Diagnostic Bubbles

Pedro Bordalo, Nicola Gennaioli, Spencer Yongwook Kwon, Andrei Shleifer[1]

September 2019

Abstract

We introduce diagnostic expectations into a standard setting of price formation in which investors learn about the fundamental value of an asset and trade it. We study the interaction of diagnostic expectations with two well-known mechanisms: learning from prices and speculation (buying for resale). With diagnostic (but not with rational) expectations, these mechanisms lead to price paths exhibiting three phases: initial underreaction, followed by overshooting (the bubble), and finally a crash. With learning from prices, the model generates price extrapolation as a byproduct of fast moving beliefs about fundamentals, which lasts only as the bubble builds up. When investors speculate, even mild diagnostic distortions generate substantial bubbles.

[1] Saïd Business School, University of Oxford, Bocconi University and IGIER, Harvard Business School and Harvard University. Gennaioli thanks the European Research Council (GA 647782) for financial support. We are grateful to Nicholas Barberis, Lawrence Jin, Josh Schwartzstein, Alp Simsek, and an anonymous referee for helpful comments.

1. Introduction

The financial crisis of 2007-2008 has revived academic interest in price bubbles. Shiller (2015) created a famous graph of home prices in the United States over the course of a century, which shows prices being relatively stable during most of the 20th century, then doubling over the ten year period after 1996, only to collapse in the crisis and begin recovering after 2011. There is also growing evidence of speculation such as buying for resale in the housing market (DeFusco, Nathanson, and Zwick 2018, Mian and Sufi 2018) and of increasing leverage of both homeowners and financial institutions tied to rapid home price appreciation. The collapse of the housing bubble is at the heart of every major narrative of the financial crisis and the Great Recession because it entailed massive losses for homeowners, holders of mortgage backed securities, and financial institutions (Mian and Sufi 2014). Nor is the U.S. experience unique. Leverage expansions and subsequent crises are often tied to bubbles in housing and other markets (Jorda, Schularick and Taylor 2015).

Despite the revival of academic interest, asset price bubbles remain controversial. Although economic historians tend to see bubbles as self-evident (Mackay 1841, Bagehot 1873, Galbraith 1954, Kindleberger 1978, Shiller 2000), Fama (2014) raised the critical question of whether they even exist in the sense of predictability of future negative returns after prices have increased substantially. Interestingly, since Smith, Suchanek and Williams (1988), predictably negative returns are commonly found in laboratory experiments even when markets have finite horizons. Greenwood, Shleifer, and You (2019) address Fama's challenge in industry stock return data around the world, and find that although past returns alone are very noisy indicators of bubbles, other measures of over-pricing do forecast poor returns going forward[2].

Early theoretical research has focused on rational price bubbles that do not violate (some definitions of) market efficiency (Blanchard and Watson 1982, Tirole 1985, Martin and Ventura 2012), but these models are not consistent with the available evidence on prices (Giglio, Maggiori, and Stroebel 2016). They are also rejected by the striking evidence of excessively optimistic investor expectations in bubble episodes (Case et al. 2012, Greenwood, Shleifer, and You 2019), which also shows up in experimental data (e.g., Haruvy et al. 2007). Because of such evidence, research has moved to behavioral models of bubbles, which emphasize

[2] A closely related literature examines the overpricing of small growth stocks with extremely optimistic analyst forecasts of future growth, and predictably poor returns (Lakonishok et al. 1994, La Porta 1996, Bordalo et al. 2019).

factors such as overconfidence and short sales constraints (Scheinkman and Xiong 2003), neglect of entry (Glaeser et al 2008, Glaeser and Nathanson 2017, Greenwood and Hanson 2014), and price extrapolation (e.g., Cutler et al. 1990, DeLong et al 1990b, Barsky and DeLong 1993, Hong and Stein 1999, Barberis and Shleifer 2003, Hirshleifer et al. 2015, Glaeser and Nathanson 2017, Barberis et al. 2015, 2018).

The focus on price extrapolation neglects one further potentially important driver of bubbles: excessively optimistic beliefs about the fundamental value of an asset. This mechanism features prominently in historical accounts of bubbles, which stress the euphoria accompanying new technologies such as the railroads and the internet. This paper explores the connection between excessive optimism about fundamentals and asset price bubbles. We introduce non-rational beliefs about fundamentals in an otherwise standard asset pricing model and ask three questions. First, when does excessive optimism arise and how far can it go in accounting for inflated asset prices? Second, can this mechanism create self-reinforcing price growth when traders seek to learn fundamental values from prices themselves? And third, how does optimism about fundamentals interact with traders' speculation for quick resale in sustaining price growth?

To address these questions, we introduce diagnostic expectations (see Bordalo, Gennaioli, and Shleifer, hereafter BGS 2018, Bordalo, Gennaioli, La Porta and Shleifer, hereafter BGLS 2019, and Bordalo, Gennaioli, Ma, and Shleifer, hereafter BGMS 2019), into a standard finite horizon model of a market for one asset, in which a continuum of investors receive noisy private information every period about the termination value of that asset. The asset is valuable (above the prior) so that traders on average receive good news about fundamentals every period. Because traders receive different noisy signals, they hold heterogeneous beliefs. This generates trading volume, which is an important feature of bubbles (Scheinkman and Xiong 2003, Hong and Stein 2007). Unlike some previous models, such as Harrison and Kreps (1978) and Scheinkman and Xiong (2003), we do not need to assume short sale constraints. In this setup, we assess how distorted beliefs about fundamentals interact with two key mechanisms: learning from prices and speculation.[3] With rational expectations, in this model the average price path rises from the prior to the fundamental value, without overshooting. There are no price bubbles in equilibrium.

[3] We do not examine leverage and other factors that link the collapses of bubbles to financial fragility and economic recessions (see, e.g., Gennaioli, Shleifer, and Vishny 2012, Reinhart and Rogoff 2009, Gennaioli and Shleifer 2018). We leave the analysis of the role of leverage to future work.

Under diagnostic expectations, traders update their beliefs excessively in the direction of the states of the world whose objective likelihood has increased the most in light of recent news. After good news, right-tail outcomes become representative and are overweighed in expectations, while left-tail outcomes become non-representative and are neglected. But how does this affect the price of the asset over time?

We find that, in line with Kindleberger's (1978) narrative, the equilibrium price evolves over three stages. The first stage is displacement. A beneficial economic innovation (see, e.g, Pastor and Veronesi 2006), entails a sequence of good fundamental news. Such good news eventually leads to excess optimism. As traders grow more optimistic, the demand for the asset soars, creating a sustained price increase.

The second and crucial stage is the acceleration of price growth. In this stage, price increases themselves encourage buying that leads to further price increases, and prices reach levels substantially above fundamental values. This second stage is not due to diagnostic expectations alone, but to their interaction with learning from prices and speculation. Consider learning from prices first. As good news arrives and prices rise, diagnostic traders act more aggressively on their private signals, which makes prices more informative than under the rational benchmark. As a consequence, traders react even more aggressively to prices, which quickly swamp the less informative private signals. Through rising prices, the over-optimism of some traders infects the entire market, causing price growth to accelerate, resulting in convex price paths. It looks like investors are extrapolating price trends, even though they are not. Instead, recent price increases lead traders to upgrade (too much) their expectations of fundamental value, and thus of the future price.

Speculation, when interacted with diagnostic expectations, adds fuel to the bubble. Buying for resale compounds overreaction: traders are not only too optimistic about fundamentals, but also exaggerate the possibility of reselling the asset to traders who are even more optimistic than them. Following a beauty context logic, a trader who believes that the asset is the next Google thinks that future diagnostic traders will receive extremely positive signals and will thus be even more optimistic about the asset. The expectation of reselling the asset to very bullish traders further inflates the price today. As a result, even a small degree of diagnosticity compounds into strong price extrapolation and large price dislocations.

The third and final phase is the collapse of the bubble. Under diagnostic expectations, two mechanisms are responsible for the crash. First, because good news become increasingly marginal over time, they cannot sustain the extent of over-reaction. The bubble collapses not because of bad news, but because traders' over-reaction to such news eventually runs out of steam. Second, as the terminal date approaches, there are fewer opportunities to resell to an overly optimistic trader. This reduces the current demand for the asset, causing the bubble to collapse today.

Our analysis shows how diagnostic expectations interact with learning from prices and speculation to create fundamentals-based boom bust price dynamics. The spreading of optimism about fundamentals through the market endogeneously generates, as a by-product, key previously recognized features of bubbles, such as a form of price extrapolation, growing trading volume, and high price volatility near the peak.

The paper proceeds as follows. In Section 2 we present the model of diagnostic expectations and show its implications for the dynamics of beliefs, absent any market mechanism. In Section 3 we isolate the role of learning from prices by introducing diagnostic expectations into a standard Grossman-Stiglitz setup in which traders there have no speculative motives. In Section 4 we isolate the role of speculation by introducing diagnostic expectations into a standard beauty context model in which learning from prices is absent. In Section 5 we present and simulate the full model with diagnostic expectations, learning from prices, and speculation. Section 6 concludes.

2. Learning from Good Shocks and Diagnostic Expectations

Traders learn about the value of a new asset over a finite number of periods $t = 0, \dots, T$. The asset yields a payoff V. The value of V is drawn from a normal distribution with mean 0 and variance σ_V^2 at $t = 0$ but it is only revealed at the terminal date T. In line with Kindleberger's (1978) description of a positive displacement as the trigger of bubble episodes, we focus on the case of a valuable innovation, $V > 0$. Our model is entirely symmetric for $V < 0$, in which case negative bubbles arise. Symmetry naturally breaks down if one introduces short sales constraints, which we abstract from here.

In each period t, each trader i (in measure one) receives a private signal $s_{it} = V + \epsilon_{it}$ of the asset's value. Noise ϵ_{it} is i.i.d. across traders and over time, and normally distributed with mean zero and variance σ_ϵ^2. Because the new asset is valuable, $V > 0$, traders are repeatedly exposed to good news, in that signals are on average positive relative to their priors, capturing the initial displacement. Moreover, the assumption of dispersed information generates variation in expectations and creates a motive for trading.

In this Section, we do not consider trading, and describe instead how Diagnostic Expectations about fundamentals evolve solely based on the arrival of noisy private signals. This is useful for two reasons. First, in this setting with learning and dispersed information, Diagnostic Expectations behave differently than in prior finance applications (BGS 2018, BGLS 2018). Second, by separately characterizing the dynamics of expectations about fundamentals, we can better understand their interaction with market forces such as trading, learning from prices, and speculation, which we introduce in Sections 3 and 4.

A rational trader observing a history of signals $(s_{i1}, \dots, s_{it})$ forms an expectation about V given by:

$$\mathbb{E}_{it}(V) = \pi_t \frac{\sum_{r=1}^{t} s_{ir}}{t},$$

where $\pi_t \equiv \frac{t/\sigma_\epsilon^2}{t/\sigma_\epsilon^2 + 1/\sigma_V^2}$ is the signal to noise ratio. The consensus expectation under rationality is given by

$$\int \mathbb{E}_{it}(V) di = \pi_t V, \tag{1}$$

lower than the full information benchmark. The rational consensus has two important properties. First, it is always below the rational benchmark because $\pi_t \leq 1$. Rational traders discount their noisy signals, which implies that average information, always equal to V, is also discounted. Second, the rational consensus gradually improves over time, because the signal to noise ratio π_t rises, in a concave way, toward one. As traders see more signals, their uncertainty falls, inducing them to weigh their evidence more heavily.

As in rational inattention models (Woodford 2003), optimal information processing by individuals who observe noisy signals creates sluggishness in consensus expectations. This sluggishness is central in thinking about price formation in rational expectations models. As we show in Sections 3 and 4, even with learning from prices and speculation, rational updating causes the consensus expectation as well as the price of the asset to under-react to the fundamental value V, monotonically converging to it from below.

Consider now updating under diagnostic expectations (DE). This model of belief formation captures Kahneman and Tversky's (1972) representativeness heuristic.[4] Representativeness refers to the notion that, in forming probabilistic assessments, decision makers put too much weight on outcomes that are likely not in absolute terms, but rather relative to some reference or baseline level. For example, many people significantly over-estimate the probability that a person's hair is red when told that the person is Irish. The share of red-haired Irish, at 10%, is a small minority, but red hair is much more common among the Irish than among other Europeans, let alone in the world as a whole. The over-estimation of the prevalence of representative types distorts beliefs and accounts for many systematic errors in probabilistic judgments documented experimentally (Gennaioli and Shleifer 2010). It also delivers a theory of stereotypes consistent with both field and experimental evidence, including gender stereotypes in assessments of ability (Bordalo et al. 2018), racial stereotypes in decisions about bail (Arnold, Dobbie and Yang 2018), and popular beliefs about immigrants (Alesina, Miano, and Stantcheva 2018).

In an intertemporal setting like the current one, DE capture the idea that investors overweight the probability of events that have become more likely in light of recent news. For instance, after observing a period of positive earnings growth, DE overweight the probability that the firm may be the next Google. This event is highly unlikely in absolute terms, but it has become more likely in light of the strong earnings growth. As a consequence, the perceived probability of such an event is inflated.

To see this formally, consider an agent forecasting at t random variable X_{t+1}. As shown in BGS (2018), if X_{t+1} is conditionally normal, the diagnostic distribution of beliefs is also normal.[5] Furthermore, if $\mathbb{E}_s(X_{t+1})$ is the rational expectation at time s, then the diagnostic expectation at time t is:

$$\mathbb{E}_t^{\theta}(X_{t+1}) = \mathbb{E}_{t-k}(X_{t+1}) + (1+\theta)[\mathbb{E}_t(X_{t+1}) - \mathbb{E}_{t-k}(X_{t+1})], \quad (2)$$

[4] According to Kahneman and Tversky (1972), the reliance on representativeness as a proxy for likelihood is a central feature of probabilistic judgments. An outcome "*is representative of a class if it's very diagnostic*", that is, if its "*relative frequency is much higher in that class than in the relevant reference class*" (Tversky and Kahneman 1983).

[5] Following BGS (2018), in a Markov process with density $f(X_{t+1}|X_t)$, after a particular current state $\hat{X}_t$ is realized and in light of the past expectation for the current state $\mathbb{E}_{t-k}(X_t)$, the distorted distribution is equal to:

$$f^{\theta}(X_{t+1}|X_t) = f(X_{t+1}|X_t)\left[\frac{f(X_{t+1}|\hat{X}_t)}{f(X_{t+1}|\mathbb{E}_{t-k}(X_t))}\right]^{\theta} Z_t,$$

where both the target $f(X_{t+1}|\hat{X}_t)$ and reference distribution $f(X_{t+1}|\mathbb{E}_{t-k}(X_t))$ have equal variance. Roughly speaking, beliefs overweigh states that have become more likely in light of the surprise $\hat{X}_t - \mathbb{E}_{t-k}(X_t)$ relative to k periods ago.

DE are forward looking: they update in the correct direction and nest rational expectations as a special case for $\theta = 0$. Crucially, however, DE overreact to information by exaggerating the difference between current conditions $\mathbb{E}_t(X_{t+1})$ and normal conditions $\mathbb{E}_{t-k}(X_{t+1})$ by a factor of $(1+\theta)$. As good news arrives, the right tail of X_{t+1} becomes fatter; while still unlikely, it is very representative because its prior probability was so low. As a result, investors overweight the right tail and neglect the risk in the left tail. For normal distributions, this reweighting results in an excessive rightward shift of the believed mean.

In Equation (2), lag k defines which recent news investors over-react to. For $k = 1$, investors only overreact to news received in the current period. For $k > 1$, the investor over-reacts to the last k pieces of news. This captures a realistic sluggishness in the perception of normal conditions: new evidence becomes normal only after enough time has passed. Put differently, the investor observing a sequence of good news takes a while to adapt to them.

BGLS (2018) show that in a setting where traders learn from homogeneous information, DE obtained from Equation (2) account well for the link between listed firms' performance and equity analysts' expectations of their future earnings growth, as well as, crucially, for the link between expectations and the predictability of their stock returns. They estimate the model and find that, with quarterly data, $\theta \approx 1$ and $k \approx 3$ years. A similar value of θ has been estimated using expectations of credit spreads by BGS, and using macroeconomic forecasts by BGMS. Later we use $k \approx 3$ years in our simulation exercises.

