Journal of Financial Economics 129 (2018) 203–227

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

Extrapolation and bubbles[☆]

Nicholas Barberis[a], Robin Greenwood[b,*], Lawrence Jin[c], Andrei Shleifer[d]

[a] Yale School of Management, P.O. Box 208200, New Haven, CT 06520, USA
[b] Harvard Business School, Soldiers Field, Boston, MA 02163, USA
[c] California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125, USA
[d] Harvard University, 1805 Cambridge Street, Cambridge, MA 02138, USA

ARTICLE INFO

Article history:
Received 25 January 2016
Revised 6 February 2017
Accepted 30 May 2017
Available online 4 May 2018

JEL classification:
G02
G11
G12

Keywords:
Bubble
Extrapolation
Volume

ABSTRACT

We present an extrapolative model of bubbles. In the model, many investors form their demand for a risky asset by weighing two signals—an average of the asset's past price changes and the asset's degree of overvaluation—and "waver" over time in the relative weight they put on them. The model predicts that good news about fundamentals can trigger large price bubbles, that bubbles will be accompanied by high trading volume, and that volume increases with past asset returns. We present empirical evidence that bears on some of the model's distinctive predictions.

1. Introduction

In classical accounts of financial market bubbles, the price of an asset rises dramatically over the course of a few months or even years, reaching levels that appear to far exceed reasonable valuations of the asset's future cash flows. These price increases are accompanied by widespread speculation and high trading volume. The bubble eventually ends with a crash, in which prices collapse even more quickly than they rose. Bubble episodes have fascinated economists and historians for centuries (e.g., Mackay, 1841; Bagehot, 1873; Galbraith, 1954; Kindleberger, 1978; Shiller, 2000), in part because human behavior in bubbles is so hard to explain, and in part because of the devastating side effects of the crash.

At the heart of the standard historical narratives of bubbles is the concept of extrapolation—the formation of expected returns by investors based on past returns. In these narratives, extrapolators buy assets whose prices have risen because they expect them to keep rising. According to Bagehot (1873), "owners of savings … rush into anything that promises speciously, and when they find that these specious investments can be disposed of at a high profit, they rush into them more and more." These historical narratives are supported by more recent research on investor expectations, using both survey data and lab experiments. Case et al. (2012) show that in the U.S. housing market, homebuyers' expectations of future house price appreciation are closely related to lagged house price appreciation. Greenwood and Shleifer (2014) present survey

☆ We are grateful to Bill Schwert (the editor), an anonymous referee, Marianne Andries, Alex Chinco, Charles Nathanson, Alp Simsek, Adi Sunderam, and seminar participants at Berkeley, Caltech, Cornell, Northwestern, NYU, Ohio State, Yale, the AEA, the Miami Behavioral Finance Conference, the NBER, and the WFA for very helpful comments.

* Corresponding author.
E-mail addresses: nick.barberis@yale.edu (N. Barberis), rgreenwood@hbs.edu (R. Greenwood), jin@caltech.edu (L. Jin), shleifer@fas.harvard.edu (A. Shleifer).

https://doi.org/10.1016/j.jfineco.2018.04.007
0304-405X/

evidence of expectations of stock market returns and find strong evidence of extrapolation, including during the Internet bubble. Extrapolation also shows up in data on expectations of participants in experimental bubbles, where subjects can be explicitly asked about their expectations of returns. Both the classic study of Smith et al. (1988) and more recent experiments such as Haruvy et al. (2007) find direct evidence of extrapolative expectations during a well-defined experimental price bubble.

In this paper, we present a new model of bubbles based on extrapolation. In doing so, we seek to shed light on two key features commonly associated with bubbles. The first is what Kindleberger (1978) called "displacement"—the fact that nearly all bubbles from tulips to South Sea to the 1929 U.S. stock market to the late 1990s Internet occur on the back of good fundamental news. We would like to understand which patterns of news are likely to generate the largest bubbles, and whether a bubble can survive once the good news comes to an end. Second, we would like to explain the crucial fact that bubbles feature very high trading volume (Galbraith, 1954; Carlos et al., 2006; Hong and Stein, 2007). At first sight, it is not clear how extrapolation can explain this: if, during a bubble, all extrapolators hold similarly bullish views, they will not trade with each other.

To address these questions, we present a model in the spirit of earlier work by Cutler et al. (1990), De Long et al. (1990), Hong and Stein (1999), Barberis and Shleifer (2003), and Barberis et al. (2015), but with some significant new elements.[1] There is a risk-free asset and a risky asset that pays a liquidating cash flow at a fixed time in the future. Each period, news about the value of the final cash flow is publicly released. There are two types of investors. The first type is extrapolators, who form their share demand based on an extrapolative "growth signal," which is a weighted average of past price changes. In a departure from prior models, extrapolators also put some weight on a "value signal" which measures the difference between the price and a rational valuation of the final cash flow. The two signals, which can be interpreted as "greed" and "fear," give the extrapolator conflicting messages. If prices have been rising strongly and the asset is overvalued, the growth signal encourages him to buy ("greed") while the value signal encourages him to sell ("fear").

Our second departure from prior models is to assume that, at each date, and independently of other extrapolators, each extrapolator slightly but randomly shifts the relative weight he puts on the two signals. This assumption, which we refer to as "wavering," reflects extrapolators' ambivalence about how to balance the conflicting signals they face. Such wavering has a biological foundation in partially random allocation of attention to various attributes of choice, which in our case are growth and value signals (see Fehr and Rangel, 2011). Importantly, the degree of wavering is constant over time. We show that wavering can plausibly account for a good deal of evidence other models have trouble with.

As in earlier models, extrapolators are met in the market by fundamental traders who lean against the wind, buying the asset when its price is low relative to their valuation of the final cash flow and selling when its price is high. Both extrapolators and fundamental traders face short-sale constraints.

In line with the Kindleberger (1978) notion of displacement, a bubble forms in our model after a sequence of large positive cash-flow shocks. The bubble evolves in three stages. In the first stage, the cash-flow news pushes up the price of the risky asset; extrapolators sharply increase their demand for the asset, buying from fundamental traders. In the second stage, the asset becomes sufficiently overvalued that the fundamental traders exit the market, leaving the asset in the hands of the exuberant extrapolators who trade with each other because of wavering. Once the good cash-flow news subsides, prices stop rising as rapidly, extrapolator enthusiasm abates, and the bubble begins its collapse. In the third stage, prices fall far enough that fundamental traders re-enter the market, buying from extrapolators.

In our model, the largest bubbles arise from sequences of cash-flow shocks that first increase in magnitude, and then decrease. Wavering can significantly increase the size of a bubble through a novel mechanism that we call a "price spiral." During a bubble, the asset can become so overvalued that even some extrapolators hit their short-sale constraints. The bubble selects only the most bullish investors as asset holders, which leads to an even greater overvaluation, causing even more extrapolators to leave. The bubble takes on a life of its own, persisting well after the end of the positive cash-flow news.

The model predicts substantial volume in the first and third stages of a bubble, as fundamental traders sell to extrapolators and vice versa. But it predicts particularly intense trading during the height of the bubble as extrapolators, as a consequence of wavering, trade among themselves. During normal times, wavering has very little impact on trading volume because it is minor. During bubbles, in contrast, the same small degree of wavering that generates little volume in normal times endogenously generates intense volume: the growth and value signals that extrapolators attend to are now so large in magnitude that even tiny shifts in their relative weights lead to large portfolio adjustments. One manifestation of such adjustments, exemplified by Isaac Newton's participation in the South Sea bubble, is extrapolators getting in, out, and back in the market.

After presenting the model, we compare it to two standard approaches to modeling bubbles: rational bubbles (Blanchard and Watson, 1982; Tirole, 1985) and disagreement (Harrison and Kreps, 1978; Scheinkman and Xiong, 2003). Models of rational bubbles assume homogeneous investors and therefore cannot explain any volume, let alone highly specific patterns of volume documented in the literature. In addition, direct tests of the key prediction of rational bubbles—that payoffs in the infinite future have positive present value—reject that prediction (Giglio et al., 2016). Disagreement-based models can explain high

[1] These earlier papers use models of return extrapolation to examine excess volatility, return predictability, and nonzero return autocorrelations. They do not discuss bubbles. Glaeser and Nathanson (2017) analyze housing bubbles using a return extrapolation framework, albeit one that is different from ours.

volume during bubbles with an exogenous increase in disagreement. In our model, in contrast, the increase in volume is due to disagreement that grows endogenously over the course of the bubble. Indeed, in our model, volume during a bubble is predicted by past returns, a new prediction that other bubble models do not share. Our framework is also more successful at matching the extrapolative expectations that many investors hold during bubble periods.

We examine empirically some of the model's predictions. Using data from four historical bubbles, we document that trading volume during a bubble is strongly predicted by the past return. For the technology bubble of the late 1990s, we also show that, as the bubble progresses, a larger fraction of trading volume is due to investors with extrapolator-like characteristics. Finally, we present direct evidence of wavering for both mutual funds and hedge funds invested in technology stocks.

Some recent research has questioned whether bubble-like price episodes are actually irrational (Pastor and Veronesi, 2006) or whether bubbles in the sense of prices undeniably and substantially exceeding fundamentals over a period of time ever exist (Fama, 2014). Although the existence of bubbles in this sense appears uncontroversial in experimental (Smith et al., 1988) or some unusual market (Xiong and Yu, 2011) settings, our paper does not speak to these controversies. Rather, we show how a simple model of extrapolative bubbles explains a lot of evidence and makes new predictions.

In the next section, we present our model. Sections 3 and 4 describe circumstances under which bubbles occur and present our findings for price patterns and volume. Section 5 considers the possibility of negative bubbles. Section 6 compares our model to other models of bubbles while Section 7 presents the empirical evidence. Section 8 concludes. Appendix A contains all the proofs.

2. A model of bubbles

We consider an economy with $T+1$ dates, $t=0,1,\ldots,T$. There are two assets: one risk-free and one risky. The risk-free asset earns a constant return which we normalize to zero. The risky asset has a fixed supply of Q shares, and each share is a claim to a dividend $\widetilde{D}_T$ paid at the final date, T. The value of $\widetilde{D}_T$ is given by

$$\widetilde{D}_T = D_0 + \widetilde{\varepsilon}_1 + \cdots + \widetilde{\varepsilon}_T, \tag{1}$$

where

$$\widetilde{\varepsilon}_t \sim N(0, \sigma_\varepsilon^2), \text{ i.i.d. over time.} \tag{2}$$

The value of D_0 is public information at time 0, while the value of $\widetilde{\varepsilon}_t$ is realized and becomes public information at time t. The price of the risky asset, P_t, is determined endogenously.

