = 1$ and $m = 2$ as well, as the first three autocorrelations determine the constants A_γ , A_β , and ω .

Proof of Proposition 5. We know from Proposition 3 that the posterior on $D_{t-\delta}$ equals $rp_{t-\delta}$. For $g_{t-\delta}$, note that $D_{i,t}$ is too noisy to provide information, and from the equation $\mathbf{x}_t = \mathbf{F}\mathbf{x}_{t-\delta} + \mathbf{w}_t$, we have $\Delta D_{t-\delta} = g_{t-2\delta}(1 - e^{-\delta\lambda})/\lambda + w_{t-2\delta}^D$. Therefore, as the naive buyer sets $rp_{t'} = D_{t'}$, her posterior on the growth rate is $\mathbf{E}(g_{t-\delta} \mid \Omega'_{i,t}) = r\Delta p_{t-\delta}\lambda e^{-\delta\lambda}/(1 - e^{-\delta\lambda})$. To arrive at the contemporaneous estimates, we again use the law of motion for \mathbf{x} and get $\mathbf{E}(D_t \mid \Omega'_{i,t}) = rp_{t-\delta} + e^{-\delta\lambda}r\Delta p_{t-\delta}$ and $\mathbf{E}(g_t \mid \Omega'_{i,t}) = r\Delta p_{t-\delta}\lambda e^{-2\delta\lambda}/(1 - e^{-\delta\lambda})$. Substituting these expressions into (5) and differencing yields the equation in the Proposition. The growth rate is an AR(1) as $g_t = e^{-\delta\lambda}g_{t-\delta} + w_t^g$, and $\text{Cov}(w_t^g, w_t^D) > 0$ by the formula for \mathbf{Q} , so the innovation constitutes an AR(1) plus noise under the definition in Lemma 4. Finally, to show that the condition $(\beta_1/2)^2 < \beta_2 < 1$ holds for some parameters, use $\delta = 1$, $\lambda = 1$, $r = 0.04$, $\mu = 0.075$, and $\phi = 1$, which yields $\beta_1 = 1.08$ and $\beta_2 = 0.43$.

Identification of Demand Parameters. We observe the following three covariances from the data: $\gamma_0 = \text{Var}(\Delta D_t)$, $\gamma_1 = \text{Cov}(\Delta D_t, \Delta D_{t-1})$, and $\gamma_2 = \text{Cov}(\Delta D_t, \Delta D_{t-2})$. These identify σ_D , σ_g , and λ as follows. First, note that $D_{t+1} = D_t + g_t(1 - e^{-\lambda})/\lambda + w_t^D$, where w^D is the error defined in the proof of Lemma 2. This equation comes from applying the law of motion for \mathbf{x}_t , which is $\mathbf{x}_{t+1} = \mathbf{F}\mathbf{x}_t + \mathbf{w}_t$, where \mathbf{w}_t and \mathbf{x}_t are independent and the covariance matrix of \mathbf{w} is \mathbf{Q} . As a result, $\Delta D_t = g_t(1 - e^{-\lambda})/\lambda + w_t^D$. It follows that the variance of the yearly price change equals

$$\gamma_0 = \left(\frac{1 - e^{-\lambda}}{\lambda}\right)^2 \frac{\sigma_g^2}{2\lambda} + \text{Var}(w_t^D) = \sigma_D^2 + \frac{\sigma_g^2}{\lambda^3}(e^{-\lambda} - 1 + \lambda).$$

The first covariance of price changes equals

$$\gamma_1 = \text{Cov}\left(\frac{1 - e^{-\lambda}}{\lambda}g_t + w_t^D, \frac{1 - e^{-\lambda}}{\lambda}(e^{-\lambda}g_t + w_t^g) + w_{t+1}^D\right) = \frac{(1 - e^{-\lambda})^2\sigma_g^2}{2\lambda^3}.$$

The second covariance equals

$$\gamma_2 = \text{Cov}\left(\frac{1 - e^{-\lambda}}{\lambda}g_t + w_t^D, \frac{1 - e^{-\lambda}}{\lambda}(e^{-\lambda}(e^{-\lambda}g_t + w_t^g) + w_{t+1}^g) + w_{t+2}^D\right) = \frac{e^{-\lambda}(1 - e^{-\lambda})^2\sigma_g^2}{2\lambda^3}.$$

These three equations identify the parameters. Note that $\gamma_2/\gamma_1 = e^{-\lambda}$, so this ratio determines λ . Conditional on λ , γ_1/γ_0 uniquely determines the ratio σ_g/σ_D . Finally, γ_0 pins down the level of these volatilities.