The case $k > 1$ is not only quantitatively, but also qualitatively different from the case $k = 1$. When $k = 1$ the law of iterated expectations holds, in the sense that $\mathbb{E}_t^{\theta}(X_{t+s}) = \mathbb{E}_t^{\theta}\left[\mathbb{E}_{t+1}^{\theta}(X_{t+s})\right]$. When $k > 1$, the law of iterated expectations fails, so it matters whether expectations are computed following the "short route", as a sequence of one step diagnostic forecasts, or the "long route", as a long term diagnostic forecast. We assume that expectations are computed following the "long route", so that the expectation of terminal value at time t is computed as $\mathbb{E}_t^{\theta}(V)$. This is intuitive and analytically convenient. We believe, however, that results are qualitatively similar if the short route is followed.

The assumption that $k > 1$ is important. It implies that investors over-react to news accrued over several periods, and thus allows for persistent over-valuation of the asset. When $k = 1$ investors over-react only to the most recent news, so over-valuation is temporary.[6] Persistent over-valuation is consistent with the evidence on price bubbles, which display a sustained buildup. It also squares with the evidence from the cross section of stock returns and beliefs about fundamentals (BGLS 2019). Evidence from both beliefs and asset prices also points to violations of the law of iterated expectations. BGMS (2019) and D'Arienzo (2019) show that analyst expectations about long term interest rates over-react more than their expectations about short term rates. Giglio and Kelly (2017) document that long term rates are excessively volatile relative to short term rates. In a conventional affine term structure model, D'Arienzo (2019) shows that these findings point to a violation of the law of iterated expectations. Indeed, because over-reaction is stronger at long horizons, computing expectations over the long vs short route will typically yield different results.

Relative to the earlier finance applications of DE, and in particular relative to BGLS (2019), the current model introduces two new ingredients. First, each trader observes a different noisy signal of the truth V. Second, the state V does not change over time, reflecting learning about, say, the potential of a new technological innovation. This perspective is central to thinking about bubbles.

Given the heterogeneity of information at time t each trader i has a different diagnostic expectation $\mathbb{E}_{it}^{\theta}(V)$. As before, we focus on the consensus diagnostic expectation. This is given by:

$$\int \mathbb{E}_{it}^{\theta}(V)di = \begin{cases} (1+\theta)\pi_t V & for \quad t \leq k \\ [\pi_t + \theta(\pi_t - \pi_{t-k})]V & for \quad t > k \end{cases}. \tag{3}$$

Equation (3) implies that, under DE, consensus expectations exhibit boom bust dynamics.

Proposition 1 *If $\theta \in \left(\frac{1}{k}\frac{\sigma_\epsilon^2}{\sigma_v^2}, \frac{\sigma_\epsilon^2}{\sigma_v^2}\right)$, the consensus diagnostic expectation $\mathbb{E}_t^{\theta}(V)$ exhibits three phases:*

1) Delayed over-reaction: $\mathbb{E}_t^{\theta}(V)$ starts below V, then increases to its peak $\mathbb{E}_t^{\theta}(V) = (1+\theta)\pi_k V$ at $t = k$.

2) Bust: There is a time $t^* > k+1$ such that $\mathbb{E}_t^{\theta}(V)$ falls between $t = k+1$ and t^*, reaching its minimum of $\mathbb{E}_{t^*}^{\theta}(V) < V$. The length of the bust phase, $t^* - k$, is increasing in θ.

3) Recovery: $\mathbb{E}_t^{\theta}(V)$ gradually recovers for $t > t^*$, asymptotically converging to the fundamental V.

[6] In particular, in the current setting overvaluation would be strongest in the first period, when the news is biggest and then monotonically decline over time.

In a noisy environment, adding a modicum of overreaction θ to recent signals upsets the monotone convergence of rational expectation, yielding rich beliefs dynamics. Early on, consensus opinion under-reacts to the fundamental displacement, $\mathbb{E}_t^\theta(V) < V$, so that in this range the behavior of DE is qualitatively similar to rational learning. Because diagnostic traders are forward looking, they discount the noise in their signals. Initially, uncertainty about V is large, so this discounting is sufficiently strong that it counteracts the tendency of each individual to overreact (as long as overreaction is moderate, $\theta < \frac{\sigma_\epsilon^2}{\sigma_v^2}$). As a result, in early stages, consensus beliefs about V improve sluggishly, gradually incorporating the good signals of traders.

The possibility that in a noisy environment individual over-reaction is consistent with sluggishness of consensus expectations is not just theoretical. BGMS (2018) document this phenomenon in professional forecasts of macroeconomic variables.

As traders receive good signals, however, they grow more confident about the value of the asset. As a result, they incorporate their signals more aggressively into their beliefs. At some point, their signal to noise ratio π_t becomes sufficiently high that, for a given amount of diagnosticity θ we have:

$$(1+\theta)\pi_t > 1.$$

The condition $\theta > \frac{1}{k}\frac{\sigma_\epsilon^2}{\sigma_v^2}$ ensures that this occurs at least at the peak, when $t = k$. This implies that at some point consensus under-reaction turns to overreaction. Displacement causes traders to be so confident that beliefs overshoot the fundamental, $\mathbb{E}_t^\theta(V) > V$. Overshooting of fundamentals stands in stark contrast not only to the rational benchmark but also to any model of mis-specified learning in which beliefs are a convex combination of priors and new information, including overconfidence (Daniel, Hirshleifer and Subrahmanyam 1998).[7] This distinctive feature reflects the fact that diagnostic expectations generate disproportional and asymmetric weight on tail events: if traders focus on the right tail and neglect the left tail, then in some sense "the sky is the limit": sufficiently many good signals about V bring to mind stratospheric values. Each fast-growing firm is believed to be a new Google, and trees are expected to grow

[7] Overconfidence is a different form of over-reaction to private news in which traders exaggerate the precision of their private signals. It implies an inflated signal to noise ratio relative to rational expectations, but still be below 1. This generates a price path lying above the rational benchmark, but never overshooting or exhibiting a boom-bust path.

to the sky. This is in line with standard narratives of bubbles, in which displacement leads investors to believe in a "paradigm shift" capturing the most optimistic scenarios that could result from the innovation.[8]

Beliefs revert after k periods, when over-reaction to early signals wanes. After a while, traders view these signals as normal and focus on the information contained in the new, most recent, signals. Because these signals have a smaller and smaller incremental value (V is finite), they cannot sustain the exorbitant optimism of the boom. As a result, beliefs start deflating. The bust here is due not to bad news, but to the declining pace at which good news arrive, which causes optimism to run out of steam. After the bust, when expectations reach their trough and get close to rational beliefs, over-reaction to good news is negligible (because good news are minor), and the consensus converges to V from below.

The condition $\theta \in \left(\frac{1}{k}\frac{\sigma_\epsilon^2}{\sigma_v^2}, \frac{\sigma_\epsilon^2}{\sigma_v^2}\right)$ entailing this path for consensus beliefs is intuitive. If diagnosticity is very strong, $\theta > \frac{\sigma_\epsilon^2}{\sigma_v^2}$, the consensus opinion is excessively optimistic from the start at $t = 1$. Here the initial under-reaction phase is absent, but boom bust dynamics are preserved. If instead diagnosticity is very weak, $\theta < \frac{1}{k}\frac{\sigma_\epsilon^2}{\sigma_v^2}$, over-shooting never occurs. In this case, the consensus belief never exhibits a bubble; it slowly converges to fundamentals from below (see the proof of Proposition 1 for details).

In sum, by introducing some over-reaction to recent news in an otherwise standard noisy information model, diagnostic expectations can account for initial rigidity of consensus expectations, delayed over-reaction of beliefs to fundamental news, and subsequent reversals as dramatic good news stop coming. This mechanism seems promising for thinking about bubbles. Insofar as prices reflect consensus beliefs, diagnostic expectations may account for sluggish boom bust price dynamics that cannot be obtained under rationality. Still, some features of Equation (3) are hard to reconcile with bubbles. First, expectations of fundamentals improve in a concave way, which is hard to square with the observed convex price paths during bubbles (Greenwood, Shleifer, and Yang 2019). Second, for realistic parameter values over-optimism about fundamentals is small relative to the price inflation observed in bubble episodes. Using $\theta = 1$ as estimated using expectations of earnings growth (BGLS 2019), and of macroeconomic time series (BGMS

[8] In Pastor and Veronesi (2009), a successful innovation is not initially overpriced, but instead becomes central enough to the economy that the risk associated with it becomes systematic, which in turn depresses prices.

2019), suggests that valuation at the peak is bounded above by $2V$. In some historical episodes, such as the internet bubble, prices reached multiple times the plausible measures of fundamentals.

The key question is whether the more realistic feature of price behavior during bubbles can be obtained once diagnostic beliefs are combined with standard market mechanisms such as learning from prices and speculative trading. In the rest of the paper we show that this is indeed the case. Price growth becomes convex and disconnected from fundamentals, and the bubble can become very large even for small degrees of diagnosticity θ. To illustrate these results most clearly, we break down the analysis. In Section 3 we consider learning from prices but abstract from speculative motives. In Section 4 we consider speculation without learning from prices. In Section 5 we combine the two ingredients.

3. Diagnostic Learning from Prices

We now analyze trading and price formation. Learning from prices enables traders to extract from price changes the information that other traders have about fundamentals. Starting with Grossman (1976) and Grossman and Stiglitz (1980), learning from prices plays an important role in formal analyses of rational expectations equilibria in financial markets. Here we study its consequences under diagnostic expectations.

At each $t = 0,1 \ldots$, traders exchange the asset and determine its price. They learn about the fundamental V from current and past private signals as well as from all prices observed up to the last period.[9] Traders are risk averse with CARA utility $u(c) = -e^{-\gamma c}$, and they have long horizons, in that they value the asset by their assessment of its fundamental value V. There is no time discounting.

To determine the demand for the asset at time t, suppose that trader i believes that V is normally distributed with mean $\mathbb{E}_{it}^{\theta}(V)$ and variance $\sigma_t^2(V)$, where θ denotes diagnostic expectations.[10] With exponential utility and normal beliefs, his preferences are described in terms of mean and variance. Trader i's demand D_{it} of the asset maximizes the mean-variance objective function:

[9] If investors learn also from the current price, the equilibrium may fail to exist for $t > k$. If the equilibrium exists, the model behaves very similarly to our benchmark case here. To study speculation, because we set k to be high, existence is guaranteed and so we allow investors to learn from the current price.

[10] As we discussed in Section 2, diagnostic beliefs are normal. This is shown in the appendix, where we also show that the variance $\sigma_t^2(V)$ is not distorted under our specification.

$$D_{it} = \underset{\widehat{D}_{it}}{\text{argmax}}[\mathbb{E}^{\theta}_{it}(V) - p_t]\widehat{D}_{it} - \frac{\gamma}{2}\,\sigma^2_t(V)\widehat{D}^2_{it},$$

where γ captures risk aversion. Trader i's demand D_{it} is then given by:

$$D_{it} = \frac{\mathbb{E}^{\theta}_{it}(V) - p_t}{\gamma\sigma^2_t(V)}, \tag{4}$$

Intuitively, demand increases in the difference between the trader's valuation and the market price.

To make learning from prices non-degenerate, we follow the literature and assume that that the supply S_t of the asset is random, i.i.d. normal with mean zero and variance σ^2_S (without common supply shocks, V is learned in one period by the law of large numbers). The classical justification here is the presence of noise or liquidity traders who demand/supply the assets for non-fundamental reasons (Black 1986, Grossman and Miler 1988, DeLong et al. 1990a). The implication is that price is no longer fully revealing: high price today may be due either to a low unobserved supply S_t shock or to a high average private signal V.

By aggregating individual demands in Equation (4) and by equating to supply we find:

$$p_t = \int \mathbb{E}^{\theta}_{it}(V)di - \gamma\sigma^2_t(V)S_t. \tag{5}$$

To solve for the equilibrium in Equation (5), we must compute the diagnostic consensus expectation at time t, recognizing that it depends both private signals and past prices. Because DE are forward looking, it is possible to amend the consensus beliefs in Equation (3) to reflect diagnostic learning from prices. As in rational expectations models (Grossman and Stiglitz 1980), we first conjecture that, at each time t, price is a linear function of the state variables of the economy, which include the fundamental V. Second, we assume that traders use this linear rule to make inferences about V in light of the current and past prices. Third, we determine at each time t the coefficients of the pricing function that equilibrate demand and supply, so that the resulting rule yields the equilibrium price.

Denote by $\mathbb{E}(V|P_t)$ the rational expectation of V based solely on the history of prices up to t, formally $P_t \equiv (p_1, \dots, p_{t-1})$. Then, our conjectured pricing rule takes the form:[11]

$$p_t = a_{1t} + a_{2t}\mathbb{E}(V|P_t) + a_{3t}\mathbb{E}(V|P_{t-k}) + b_t\left(V - \frac{c_t}{b_t}S_t\right). \tag{6}$$

Equation (6) is reminiscent of rational expectations models. The current price reflects consensus expectations derived from all prices up to date t, as well as the average private signal. Because diagnostic expectations combine current and lagged rational forecasts, the lagged forecast is also added as a state variable.

To solve for the diagnostic expectations equilibrium (DEE), we must find the coefficients $(a_{1t}, a_{2t}, a_{3t}, b_t, c_t)_{t\geq 1}$ that equate supply with demand when traders make diagnostic inferences from prices. We now sketch the logic of the result, and leave a fuller account to the proof in Appendix A.

First, consider how traders learn in light of the pricing rule. Because diagnostic traders over-react to news, they over-react also to the shared news coming from prices. To compute the diagnostic expectations with learning from prices, we proceed in two steps. We compute the rational expectations when prices are generated by Equation (6), and then apply the diagnostic transformation of Equation (2).[12]

The rational news conveyed by price at time r is captured by the term $p_r - a_{1r} - a_{2r}\mathbb{E}(V|P_r) - a_{3r}\mathbb{E}(V|P_{r-k})$ as well as by the coefficients b_r and c_r, which are known by all (in equilibrium). That is, observing p_r effectively endows all traders with the following unbiased public signal of V:

$$s_r^p = V - \left(\frac{c_r}{b_r}\right)S_r. \tag{7}$$

The variance of the signal, which is the inverse of its precision, is equal to $(c_r/b_r)^2\sigma_S^2$. Intuitively, the price is more informative when it is more sensitive to the persistent fundamental than to the transient supply shock, namely when c_r/b_r is lower. This price sensitivity is an endogenous part of the equilibrium, and later we characterize it in terms of primitives.

---

[11] The price could equivalently be assumed to be linear in the diagnostic expectation $\mathbb{E}^\theta(V|P_t, s_{i1} = \cdots = s_{it} = 0)$ held by traders held solely on the basis of prices (i.e. assuming that private signals are neutral).

[12] In the appendix we show that the diagnostic expectations so obtained are equivalent to those obtained by applying the distorted distribution of footnote 4 to the true distribution $f(V|s_{i1},,, s_{it}, P_t)$ conditional on private signals and prices.

Since private and public signals s_{it} and s_t^p are normally distributed, conditional on a history of signals $\left(s_{i1}, \dots s_{it}, s_1^p, \dots s_t^p\right)$, a rational trader's beliefs about V are normal with mean:

$$\mathbb{E}_{i,t}(V) = \bar{s}_{i,t} z_t + \mathbb{E}(V|P_t)(1 - z_t), \tag{9}$$

where $\bar{s}_{i,t} = \frac{1}{t}\sum_{r=1}^{t} s_{ir}$ and $\mathbb{E}(V|P_t) = \left(\frac{1}{\sigma_V^2} + \sum_{r=1}^{t} \frac{1}{\sigma_S^2}\left(\frac{b_r}{c_r}\right)^2\right)^{-1} \sum_{r=1}^{t} \frac{1}{\sigma_S^2}\left(\frac{b_r}{c_r}\right)^2 s_r^p$. The rational inference combines the private signals with the price signals embodied in $\mathbb{E}(V|P_t)$. The weight z_t attached to the private signals is given by:[13]

$$z_t = \frac{t}{\sigma_\epsilon^2}\left[\frac{t}{\sigma_\epsilon^2} + \frac{1}{\sigma_{p,t}^2(V)}\right]^{-1}, \tag{10}$$

where $\sigma_{p,t}^2(V)$ is the variance of the fundamental conditional only on prices. The weight z_t attached to private signals is higher when the informativeness of prices is low (when $\sigma_{p,t}^2(V)$ is high).