There are two types of traders in the economy: fundamental traders and extrapolators. The time t per capita demand of fundamental traders for shares of the risky asset is

$$\frac{D_t - \gamma\sigma_\varepsilon^2(T-t-1)Q - P_t}{\gamma\sigma_\varepsilon^2}, \tag{3}$$

where $D_t = D_0 + \sum_{j=1}^{t} \varepsilon_j$ and γ is fundamental traders' coefficient of absolute risk aversion.

In Appendix A, we show how this expression can be derived from utility maximization. In brief, it is the demand of an investor who, at each time, maximizes a constant absolute risk aversion (CARA) utility function defined over next period's wealth, and who is boundedly rational: he uses backward induction to determine his time t demand, but, at each stage of the backward induction process, he assumes that, in future periods, the other investors in the economy will simply hold their per capita share of the risky asset supply. In other words, he does not have a detailed understanding of how other investors in the economy form their share demands. For this investor, the expression $D_t - \gamma\sigma_\varepsilon^2(T-t-1)Q$ in the numerator of (3) is the expected price of the risky asset at the next date, date $t+1$. The numerator is therefore the expected price *change* over the next period, and the fundamental trader's demand is this expected price change scaled by the trader's risk aversion and by his estimate of the risk he is facing. If all investors in the economy were fundamental traders, then, setting the expression in (3) equal to the risky asset supply of Q, the equilibrium price of the risky asset would be

$$D_t - \gamma\sigma_\varepsilon^2(T-t)Q. \tag{4}$$

We call this the "fundamental value" of the risky asset and denote it by P_t^F.[2]

Extrapolators are the second type of trader in the economy. There are I types of extrapolators, indexed by $i \in \{1, 2, \ldots, I\}$; we explain below how one type of extrapolator differs from another. We build up our specification of extrapolator demand for the risky asset in three steps. An initial specification of per capita extrapolator share demand at time t is

$$\frac{X_t}{\gamma\sigma_\varepsilon^2}, \quad \text{where} \quad X_t \equiv (1-\theta)\sum_{k=1}^{t-1}\theta^{k-1}(P_{t-k} - P_{t-k-1}) + \theta^{t-1}X_1, \tag{5}$$

and where $0 < \theta < 1$.

In Appendix A, we show that this is the demand of an investor who, at each time, maximizes a CARA utility function defined over next period's wealth, and whose belief about the expected price change of the risky asset over the next period is a weighted average of past price changes, with more recent price changes weighted more heavily. The parameter X_1 is a constant that measures extrapolator enthusiasm at time 1; in our numerical analysis, we assign it a neutral, steady-state value. The specification in (5) is similar to that in previous models of extrapolative beliefs, which have been used to shed light on asset pricing anomalies (Cutler et al., 1990; De Long et al., 1990; Hong and Stein, 1999; Barberis and Shleifer, 2003; Barberis et al., 2015).[3]

[2] We assume, for simplicity, that fundamental traders' estimate of the risk they are facing is given by fundamental risk σ_ε^2 rather than by the conditional variance of price changes. When fundamental traders are the only traders in the economy, this approximation is exact.

[3] The form of bounded rationality we have assumed for fundamental traders means that these traders expect the price of the risky asset to

We modify the specification in (5) in two quantitatively small but conceptually significant ways. First, we make extrapolators pay at least some attention to how the price of the risky asset compares to its fundamental value. Specifically, we change the demand function in (5) so that the demand of extrapolator i takes the form

$$w_i\left(\frac{D_t - \gamma\sigma_\varepsilon^2(T-t-1)Q - P_t}{\gamma\sigma_\varepsilon^2}\right) + (1-w_i)\left(\frac{X_t}{\gamma\sigma_\varepsilon^2}\right). \quad (6)$$

Extrapolator i's demand is now a weighted average of two components. The second component is the expression we started with in (5), while the first component is the fundamental trader demand from (3); w_i is the weight on the first component. Our framework accommodates any $w_i \in (0, 1]$, but we maintain $w_i < 0.5$ for all i so that the extrapolative component is weighted more heavily. In our numerical work, the value of w_i is approximately 0.1. The motivation for (6) is that even extrapolators worry about how the price of the risky asset compares to its fundamental value. A high price relative to fundamental value exerts some downward pressure on their demand, counteracting the extrapolative component.

In what follows, we refer to the two components of the demand function in (6) as "signals": the first component, the expression in (3), is a "value" signal; the second component, the expression in (5), is a "growth" signal. These signals typically point in opposite directions. If the price of the risky asset is well above fundamental value, it has usually also been rising recently. The value signal then takes a large negative value, telling the investor to reduce his position, while the growth signal takes a large positive value, telling the investor to raise it. The signals can be informally interpreted in terms of "fear" and "greed." If the price has recently been rising, the value signal captures extrapolators' fear that it might fall back to fundamental value, while the growth signal captures greed, their excitement at the prospect of more price rises. If the price has recently been falling, the growth signal captures extrapolators' fear of further price declines, and the value signal, their greed—their excitement at the thought of prices rebounding toward fundamental value.[4]

Our second modification is to allow the weight w_i to vary *slightly* over time, and independently so for each extrapolator type, so that the demand function for extrapolator i becomes

$$w_{i,t}\left(\frac{D_t - \gamma\sigma_\varepsilon^2(T-t-1)Q - P_t}{\gamma\sigma_\varepsilon^2}\right) + (1-w_{i,t})\left(\frac{X_t}{\gamma\sigma_\varepsilon^2}\right), \quad (7)$$

where (7) differs from (6) only in the t subscript added to $w_{i,t}$. Since the demand function in (6) is based on two signals that often point in opposite directions, the investor is likely to be unsure of what to do—and, in particular, unsure about how to weight the signals at any point in time. As we model it, the weight an extrapolator puts on each signal shifts or "wavers" over time, to a small extent. Fehr and Rangel (2011) and Towal et al. (2013) argue that individual decisions are shaped by computations in the brain, which are mediated by the allocation of attention to various attributes of choice that is in part random. We can think of wavering as resulting from such random allocation of extrapolators' attention to growth and value signals.

To model wavering, we set

$$\begin{aligned} w_{i,t} &= \overline{w}_i + \widetilde{u}_{i,t} \\ \widetilde{u}_{i,t} &\sim N(0, \sigma_u^2), \text{ i.i.d. over time and across extrapolators.} \end{aligned} \quad (8)$$

Here, $\overline{w}_i \in (0,\ 1]$ is the *average* weight that extrapolator i puts on the value signal; in our numerical analysis, we set $\overline{w}_i = 0.1$ for all extrapolator types. The *actual* weight that extrapolator i puts on the value signal at time t is the mean weight $\overline{w}_i$ plus Normally-distributed noise. To ensure that $w_{i,t}$ stays in the (0,1] interval, we truncate the distribution of $\widetilde{u}_{i,t}$.[5] Under our assumptions, the I types of extrapolator differ only in the weight $w_{i,t}$ that they put on the value signal in each period. The values of the two signals themselves are identical across extrapolators.[6]

We also impose a short-sale constraint, so that the final risky asset share demand of the fundamental traders, N_t^F, and of extrapolator $i \in \{1, 2, \ldots, I\}$, $N_t^{E,i}$, are given by

$$N_t^F = \max\left[\frac{D_t - \gamma\sigma_\varepsilon^2(T-t-1)Q - P_t}{\gamma\sigma_\varepsilon^2}, 0\right] \quad (9)$$

and

$$\begin{aligned} N_t^{E,i} = \max\Bigg[& w_{i,t}\left(\frac{D_t - \gamma\sigma_\varepsilon^2(T-t-1)Q - P_t}{\gamma\sigma_\varepsilon^2}\right) \\ & + (1-w_{i,t})\left(\frac{X_t}{\gamma\sigma_\varepsilon^2}\right), 0\Bigg]. \end{aligned} \quad (10)$$

As we explain in Section 4, the short-sale constraint is not needed for our most important results. In contrast, the assumption that extrapolators waver slightly between a growth and a value signal is crucial.

In the Online Appendix, we show that our principal results also hold in a model with fundamental traders who are more fully rational in that they understand how extrapolators form their demand. For tractability, we assume both in the main text and the Online Appendix that investors maximize the expected utility of next period's

revert to fundamental value within one period. This, in turn, means that they trade aggressively against any mispricing—more aggressively than if they were fully rational. In the latter case, they would recognize that extrapolator demand is persistent and trade more conservatively against it, or even in the same direction as extrapolators (De Long et al., 1990; Brunnermeier and Nagel, 2004).

[4] We use the term "growth signal" both for $X_t/\gamma\sigma_\varepsilon^2$ and for X_t itself. When it is important to clarify which of the two quantities is being referred to, we do so.

[5] Specifically, we truncate $\widetilde{u}_{i,t}$ at $\pm 0.9\min(\overline{w}_i, 1-\overline{w}_i)$, a formulation that allows the fundamental trader demand in (3) to be a special case of the more general demand function in (7) and (8), namely, the case where $\overline{w}_i = 1$. The exact form of truncation is not important for our results.

[6] We think of extrapolator i's beliefs at time t about the price change $P_{t+1} - P_t$ as being equal to $w_{i,t}(D_t - \gamma\sigma_\varepsilon^2(T-t-1)Q - P_t) + (1-w_{i,t})X_t$, in other words, a weighted average of the beliefs of a fundamental trader and of a "pure" extrapolator. When coupled with a CARA utility function defined over next period's wealth, these beliefs lead to the demand function in Eq. (7).

wealth. Barberis et al. (2015) study asset prices in an economy where extrapolators and rational traders maximize *lifetime* consumption utility. While they do not address facts about bubbles, their predictions for prices are similar to ours, which suggests that our assumption of myopic demand is innocuous.