To compute diagnostic consensus beliefs, we need to: i) compute diagnostic beliefs by transforming rational beliefs in Equation (10) according to Equation (2), and ii) aggregate the resulting beliefs into the consensus. The implied consensus dynamics works as follows:

$$\int \mathbb{E}_{it}^{\theta}(V)di = \begin{cases} (1+\theta)[(1-z_t)\mathbb{E}(V|P_t) + Vz_t] & for \quad t \leq k \\ (1+\theta)(1-z_t)\mathbb{E}(V|P_t) - \theta(1-z_{t-k})\mathbb{E}(V|P_{t-k}) + [(1+\theta)z_t - \theta z_{t-k}]V & for \quad t > k \end{cases}.$$

Diagnostic expectations exaggerate the information revealed by prices, not only private information. This exaggeration comes from the amplification $(1+\theta)$ of the impact of the current price-based estimate $\mathbb{E}(V|P_t)$, and from the reversal of past price inferences $\mathbb{E}(V|P_{t-k})$. This is another difference from overconfidence, in which public information – including that coming from prices – is underweighted. Because a price increase says that positive information about fundamentals is dispersed in the economy, it renders the right tail representative, causing over-reaction in beliefs.

[13] The variance of V is equal to $\sigma_t^2(V) = \left[\frac{t}{\sigma_\epsilon^2} + \frac{1}{\sigma_{p,t}^2(V)}\right]^{-1}$. $\sigma_{p,t}^2(V)$ decreases in the precision of the public signals observed up to t. See the appendix for details.

If prices become very informative over time, the weight z_t attached to private signals falls and that attached to prices rises. All traders then over-react to the common market signals, the effects of information dispersion subside, so the dynamics of beliefs and prices may be very different from that in Section 2.

3.1 The Boom Phase

To assess the consequences of learning from prices, we need to solve for the coefficient of price informativeness b_r/c_r. The key question is whether diagnosticity increases or reduces price informativeness, boosting or moderating the reaction to price signals. To understand this, we need to find a fixed point at which consensus beliefs are consistent with market equilibrium in Equation (5). Here the key result comes from considering the equilibrium for the boom phase of the bubble, $t \leq k$.

Proposition 2 *With learning from prices and $t \leq k$, the average consensus belief about V is higher than the average consensus belief formed when investors learn from private information alone. The precision of the equilibrium price signal increases over time as:*

$$\frac{b_t}{c_t} = \left(\frac{1+\theta}{\gamma\sigma_\epsilon^2}\right)t, \tag{11}$$

and the average equilibrium price path (for $S_1 = S_2 = \cdots = S_t = 0$) is given by:

$$p_t = (1+\theta)\left[\frac{\frac{t}{\sigma_\epsilon^2} + \left(\frac{1+\theta}{\sigma_S\sigma_\epsilon^2}\right)^2 \frac{t(t-1)(2t-1)}{6}}{\frac{t}{\sigma_\epsilon^2} + \frac{1}{\sigma_V^2} + \left(\frac{1+\theta}{\sigma_S\sigma_\epsilon^2}\right)^2 \frac{t(t-1)(2t-1)}{6}}\right]V. \tag{12}$$

As in the consensus expectations described in Section 2, the equilibrium price grows in the boom phase $t \leq k$, before investors adapt to the displacement. The price in Equation (12) increases over time. If θ in in the range of Proposition 1, the price starts below the fundamental V. This initial under-reaction is due to the same reason that consensus beliefs underreact: traders discount the noise in their signals, so the good news they on average receive is not incorporated into prices.

As time goes by, the price increases due to two forces. First, as traders accumulate private signals, they gain confidence about the innovation and revise their beliefs upward. This effect is captured by the term t/σ_ϵ^2 in Equation (12). Second, the observation of prices provides on average additional good news about displacement, which makes all traders more confident at the same time. As a result, the consensus estimate of V increases relative to the case in which only private signals are observed, and the price booms. The effect of learning from prices is captured by the second term in the numerator of Equation (12). At some point, traders become so confident that the price over-reacts, overshooting the fundamental V.

How does diagnosticity interact with learning from prices in shaping beliefs about fundamentals? And how does learning from prices contribute to the price path? To answer the first question, consider Equation (11). Stronger diagnosticity θ increases the informativeness of prices. When θ is higher, investors are aggressive both in revising their beliefs and in trading on the basis of their private signals. Because these signals are informative about V, price informativeness increases. In turn, greater price informativeness implies that diagnostic traders over-react faster to the price signals, which implies that the bubble arises earlier. Unlike non-fundamental based behavioral biases, such as mechanical extrapolation, which always disconnects prices from fundamentals, diagnostic expectations exaggerate this link, creating faster and accelerating overreaction to the initial displacement.

To address the second question, and thus to see how learning from prices influences the price path, consider again Equation (11). It says that, for a given θ, the price signal becomes more informative over time. As the i.i.d. supply shocks average out over time, a path of consistent price increases is highly informative of good fundamentals. Consequently, learning from prices progressively gains ground, and swamps private signals as prices rise up to the peak of the bubble. As shown in Equation (12) the precision of price signals grows with the cubic power of t, which eventually swamps the linear precision of private signals, t/σ_ϵ^2. This has the following implication.

Proposition 3 *With learning from prices, there exists a* $\sigma_V^{2*} > 0$ *such that for* $\sigma_V^2 < \sigma_V^{2*}$ *there exists a* $t^* > 0$ *such the average price path is convex for* $t < t^*$ *and concave for* $t \in (t^*, k)$. *When the supply shocks are so volatile (* $\sigma_S \to \infty$*) that prices convey no information, the price path is concave:* $\sigma_V^{2*} \to 0$.

When fundamental displacement is sufficiently large relative to prior expectations ($\frac{V}{\sigma_V}$ sufficiently high, or normalizing $V = 1$, σ_V sufficiently low), learning from prices generates a price path that is initially convex and then slows down as the bubble reaches its peak. This occurs because all traders, regardless of their private signals, aggressively infer fundamentals from the common price signal, so the price informativeness increases at first. After enough information gets incorporated into expectations and prices, the value of additional signals diminishes, which causes a price growth slowdown.

Under rational expectations, learning from prices would also coordinate traders' beliefs and generate convexity in prices, but not lead to overvaluation. Under diagnostic expectations convex price growth eventually overshoots the fundamentals, creating a bubble. As an endogenous signal observed by all traders, the price spreads overly optimistic beliefs through markets, supporting the convex growth path typical of real world bubbles.

3.2 Model Simulation: Boom, Bust and Price Extrapolation

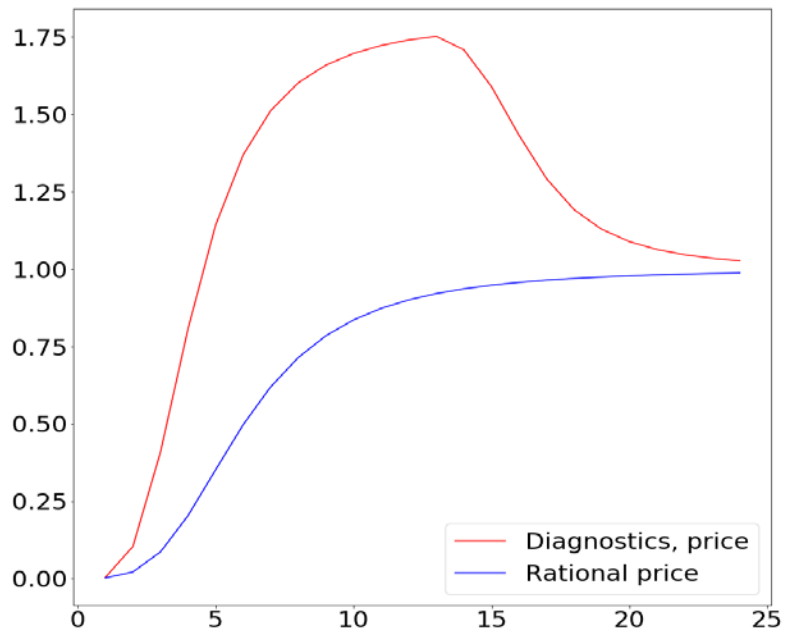
To explore the entire path of the bubble, we resort to simulations, because the model cannot be analytically solved for $t > k$. This analysis shows that diagnostic expectations can produce the boom and bust phases of bubbles and the dynamics of investor disagreement (and hence trading). Most importantly, diagnostic expectations can endogenously produce extrapolative expectations of prices, but in a way that distinguishes them from adaptive expectations or from alternative formulations (Hong and Stein 1999, Glaeser and Nathanson 2017). The predictions may also suggest how to detect bubbles in the data.

We now describe our choice of parameters. We normalize V to 1. To capture that displacement is a fairly rare event, we assume $\sigma_V = 0.5$, so V is two standard deviations away from the mean. The dispersion of trader's private signals is set at $\sigma_\epsilon = 12.5$, which is in the ballpark of estimates obtained from the quarterly dispersion of professional forecasts (BGMS 2018).[14] We set $\theta = 0.8$, in line with the quarterly estimates from macroeconomic and financial survey data. For the model without speculation, we take a time period to be a quarter, set the sluggishness of diagnostic beliefs at $k = 12$ (in line with BGLS 2019) and run the model for $T = 24$ periods, i.e. 6 years. We set the volatility of supply shocks at $\sigma_S = 0.3$.

[14] In BGMS (2019), the estimated signal to noise ratio of the average macroeconomic series was between 3.5 and 4. In the current setting, this should be compared to the precision $\frac{\sigma_\epsilon}{V\sqrt{\tau}}$ of the signals received by the traders over some natural time scale τ. Picking this time scale to be around k, we get σ_ϵ between 12 and 14.

Figure 1 reports the actual price for the average path (no supply shocks) both under diagnostic expectations ($\theta = 0.8$) and under rational expectations ($\theta = 0$). Under diagnostic expectations, the equilibrium price exhibits the typical boom bust pattern, where the boom is driven by overreaction to private signals and prices, while the bust is due to the reversal of expectations at $t = k = 12$. In the rational model, by contrast, the price monotonically converges to V from below.

Figure 1. Average Price Path

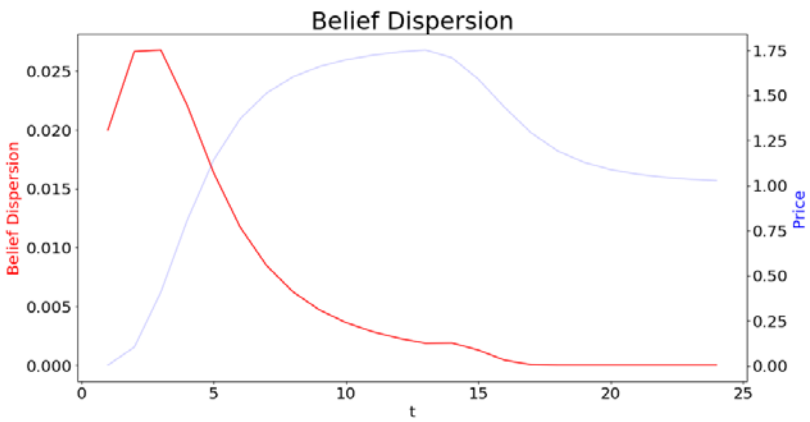


As we discussed in Section 2, rational expectations cannot produce over-reaction and price inflation because they constrain assessed fundamentals to always stay between the prior of zero and the true value V. The same is true for overconfidence (which generates bubbles only in the presence of short sales constraints, e.g. Harrison and Kreps 1978). In our model, a displacement drives continued good news, resulting in a price boom. This leads traders to focus on the right tail of V and think that the innovation is truly exceptional, causing prices to overreact.

The bust occurs at time $t = k + 1$, when investors adapt to the displacement, starting to view the innovative asset or technology as the new norm. Here the length k of the boom phase is deterministic, but the model could be made more realistic by having k stochastic (and even heterogeneous across investors). As in Proposition 1, adaptation to early news causes excess optimism to run out of steam, generating the bust. Reversal of expectations and prices due to disappointment of prior optimism can help account for the slowdown of some bubbles, but it is not the only mechanism behind a bust; other factors including bad news (the housing bubble deflating from 2006 onward), as well as the proximity of a terminal trading date (crucial in experimental findings), are surely significant. We consider the latter mechanism in Section 4.

Because traders observe independent signals, they have heterogeneous beliefs about the value of the asset. This creates room for disagreement and trading (Scheinkman and Xiong 2003). Barberis et al. (2018) show that disagreement tends to rise in the boom phase. Our model can create very similar effects. As time goes by, traders become more confident in their information, which causes them to place stronger weight on private signals. This effect tends to foster disagreement. At some point, the common price shock becomes so strong that disagreement declines. Figure 2 plots the standard deviation of investor beliefs: disagreement rises in the early part of the boom, but falls as the public signal dominates the private information.

Figure 2. Model-Implied Belief Dispersion



We can also use simulations to describe the dynamics of expectations of future prices. Under mechanical extrapolation, traders project past price increases into the future using the updating rule:

$$\mathbb{E}_t(p_{t+1}) = p_t + \beta(p_t - p_{t-1}), \tag{13}$$

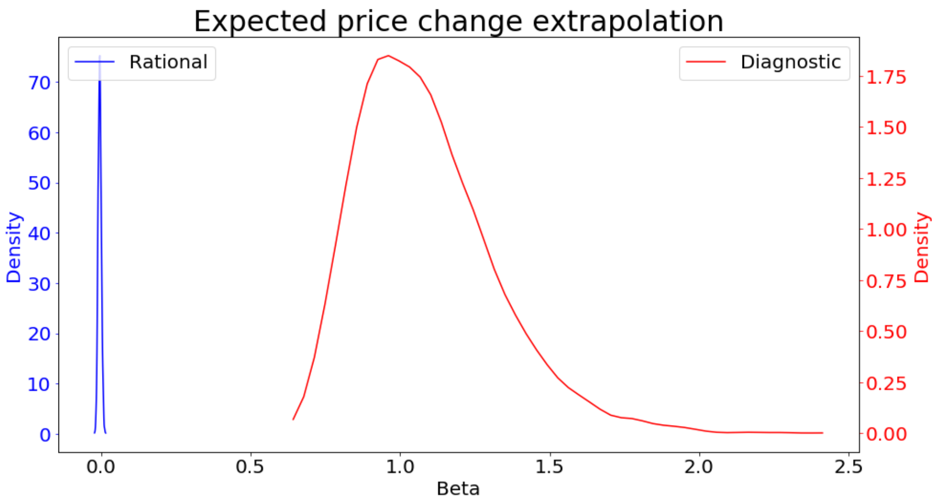
where $\beta > 0$ captures the fixed degree of price extrapolation. In our model, in contrast, traders watch prices in order to infer fundamentals. As a result, price extrapolation arises because high past prices signal high fundamentals and hence even higher future expected prices.

In Hong and Stein (1999), extrapolation is due to under-reaction, which makes it optimal for momentum traders to chase the upward trend in prices. In that model, momentum traders form expectations of future price changes by running simple univariate regressions of current on past price growth. In Glaeser and Nathanson (2017), investors believe that the price reflects fundamental value. An increase in price is then interpreted as stronger fundamentals, and leads to extrapolation of high prices into the future. In both models, as in adaptive expectations, extrapolation is due to the use of simplified (or wrong) models.

This logic suggests a testable difference between mechanical extrapolation and price learning under diagnostic expectations. Whereas mechanical extrapolation models impose a constant β, price extrapolation arises endogenously in our model as a product of the distorted inference process. Thus, the degree of extrapolation depends on the degree of fundamental uncertainty. In terms of Equation (13), our model predicts that the updating coefficient β should be positive at the early stages of the bubble when price movements convey information about fundamentals, but fall to 0 as learning accelerates.

We evaluate these ideas by simulating the model. We run regression (13) using a time series of the model simulated using the parameters above. We produce 2000 such time series and plot in Figure 3 the histogram of estimated coefficients for both the diagnostic and the rational model.

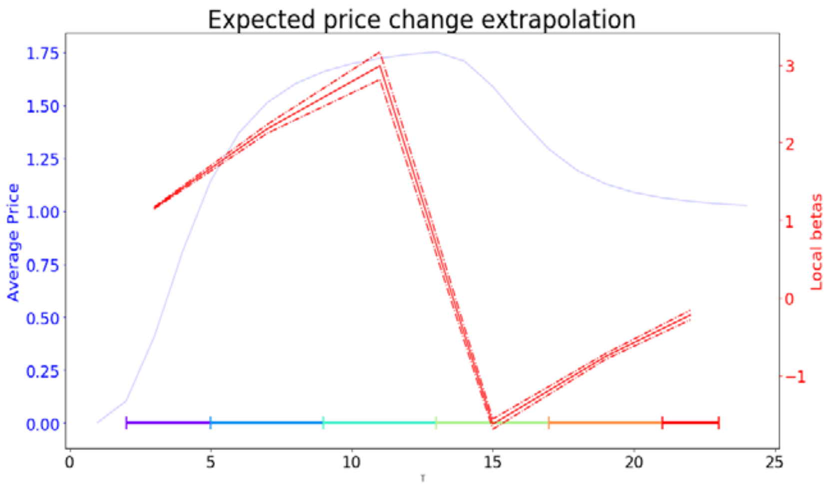
Figure 3. Model-Implied Extrapolation Coefficient



The coefficient of price extrapolation implied by the model is positive, between .5 and 1.5. Even though our investors are entirely forward looking, they appear to mechanically extrapolate past prices. This is not the case under rational expectations, where the coefficient is around 0, and even slightly negative because of rational supply shock effects that are dominated by extrapolation in the diagnostic case.

While diagnostic expectations entail a positive extrapolation coefficient β on average across the entire bubble episode (as does mechanical extrapolation), that coefficient is the highest when prices are most informative, which in this case is at the peak of the bubble. Figure 4 reports the average estimates of β in each of six buckets, capturing growth, overshooting and collapse.[15] The results confirm that price extrapolation is strongest in the making of the bubble when learning is rapid (the first phase highlighted in blue). This occurs because prices are most informative (relative to the private signal) in that range, inducing diagnostic traders to update more aggressively after a price rise. At the peak of the bubble, expectations of future prices are significantly above actual prices. After the bubble bursts, traders adjust their expectations down significantly, but not fast enough to converge to the actual prices. Thus, in this period extrapolation appears negative. Finally, as learning subsides, extrapolation goes to zero, just as in the rational case.

Figure 4. Time-Dependent Extrapolation



As Figures 3 and 4 show, this model can produce some price convexity and moderate overvaluation. However, this model precludes large bubbles because for reasonable values of θ prices are tethered to V. In contrast, prices sometimes strongly overshoot sensible measures of fundamentals. In addition, while learning from prices generates some convexity in the price path, it does not create enough acceleration to generate increasing growth rate of prices (accelerating returns) seen in the data (Greenwood, Shleifer, and You 2019). We next show that both features can be attained by adding speculation to our model.

[15] To build Figures 3 and 4, we simulate 5000 price paths and expected future price paths. We pool simulations and compute the regression $\mathbb{E}_t(p_{t+1}) = p_t + \beta(p_t - p_{t-1})$ *within* buckets of 4 time periods. Figure 3 reports the resulting βs and their confidence intervals (running regressions for individual paths and averaging the βs yields similar results).

4. Speculation

To introduce speculation, we assume that traders have short horizons in the sense that their objective function at each time t is to resell the asset at time $t + 1$. In the model of Section 5 we relax this assumption and allow traders to choose optimally whether to sell the asset or not. The trading game lasts for T rounds. The traders holding the asset in the terminal date receive V. We take T to be exogenously given and deterministic, as in laboratory experiments of bubbles. In real markets, there is no such thing as a terminal date, but taking a fixed T is a convenient approximation to a setting in which there is a certain probability that at some point the “speculation game” ends in the sense that most traders attend to fundamentals.

With speculation, diagnostic expectations generate price paths with significantly larger overvaluation than in the previous models, followed by a price collapse as the terminal date approaches. This occurs because speculators not only overreact to good fundamental news, but also expect to resell to overreacting buyers in the future, which drives the price today higher. As T approaches, the prospects for re-trading fade and the bubble bursts. These dynamics are very different from those obtained under rationality.

Traders continue to have mean-variance preferences. Away from the terminal date, $t < T$, trader i chooses demand D_{it} to maximize $[\mathbb{E}_{it}(p_{t+1}) - p_t]D_{it} - \frac{\gamma}{2}\, Var_t(p_{t+1})D_{it}^2$, while his objective at time T is fundamental-based as before. Demand in each period is then given by:

$$D_{i,t} = \frac{\left[\mathbb{E}_{i,t}^{\theta}(p_{t+1}) - p_t\right]}{\gamma Var_t(p_{t+1})}, \quad for\ t = 1, \dots, T-1, \tag{14}$$

$$D_{i,t} = \frac{\left[\mathbb{E}_{i,t}^{\theta}(V) - p_t\right]}{\gamma Var_t(p_{t+1})}, \qquad for\ t = T. \tag{15}$$

With speculation, demand increases in the expected capital gain $\mathbb{E}_{i,t}^{\theta}(p_{t+1}) - p_t$ except in the last period $t = T$, in which traders buy the asset to hold it.

We illustrate the key consequences of speculation this Section by ruling out learning from prices, which we then re-introduce in Section 5 along with dynamic optimization. We also simplify the analysis by assuming that the diagnostic reference is very sluggish, $k > T$, so that information about the asset’s value is always assessed compared to the prior $V = 0$. The reason is that, as we will see, speculation itself creates a reason for the bubble to deflate as the terminal date T approaches.

Without learning from prices, we do not need a supply shock, so we assume that supply is equal to zero. Aggregating the individual demand functions, prices are pinned down by the conditions:

$$p_t = \int \mathbb{E}_{i,t}^{\theta}(p_{t+1})\, di, \quad for\ t = 1, \dots, T-1, \tag{16}$$

$$p_T = \int \mathbb{E}_{i,T}^{\theta}(V)\, di. \tag{17}$$

In the final period T, the consensus fundamental value is $\mathbb{E}_T^{\theta}(V) = (1+\theta)\pi_T V$, as per Equation (3), leading to the terminal price $p_T = (1+\theta)\pi_T V$. Under the assumption $\theta \in \left(\frac{1}{T}\frac{\sigma_\epsilon^2}{\sigma_v^2}, \frac{\sigma_\epsilon^2}{\sigma_v^2}\right)$ of Proposition 1, which we maintain, this price is above the fundamental, $(1+\theta)\pi_T > 1$.[16]

Consider now the price at $T-1$. By Equation (16), this price is the consensus expectation as of $T-1$ of the terminal price p_T. To compute this consensus, consider first the expectation held at $T-1$ by a generic trader j. When forecasting the terminal price, this trader must make two assessments. First, he must assess the fundamental value V. Second, he must forecast how traders at T will react to noisy signals *of the same* fundamental value. Because the beliefs of future traders are a random variable, trader j forecasts them using the very same diagnostic formula of Equation (2). One can interpret this forecasting process in two ways. First, one can view trader j as placing himself in the shoes of future traders receiving different signals, predicting that these traders will behave the way he would behave in light of the same signals. Alternatively, one can view trader j as forecasting the behavior of others with the understanding that they will update diagnostically. In both cases, we continue to rule out the possibility that any trader is sophisticated enough to be aware of his own diagnosticity, otherwise he would de-bias his beliefs about self and others.

Consider how trader j forecasts the beliefs at T of a generic trader i who has observed an average signal $\frac{\sum_{r=1}^{T} s_{ir}}{T}$ from the initial date to the terminal period. Trader j knows that trader i overreacts to all

[16] Recall that in this Section we shut down the adaptation of diagnostic expectations by setting $k > T$.

signals received, forming a terminal estimate $\mathbb{E}_{iT}^{\theta}(V) = (1+\theta)\pi_T \frac{\sum_{r=1}^{T} s_{ir}}{T}$. By averaging across all traders i, trader j knows that, if the fundamental value is V, the consensus estimate, and hence the equilibrium price at T is equal to:

$$p_T = (1+\theta)\pi_T V.$$

This prediction is based on the fact that trader *j* knows that, whichever signals are received by individual traders, they will average out to the true V. Of course, trader *j* does not know the true value of V at $T-1$; he only has an estimate of it, based on the signals $\frac{\sum_{r=1}^{T-1} s_{jr}}{T-1}$ observed up to that period. This $T-1$ estimate of fundamentals is of course diagnostic and equal to:

$$\mathbb{E}_{jT-1}^{\theta}(V) = (1+\theta)\pi_{T-1}\frac{\sum_{r=1}^{T-1} s_{jr}}{T-1}.$$

The diagnostic expectation held at $T-1$ by trader *j* about the terminal price is then given by:

$$\mathbb{E}_{jT-1}^{\theta}\left[\mathbb{E}_{T}^{\theta}(V)\right] = (1+\theta)\pi_T \mathbb{E}_{jT-1}^{\theta}(V)$$

$$= (1+\theta)^2\pi_T\pi_{T-1}\frac{\sum_{r=1}^{T-1} s_{jr}}{T-1}.$$

Trader j at $T-1$ uses his signals but he compounds diagnosticity twice. First, he diagnostically overreacts to his signal, creating an inflated estimate of fundamentals. Second, he expects future traders to over-react to signals generated by the inflated fundamentals. To see the intuition, imagine that *j* overestimates the share of future Googles in the population of tech firms to be 7%. He then expects future traders to overreact relative his assessment and estimate the share of Googles to be, say, 10%. In this way, overreaction to news compounds as the current forecast is projected into the future.

Because every trader *j* repeats the same logic, by averaging across all of them, the consensus forecast held at time $T-1$ about the terminal price, and thus the equilibrium price at $T-1$ is given by:

$$p_{T-1} = (1+\theta)^2\pi_T\pi_{T-1}V. \quad (18)$$

To gauge the role of diagnostic expectations, suppose that traders are rational, so $\theta = 0$. In this case, the price at $T-1$ is equal to $p_{T-1}^{rat} = \pi_T\pi_{T-1}V$ while the price at T is equal to $p_T^{rat} = \pi_T V$. Critically, because $\pi_{T-1} < 1$, the price at $T-1$ is lower than the terminal price, $p_{T-1}^{rat} < p_T^{rat}$. This captures a broader and well-known point (Allen et al. 2006): under rationality, speculation leads the price to initially increase slowly over time, and then to increase faster as the terminal date approaches. However, there is neither over-valuation nor collapse. The intuition is that rational traders discount their signals and expect future traders to do the same. As a result, they do not expect to be able to resell the asset for a very high price, which keeps the current price low. As the terminal date approaches, this mechanism becomes weaker so the price increases faster to the consensus terminal belief about V.

Crucially, even a modicum of diagnosticity $\theta > 0$ dramatically changes the calculus. To begin, note that when $\theta > 0$, it is entirely possible that the price drops at the terminal date. This is true if and only if:

$$p_{T-1} > p_T \Leftrightarrow (1+\theta)\pi_{T-1} > 1.$$

If traders overestimate the fundamental value at time $T-1$, i.e. $(1+\theta)\pi_{T-1} > 1$, then the price at $T-1$ is above both fundamentals *and* the terminal price. The intuition goes as follows. By overestimating V, traders at $T-1$ believe that future traders will overreact to this estimate, compounding overreaction twice. But then, the expectation to sell to these bullish traders in the future raises the price of the asset in $T-1$ itself.

This leads to a first important remark: in sharp contrast with the rational case, which leads to a monotone rising price path, diagnostic expectations introduce the opposite effect. By creating overreaction, they imply that prices decline toward the terminal date, reflecting an initial strong overvaluation of the asset.

To study the implications of $\theta > 0$ fully, we need to iterate the same logic backward to earlier periods until the initial date $t = 1$. It is immediate to see that the full path of equilibrium prices obtained by iterating Equation (18) backwards is described by:

$$p_t = (1+\theta)^{T-t+1}\left[\prod_{r=t}^{T}\pi_r\right]V, \tag{19}$$

which implies the following result.

Proposition 4 *Define the geometric average of all signal to noise ratios* $\hat{\pi} \equiv [\prod_{r=1}^{T}\pi_r]^{\frac{1}{T}}$. *Then, if* $\theta \in \left(\frac{1}{T}\frac{\sigma_\epsilon^2}{\sigma_v^2}, \frac{1-\hat{\pi}}{\hat{\pi}}\right)$, *where* $\frac{1-\hat{\pi}}{\hat{\pi}} < \frac{\sigma_\epsilon^2}{\sigma_v^2}$, *the speculative price dynamics exhibit the three bubble phases. In particular:*

1. The price starts below fundamental, $p_1 < V$, *and gradually increases above fundamentals, reaching its maximum at the smallest time* $\hat{t}$ *for which* $(1+\theta)\pi_{\hat{t}} > 1$.

2. From $t = \hat{t}$ *onwards the price monotonically declines toward* p_T.

With diagnostic expectations, speculative dynamics can generate both the sluggish upward price adjustment typical of underreaction (provided θ is not too large), the price inflation relative to fundamentals typical of overreaction (provided θ is not too small), and the bust phase in which prices collapse, which here is driven by the reduction in the available rounds of reselling.

Because $(1+\theta)\hat{\pi}^T < 1$, individual traders underreact to the aggregate information in the first period. The logic is the same as before: individual uncertainty about V is still very large. Traders are not only cautious in estimating V, but also think that next period buyers will be cautious as well. This effect curtails the expected resale price and demand for the asset today, keeping its price low. As time goes by, traders acquire more information, become more confident, and start using it more aggressively. They become more optimistic about the signals future buyers will get, more confident about future buyers' over-optimism, and the price starts increasing. As traders gain confidence, the possibility of multiple rounds of reselling to over-reacting traders dramatically boosts price, which overshoots V. The price then starts declining as the terminal date T approaches, because there are fewer and fewer rounds of trading and thus less scope for reselling to overreacting buyers. Once again, the dynamics of speculation under rationality are very different: they display momentum but not overshooting or reversal (Allen, Morris, and Shin 2006).

Another important consequence of speculation is that it can greatly exacerbate the overshooting of fundamentals, relative to the benchmark model of Section 3. Equation (19) shows how speculation fuels bubbles under diagnostic expectations, and can cause strong price inflation even with small diagnostic distortions θ. Consider the ratio of price under speculation to consensus expectations of fundamentals (which equals price in the absence of speculation). At the peak of the bubble, which occurs at $\hat{t} = \frac{1}{\theta}\frac{\sigma_\epsilon^2}{\sigma_v^2}$, this ratio is inflated relative to the rational benchmark as follows:

$$\frac{p_{\hat{t}}(\theta)}{\mathbb{E}_{\hat{t}}^{\theta}(V)} = (1+\theta)^{T-\frac{1}{\theta}\frac{\sigma_\epsilon^2}{\sigma_v^2}} \frac{p_{\hat{t}}(0)}{\mathbb{E}_{\hat{t}}^{0}(V)}$$

While under diagnostic expectations beliefs about fundamentals are inflated by a linear factor of θ, namely $\frac{\mathbb{E}_{\hat{t}}^{\theta}(V)}{\mathbb{E}_{\hat{t}}^{0}(V)} = 1+\theta$, when speculation is included the inflation of price relative to beliefs grows as a power of $1+\theta$. Even a small departure θ from rationality can fuel large bubbles. This much stronger growth reflects the compounding effect of over-optimism about selling to overreacting investors until the horizon T. Increasing θ increases optimism, which also implies that the peak of the bubble is reached earlier, which in turn implies a stronger compound effect of diagnostic expectations. A small amount of diagnostic distortions can be amplified by forward looking speculators into large bubbles.

5. The Full Model: Forward Looking Speculation and Learning from Prices

We now combine all the ingredients by introducing diagnostic expectations into the dynamic trading model of He and Wang (1995). In this setting, at each time t each investor i chooses how many units of the asset to buy or sell so as to maximize the expected utility of final consumption:

$$\max_{s.t. W_{t+1}^i = W_t^i + X_t^i (p_{t+1} - p_t)} \mathbb{E}_{it}^{\theta}\left[-e^{-\gamma W_T^i}\right], \tag{20}$$

where the law of motion of individual wealth W_t^i depends on X_t^i, the holdings of the risky asset optimally chosen by the investor at time t, and on $(p_{t+1} - p_t)$, which is next period's capital gain on this asset (remember that the dividend is paid at the end and $p_T = V$).

The diagnostic expectation $\mathbb{E}_{it}^{\theta}(.)$ is formed on the basis of the private signals described in Section 2, and the public signals obtained from price movements as shown in Section 3 (again, to obtain gradual learning from prices we allow for i.i.d supply shocks S_t). The trader's speculative motive comes from his individual expectations of capital gains and capital losses, which affects the evolution of his wealth. Critically, the trader is not forced to sell all of the risky asset at $t+1$. This flexibility is valuable. For instance, it allows a trader expecting strong price increases in the distant future to buy now and sell then.