In Proposition 1 in Appendix A, we show that, in the economy described above, a unique market-clearing price always exists and is given by

$$P_t = D_t + \frac{\sum_{i\in I^*}\mu_i(1-w_{i,t})}{\sum_{i\in I^*}\mu_i w_{i,t}}X_t - \gamma\sigma_\varepsilon^2 Q\frac{(\sum_{i\in I^*}\mu_i w_{i,t})(T-t-1)+1}{\sum_{i\in I^*}\mu_i w_{i,t}}, \quad (11)$$

where μ_0 and μ_i are the fraction of fundamental traders and of extrapolators of type i in the population, respectively, so that $\sum_{i=0}^{I}\mu_i = 1$, and where I^* is the set of $i \in \{0, 1, \ldots, I\}$ such that trader i has strictly positive demand for the risky asset at time t. The statement of Proposition 1 explains how I^* is determined at each time t.[7]

The first term on the right-hand side of (11) shows that the price of the risky asset is anchored to the expected value of the final cash flow. The second term reflects the impact of extrapolator demand: if past price changes have been high, so that X_t is high, extrapolator demand at time t is high, exerting upward pressure on the price. The third term is a price discount that compensates the holders of the risky asset for the risk they bear.

We define the "steady state" of our economy as the state to which the economy converges after many periods in which the cash-flow shocks equal zero. It is straightforward to show that, in this steady state: the fundamental traders and *all* the extrapolators are in the market, with each trader holding the risky asset in proportion to his weight in the population; the price of the risky asset equals the fundamental value in (4); the change in price from one date to the next is constant and equal to $\gamma\sigma_\varepsilon^2 Q$; and the growth signal X_t, defined in (5), is also equal to $\gamma\sigma_\varepsilon^2 Q$.

2.1. Parameter values

In Sections 3 and 4, we explore the predictions of our model through both analytical propositions and numerical analysis. We now discuss the parameter values that we use in the numerical analysis. The asset-level parameters are D_0, Q, σ_ε, and T. The investor-level parameters are I, μ_i and $\overline{w}_i$ for $i \in \{0, 1, \ldots, I\}$, γ, θ, and σ_u.

We begin with θ, which governs the weight extrapolators put on recent as opposed to distant past price changes when forming beliefs about future price changes; as such, it determines the magnitude of the growth signal X_t. In setting θ, we are guided by the survey evidence analyzed by Greenwood and Shleifer (2014) on the beliefs of actual stock market investors about future returns. If we assume that the time period in our model is a quarter, the evidence and the calculations in Barberis et al. (2015) imply $\theta \approx 0.9$.[8]

We set μ_0, the fraction of fundamental traders in the economy, to 0.3, so that fundamental traders make up 30% of the population, and extrapolators, 70%. The survey evidence in Greenwood and Shleifer (2014) suggests that many investors in the economy are extrapolators. We have $I = 50$ types of extrapolators, where each type has the same population weight, so that $\mu_i = (1-\mu_0)/I$ for $i = 1, \ldots, I$. As discussed earlier, we set $\overline{w}_i$ to the same low value of 0.1 for all extrapolators. And we set γ to 0.1. We do not have strong priors about the value of σ_u, which controls the degree of wavering. We assign it a low value—specifically, 0.03—so as to show that even a small degree of wavering can generate interesting results. This value of σ_u implies that, about two-thirds of the time, the weight $w_{i,t}$ extrapolator i puts on the value signal is in the interval (0.07,0.13), a very small amount of wavering relative to the base weight $\overline{w}_i = 0.1$.

We set the initial expected dividend D_0 to 100, the standard deviation of cash-flow shocks σ_ε to 3, the risky asset supply Q to 1, and the number of dates T to 50. Since the interval between dates is a quarter, this value of T means that the life span of the risky asset is 12.5 years.

3. Asset prices in a bubble

Our model can generate the most essential feature of a bubble, namely, a large and growing overvaluation of the risky asset, where, by overvaluation, we mean that the price exceeds the fundamental value in (4). In the model, bubbles are initiated by a sequence of large, positive cash-flow shocks, which here are news about the future liquidating dividend. Fig. 1 illustrates this. It uses the parameter values from Section 2 and Eqs. (1), (4), (5), and (11) to plot the price (solid line) and fundamental value (dashed line) of the risky asset for a particular 50-period sequence of cash-flow shocks, in other words, for a particular set of values of $\tilde{\varepsilon}_1$, $\tilde{\varepsilon}_2$,..., $\tilde{\varepsilon}_{50}$. The first ten shocks, $\tilde{\varepsilon}_1$ through $\tilde{\varepsilon}_{10}$, are all equal to zero. These are followed by four positive shocks, namely, 2, 4, 6, and 6; these are substantial shocks: the last two are two-standard deviation shocks. These are followed by 36 more shocks of zero.[9]

Once the positive shocks arrive, a large and sustained overpricing follows. The positive cash-flow news pushes prices up, which leads the extrapolators to sharply increase their share demand in subsequent periods; this, in turn, pushes prices well above fundamental value. Over the four periods of positive cash-flow news, starting at date 11, the expected final dividend increases by 18, the sum of 2, 4, 6,

[7] Here and elsewhere, we index fundamental traders by the number "0." If $i = 0$ is in the set I^*, the expression in (11) requires the value of $w_{0,t}$, in other words, the weight fundamental traders put on the value signal. By definition, $w_{0,t} = 1$.

[8] Specifically, $\theta = \exp(-(0.5)(0.25)) \approx 0.9$, where 0.5 is the Barberis et al. (2015) estimate of the extrapolation parameter in a continuous-time framework, and 0.25 corresponds to the one-quarter interval between consecutive dates in our model.

[9] We set the growth signal at time 1, X_1, equal to the steady-state value of X, namely, $\gamma\sigma_\varepsilon^2 Q = 0.9$. This, together with the fact that the first ten cash-flow shocks are equal to zero, means that the price of the risky asset equals the asset's fundamental value for the first ten periods.

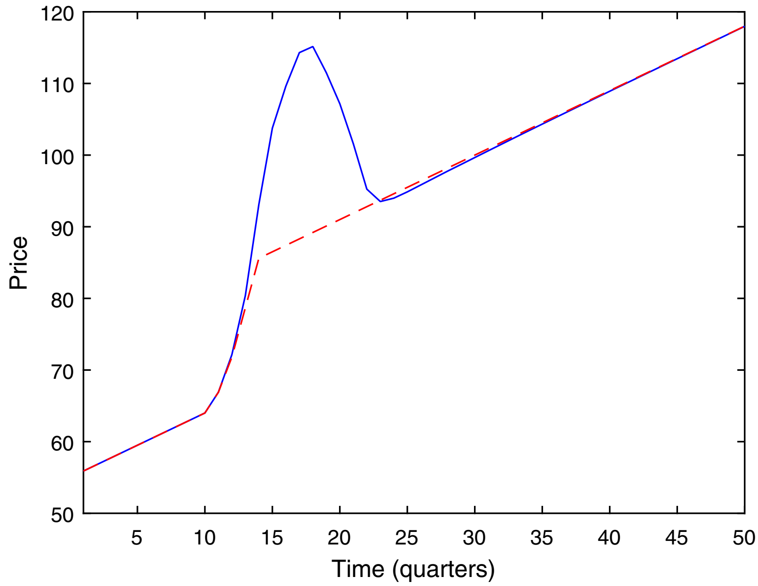


Fig. 1. Prices in a bubble. The solid line plots the price of the risky asset for the following sequence of 50 cash-flow shocks: 10 shocks of zero, followed by shocks of 2, 4, 6, 6, followed by 36 shocks of zero. 30% of the investors are fundamental traders; the remainder are extrapolators with an extrapolation parameter θ of 0.9 and who also put a base weight $\overline{w}_i = 0.1$ on a value signal. The dashed line plots the fundamental value of the asset for the same cash-flow sequence. The other parameters are $D_0 = 100$, $T = 50$, $\sigma_\varepsilon = 3$, $Q = 1$, $\gamma = 0.1$, $\sigma_u = 0.03$, and $I = 50$.

and 6. The figure shows, however, that between dates 11 and 18, prices rise by more than double this amount. After the cash-flow shocks drop back to zero at date 15, prices stop rising as rapidly; this, in turn, cuts off the "fuel" driving extrapolator demand. These investors eventually start reducing their demand and the bubble collapses.

This bubble has three distinct stages defined by the composition of the investor base. In the first stage, the fundamental traders are still in the market: even though the risky asset is overvalued, the overvaluation is sufficiently mild that the short-sale constraint does not bind for the fundamental traders. In our example, this first stage spans just one date, date 12. Fig. 1 shows that, during this stage, the overvaluation is small in magnitude: precisely because the fundamental traders are present in the market, they absorb much of the demand pressure from extrapolators by selling to them.

The second stage of the bubble begins when the risky asset becomes so overvalued that the fundamental traders exit the market. In our example, this occurs at date 13. During this stage, extrapolators alone trade the risky asset, which becomes progressively more overvalued: the high past price changes make the extrapolators increasingly enthusiastic, and there is no countering force from fundamental traders. In the absence of cash-flow news, however, the price increases eventually decline in magnitude, extinguishing extrapolator enthusiasm and causing the bubble to deflate.

To see how the bubble in Fig. 1 bursts, note that, from Eq. (11), the size of the bubble depends on the magnitude of the growth signal X_t, itself a measure of extrapolator enthusiasm. From Eq. (5), this signal evolves as

$$X_{t+1} = \theta X_t + (1-\theta)(P_t - P_{t-1}). \tag{12}$$

The first term on the right-hand side, θX_t, indicates that the bubble has a natural tendency to deflate; recall that $0 < \theta < 1$. As time passes, the sharply positive price changes that excited the extrapolators recede into the past; they are therefore downweighted, by an amount θ, reducing extrapolator enthusiasm. However, if the recent price change $P_t - P_{t-1}$ is sufficiently positive, both the growth signal and the bubble itself can maintain their size. Once the good cash-flow news comes to an end—after date 14 in our example—it becomes increasingly unlikely that the recent price change is large enough to offset the bubble's tendency to deflate, in other words, that the second term on the right-hand side of (12) dominates the first. As a consequence, the price level starts falling, sharply reducing extrapolator enthusiasm and setting in motion the collapse of the bubble.[10]

The third stage of the bubble begins when the bubble has deflated to such an extent that the fundamental traders re-enter the market. In our example, this occurs at date 23. In this example, both the fundamental traders *and* the extrapolators are present in the market in this stage. For other cash-flow sequences, the price declines during the collapse of the bubble can be so severe as to cause the extrapolators to exit the market, leaving the asset in the hands of the fundamental traders for some period of time.