To solve this model under diagnostic expectations, we follow the methodology of He and Wang, adapting it only to the fact that in our model expectations are diagnostic. To quickly review, He and Wang show that under rational expectations solving (20) entails maximizing the expectation of the value function:

$$J\left(W_{t+1}^{i}, \Psi_{t+1}^{i}\right) = -\exp\left[-\gamma W_{t+1}^{i} - \frac{1}{2}\left(\Psi_{t+1}^{i}\right)' U_{t+1} \Psi_{t+1}^{i}\right], \tag{21}$$

which depends on the trader's current wealth W_{t+1}^{i} and on the vector of trader-specific beliefs:

$$\Psi_{t+1}^{i} = \left[1, \mathbb{E}_{i,t+1}(V), \mathbb{E}_{p,t+1}(V), \mathbb{E}_{pt}(V)\right],$$

which include the trader's rational expectation of fundamentals computed solely on the basis of his private signals, $\mathbb{E}_{i,t+1}(V)$, the trader's as well as the market rational expectation of fundamental computed solely on the basis of public price signals $\mathbb{E}_{p,t+1}(V)$, and the past value of such public expectation $\mathbb{E}_{pt}(V)$.

The term $\left(\Psi_{t+1}^{i}\right)' U_{t+1} \Psi_{t+1}^{i}$ in Equation (21) captures dynamic trading motives. When the trader is much more optimistic than the market, $\mathbb{E}_{i,t+1}(V) \gg \mathbb{E}_{p,t+1}(V)$, the term $\left(\Psi_{t+1}^{i}\right)' U_{t+1} \Psi_{t+1}^{i}$ is high. As a result, the trader's marginal utility of current wealth is low, which makes it optimal for him to buy the asset. When instead the trader is much more pessimistic than the market the term $\left(\Psi_{t+1}^{i}\right)' U_{t+1} \Psi_{t+1}^{i}$ is low. Here the marginal utility of current wealth is high, which causes the trader to sell the asset. He and Wang further show that the matrix of coefficients U_{t+1} regulating dynamic incentives can be recursively determined.

Here we follow the same approach but allow expectations to be diagnostic. Specifically, we assume that the generic trader maximizes the objective:

$$\mathbb{E}_{it}^{\theta}\left\{J\left(W_{t+1}^{i}, \Psi_{t+1}^{i}\right)\right\} = \mathbb{E}_{it}^{\theta}\left\{-\exp\left[-\gamma W_{t+1}^{i} - \frac{1}{2}\left(\Psi_{t+1}^{i}\right)' U_{t+1} \Psi_{t+1}^{i}\right]\right\}, \tag{22}$$

where diagnostic expectations are for simplicity specified for the case in which $k > T$. To solve the problem, we use the values U_{t+1} computed in the rational expectations solution, and diagnostically distort only the random variables, in particular future wealth and beliefs. This approach greatly simplifies computation, and can be interpreted as a diagnostic perturbation of the dynamic rational expectations equilibrium.

Appendix B describes the implementation of this approach in detail, including the trader's optimal portfolio choice, the market clearing condition, and the recursive definition for U_{t+1}. The model is complex and the resulting equilibrium can only be solved numerically.[17] To characterize its properties, we simulate it at the benchmark parameter values $\sigma_V = 0.5$, $\sigma_S = 0.2$, $\sigma_\epsilon = 9.5$, $\gamma = 0.12$, $\theta = 0.8$ and $T = 15$. These parameter values are very close to those used in Section 3, except that we have increased the informativeness of private signals. This allows us to bolster the effect of dynamic trading, which would otherwise matter very little for high σ_ϵ. We examine the following key outcomes: the average price path, price volatility, investor disagreement, trading volume, and time-varying average price extrapolation. For the average price and disagreement, we derive the values directly from the equilibrium coefficients, whereas we simulate 1000 price paths to compute the time-varying volatility, volume, and extrapolation.

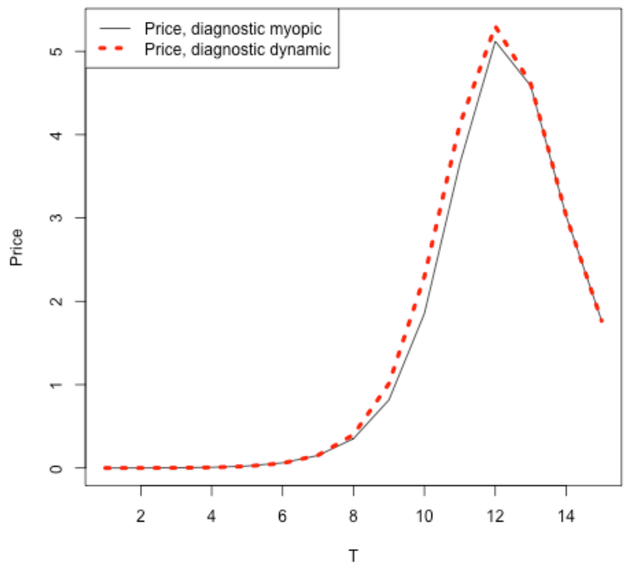
Consider first the key outcome: the average price path, which is the price path for the case in which the supply shock S_t is zero in all periods. As our equilibrium is linear in supply, this also is the unconditional expected price path $\mathbb{E}(p_t)$ for $t = 1, 2, \dots T$. To assess the role of dynamic trading in isolation, we also evaluate and report the price path in a model in which traders are myopic, as in Section 4, but learn from prices.[18] To assess the role of diagnostic beliefs, we report the price path under rational expectations (with and without dynamic trading). The simulation results are presented in Figure 5.

Figure 5. Average Price Paths

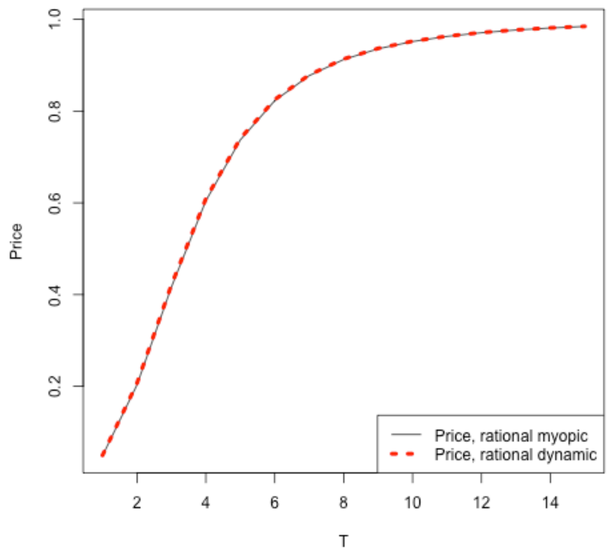
Panel A: The diagnostic case

[17] Similarly to Section 3, the full model is characterized by the equilibrium-price coefficients (a_t, b_t, c_t) for $t = 1, 2, \dots T$, where $p_t = a_t \mathbb{E}(V|P_t) + b_t \left(V - \frac{c_t}{b_t} S_t\right)$. As shown in the Appendix, one can recursively solve for the equilibrium coefficients given a guess for the final precision of the public signals $\frac{1}{\sigma_{p,T}^2} = \frac{1}{\sigma_V^2} + \frac{1}{\sigma_S^2} \sum_{r=1}^{T} \left(\frac{b_r}{c_r}\right)^2$. The equilibrium is then numerically pinned down by the final precision that implies $\frac{1}{\sigma_{p,0}^2} = \frac{1}{\sigma_V^2}$.

[18] The myopic problem is a special case of Equation (2) in which, instead of optimizing $\mathbb{E}_{it}^{\theta}\{J(W_{t+1}^i, \Psi_{t+1}^i)\} = \mathbb{E}_{it}^{\theta}\left\{-\exp\left[-\gamma W_{t+1}^i - \frac{1}{2}\left(\Psi_{t+1}^i\right)' U_{t+1} \Psi_{t+1}^i\right]\right\}$, the individual optimizes $\mathbb{E}_{it}^{\theta}\{J^{myop}(W_{t+1}^i, \Psi_{t+1}^i)\} = \mathbb{E}_{it}^{\theta}\{-\exp[-\gamma W_{t+1}^i]\}$, or equivalently myopically maximizes next period wealth assuming immediate resale of asset.



Panel B: The rational case



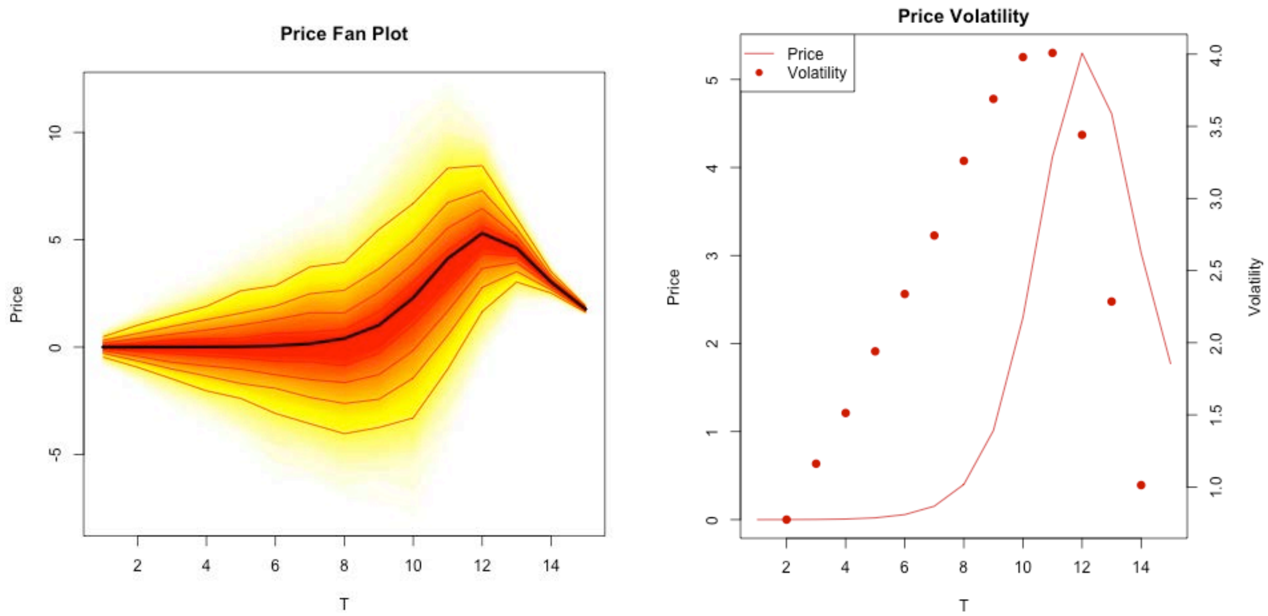
Panel A reports the outcome under diagnostic expectations, plotting the dynamic trading case with the dashed line and the myopic case with the solid line. With both dynamic and myopic trading, the price path displays the three Kindleberger phases: a price growth phase, an acceleration and overshooting phase, and then a collapse. Dynamic trading changes matters a bit, and in particular causes the bubble to occur slightly faster and to reach a peak that is slightly higher (by roughly 20 percent of the fundamental). Intuitively, as traders expect the future price to go up, they buy more aggressively today, which pushes up prices earlier and more. The convexity of the price path is due to the interaction of DE with speculation and learning from prices, in line with the analyses of Sections 3 and 4. On the one hand, overreaction to news causes prices to be more informative, fueling common over-reaction by traders. On the other hand, traders expect future buyers to over-react even more, causing high speculative demand and hence a high price today.

As we saw in Section 3 and 4, under rational expectations these boom bust patterns do not emerge. The price gradually converges to the truth from below. Under DE, in contrast, even a modest departure from rationality creates substantial bubbles. A degree of distortions of $\theta = 0.8$ generates in the above simulation a price peak that is 5.3 times the fundamental, and about 3 times higher than the price peak in the absence of speculation. Mechanical price extrapolation is not necessary to generate such overpricing. The roles of Kindleberger's fundamental displacement and the resulting belief distortions in generating bubbles may be larger than usually thought.

To assess price volatility, we simulate the model for many different realizations of the supply shocks S_t. We present the fan-plot of the resulting price paths and the simulated standard deviation of the price innovations $p_{t+1} - p_t$. The result is reported in Figure 6.

Figure 6. Price Volatility

Left: the fan plot of 1000 simulated price paths. The quantiles and the average of the price paths are indicated by the red and black curves respectively. Right: a plot of $\sigma(p_{t+1} - p_t)$ for $t = 1, 2, \dots T$.

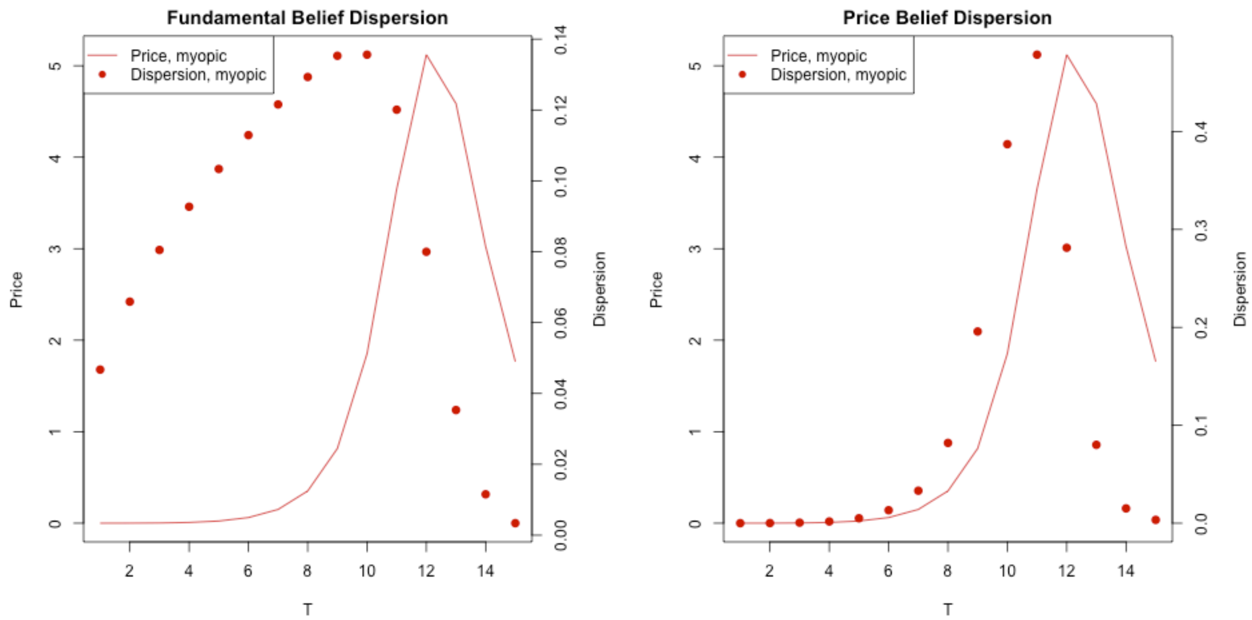


Intuitively, the price of the asset is most volatile in periods leading up to the peak of the bubble, and it becomes extremely stable as we converge to the final period. Around the peak of the bubble, in fact, prices are most informative and even a slight supply shock may lead to a dramatic change in beliefs, trading strategies and prices. This outcome underscores the fragility of bubbles in our model, and how – through sudden changes in expectations – they can turn into busts even with small changes in market conditions.

Consider next disagreement. We compute the dispersion of traders' beliefs from the equilibrium coefficients directly.[19] There are two relevant types of disagreement. First, individuals may disagree about fundamentals. Second, this disagreement translates to a disagreement about next period prices, which is more directly relevant for trading. The two measures are not redundant: disagreement about next period prices depends on both disagreement about fundamentals, as well as how much this disagreement is amplified into next period prices. Figure 7 reports the result on both fundamental and next period price disagreement.

Figure 7. Disagreement

Plot of the average belief dispersion across i. Left: the dispersion of $\mathbb{E}_{it}^{\theta}[V]$. Right: the dispersion of $\mathbb{E}_{it}^{\theta}[p_{t+1}]$.