Our prediction that, in the presence of extrapolators, a sequence of strongly positive cash-flow news leads to a large overvaluation holds for a wide range of parameter values. Fig. 2 illustrates this. The four graphs in the figure correspond to four important model parameters: μ_0, the

[10] The use of leverage can amplify the effects of extrapolation, leading to larger bubbles and more dramatic collapses. See Simsek (2013) and Jin (2015) for analyses of this idea.

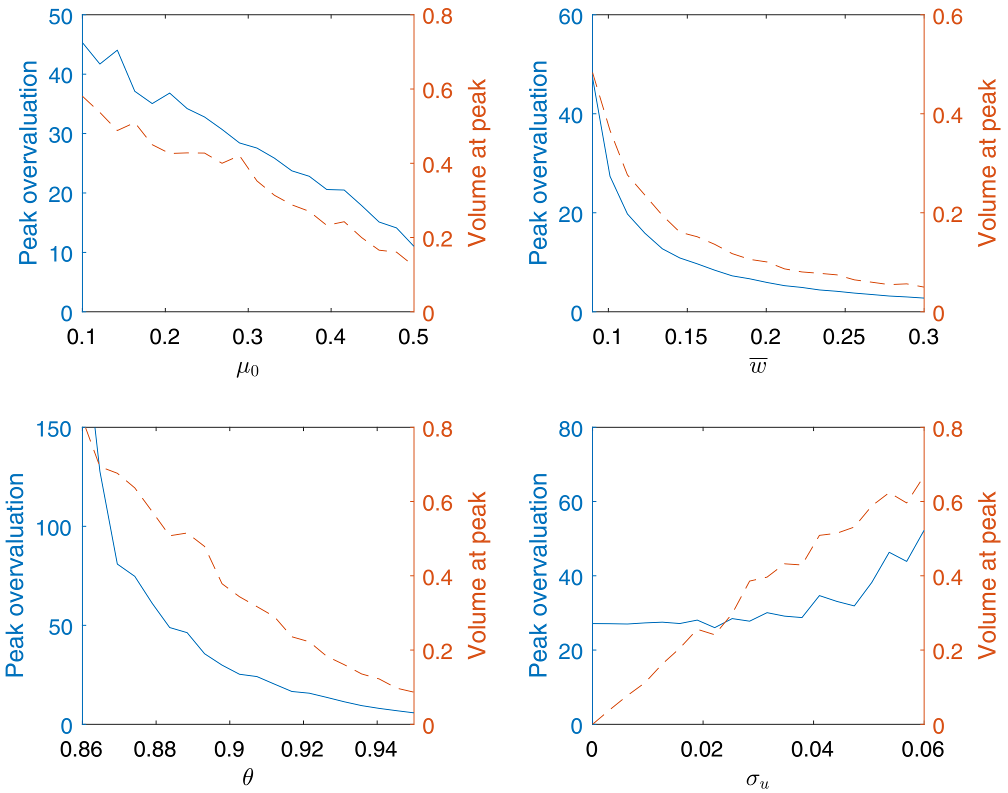


Fig. 2. Comparative statics. The four graphs correspond to four model parameters: μ_0, $\overline{w}$, θ, and σ_u. In each graph, the solid line plots the maximum overvaluation of the risky asset across the $T = 50$ dates for the same sequence of cash-flow shocks used in Fig. 1. The dashed line plots trading volume on the date of peak overvaluation. The solid and dashed lines are computed by varying the value of the parameter on the horizontal axis while keeping the values of the other parameters at the benchmark levels specified in Section 2.

fraction of fundamental traders in the population; $\overline{w}$, the average weight that each extrapolator puts on the value signal; θ; and σ_u, the degree of wavering. In each graph, the solid line plots the maximum overvaluation of the risky asset across the 50 dates for the cash-flow sequence used in Fig. 1, where "overvaluation" means the difference between price and fundamental value. The dashed line, which we discuss in Section 4, plots the volume of trading in the risky asset at the moment of peak overvaluation. For each graph, we generate the solid and dashed lines by varying the value of the parameter on the horizontal axis while keeping all other parameter values at the benchmark levels listed in Section 2.[11]

The figure confirms that our model generates a large overvaluation for a wide range of parameter values. Not surprisingly, lower values of μ_0 and $\overline{w}$ increase the magnitude of overvaluation. More interestingly, lower values of θ also generate bubbles that are larger in size. To understand this, suppose that there is good cash-flow news at time $t-1$ that pushes up the asset price. When θ is low, extrapolators become much more bullish at time t, precisely because they put a lot of weight on the most recent price change. This means that $P_t - P_{t-1}$ is high, which, from (12), means that X_{t+1} is high, and hence that X_{t+2} is also high. Since the growth signal X is an important determinant of bubble size, this explains why a low θ generates a large bubble. However, the fact that X_t in Eq. (12) is scaled by θ also means that, when θ is low, the bubble deflates faster after reaching its peak. A low θ therefore generates bubbles that are more "intense"—they feature a high degree of overvaluation but are short-lived.

Wavering does not play a significant role in the evolution of the price path in Fig. 1. If we replaced the extrapolators in our model with extrapolators who all put the same, invariant weight of 0.1 on the value signal, we would obtain a price path almost identical to that in Fig. 1. The reason is that, for the particular sequence of cash-flow shocks used in Fig. 1, virtually *all* of the extrapolators are present in the market during all three stages of the bubble. By the law of large numbers, the aggregate demand of $I = 50$ extrapolators whose weight on the value signal equals 0.1 is approximately equal to the aggregate demand of $I = 50$ extrapolators whose weight on the value signal is drawn from a distribution with *mean* 0.1. As a result, the pricing of the risky asset is very similar whether the extrapolators are homogeneous or waver.

The reasoning in the previous paragraph explains why, in the bottom-right graph in Fig. 2, the degree of

[11] The unevenness in the lines in Fig. 2 is due to the randomness introduced by wavering.

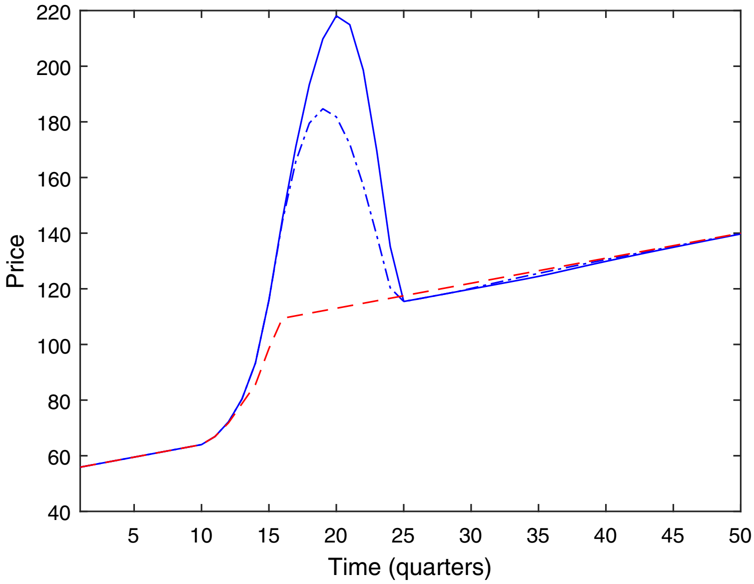


Fig. 3. Price spiral. The solid line plots the price of the risky asset for the following sequence of 50 cash-flow shocks: 10 shocks of zero, followed by shocks of 2, 4, 6, 6, 12, 10, followed by 34 shocks of zero. 30% of the investors are fundamental traders; the remainder are extrapolators with an extrapolation parameter θ of 0.9 and who also put a base weight $\overline{w}_i = 0.1$ on a value signal. The dashed line plots the fundamental value of the asset for the same cash-flow sequence. The dash–dot line plots the price in an economy where the extrapolators are homogeneous, placing the same invariant weight of $\overline{w}_i = 0.1$ on the value signal. The other parameters are $D_0 = 100$, $T = 50$, $\sigma_\varepsilon = 3$, $Q = 1$, $\gamma = 0.1$, $\sigma_u = 0.03$, and $I = 50$.

overvaluation remains unchanged as we increase the level of wavering from zero to 0.03. However, the graph shows that additional increases in the level of wavering do lead to higher overvaluation. This is due to a novel bubble mechanism that we call a "price spiral." If the level of wavering is sufficiently high or the cash-flow shocks are sufficiently large, then, during the second stage of the bubble, when the fundamental traders are out of the market, the asset can become so overvalued that even some extrapolators exit the market—specifically, those who put the highest weight $w_{i,t}$ on the value signal. Once these extrapolators leave the market, the asset is left in the hands of the more enthusiastic extrapolators, who put more weight on the growth signal. This generates an even larger overvaluation, causing yet more extrapolators to hit their short-sale constraints and leaving the asset in the hands of an even more enthusiastic group of extrapolators. Eventually, in the absence of positive cash-flow shocks, the price increases become less dramatic and extrapolator demand abates, causing the bubble to deflate. At this point, extrapolators who had previously exited the market begin to re-enter.

Fig. 3 depicts a price spiral. The parameter values are the same as for Fig. 1, but we now use the cash-flow sequence 2, 4, 6, 6, 12, 10 in place of 2, 4, 6, 6. The dashed line plots the asset's fundamental value, while the solid line plots its price. For comparison, the dash–dot line plots the price in an economy where the extrapolators are homogeneous, placing the same, invariant weight of 0.1 on the value signal. For this cash-flow sequence, wavering significantly amplifies the degree of overpricing: the solid line rises well above the dash–dot line. As explained above, this is due to some extrapolators exiting the market, starting at date 15; at the peak of the price spiral around date 20, about half of the extrapolators are out of the market.[12]

Price spirals typically deflate within a few periods. In some cases, however—specifically, for sequences of very large, positive cash-flow shocks—the price spiral can lead to extremely high prices. We do not put much weight on this prediction. First, the cash-flow shocks required for such out-of-control spirals are so large as to be unlikely in reality. Second, factors absent from our model, such as equity issuance by firms, are likely to prevent these extreme spirals from arising.