[19] Given the precision of prices $\zeta_t = \frac{1}{(\sigma_p^t)^2}$, standard Bayesian analysis implies that the dispersion of diagnostic beliefs about V is $\frac{(1+\theta)\sigma_\epsilon}{\sqrt{t}} \cdot \frac{\frac{t}{\sigma_\epsilon^2}}{\frac{t}{\sigma_\epsilon^2}+\zeta_t}$. We can then feed this into expectations of next period prices.

Disagreement about both fundamentals and next period prices peaks near the peak of the bubble and drops toward zero as traders progressively learn. This is due to two counteracting forces. On the one hand, as traders acquire information their uncertainty falls. As a result, they update more aggressively on their private signals and disagreement increases. On the other hand, information becomes more common across traders, also because more price signals are observed over time. This causes disagreement to fall, and to eventually disappear altogether. Interestingly, whereas disagreement on fundamentals rise relatively quickly, disagreement on prices rises with a lag, as fundamentals are weakly reflected into prices at the beginning of the bubble. Thus, disagreement about prices is decoupled from disagreement about fundamentals.

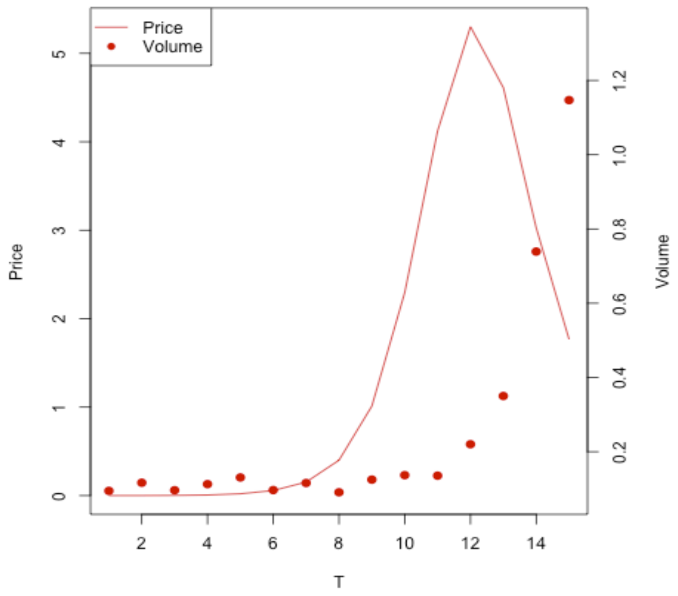
We next investigate trading volume. As our equilibrium is linear, market clearing implies that the portfolio holding of individual i is given by $X_t^i = g_t\left(\bar{s}_t^i - V\right) + S_t$, where g_t captures the strength of the strategic trading motives and it is determined in equilibrium. This implies that the individual portfolio adjustment ΔX_{t+1}^i can be broken down into three components:

$$\Delta X_{t+1}^i = X_{t+1}^i - X_t^i = \underbrace{S_{t+1} - S_t}_{supply\ adjustment} + \underbrace{g_{t+1}\left(\bar{s}_{t+1}^i - \bar{s}_t^i\right)}_{learning\ adjustment} + \underbrace{(g_{t+1} - g_t)\left(\bar{s}_t^i - V\right)}_{strategic\ adjustment}. \tag{23}$$

The first component is due to an overall shift in portfolio holdings due to supply adjustment. The second learning component captures the standard trading motives explained directly by the change in a trader's beliefs. The final term captures strategic adjustment motives: controlling for relative optimism/pessimism, the trader faces time-varying incentives to trade that depend on expectations of risk and return of speculation. Figure 8 reports the evolution of aggregate trading volume over time.[20]

Figure 8. Trading Volume

[20] We directly compute g_t from equilibrium coefficients, and estimate $Volume_t = \frac{1}{2}\ \mathbb{E}\ |\Delta X_t|$ by simulating the supply shocks and the average signals in Equation (23).



Trading volume monotonically increases over time. This is due to two forces. First, and crucially, disagreement increases as we approach the peak of the bubble. Greater disagreement induces pessimistic traders to sell and optimistic traders to buy, raising the volume. Furthermore, the upward price path increases the returns to speculation, which contributes to the strategic adjustment component of Equation (22), also creating a stronger incentive to trade. This mechanism continues for a while. Eventually, however, one would expect trading volume to fall due to falling disagreement after the peak of the bubble (see Figure 7).

This is not what is shown in Figure 8, in which trading volume monotonically increases, even after the collapse of the bubble. This is due to a countervailing effect: price volatility. After the peak of the bubble, price volatility swiftly drops, as shown in Figure 8. This reduction in price risk encourages risk averse traders to aggressively trade on the basis of their heterogeneous beliefs. In other words, lower price volatility has reduced the risk of speculation, which raises the strategic adjustment term. It is true that disagreement has fallen, but in our simulation the reduction in price volatility dominates, leading to monotonically increasing trading volume over time.

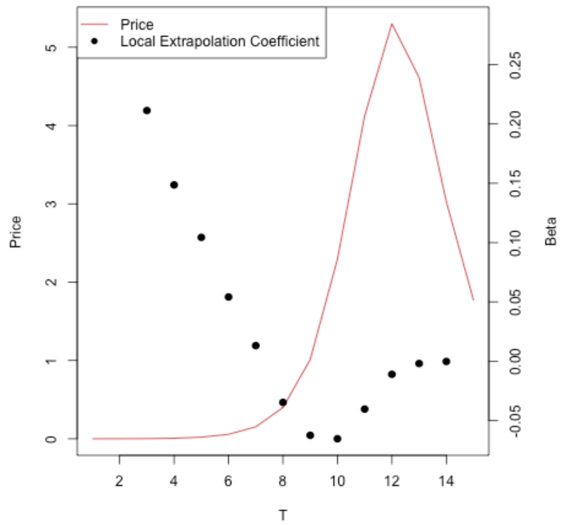
Due to this second force, the current simulation of the model does not yield the decline in trading volume observed after bubbles collapse (Barberis et al. 2018).[21] One solution to this problem would be to

[21] While volume rising at the end of the bubble seems counterfactual, there is also documentation of elevated volume during market crashes (Galbraith 1954, Kindleberger 1978).

introduce traders with heterogenous risk aversion: the drop in the price of the bubble can reduce the wealth of risk seeking traders, thereby reducing their ability to aggressively trade. Progressive entry of more risk-seeking and less informed market participants seems relevant in several bubble episodes, as narratively documented (Galbraith 1954), and it may help our model generate richer dynamics of trading volume.

Finally, consider time-varying extrapolation. We simulate coefficient β_t in Equation (13) for the full model.[22] Figure 9 below shows the coefficients for both models.

Figure 9. Time-Varying Extrapolation



As in Section 3, the model generates positive extrapolation in the early stages of the bubble. Forward looking traders react strongly to their private signals, as they foresee large opportunities to speculate in the future. Hence, price informativeness increases, causing diagnostic traders to expect a large future price increase after observing a current price increase. This mechanism, which creates a semblance of price extrapolation, becomes weaker over time, because learning slows down. Furthermore, unlike in Section 3, extrapolation is strongest at the early stages of the bubble. This is because, as we approach the peak, speculation causes prices to adjust contemporaneously to expectations of future prices (see Section 4), causing the difference between the two to be less predictable by recent price innovation. In contrast with mechanical price extrapolation models, our model demonstrates that apparent price extrapolation can arise

[22] We rely on simulations to compute the extrapolation coefficients, generating 1000 sample paths. There are two main approaches to compute β_t. First, one could pool $(p_t^n - p_{t-1}^n, \mathbb{E}_t^\theta[p_{t+1}^n] - p_t^n)$ over n for each t and compute the regression coefficients (pooling). Alternatively, one can subdivide $\{1, 2, \dots T\}$ into contiguous buckets $b_1, \dots b_K$, compute the regression coefficients β_k^n, and average over n to obtain $\overline{\beta_k} = \frac{1}{1000}\sum_n \beta_k^n$.(averaging) The two approaches produce mostly similar coefficients. We report the results from the first approach.

endogenously from inference about fundamentals, and that the exact dynamics of price extrapolation may depend on the horizon and trading behavior of the speculators.

To summarize, we have examined the implications of our model on a rich set of observables, including average prices, volatility, trading, disagreement, and extrapolation. By combining diagnostic expectations of fundamentals with standard mechanisms such as learning from prices and speculation, our model generates some salient aspects of the Kindleberger narrative and other realistic features of bubbles.

6. Conclusion

In this paper, we brought a micro-founded model of beliefs, diagnostic expectations, to the problem of modeling bubbles. We have considered two formulations: belief formation about fundamentals with learning from prices and also speculation, whereby investors focus on reselling the asset next period. We showed that both of these formulations exhibit the central features of bubbles as conceived by Kindleberger (1978): displacement, price acceleration, and a crash. Moreover, these models deliver extrapolative beliefs and overreaction to information during the later stages of the bubble that are so central both to the Kindleberger narrative and empirical facts about bubbles.

Our micro-founded model of beliefs, based on expectations about fundamentals, delivers two further insights into the anatomy of bubbles. First, it connects over-reaction to fundamental news, which is the central implication of diagnostic expectations, to price extrapolation, which has been increasingly seen as a key feature of bubbles (see Barberis et al. 2018). In our model, price extrapolation is far from constant over the course of the bubble, as in mechanical models of adaptive expectations, but in fact emerges as a byproduct of diagnostic expectations during the rapid price growth stage of the bubble. In fact, the bubble collapses in part because the psychological mechanisms that entail price extrapolation run out of steam.

Second, our model illustrates the centrality of speculation for bubbles. Bubbles exist in specifications where traders focus on the final liquidation value of the asset. But bubbles are much more dramatic when traders focus on the resale next period because their valuations are no longer tethered to liquidation values, and they bid up the asset's price based on the expectation of other trader's optimism next period. Indeed, we show that in a model with speculation, but not otherwise, even a small amount of diagnosticity in belief formation can lead to extremely large overvaluation during the rapid growth stage of the bubble. Even a mild departure from rational expectations, when combined with speculation, can entail extreme overvaluation.

These insights into the structure of asset price bubbles would not be obtained without modeling beliefs explicitly from fundamental psychological assumptions, and combining this with standard neoclassical mechanisms, such as learning from prices and speculation. But while this approach advances our understanding of the anatomy of price bubbles, it is only a first step. On the one hand, our theoretical setup is quite simple and we did not perform a full quantitative assessment of our mechanism. On the other hand, we have not considered further critical features of price bubbles, emphasized by Kindleberger but also obviously critical to the financial crisis of 2008 as well as other crises (Gennaioli and Shleifer 2018). These include leverage as well as the central involvement of banks and other financial institutions in the bubble episode. Introducing these elements into a model of bubbles with diagnostic expectations would get us closer to understanding the structure of financial fragility, beginning with basic features of expectations.

References

Alesina, Alberto, Armando Miano, and Stefanie Stantcheva. 2018. "Immigration and Redistribution," NBER Working Paper no. 24733.

Allen, Franklin, Stephen Morris, and Hyun Song Shin. 2006. "Beauty Contests and Iterated Expectations in Asset Markets." *Review of Financial Studies* 19 (3): 719–752.

Arnold, David, Will Dobbie and Crystal Yang. 2018. "Racial Bias in Bail Decisions," *Quarterly Journal of Economics* 133 (4): 1885–1932.

Bagehot, Walter. 1873. *Lombard Street: A Description of the Money Market*. New York: Wiley.

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.

Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer. 2018. "Extrapolation and Bubbles," *Journal of Financial Economics* 129 (2): 203–227.

Barberis, Nicholas, and Andrei Shleifer. 2003. "Style Investing." *Journal of Financial Economics* 68 (2): 161–199.

Barsky, Robert and Bradford De Long. 1993. "Why Does the Stock Market Fluctuate?" *Quarterly Journal of Economics* 108 (2): 291-311.

Black, Fischer. 1986. "Noise," *Journal of Finance* 41 (3): 528–543.

Blanchard, Olivier and Mark Watson. 1982. "Bubbles, Rational Expectations and Financial Markets," in *Crises in the Economic and Financial Structure*, Paul Wachtel, editor, 296-316.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer. 2016. "Stereotypes," *Quarterly Journal of Economics* 131 (4): 1753-1794.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer. 2018. "Beliefs about Gender," *American Economic Review* 109 (3): 739-773.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer. 2018. "Diagnostic Expectations and Credit Cycles," *Journal of Finance* 73 (1): 199-227.

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer. 2019. "Diagnostic Expectations and Stock Returns," *Journal of Finance*, forthcoming.

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer. 2019. "Over-reaction in Macroeconomic Expectations," NBER Working Paper no. 24932.

Case, Karl and Robert Shiller. 2003. "Is There a Bubble in the Housing Market?" *Brookings Papers on Economic Activity*, 2: 299-362.

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* 45: 265–315.

Cutler, David M., James M. Poterba, and Lawrence H. Summers. 1990. "Speculative Dynamics and the Role of Feedback Traders." *American Economic Review Papers and Proceedings* 80 (2): 63–68.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam. 1998. "Investor Psychology and Security Market Under- and Over-Reactions." *Journal of Finance* 53 (6): 1839–1885.

DeFusco, Anthony, Charles Nathanson, and Eric Zwick. 2018. "Speculative Dynamics of Prices and Volume," NBER Working Paper no. 23449.

DeLong, Bradford, Andrei Shleifer, Lawrence Summers, and Robert Waldmann. 1990a. "Noise Trader Risk in Financial Markets," *Journal of Political Economy* 98 (4): 703-738.

DeLong, Bradford, Andrei Shleifer, Lawrence Summers, and Robert Waldmann 1990b. "Positive Feedback Investment Strategies and Destabilizing Rational Speculation," *Journal of Finance* 45 (2): 379-395.

Fama, Eugene. 2014. "Two Pillars of Asset Pricing," *American Economic Review* 104 (6): 1467-1485.

Galbraith, John K. 1954. *The Great Crash: 1929*. Houghton Mifflin Co., Boston, MA.

Gennaioli, Nicola, and Andrei Shleifer. 2010. "What Comes to Mind," *Quarterly Journal of Economics* 125 (4): 1399-1433.

Gennaioli, Nicola, and Andrei Shleifer. 2018. *A Crisis of Beliefs: Investor Psychology and Financial Fragility*, Princeton, NJ: Princeton University Press.

Gennaioli, Nicola, Andrei Shleifer, and Robert Vishny. 2012. "Neglected Risks, Financial Innovation, and Financial Fragility." *Journal of Financial Economics* 104 (3): 452–68.

Giglio, Stefano, Matteo Maggiori, and Johannes Stroebel. 2016. "No-Bubble Condition: Model-Free Tests in Housing Markets," *Econometrica* 84 (3): 1047-1091.

Glaeser, Edward, Joseph Gyourko, and Albert Saiz. 2008. "Housing Supply and Housing Bubbles," *Journal of Urban Economics* 64 (2): 198-217.

Glaeser, Edward, and Charles Nathanson. 2017. "An Extrapolate Model of House Price Dynamics," *Journal of Financial Economics* 126 (1): 147-170.

Greenwood, Robin, and Samuel G. Hanson. 2014. "Waves in Ship Prices and Investment," *Quarterly Journal of Economics* 130 (1): 55-109.

Greenwood, Robin, and Andrei Shleifer. 2014. "Expectations of Returns and Expected Returns," *Review of Financial Studies* 27 (3): 714-746.

Greenwood, Robin, Andrei Shleifer, and Yang You. 2019. "Bubbles for Fama," *Journal of Financial Economics* 130 (1): 20-43.

Grossman, Sanford. 1976. "On the Efficiency of Competitive Stock Markets where Trades have Diverse Information," *Journal of Finance* 31 (2): 573–585.

Grossman, Sanford and Joseph Stiglitz. 1980. "On the Impossibility of Informationally Efficient Markets," *American Economic Review* 70 (3): 393–408.

Grossman, Sanford and Merton Miller. 1988. "Liquidity and Market Structure," *Journal of Finance* 43 (3): 617–633.

He, Hua, Jiang Wang. 1995. "Differential information and dynamic behavior of stock trading volume." *Review of Financial Studies* 8 (4): 919–972.

Harrison, Michael, and David Kreps. 1978. "Speculative Investor Behavior in a Stock Market with Heterogenous Expectations," *Quarterly Journal of Economics* 92 (2): 323-336.