In Proposition 2 in Appendix A, we show how the magnitude of the asset's overvaluation at time t can be expressed as a function of the cash-flow shocks that have been realized up until that time. Suppose that the economy has been in its steady state up to time $l-1$ and that there is then a sequence of positive shocks $\varepsilon_l, \varepsilon_{l+1}, \ldots, \varepsilon_n$ that move the economy from the first stage of the bubble to the second stage of the bubble at some intermediate date j with $l < j < n$. Suppose also that the bubble remains in the second stage through at least date $N > n$, and that all the extrapolators are in the market in the second stage, so that there is no price spiral. The proposition shows that, in this case, the overvaluation at time t in the second stage,

[12] The price spiral we have just described can also result from a type of heterogeneity that is simpler than wavering, one where extrapolators differ in the weight they put on the value signal, but this weight is constant over time for each extrapolator, so that $w_{i,s} = w_{i,t}$ for all s, t. While the stochasticity embedded in wavering is not required for price spirals to occur, it is crucial for the volume predictions in Section 4.

where $j \le t \le N$, is approximately equal to

$$\sum_{m=j}^{t-1} \mathcal{L}_2(t-m)\varepsilon_m, \tag{13}$$

where the coefficients $\mathcal{L}_2(t-m)$ depend only on the model parameters and not on the values of the shocks. Indeed, the coefficients depend on just two parameters: θ, which governs the relative weight extrapolators put on recent as opposed to distant past price changes when forming beliefs, and $\overline{w}$, the mean weight that extrapolators put on the value signal. The "2" subscript in $\mathcal{L}_2(\cdot)$ indicates that the coefficients are applied to cash-flow news that arrives during the second stage of the bubble: the summation in (13) starts at time j, when the second stage begins.[13]

The expression in (13) shows that, to a first approximation, the degree of overvaluation in the second stage has a simple linear structure: it is approximately a weighted sum of the cash-flow news in the second stage, where the weights are constant. For example, if there have been eight cash-flow shocks during the second stage of the bubble, namely, $\varepsilon_{t-8}, \varepsilon_{t-7}, \ldots, \varepsilon_{t-1}$, then, for the parameter values we are using, the degree of overvaluation at time t is approximately

$$\begin{aligned}&\mathcal{L}_2(1)\varepsilon_{t-1} + \mathcal{L}_2(2)\varepsilon_{t-2} + \cdots + \mathcal{L}_2(7)\varepsilon_{t-7} + \mathcal{L}_2(8)\varepsilon_{t-8}\\ &= 0.9\varepsilon_{t-1} + 1.62\varepsilon_{t-2} + 2.11\varepsilon_{t-3} + 2.33\varepsilon_{t-4} + 2.3\varepsilon_{t-5}\\ &\quad +2.05\varepsilon_{t-6} + 1.61\varepsilon_{t-7} + 1.06\varepsilon_{t-8}.\end{aligned} \tag{14}$$

This expression shows that the cash-flow news that contributes the most to time t overvaluation—the shock with the largest coefficient—is the news from four periods earlier, ε_{t-4}. This news causes a price increase at time $t-4$, which increases extrapolator enthusiasm at time $t-3$, thereby causing a large price increase at that time as well; this, in turn, increases extrapolator enthusiasm at time $t-2$, and so on. Through its accumulated effect on prices over several periods, the cash-flow news ε_{t-4} has a large impact on time t overvaluation. By contrast, the most recent cash-flow news, ε_{t-1}, has a smaller effect on time t overvaluation: much of its impact will come *after* time t. The more distant cash-flow news ε_{t-8} also has a small effect on time t overvaluation. While that shock contributed to price increases at time $t-8$ and on a few subsequent dates, those price increases are now so far in the past that they have little impact on extrapolator beliefs at time t.

The expression in (14) helps us understand what kinds of cash-flow sequences generate the largest bubbles. More concretely, which sequence $\{\varepsilon_{t-8}, \ldots, \varepsilon_{t-1}\}$ leads to the largest overvaluation at time t? To generate a large bubble at time t, we want to associate the highest value of ε with the highest value of $\mathcal{L}_2(\cdot)$, namely, 2.33; the second highest value of ε with the second highest value of $\mathcal{L}_2(\cdot)$, namely, 2.3, and so on. Since the highest values of the $\mathcal{L}_2(\cdot)$ coefficients are for lags that are temporally close—specifically, for lags 3, 4, 5, and 6—this means that the largest bubbles occur when the biggest cash-flow shocks are clumped together in time. More generally, since the $\mathcal{L}_2(\cdot)$ coefficients rise to a peak and then decline, the largest bubble is created by a sequence of cash-flow news that itself rises to a peak and then declines. For example, if ε_{t-8} through ε_{t-1} take the values 1, 2, 3, 4, 5, 6, 7, 8, in some order, the above discussion suggests that the largest time t overvaluation is generated by the ordering 2, 3, 5, 7, 8, 6, 4, 1—and this is indeed the case.

To compute the frequency of large bubbles in our model, we use the cash-flow process in (1) and the price process in (11) to simulate a $T = 40,000$-period price sequence and record the number of bubbles for which the level of overvaluation exceeds a threshold such as 10 or 20, and also the length of time for which this threshold is exceeded. To put these bubble sizes in context, recall that, in non-bubble times, a one-standard deviation cash-flow shock increases the asset's price by approximately 3.

In our model, large bubbles are rare. For our benchmark parameter values, a bubble whose size exceeds 10 occurs once every 17 years, on average, with the overvaluation exceeding 10 for approximately one year. A bubble of size 20 occurs just once every 50 years, on average, and maintains this size for approximately three quarters. Bubbles are rare for two reasons. First, for a sizable bubble to occur, the cash-flow shocks need to be large enough to cause the fundamental traders to exit. Second, for a large bubble to form, the cash-flow shocks need to be both large *and* clumped together in time. The probability of this happening is low. Fig. 2 suggests that large bubbles arise more frequently for lower values of μ_0 and $\overline{w}$, and, more interestingly, for lower values of θ. Our simulations confirm this.

To conclude our analysis of prices, we verify, again through simulations, that the model also captures the basic asset pricing patterns that the previous generation of extrapolation models was designed to explain. Specifically, we confirm that the model generates excess volatility (the standard deviation of price changes exceeds the standard deviation of changes in fundamental value); predictability (the difference between the price and the fundamental value at time t, $P_t - P_t^F$, predicts the change in price over the next 12 periods, $P_{t+12} - P_t$, with a negative sign); and positive (negative) autocorrelations in price changes at short (long) lags. It is not surprising that our framework can generate these patterns: while we modify the earlier extrapolation models in qualitatively significant ways, these modifications are quantitatively small.

4. Volume in a bubble

Bubbles feature very high trading volume (Ofek and Richardson, 2003; Hong and Stein, 2007). A central goal of our paper is to propose a way of understanding this fact.[14]

Fig. 4 plots the share demand N_t^F of the fundamental traders (dashed line) and the share demands $N_t^{E,i}$ of the I

[13] The proposition presents analogous results for the first stage of the bubble and also for the second stage in the case of a price spiral. For tractability, it assumes a continuum of extrapolators rather than a finite number of them.

[14] A small fraction of bubbles, often those associated with debt securities, do not exhibit very high trading volume. Hong and Sraer (2013) explain this by noting that, if investors are over-optimistic about the value of the asset underlying a debt security and also differ in how optimistic they are, they overvalue the debt security but do not disagree about its

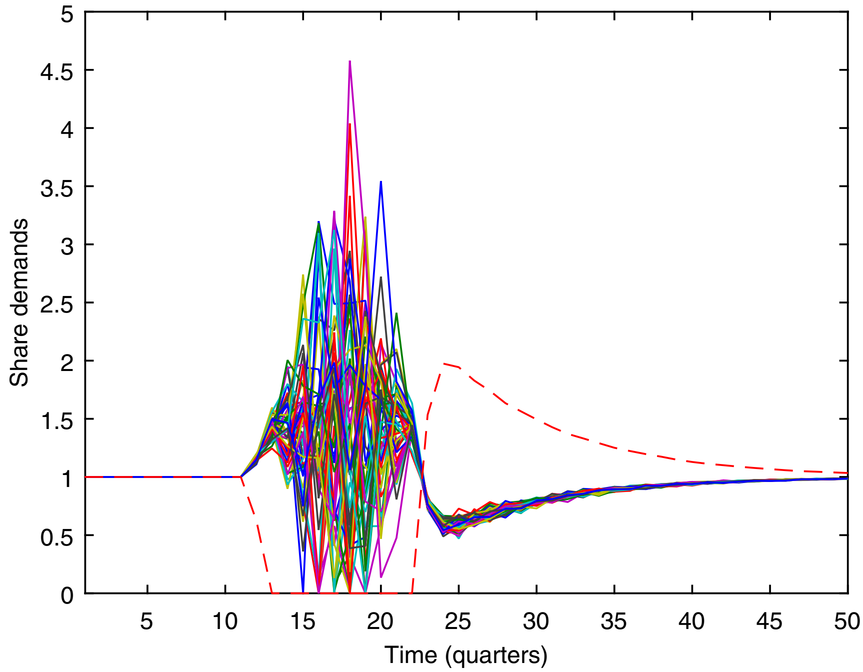


Fig. 4. Share demands in a bubble. The solid lines plot the risky asset share demands of extrapolators for the following sequence of 50 cash-flow shocks: 10 shocks of zero, followed by shocks of 2, 4, 6, 6, followed by 36 shocks of zero. The dashed line plots the share demand of the fundamental traders for the same cash-flow sequence. 30% of the investors are fundamental traders; the remainder are extrapolators with an extrapolation parameter θ of 0.9 and who also put a base weight $\overline{w}_i = 0.1$ on a value signal. The other parameters are $D_0 = 100$, $T = 50$, $\sigma_\varepsilon = 3$, $Q = 1$, $\gamma = 0.1$, $\sigma_u = 0.03$, and $I = 50$.

types of extrapolator (solid lines) for the same 50-period cash-flow sequence that we used in Fig. 1, namely, ten shocks of zero, followed by four positive shocks of 2, 4, 6, and 6, followed by 36 more shocks of zero. Recall from Fig. 1 that this sequence of cash-flow shocks generates a large bubble between dates 15 and 21.

Fig. 4 shows that, during the bubble, share demands of extrapolators become very volatile, suggesting a large increase in volume. Fig. 5 confirms this. The solid line in this figure plots total trading volume at each of the 50 dates, where volume at time t is defined as

$$0.5\left(\mu_0|N_t^F - N_{t-1}^F| + \sum_{i=1}^{I}\mu_i|N_t^{E,i} - N_{t-1}^{E,i}|\right). \tag{15}$$

The figure shows a dramatic increase in volume between dates 12 and 25. In particular, it shows that our model predicts three "peaks" in volume which correspond to the three bubble stages outlined in Section 3: a small peak centered around date 13 in the first stage, a much wider peak centered around date 17 in the second stage, and a thin but tall peak centered around date 23 in the third stage. Total volume at each date is the sum of two components: trading that takes place within the set of I extrapolators, and trading that takes place between the extrapolators in aggregate and fundamental traders. The dashed line in Fig. 5 plots the first component—trading volume within the set of I extrapolators.

value—its value does not depend on beliefs about right-tail outcomes for the underlying asset. Trading in the debt security is therefore muted.

The first peak in Fig. 5 centered around date 13 arises during the first stage of the bubble and reflects trading between the extrapolators in aggregate and fundamental traders. Arrival of the good cash-flow news pushes prices up, which, in turn, leads extrapolators to buy and fundamental traders to sell the asset. Before long, however, all the fundamental traders are out of the market and this first wave of trading subsides.

During the second stage, the bubble keeps growing and trading volume rises again, as indicated by the wide second peak centered around date 17 in Fig. 5. In this stage, all of the trading takes place among the I extrapolators. This potentially large volume generated by our model is due to wavering. It is not surprising that, in general, wavering leads to trading volume. What is more interesting is that, even though the degree of wavering remains *fixed* over time—the value of σ_u in Eq. (8) is constant—the model endogenously generates much greater volume during bubble periods than non-bubble periods.

To understand this, we write the share demand of extrapolator i in Eq. (10) more simply as $w_{i,t}V_t + (1 - w_{i,t})G_t$, where V_t and $G_t = X_t/\gamma\sigma_\varepsilon^2$ are the value and growth signals, respectively, at time t. We ignore the short-sale constraint because it is not important for the intuition. A unit change in the weight $w_{i,t}$ on the value signal changes the extrapolator's share demand by $V_t - G_t$. In "normal" times, when the cash-flow shocks are neither abnormally high nor abnormally low, the value and growth signals are both small in absolute magnitude: since the risky asset is neither particularly overvalued nor undervalued, the value signal V_t is close to zero in absolute magnitude; and since

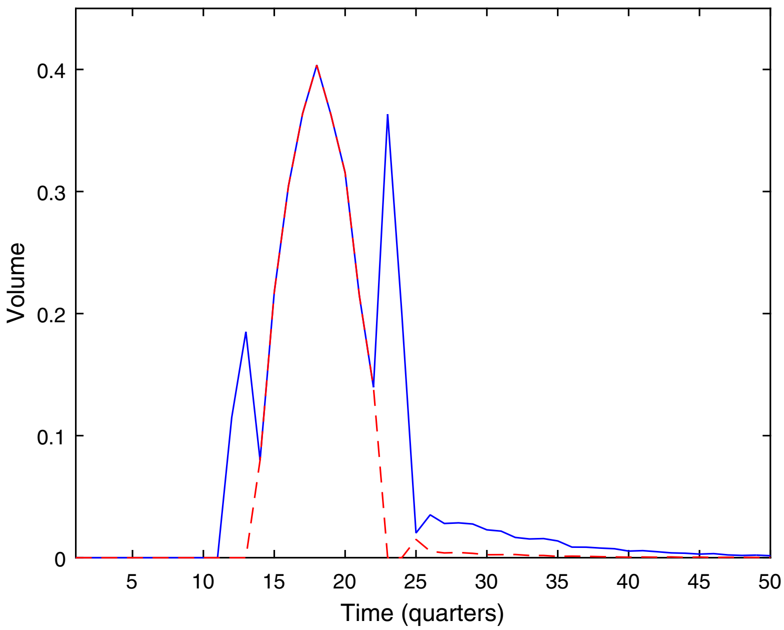


Fig. 5. Volume in a bubble. The solid line plots the total trading volume in the risky asset for the following sequence of 50 cash-flow shocks: 10 shocks of zero, followed by shocks of 2, 4, 6, 6, followed by 36 shocks of zero. The dashed line plots the trading volume between the extrapolators for the same cash-flow sequence. 30% of the investors are fundamental traders; the remainder are extrapolators with an extrapolation parameter θ of 0.9 and who also put a base weight $\overline{w}_i = 0.1$ on a value signal. The other parameters are $D_0 = 100$, $T = 50$, $\sigma_\varepsilon = 3$, $Q = 1$, $\gamma = 0.1$, $\sigma_u = 0.03$, and $I = 50$.

prices have not been rising or falling particularly sharply in recent periods, the growth signal G_t is also close to zero in absolute magnitude. In this case, $V_t - G_t$ is itself low in absolute magnitude, implying that, in normal times, wavering does not induce much variation in extrapolator demand.[15]

During a bubble, the situation is very different. At that time, the value signal V_t is large and *negative* (the asset is highly overvalued), and the growth signal is large and *positive* (the asset's price has been rising sharply in recent periods). As a result, $V_t - G_t$ is *very* large in absolute value, and the same degree of wavering that generates low trading volume in normal times now generates very high trading volume. This is the mechanism behind the high trading volume represented by the wide peak centered around date 17 in Fig. 5.

To put this more simply, during the bubble, the extrapolators holding the risky asset are subject to two powerful but conflicting investment signals. On the one hand, they see that prices are far above fundamental value; this makes them fearful of a crash and encourages them to sell. On the other hand, prices have recently been rising sharply, which makes extrapolators expect continued price appreciation and encourages them to buy. These two signals are so strong that even small shifts in their relative weight lead to large—and independent across traders—portfolio adjustments, and hence trading volume.

Once the bubble starts collapsing, the second wave of trading volume begins to subside: as the bubble deflates, both the value and growth signals decline in absolute magnitude; the quantity $V_t - G_t$ then also declines in absolute magnitude, and the impact of wavering on extrapolator share demands is reduced. Fig. 5 shows that once the bubble's collapse is well under way, there is a third wave of trading, represented by the thin third peak centered around date 23, between the selling extrapolators and the fundamental traders who re-enter the market. The third peak is taller than the first peak. The reason is that the first peak consists of extrapolators shifting from moderate holdings of the risky asset to large holdings of the asset, while the third peak consists of extrapolators shifting from large holdings of the risky asset to *low* holdings of the asset as they extrapolate price declines into the future and sell. This third volume peak thus represents more intense trading than the first one.

The central message in the discussion above is that a fixed amount of wavering can endogenously generate much higher trading volume during bubble periods. Proposition 3 below formalizes this idea in the following way. The change in extrapolator i's share demand between time t and time $t+1$ has two components. The first is unrelated to wavering; it is present even if $w_{i,t+1} = w_{i,t}$. Specifically, in the first stage of the bubble, the extrapolator buys from fundamental traders as the bubble grows, even in the absence of wavering; and as the bubble grows further in its second stage, he buys from less bullish extrapolators if he has a relatively low value of $\overline{w}_i$ or sells to more bullish extrapolators if his $\overline{w}_i$ is relatively high—again, even in the absence of wavering.

[15] If the growth signal G_t rises in value, this increases the aggregate demand for the risky asset. To counteract this increase and thus ensure that the market clears, the value signal V_t must decline. The two signals are therefore related: the more positive one of them is, the more negative the other must be.

The second component of the change in the extrapolator's share demand between time t and time $t+1$ is driven by wavering: it reflects his buying at time $t+1$ during the bubble if $w_{i,t+1}$ shifts down at that time, or his selling if $w_{i,t+1}$ shifts up. We sum the absolute value of this second component across all extrapolators and label the sum "wavering-induced trading volume," $V^W(X_t)$, a quantity that depends on X_t. Proposition 3 shows that $V^W(X_t)$ is typically *increasing* in X_t, a measure of bubble size. This is the formal sense in which wavering leads to more trading volume as the bubble grows.

Proposition 3. Suppose that there is a continuum of extrapolators and that each extrapolator draws an independent weight $w_{i,t}$ at time t from a bounded and continuous density function $g(w)$, $w \in [w_l, w_h]$, with mean $\overline{w}$ and with $0 < w_l < w_h < 1$. The sensitivity of per capita wavering-induced trading volume $V^W(X_t)$ to the growth signal X_t, denoted by $\partial V^W(X_t)/\partial X_t$, is given by

$$\frac{\partial V^W(X_t)}{\partial X_t} = \begin{cases} \dfrac{\text{sign}(X_t - \gamma\sigma_\varepsilon^2 Q)\overline{\Delta}_0}{(\mu_0 + (1-\mu_0)\overline{w})\gamma\sigma_\varepsilon^2} & \text{for } -\dfrac{w_l\gamma\sigma_\varepsilon^2 Q}{\mu_0(1-w_l) + (1-\mu_0)(\overline{w} - w_l)} \le X_t < \dfrac{\gamma\sigma_\varepsilon^2 Q}{(1-\mu_0)(1-\overline{w})} \\ \dfrac{\overline{\Delta}_0}{\overline{w}\gamma\sigma_\varepsilon^2} & \text{for } \dfrac{\gamma\sigma_\varepsilon^2 Q}{(1-\mu_0)(1-\overline{w})} \le X_t \le \dfrac{w_h\gamma\sigma_\varepsilon^2 Q}{(w_h - \overline{w})(1-\mu_0)} \end{cases}, \tag{16}$$

where

$$\overline{\Delta}_0 \equiv \int_{w_l}^{w_h}\int_{w_l}^{w_h} |w_1 - w_2| g(w_1) g(w_2) dw_1 dw_2. \tag{17}$$

If $X_t > w_h\gamma\sigma_\varepsilon^2 Q/((w_h - \overline{w})(1-\mu_0))$, $\partial V^W(X_t)/\partial X_t$ may become smaller and even turn negative as extrapolators exit the market.

The key part of Proposition 3 is the fourth row of (16). It says that, in the less extreme part of the second stage of the bubble, when all extrapolators are in the market, wavering induces more trading volume, the larger the size of the bubble: $\overline{\Delta}_0$ is a positive quantity. The same is true during the first stage of the bubble—see the first row of (16)—although the relationship is weaker; moreover, wavering-induced volume is here a relatively small part of overall trading volume. If, during its second stage, the bubble becomes so large that even some extrapolators exit the market, then wavering-induced volume increases more slowly as a function of X_t, and can even decrease, simply because there are fewer extrapolators available to trade.