Haruvy, Ernan, Yaron Lahav, and Charles N. Noussair. 2007. "Traders' Expectations in Asset Markets: Experimental Evidence." *American Economic Review* 97 (5): 1901-1920.

Hirshleifer, David, Jun Li, and Jianfeng Yu. 2015. "Asset Pricing in Production Economies with Extrapolative Expectations." *Journal of Monetary Economics* 76 (2015): 87–106.

Hong, Harrison, and Jeremy C. Stein. 1999. "A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets." *Journal of Finance* 54 (6): 2143–84.

Hong, Harrison, and Jeremy C. Stein. 2007. "Disagreement and the Stock Market." *Journal of Economic Perspectives* 21 (2): 109-128.

Jorda, Oscar, Moritz Schularick, and Alan Taylor. 2015. "Leveraged Bubbles," *Journal of Monetary Economics* 76S: 1–20.

Kahneman, Daniel, and Amos Tversky. 1972. "Subjective Probability: A Judgment of Representativeness," *Cognitive Psychology* 3 (3): 430–454.

Kahneman, Daniel, and Amos Tversky. 1983. "Extensional versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment." *Psychological Review* 90 (4): 293–315.

Kindleberger, Charles. 1978. *Manias, Panics, and Crashes: A History of Financial Crises*, New York: Basic Books.

La Porta, Rafael. 1996. "Expectations and the Cross-Section of Stock Returns." *Journal of Finance* 51 (5): 1715–42.

Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny. 1994. "Contrarian Investment, Extrapolation, and Risk." *Journal of Finance* 49 (5): 1541–78.

Mackay, Charles. 1841. *Extraordinary Popular Delusions and the Madness of Crowds*. New York: Harmony Books.

Martin, Alberto and Jaume Ventura. 2012. "Economic Growth with Bubbles." *American Economic Review* 102 (6): 3033-58.

Mian, Atif, and Amir Sufi. 2014. *House of Debt: How They (and You) Caused the Great Recession, and How We Can Prevent It from Happening Again*. Chicago, IL: University of Chicago Press.

Mian, Atif, and Amir Sufi. 2018. "Finance and Business Cycles: The Credit-Driven Household Demand Channel," *Journal of Economic Perspectives* 32 (3): 31–58.

Pastor, Lubos and Pietro Veronesi. 2006. "Was There a Nasdaq Bubble in the Late 1990s?" *Journal of Financial Economics* 81 (1): 61-100.

Pastor, Lubos and Pietro Veronesi. 2009. "Technological Revolutions and Stock Prices," *American Economic Review* 99 (4): 1451-1483.

Reinhart, Carmen M., and Kenneth S. Rogoff. 2009. *This Time Is Different: Eight Centuries of Financial Folly*. Princeton, NJ: Princeton University Press.

Scheinkman, José and Wei Xiong. 2003. "Overconfidence and Speculative Bubbles," *Journal of Political Economy* 111 (6): 1183-1220.

Shiller, Robert J. 2000. *Irrational Exuberance.* Princeton, NJ: Princeton University Press.

Shiller, Robert. 2015. *Irrational Exuberance, 3rd edition.* Princeton, NJ: Princeton University Press.

Smith, Vernon, Gerry Suchanek, and Arlington Williams. 1988. "Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets," *Econometrica* 56 (5): 1119-1151.

Tirole, Jean. 1985. "Asset Bubbles and Overlapping Generations." *Econometrica* 53 (5): 1071–1100.

Woodford, Michael. 2003. "Imperfect Common Knowledge and the Effects of Monetary Policy," in *Knowledge, Information and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, edited by P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford, 25-58. Princeton, NJ: Princeton Univ. Press.

Appendix A: Proofs.

Lemma 1. At time t, each trader has access to t signals of precision $\frac{1}{\sigma_\epsilon^2}$, as well as a prior with precision $\frac{1}{\sigma_V^2}$ around 0. By standard results in normal posterior updating, the trader's rational posterior is a normal distribution, with the following mean and variance.

$$\mathbb{E}_{i,t}(V) = \frac{t/\sigma_\epsilon^2}{t/\sigma_\epsilon^2 + 1/\sigma_V^2}\,\overline{s_{i,t}} = \pi_t \bar{s}_{it}$$

$$\sigma_t^2(V) = \frac{1}{\frac{t}{\sigma_\epsilon^2} + \frac{1}{\sigma_V^2}} = (1-\pi_t)\sigma_V^2$$

The diagnostic distribution then reads (up to normalization constants):

$$\exp\left\{-\frac{1}{2\sigma_t^2(V)}\left[\left(V - \mathbb{E}_{i,t}(V)\right)^2(1+\theta) - \theta\left(V - \mathbb{E}_{i,t-1}(V)\right)^2\right]\right\}$$

The quadratic and linear terms in V are given by (the constant terms are absorbed by normalization):

$$\exp\left\{-\frac{1}{2\sigma_t^2(V)}\left[V^2 - 2V\left(\mathbb{E}_{i,t}(V)(1+\theta) - \theta\mathbb{E}_{i,t-1}(V)\right)\right]\right\}$$

It follows that the diagnostic distribution is also a normal $\mathcal{N}\left(\mathbb{E}_{i,t}^\theta(V), Var_t^\theta(V)\right)$ with mean:

$$\mathbb{E}_{i,t}^\theta(V) = \mathbb{E}_{i,t}(V) + \theta\left[\mathbb{E}_{i,t}(V) - \mathbb{E}_{i,t-1}(V)\right]$$

and variance:

$$\left(\sigma_t^\theta\right)^2(V) = \sigma_t^2(V)$$

from which the result follows.

Proposition 1. For simplicity, denote $\alpha = \frac{\sigma_\epsilon^2}{\sigma_V^2}$. The consensus is then equal to $\frac{t}{t+\alpha}(1+\theta)V$ for $t \leq k$. Thus, the consensus is increasing from $t = 1$ to k. At $t = 1$, the price is below the fundamental iff $\frac{1}{1+\alpha}(1+\theta) < 1$ if $\theta < \alpha$. At $t = k$, the consensus overvalues the asset, namely $\frac{k}{k+\alpha}(1+\theta) > 1$ if $\theta > \alpha/k$. Next, let us show that $\mathbb{E}_{k+1}^\theta(V) < \mathbb{E}_k^\theta(V)$. It suffices to show:

$$\frac{(1+\theta)k}{k+\alpha} > \frac{(1+\theta)(k+1)}{k+1+\alpha} - \frac{\theta}{1+\alpha}$$
$$\Leftrightarrow \theta > \frac{\alpha}{k} \cdot \frac{\alpha+1}{k+2\alpha+1}$$

Whenever the consensus overvalues the asset at time t, namely $\theta > \alpha/k$, the above condition is satisfied, so the price declines at $t = k+1$. What happens from there on? Taking the derivative of the consensus with respect to t, one obtains:

$$\frac{d\mathbb{E}_t^{\theta}(V)}{dt} \propto \frac{(1+\theta)}{(t+\alpha)^2} - \frac{\theta}{(t-k+\alpha)^2}$$

The above expression is negative if and only if:

$$\frac{d\mathbb{E}_t^{\theta}(V)}{dt} < 0 \Leftrightarrow \left(\frac{t-k+\alpha}{t+\alpha}\right)^2 < \frac{\theta}{1+\theta} \Leftrightarrow t < \left(1 - \sqrt{\frac{\theta}{1+\theta}}\right)^{-1} \cdot k - \alpha$$

As a result, after $t^* \equiv max\left[k+1, \left(1-\sqrt{\frac{\theta}{1+\theta}}\right)^{-1} \cdot k - \alpha\right]$, the consensus increases monotonically. To conclude, observe that $[\pi_t + \theta(\pi_t - \pi_{t-k})]V \mapsto V$, because $\pi_t \mapsto 1$ as $t \mapsto \infty$.

Consider the other parametric cases. If $\theta > \alpha$ the consensus is immediately overvalued, and follows the same boom bust path as in the leading case. If $\theta < \alpha/k$ the consensus is below fundamental at $t = k$. From there on, it either monotonically increases toward the fundamental value, or it drops a bit but then increases toward the fundamental value. In either case, though, the consensus stays below the fundamental.

Proposition 2. We start by assuming a linear price formula of the form:

$$p_t = a_{2t}\mathbb{E}(V|P_t) + +b_t\left(V - \frac{c_t}{b_t}S_t\right)$$

for $t \leq k$. Denote $s_t^p = \frac{1}{b_t}(p_t - a_{t2}E[V|P_t]) = V - \frac{c_t}{b_t}S_t$ as the public signal obtained about V from the prices. Furthermore, let ζ_t be the precision of the public distribution, i.e. $\zeta_t = \frac{1}{\sigma_{p,t}^2}$, and use the shorthand $E_t^p = E[V|P_t]$. Using standard results of normal posteriors, we obtain:

$$E_t^p = \frac{1}{\sigma_S^2}\sum_{r=1}^{t-1} s_r^p \left(\frac{b_r}{c_r}\right)^2 \frac{1}{\zeta_t}$$

$$\zeta_t = \frac{1}{\sigma_V^2} + \frac{1}{\sigma_S^2}\sum_{r=1}^{t-1}\left(\frac{b_r}{c_r}\right)^2$$

Denoting $\overline{E_t}$ as the average rational fundamental beliefs, it follows that:

$$\bar{E}_t = \int \left[\frac{1}{\sigma_V^2} + \frac{1}{\sigma_S^2}\sum_{r=1}^{t-1} s_r^p \left(\frac{b_r}{c_r}\right)^2 \ \right] \frac{1}{\zeta_t}\, di = \frac{\frac{t}{\sigma_\epsilon^2}}{\frac{t}{\sigma_\epsilon^2} + \zeta_t} V + \frac{\zeta_t}{\frac{t}{\sigma_\epsilon^2} + \zeta_t} E_t^p$$

Then, as $t \leq k$, we have $\overline{E_t}^\theta = (1+\theta)\,\bar{E}_t$, and hence our equilibrium condition

$$p_t = \overline{E_t}^\theta - \gamma \sigma_t^2(V) S_t$$

translates to:

$$p_t = (1+\theta)\left(\frac{\frac{t}{\sigma_\epsilon^2}}{\frac{t}{\sigma_\epsilon^2} + \zeta_t} V + \frac{\zeta_t}{\frac{t}{\sigma_\epsilon^2} + \zeta_t} E_t^p\right) - \gamma \frac{S_t}{\frac{t}{\sigma_\epsilon^2} + \zeta_t}$$

Matching coefficients, we obtain:

$$a_{2t} = (1+\theta)\frac{\zeta_t}{\frac{t}{\sigma_\epsilon^2} + \zeta_t}$$

$$b_t = (1+\theta)\frac{\frac{t}{\sigma_\epsilon^2}}{\frac{t}{\sigma_\epsilon^2} + \zeta_t}$$

$$c_t = \gamma\left(\frac{t}{\sigma_\epsilon^2} + \zeta_t\right)^{-1}$$

In particular, note $\frac{b_t}{c_t} = \frac{1+\theta}{\gamma} \cdot \frac{t}{\sigma_\epsilon^2}$, and plugging this into our expression for ζ_t, one obtains:

$$\zeta_t = \frac{1}{\sigma_V^2} + \frac{(1+\theta)^2}{\gamma^2 \sigma_S^2}\sum_{r=1}^{t-1}\frac{r^2}{\sigma_\epsilon^4} = \frac{1}{\sigma_V^2} + \frac{(1+\theta)^2}{\gamma^2 \sigma_S^2 \sigma_\epsilon^4}\frac{(t-1)t(2t-1)}{6}$$

Consider now the average price $\overline{p_t}$, obtained by setting the supply shocks to their average, $S_t = 0$. Plugging in E_t^p and then ζ_t we find:

$$\overline{p_t} = (1+\theta)\left(\frac{\frac{t}{\sigma_\epsilon^2} + \zeta_t - \frac{1}{\sigma_V^2}}{\frac{t}{\sigma_\epsilon^2} + \zeta_t}\right) V = \frac{(1+\theta)\left(\frac{t}{\sigma_\epsilon^2} + \frac{(1+\theta)^2}{\gamma^2 \sigma_S^2 \sigma_\epsilon^4}\frac{(t-1)t(2t-1)}{6}\right)}{\left(\frac{1}{\sigma_V^2} + \frac{t}{\sigma_\epsilon^2} + \frac{(1+\theta)^2}{\gamma^2 \sigma_S^2 \sigma_\epsilon^4}\frac{(t-1)t(2t-1)}{6}\right)} V.$$

Proposition 3. Under learning from prices, the price at time t is $p_t = \mathbb{E}_t^\theta(V | p_{t-1}, \dots, p_1; \vec{s}_t, \dots, \vec{s}_1)$. From Proposition 2, the average price path is:

$$p_t = (1+\theta)\left[\frac{\frac{t}{\sigma_\epsilon^2} + \left(\frac{1+\theta}{\sigma_S \sigma_\epsilon^2}\right)^2 \frac{t(t-1)(2t-1)}{6}}{\frac{t}{\sigma_\epsilon^2} + \frac{1}{\sigma_V^2} + \left(\frac{1+\theta}{\sigma_S \sigma_\epsilon^2}\right)^2 \frac{t(t-1)(2t-1)}{6}}\right] V$$

To explore convexity, rewrite price as:

$$p_t = (1+\theta)V \frac{f}{f+c}$$

where $f = k_1 t + k_2 t(t-1)(2t-1)$, $c = \frac{1}{\sigma_V^2}$, $k_1 = \frac{t}{\sigma_\epsilon^2}$ and $k_2 = \frac{1}{6}\left(\frac{1+\theta}{\sigma_S \sigma_\epsilon^2}\right)^2$. Note that, up to a constant, we have:

$$\partial_t^2 p_t = \partial_t \frac{f'}{(f+c)^2} = (f+c)^{-2}\left[f'' - \frac{2(f')^2}{(f+c)}\right]$$

which has the same sign as $f'' - \frac{2(f')^2}{(f+c)}$. This is positive when:

$$c > \frac{2(f')^2}{f''} - f$$

Convexity requires c to be large, that is σ_V^2 to be small. For example, convexity at $t = 1$ requires:

$$c > \frac{k_1^2 - k_1 k_2 + k_2^2}{3k_2}$$

Rewrite the condition above as:

$$f'' c > 2(f')^2 - f f''$$

that is

$$6k_2(2t-1)c > 2\left(k_1 + k_2(6t^2 - 6t + 1)\right)^2 - 6k_2(2t-1)\left(k_1 t + k_2(2t^3 - 3t^2 + t)\right)$$

As we are looking near t = 1, let us set $s = t - 1$. Then, the inequality simplifies to:

$$6k_2(2s+1)c > 48k_2^2 s^4 + 96k_2^2 s^3 + 66k_2 s^2 + 12k_1 k_2 s^2 + 18k_2^2 s + 6k_1 k_2 s + 2k_1^2 + 2k_2^2 - 2k_1 k_2$$

In particular, the right-hand side is a quartic with positive coefficients (and in particular it is convex in s), whereas the left hand side is a linear function. Hence, if the left-hand side lies above the right hand side at $t = 1$, the two will cross at $t = t^* > 1$, and never cross again. Hence, for the average price path to be convex at $t \in [1, t^*]$ and concave afterwards, it is necessary and sufficient for the above inequality to hold at $s = 0$, which is given by:

$$c > \frac{k_1^2 - k_1 k_2 + k_2^2}{3k_2} \rightarrow \sigma_V^2 < \left(\frac{k_1^2 - k_1 k_2 + k_2^2}{3k_2}\right)^{-1} = \sigma_V^{*2}$$

as desired.

Proposition 4. Consider the price path

$$p_t = (1+\theta)^{T-t+1}\left[\prod_{r=t}^{T}\pi_r\right]V$$

where $\pi_t = \frac{t}{t+\frac{\sigma_\epsilon^2}{\sigma_V^2}}$ and denote $\hat{\pi} \equiv [\prod_{r=1}^{T}\pi_r]^{\frac{1}{T}}$. First note that $p_t - p_{t-1} = p_t[1-(1+\theta)\pi_{t-1}]$. So price increases at t if $(1+\theta)\pi_{t-1} < 1$ and it decreases otherwise. Because π_t is monotonically increasing, it follows that the price path is either always increasing (if $(1+\theta)\pi_T < 1$), or always decreasing (if $(1+\theta)\pi_1 > 1$) or is first increasing and then decreasing. This holds provided $(1+\theta)\pi_1 < 1 < (1+\theta)\pi_T$, which reads $\theta \in \left(\frac{\sigma_\epsilon^2}{T\sigma_V^2}, \frac{1-\hat{\pi}}{\hat{\pi}}\right)$. In particular, we then have both $p_1 = (1+\theta)^T[\prod_{r=1}^{T}\pi_r]V < V$ and $p_T = (1+\theta)\pi_T V > V$.