The dashed lines in the four graphs in Fig. 2 show quantitatively how total trading volume at the peak of the bubble—the moment of maximum overvaluation—depends on four key model parameters: μ_0, $\overline{w}$, θ, and σ_u. Lower values of μ_0, $\overline{w}$, and θ lead to significantly higher volume. When these parameters have low values, the degree of overvaluation is higher. This means that the value and growth signals are larger in absolute magnitude, and this in turn interacts with wavering to generate more volume. As expected, higher values of σ_u lead to higher volume. Indeed, increasing σ_u leads to higher volume even if, as is the case for low values of σ_u, the degree of overvaluation remains unchanged.

Proposition 3 indicates that, during a bubble, wavering-induced volume is typically increasing in X_t. Since X_t is an average of past price changes, this suggests the following testable prediction: during a bubble, volume is strongly positively related to the asset's past return. To verify that this is a prediction of our model, we simulate a 40,000-period price sequence from the model and extract three subsamples—the subsample where the asset price differs from fundamental value by less than $\gamma\sigma_\varepsilon^2 Q = 0.9$; the subsample where the asset is overvalued by at least $\gamma\sigma_\varepsilon^2 Q = 0.9$; and the subsample where it is overvalued by at least $10\gamma\sigma_\varepsilon^2 Q = 9$.[16] We find that in these three subsamples, the correlation between volume at time $t+1$ and the price change between $t-4$ and t, a year-long interval, is −0.22, 0.41, and 0.6, respectively. These monotonically increasing correlations confirm that, in our model, the relationship between trading volume and past return is stronger during bubble episodes. In Section 7.1, we test this prediction for four historical bubbles.

We conclude our discussion of trading volume with two points. First, alternative sources of heterogeneity among extrapolators—sources other than wavering—do not generate nearly as much trading volume during the bubble period. Specifically, if we turn off wavering by setting σ_u in (8) to zero and instead allow the base weights $\overline{w}_i$ and the weighting parameter θ to differ across extrapolators, we no longer obtain a large second volume peak like the one in Fig. 5. The reason is that, after a sequence of price increases, extrapolators who do not exhibit wavering would almost all like to increase their holdings of the risky asset, even if they differ in their values of $\overline{w}_i$ and θ: regardless of the specific values of $\overline{w}_i$ and θ, a high growth signal means that most extrapolators find the risky asset more attractive. Since most extrapolators want to trade in the same direction, there is relatively little trading between them: prices, not quantities, adjust. In our model, the specific type of heterogeneity induced by wavering is uniquely able to generate heavy trading.[17]

[16] The quantity $\gamma\sigma_\varepsilon^2 Q$ is the degree of overvaluation that causes fundamental traders to exit the market. It is therefore a natural "unit" of overvaluation.

[17] In many bubble episodes, including two that we study in Section 7, the peak in volume coincides with the peak in prices. This is the pattern predicted by our model so long as the degree of wavering, governed by σ_u, is not too low. In the technology bubble of the late 1990s, the peak in volume precedes the peak in prices (Hong and Stein, 2007). While we do not have a full theory of this phenomenon, even here, wavering is helpful for understanding the evidence. Although volume at the price peak is lower than it was a few months earlier, it is still extremely high. In our model of extrapolators and fundamental traders, only wavering can generate such intense trading.

Second, our predictions about prices and volume depend primarily on the presence of extrapolators who waver between two signals; the short-sale constraint is not nearly as important. We incorporate the short-sale constraint into our model because the fundamental traders are boundedly rational. These traders do not attempt to forecast extrapolator demand, but instead simply assume that any mispricing will correct by the next date. As a consequence, they trade aggressively against the extrapolators, and this reduces the mispricing. To generate a substantial overvaluation, we need a short-sale constraint: this forces the fundamental traders out as the bubble starts to form, allowing the bubble to grow.

In an Online Appendix, we show that, in a model with fully rational traders, we can dispense with the short-sale constraint without substantially affecting our predictions. Specifically, we analyze a model in which extrapolators have the demand function in (7) and (8) and where we replace the fundamental traders with fully rational traders who are able to forecast future extrapolator demand. Moreover, all investors can short. This model makes predictions about bubbles that are similar to those of the model in Section 2. First, a sequence of strongly positive cash-flow news leads to a large overvaluation. Since the rational traders can forecast extrapolator demand, they recognize that, when extrapolators are bullish, they are likely to remain bullish for a while, and that it therefore does not pay to trade aggressively against them. This, in turn, allows a large overvaluation to build up. Second, there is again high trading volume during the bubble. Since all investors can short, some of the trading volume, even at the height of the bubble, is between extrapolators and rational traders—but a lot of it is between wavering extrapolators.[18]

In the model in the Online Appendix, the rational traders maximize the expected utility of next period's wealth. We expect that the results will be similar if these traders instead maximize the expected utility of time T wealth, so long as they adjust their portfolios in a dynamically optimal way at each date. Indeed, Barberis et al. (2015) show that, when rational dynamically optimizing traders with long horizons interact with extrapolators, there can be a substantial overvaluation even without a short-sale constraint. The reason is that, when the risky asset is overvalued, rational traders recognize that extrapolators are likely to remain optimistic for a while longer, and that they can therefore improve their long-run wealth prospects by not attacking the mispricing too aggressively.

We focus on the model of Section 2 with boundedly rational traders and a short-sale constraint because it captures our main ideas in a simpler way than do models with fully rational traders. Another advantage of the Section 2 model is that it makes more realistic predictions about downturns. We discuss these predictions in the next section.

5. Negative bubbles

The behavior of prices and volume after a sequence of negative cash-flow shocks is *not* the "mirror image" of those for the case of positive cash-flow shocks. First, our model does not generate "negative" bubbles: while the price of the risky asset falls when bad cash-flow news arrives, it does not fall much below fundamental value. After disappointing cash-flow news pushes the price of the risky asset down, extrapolators reduce their holdings of the risky asset. However, since the fundamental traders believe that any mispricing will correct by the next period, they trade aggressively against any underpricing they perceive. As a result, there is no significant undervaluation. Still, there is some. After several periods of bad cash-flow news, the risky asset is held only by fundamental traders: the short-sale constraint forces the extrapolators out of the market. For the fundamental traders to be willing to hold the entire supply of the risky asset, the price has to be lower than the fundamental value in (4), and in particular equals

$$D_t - \gamma\sigma_\varepsilon^2(T-t-1)Q - \frac{\gamma\sigma_\varepsilon^2}{\mu_0}Q,$$

which differs from fundamental value by $\gamma\sigma_\varepsilon^2 Q(1-\mu_0)/\mu_0$. For our parameter values, this wedge is approximately two dollars.[19]

Our model predicts heavy trading during bubbles, but little trading during severe downturns. When bad cash-flow news arrives, there is some trading as extrapolators sell to fundamental traders. Once the extrapolators leave the market, however, the asset is held only by fundamental traders, a homogeneous group. There is no more trading until the market recovers and extrapolators re-enter. More broadly, our model predicts higher trading volume during bull than bear markets, a prediction consistent with the available evidence (Statman et al., 2006; Griffin et al., 2007).[20]

6. Comparison with other bubble models

It is impossible to do justice here to all the important contributions in the literature on bubbles, recently surveyed by Brunnermeier and Oehmke (2013) and Xiong (2013). Instead, we focus on two classes of models—rational bubble models and disagreement-based models. The former are notable for their simplicity and long tradition; the latter, like us, deal with volume.

[18] We have also studied a model where we *keep* the boundedly rational traders, but add a third group of fully rational traders to the economy. This model preserves the same essential predictions as the model in Section 2: the asset becomes overvalued after a sequence of good cash-flow news and a large percentage of volume during the bubble is due to trading among extrapolators. But consistent with De Long et al. (1990), Brunnermeier and Nagel (2004), and others, it also captures the idea that rational traders ride the bubble: they buy the asset after good cash-flow news in the expectation of selling to extrapolators at a later date.

[19] For comparison, recall that a one-standard deviation cash-flow shock moves the risky asset price by approximately three dollars.

[20] This prediction holds even if fundamental traders waver—for example, even if they have the demand function in (7) and (8) with a base weight $\overline{w}_0 = 0.9$. Since the asset does not become very undervalued in a downturn, the growth and value signals remain low in absolute magnitude. As a consequence, wavering induces little trading volume.

6.1. Rational bubble models

In models of rational bubbles, the price of a risky asset is given by

$$P_t = P_{D,t} + B_t, \tag{18}$$

where $P_{D,t}$ is the present value of the asset's future cash flows and where B_t, the bubble component, satisfies

$$B_t = \frac{\mathbb{E}(B_{t+1})}{1+r}, \tag{19}$$

where r is the expected return. We note four points.

First, the rational bubble model does not explain how a bubble gets started in the first place. Under limited liability, the value of B must always be non-negative. Eq. (19) then implies that, if B is strictly positive in any future state of the world, it must be positive at the current time. Put simply, if a bubble exists, it must always have existed. In our framework, in contrast, bubbles are initiated in a much clearer way, as a consequence of what Kindleberger (1978) calls "displacement": a sequence of good cash-flow news leads to price increases which, in turn, cause extrapolators to raise their demand for the risky asset.

Second, the rational bubble model has nothing to say about trading volume. In its usual form, agents are assumed to be homogeneous; trading volume is therefore zero.

Third, the rational bubble model does not capture the extrapolative expectations that are often observed during bubbles. In the basic version of this framework, the return investors expect to earn on the risky asset is constant over time.

Finally, direct tests of the key prediction of rational bubble models—that payoffs in the infinite future have positive present value—reject it (Giglio et al., 2016).

6.2. Disagreement-based models

Building on Harrison and Kreps (1978), Scheinkman and Xiong (2003) present a model in which two risk-neutral investors observe two signals about the fundamental value of a risky asset, but disagree about how useful each signal is. Their disagreement leads to trading volume. With a short-sale constraint, disagreement also leads to overpricing: the price of the risky asset can be higher than the present value of its future cash flows, as perceived by the investor holding the asset. The reason is that the holder of the asset recognizes that, as more signals and cash-flow news are revealed over time, the other investor may become more optimistic than he is, allowing him to sell the asset on at an attractive price.