Appendix B: Material on Dynamic Trading

1. Recap of He and Wang Dynamic REE

We shall closely follow the notation of He and Wang 1995. Let $Q_{t+1} = P_{t+1} - P_t$ be the price innovation at time $t+1$. Let $\Psi_{t+1}^i = [1,\ \mathbb{E}_{i,t+1}(V), \mathbb{E}_{p,t+1}(V), \mathbb{E}_{pt}(V)]$ be the relevant state space. (Adding the lagged value of public expectations ensures that P_{t+1} is a linear function of the state variables.) As always, denote γ as the risk aversion parameter. We denote:

$$Q_{t+1} = A_{Q,t+1}\,\Psi_t^i + B_{Q,t+1}\epsilon_{t+1}^i$$

$$\Psi_{t+1}^i = A_{\Psi,t+1}\,\Psi_t^i + B_{\Psi,t+1}\epsilon_{t+1}^i,$$

where ϵ_{t+1}^i is a vector of shocks (including both the supply shock, the innovation to fundamentals, and signal noise). Setting X_t as the portfolio weight, we have that wealth evolves according to $W_{t+1}^i = W_t^i + X_t^i Q_{t+1}$. The individual optimizes $\mathbb{E}_{it}^{\theta}\left[-e^{-\gamma W_T^i}\right]$, where $W_{final}^i = W_T^i + X_T^i(V - P_T)$.

According to He and Wang 1995, one can apply backwards induction on the proposed value function. They show that the value function takes the following convenient form:

$$J(W_{t+1}^i;\ \Psi_{t+1}^i) = -\exp\left[-\gamma W_{t+1}^i - \frac{1}{2}(\Psi_{t+1}^i)' U_{t+1}\Psi_{t+1}^i\right],$$

Standard operations involving multivariate Gaussian integrals can be used to solve the coefficients recursively. By setting:

$$\Xi_{t+1} = \left(\Sigma_{t+1}^{-1} + B'_{\Psi,t+1}U_{t+1}B_{\Psi,t+1}\right)^{-1}$$

$$\rho_{t+1} = \sqrt{|\Xi_{t+1}|/|\Sigma_{t+1}|}$$

$$F_t = \left[B_{Q,t+1}\Xi_{t+1}B'_{Q,t+1}\right]^{-1}\left(A_{Q,t+1} - B_{Q,t+1}\Xi_{t+1}B'_{\Psi,t+1}\,U_{t+1}A_{\Psi,t+1}\right),$$

we can complete the recursion for U_t by computing

$$M_t = F_t'\left(B_{Q,t+1}\Xi_{t+1}B'_{Q,t+1}\right)F_t - \left(B'_{\Psi,t+1}U_{t+1}A_{\Psi,t+1}\right)'\Xi_{t+1}\left(B'_{\Psi,t+1}U_{t+1}A_{\Psi,t+1}\right),$$

and setting $U_t = M_t - 2\log\rho_{t+1}I_{1,1}^{4\times4}$, where $I_{1,1}^{4\times4}$ is a matrix that only has the (1, 1) element set to 1. Furthermore, one can easily show that the portfolio choice for individual i is given by:

$$X_t^i = \frac{1}{\gamma} F_t \Psi_t^i.$$

Thus, we need the following information to start computing the equilibrium coefficients:

1. The update matrix $A_{Q,t+1}, A_{\Psi,\mathrm{t+1}}, B_{Q,t+1}, B_{\Psi,\mathrm{t+1}}$.
2. The final objective matrix U_T (In the final period, Q_{T+1} should be $V - P_T$)

2. Applying dynamic REE to our setting

A. Introducing diagnostic distortions

Recall that our state space for each individual's dynamic programming problem was $\Psi_{t+1}^i = [1\ \ \mathbb{E}_{i,t+1}(V)\ \mathbb{E}_{p,t+1}(V)\ \mathbb{E}_{pt}(V)]$. A natural way to model diagnostic expectations in this notation is to assume that the individual has distorted expectations of the transition probabilities of the future state space. Under rationality and normal shocks, the distribution of the future states consists of a multivariate normal. One can show then that the diagnostic distribution of the future state $f^{\theta}(\Psi_{t+1}|F_t)$ consists of a shift of the rational distribution by $1+\theta$ of its current mean, as done in BGMS 2019. In the notation of He and Wang 1995, this is equivalent to distorting the transition matrix $A_{\Psi,\mathrm{t+1}}$ by:

$$A_{\Psi,\mathrm{t+1}}^{diag} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1+\theta & 0 & 0 \\ 0 & 0 & 1+\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot A_{\Psi,\mathrm{t+1}},$$

while maintaining $B_{\Psi,\mathrm{t+1}}$. Note that the entries corresponding to the constant term and E_{pt} are not inflated. This is because the two quantities are measurable at time t, and hence diagnostic expectations does not distort the quantity. Denoting $P_t = \Lambda_t \Psi_t$, we then obtain:

$$A_{\mathrm{Q,t+1}}^{diag} = \Lambda_{t+1} A_{\Psi,\mathrm{t+1}}^{diag} - \Lambda_t$$

$$B_{Q,t+1} = \Lambda_{t+1} B_{\Psi,\mathrm{t+1}}.$$

All that remains for our purposes are specifying U_t, which is the dynamic component of the objective function. For this, we use the U_t^{rat} computed from the above recursive equations assuming $\theta = 0$. Intuitively, the agent trades by using U_t^{rat} as sufficient statistics to account for dynamic trading. This assumption greatly simplifies computation, and can be interpreted as a diagnostic perturbation of the dynamic rational expectations equilibrium: we are preserving the rational dynamic motivation, while introducing a distortion in future state transition probabilities.

B. Market clearing and computing the equilibrium coefficients

As in Section 3, we stipulate the following equilibrium coefficients (a_t, b_t, c_t), where

$$p_t = a_t\, E_{pt} + b_t V - c_t S_t.$$

Setting $\zeta_t = \frac{1}{\sigma_{pt}^2}$ as the accumulated precision of the public price signals by time t, we have the following equations from standard Bayesian computations:

$$\zeta_t = \frac{1}{\sigma_V^2} + \frac{1}{\sigma_S^2} \sum_{r=1}^{t} \left(\frac{b_r}{c_r}\right)^2$$

$$E_{pt} = \zeta_t^{-1} \left(\frac{1}{\sigma_S^2} \sum_{r=1}^{t} \left(\frac{b_r}{c_r}\right)^2 \left(V - \frac{c_r}{b_r} S_r\right) \right)$$

$$E_{it}[V] = \frac{\frac{t}{\sigma_\epsilon^2}}{\frac{t}{\sigma_\epsilon^2} + \zeta_t} \ \ \bar{s}_t^{\,i} + \frac{\zeta_t}{\frac{t}{\sigma_\epsilon^2} + \zeta_t} E_{pt}.$$

By definition, according to individual i's beliefs, the true value V is distributed according to: $V \sim N\left(E_{i,t}[V], \left(\zeta_t + \frac{t}{\sigma_\epsilon^2}\right)^{-1}\right)$. Hence, we can express: $V = E_{i,t}[V] + u,\ u \sim N\left(0,\ \left(\zeta_t + \frac{t}{\sigma_\epsilon^2}\right)^{-1}\right)$. Furthermore, denote $s_t^i = V + \epsilon_t^i$. For simplicity, let us suppress the terms corresponding to the constant term in the state space – they remain 1. After standard algebra, one obtains the following rational transition equation for the private and public expectations of the fundamental value.

$$\begin{pmatrix} E_{p,t+1} \\ E_{i,t+1}[V] \end{pmatrix} = \begin{pmatrix} \frac{\zeta_t}{\zeta_{t+1}} & 1 - \frac{\zeta_t}{\zeta_{t+1}} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} E_{p,t} \\ E_{i,t}[V] \end{pmatrix} + \begin{pmatrix} \frac{\zeta_{t+1} - \zeta_t}{\zeta_{t+1}} \left(u - \frac{c_{t+1}}{b_{t+1}} S_{t+1}\right) \\ \frac{\frac{1}{\sigma_\epsilon^2} + \zeta_{t+1} - \zeta_t}{\frac{t+1}{\sigma_\epsilon^2} + \zeta_{t+1}} u + \frac{\frac{1}{\sigma_\epsilon^2}}{\frac{t+1}{\sigma_\epsilon^2} + \zeta_{t+1}} \epsilon_{t+1}^i - \frac{\zeta_{t+1} - \zeta_t}{\frac{t+1}{\sigma_\epsilon^2} + \zeta_{t+1}} \frac{c_{t+1}}{b_{t+1}}\ S_{t+1} \end{pmatrix}$$

This implies:

$$A_{\Psi,\text{t+1}} = \begin{pmatrix} 1 & 0 & 0 \\ 1-\dfrac{\zeta_t}{\zeta_{t+1}} & \dfrac{\zeta_t}{\zeta_{t+1}} & 0 \\ 0 & 1 & 0 \end{pmatrix} \rightarrow A_{\Psi,\text{t+1}}^{diag} = \begin{pmatrix} 1+\theta & 0 & 0 \\ (1+\theta)\left(1-\dfrac{\zeta_t}{\zeta_{t+1}}\right) & (1+\theta)\dfrac{\zeta_t}{\zeta_{t+1}} & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

$$B_{\Psi,\text{t+1}} = \begin{pmatrix} \dfrac{\dfrac{1}{\sigma_\epsilon^2} + \zeta_{t+1} - \zeta_t}{\dfrac{t+1}{\sigma_\epsilon^2} + \zeta_{t+1}} & \dfrac{\dfrac{1}{\sigma_\epsilon^2}}{\dfrac{t+1}{\sigma_\epsilon^2} + \zeta_{t+1}} & -\dfrac{\zeta_{t+1} - \zeta_t}{\dfrac{t+1}{\sigma_\epsilon^2} + \zeta_{t+1}} \dfrac{c_{t+1}}{b_{t+1}} \\ 1-\dfrac{\zeta_t}{\zeta_{t+1}} & 0 & -\left(1-\dfrac{\zeta_t}{\zeta_{t+1}}\right)\dfrac{c_{t+1}}{b_{t+1}} \\ 0 & 0 & 0 \end{pmatrix}$$

with the variance of $\epsilon_t^i = \begin{pmatrix} u \\ \epsilon \\ S \end{pmatrix}$ set to $\begin{pmatrix} \left(\zeta_t + \frac{t}{\sigma_\epsilon^2}\right)^{-1} & 0 & 0 \\ 0 & \sigma_\epsilon^2 & 0 \\ 0 & 0 & \sigma_S^2 \end{pmatrix}$. Furthermore, one obtains:

$$p_{t+1} = \Lambda_{t+1}\Psi_{t+1} = \begin{pmatrix} 0 & a_{t+1} - \dfrac{b_{t+1}}{1-\dfrac{\zeta_t}{\zeta_{t+1}}} & -b_{t+1}\dfrac{\dfrac{\zeta_t}{\zeta_{t+1}}}{1-\dfrac{\zeta_t}{\zeta_{t+1}}} \end{pmatrix} \cdot \Psi_{t+1}$$

As mentioned above,

$$A_{\text{Q,t+1}}^{diag} = \Lambda_{t+1} A_{\Psi,\text{t+1}}^{diag} - \Lambda_t$$

$$B_{Q,t+1} = \Lambda_{t+1} B_{\Psi,\text{t+1}}.$$

To compute the equilibrium coefficients, we impose market-clearing: the investor demand should equal the supply at t:

$$\frac{1}{\gamma} F_t^{diag} \int \Psi_t^{\text{i}}\, di = S_t.$$

Note:

$$\int \Psi_t^i di = \begin{pmatrix} V \cdot \dfrac{\dfrac{t}{\sigma_\epsilon^2}}{\dfrac{t}{\sigma_\epsilon^2} + \zeta_t} + \dfrac{\zeta_t}{\dfrac{t}{\sigma_\epsilon^2} + \zeta_t} E_{p,t} \\ E_{p,t} \\ E_{p,t-1} \end{pmatrix} = \begin{pmatrix} \dfrac{\zeta_t}{\dfrac{t}{\sigma_\epsilon^2} + \zeta_t} & \dfrac{\dfrac{t}{\sigma_\epsilon^2}}{\dfrac{t}{\sigma_\epsilon^2} + \zeta_t} & 0 \\ 1 & 0 & 0 \\ \dfrac{\zeta_t}{\zeta_{t-1}} & -\dfrac{1}{\sigma_S^2}\left(\dfrac{b_t}{c_t}\right)^2 \dfrac{1}{\zeta_{t-1}} & \dfrac{1}{\sigma_S^2}\dfrac{b_t}{c_t}\dfrac{1}{\zeta_{t-1}} \end{pmatrix} \begin{pmatrix} E_{p,t} \\ V \\ S_t \end{pmatrix}$$

$$= \Gamma_t(a_t, b_t, c_t) \begin{pmatrix} E_{p,t} \\ V \\ S_t \end{pmatrix}.$$

Hence, one can solve for the equilibrium coefficients (a_t, b_t, c_t) that satisfies the market-clearing conditions:

$$\underbrace{\frac{1}{\gamma}F_t^{diag}(a_t, b_t, c_t)}_{Dynamic\ diagnostic\ demand\ given\ rational\ beliefs} \cdot \underbrace{\Gamma_t(a_t, b_t, c_t)\begin{pmatrix} E_{p,t} \\ V \\ S_t \end{pmatrix}}_{average\ rational\ beliefs} = \underbrace{(0 \quad 0 \quad 1)\begin{pmatrix} E_{p,t} \\ V \\ S_t \end{pmatrix}}_{supply}.$$

Using the formula we have for $F_t^{diag} = \left[B_{Q,t+1}\Xi_{t+1}B'_{Q,t+1}\right]^{-1}\left(A^{diag}_{Q,t+1} - B_{Q,t+1}\Xi_{t+1}B'_{\Psi,t+1}\, U_{t+1}A^{diag}_{\Psi,t+1}\right)$,

one can use the above market clearing equation to numerically solve for (a_t, b_t, c_t) inductively backwards, given the boundary conditions[23]:

$$a_T = \frac{(1+\theta)\zeta_T}{\frac{T}{\sigma_\epsilon^2} + \zeta_T}$$

$$b_T = \frac{\frac{(1+\theta)T}{\sigma_\epsilon^2}}{\frac{T}{\sigma_\epsilon^2} + \zeta_T}$$

$$c_T = \gamma\left(\frac{T}{\sigma_\epsilon^2} + \zeta_T\right)^{-1}$$

In summary, the equilibrium coefficients can be computed from the following procedure:

1. Guess the final public precision ζ_T in a given grid.
2. Compute the boundary coefficients (a_T, b_T, c_T).
3. Inductively compute the coefficients (a_t, b_t, c_t) backwards.
4. Verify $\zeta_0 = \zeta_1 - \frac{1}{\sigma_S^2}\left(\frac{b_1}{c_1}\right)^2 = \frac{1}{\sigma_V^2}$.

[23] It is not entirely trivial to solve for the coefficients numerically from the above equations. First, one can deduce that the equilibrium coefficients must satisfy:

$$\Lambda_t = \Lambda_{t+1}A^{diag}_{\Psi,t+1} - B_{Q,t+1}\Xi_{t+1}B'_{\Psi,t+1}U_{t+1}A^{diag}_{\Psi,t+1} - \gamma \cdot B_{Q,t+1}\Xi_{t+1}B'_{Q,t+1}(0 \quad 0 \quad 1)\Gamma_t^{-1}.$$

As the first coefficient of Λ_t is 0, one can show that this pins down ζ_{t-1} as a univariate zero, from which one can easily compute Γ_t. The concrete coefficients (a_t, b_t, c_t) then follows from $\Lambda_t = \left(0 \quad a_t - \frac{b_t}{1-\frac{\zeta_{t-1}}{\zeta_t}} \quad -b_t\frac{\frac{\zeta_{t-1}}{\zeta_t}}{1-\frac{\zeta_{t-1}}{\zeta_t}}\right)$. The precise numerical procedure can be given upon request.