Both in Scheinkman and Xiong (2003) and in our model, the increase in volume during a bubble is due to an increase in disagreement among investors. In Scheinkman and Xiong (2003), this increase in disagreement is exogenous. In our model, disagreement grows *endogenously* over the course of the bubble. As the bubble increases in size, the growth and value signals in Eq. (10) become very large in absolute magnitude. Extrapolators who, as a consequence of wavering, differ even very slightly in the relative weight they put on the two signals disagree sharply about the expected price change on the risky asset and therefore trade in large quantities. Whereas in Scheinkman and Xiong (2003) an exogeneous increase in disagreement leads to both higher volume and overpricing, in our model, the causation is different: overpricing leads to endogenously higher disagreement and hence higher volume.

Our model differs from disagreement models in other important ways. In our model, many investors hold expectations that depend positively on past returns, consistent with survey evidence on the expectations of actual investors. In Scheinkman and Xiong (2003), however, the holder of the asset has *constant* expectations about the asset's future return. Our framework also predicts a strongly positive correlation between volume and past returns during bubble episodes, a prediction that we confirm empirically in the next section. Using simulations, we find that, in Scheinkman and Xiong (2003), this correlation is close to zero: the exogeneous process that governs disagreement and hence volume is uncorrelated with the process for fundamentals that is the main determinant of price movements.

7. Some evidence

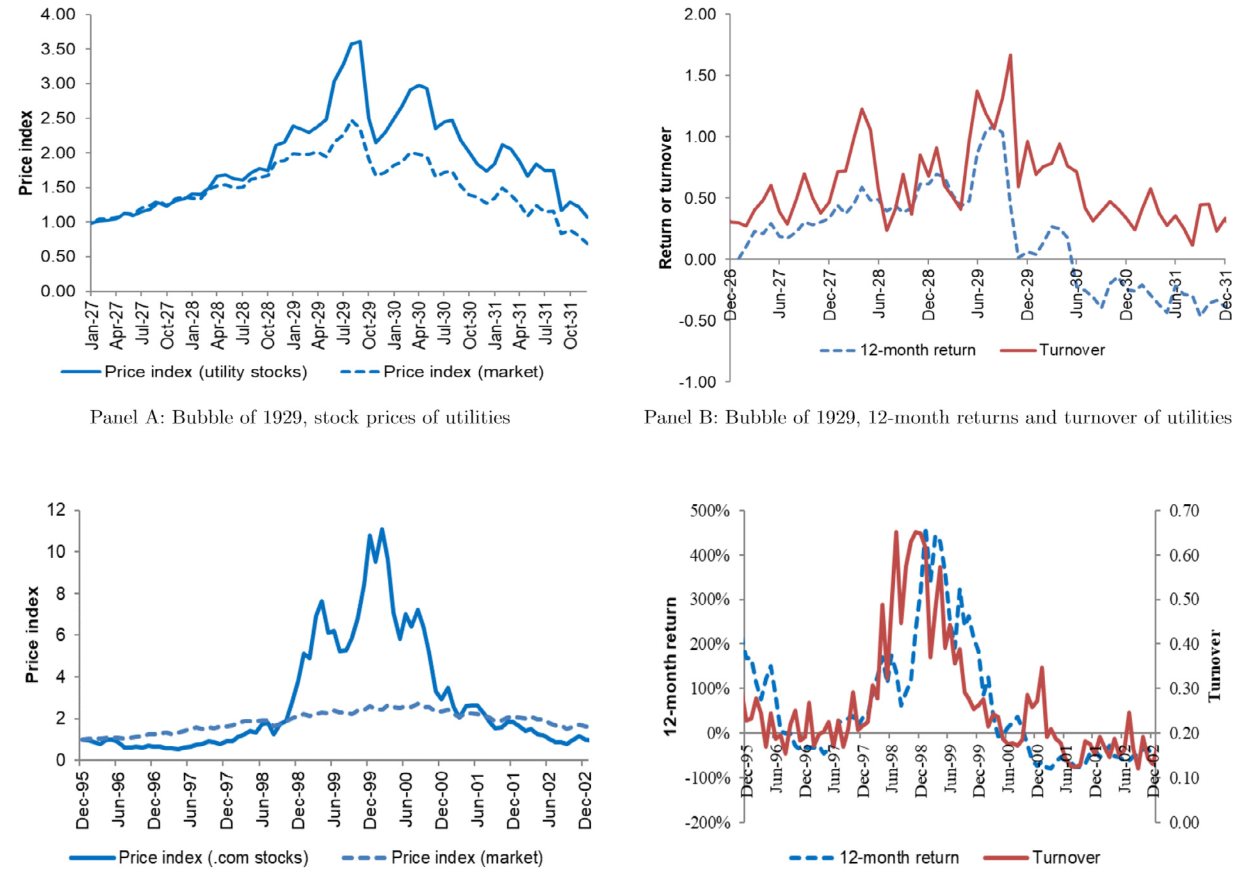
We present empirical evidence bearing on some distinctive predictions of the model. One prediction, outlined in Section 4, is that the correlation between the trading volume in an asset and its return over the previous year is higher during a bubble episode than at other times. In Section 7.1, we examine this prediction for four historical bubbles. In Section 7.2, we evaluate another central prediction of the model: as a bubble develops, a growing fraction of trading volume is due to investors with extrapolator-like characteristics. In Section 7.3, we look to see if these investors also exhibit evidence of wavering.

7.1. Volume and past returns

For four bubble episodes—the U.S. stock market in 1929, technology stocks in 1998–2000, U.S. housing in 2004–2006, and commodities in 2007–2008—we check whether, as predicted by our model, the correlation between volume and past return for the asset in question is higher during the bubble period than during the two-year period that follows the bubble's collapse.

Accounts of the stock market boom of the late 1920s suggest that the bubble began in March 1928 (Allen, 1931; Galbraith, 1954; White, 1990). White (1990) shows that new industries, especially utilities, led the stock market boom. Panel A of Fig. 6 confirms White's account. It compares the value-weighted cumulative return of public utilities listed in the Center for Research in Security Prices (CRSP) data (Standard Industrial Classification (SIC) codes 4900–4990) with the cumulative return of the broader stock market. Utilities outperformed the broader stock market by more than 80% in the March 1928–September 1929 period.[21]

[21] Wigmore (1985, p. 42), describes the market for these stocks: "There is no gainsaying the enthusiasm of the financial markets for these pub-



Panel A: Bubble of 1929, stock prices of utilities

Panel B: Bubble of 1929, 12-month returns and turnover of utilities

Panel C: Technology stock bubble, stock prices

Panel D: Technology stock bubble, 12-month returns and turnover

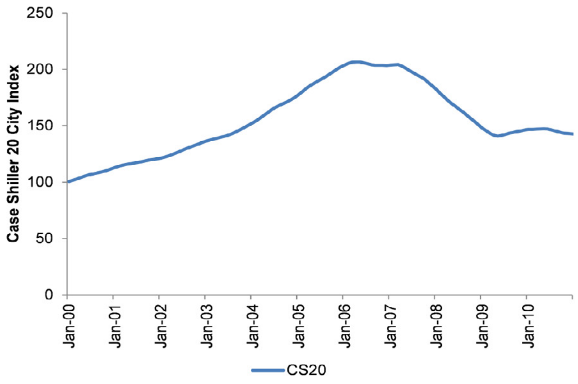
Fig. 6. Prices, past 12-month returns, and value-weighted turnover during four bubble episodes: utility stocks in 1929; technology stocks in 1998–2000; house prices in 2004–2006 as measured by the Case–Shiller 20-City Index; and oil in 2007–2008 as proxied by the price of the USO ETF. All data are monthly and value-weighted across stocks. For housing, turnover is measured as the number of existing-home sales.

If we accept that utility stocks experienced a bubble in 1928–1929, our model predicts that trading volume in these stocks during this time will be positively related to their past return. Panel B of Fig. 6 plots the value-weighted monthly turnover of utility stocks over this period alongside their value-weighted 12-month past return; turnover is defined as volume divided by shares outstanding. After a spike in April 1928, the turnover of utility stocks closely tracks their 12-month past return. For example, the second highest volume month in the series occurs in June 1929, following a 12-month cumulative return of 86%. From January 1927 to December 1930, the correlation between turnover and the 12-month past return is 0.59. Over the two-year period after the bubble ends—from January 1931 to December 1932—the correlation is −0.03.

The explosion of volume during the technology stock bubble of 1998–2000 is well-known (Ofek and Richardson, 2003; Hong and Stein, 2007). In Panels C and D of Fig. 6, we replicate and extend these findings. Panel C plots value-weighted monthly cumulative returns for the sample of .com stocks used by Ofek and Richardson (2003) and compares them to the cumulative returns of the CRSP value-weighted stock market index; returns for .com stocks are from CRSP. Panel D shows that turnover, measured as before and value-weighted, increases steadily as the bubble progresses, peaking in April 1999, the same point at which technology stocks reach their highest 12-month return of 429%. Overall, the figure shows that turnover closely tracks the past 12-month return, with a time-series correlation between the two of 0.73 between January 1998 and December 2002. In the 24-month post-bubble period from January 2003 to December 2004, the correlation is −0.14.

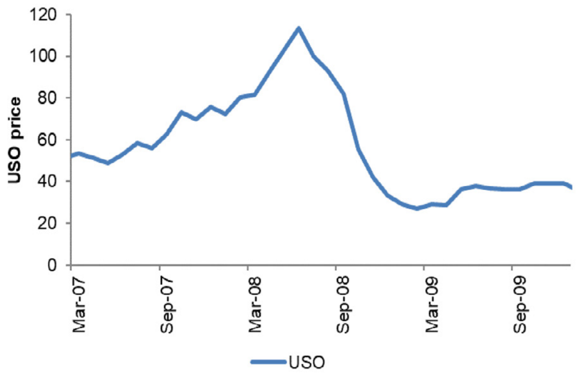
The relationship between turnover and past returns also appears during the U.S. housing bubble of the mid-2000s. In Panel E of Fig. 6, we plot the Case–Shiller 20-City Com-

lic utility holding companies, however. […] Their trading volume in 1929 exceeded 100% of their outstanding shares. At the high point in the market, their stocks averaged prices 57 times earnings per share, with Electric Bond and Share, which was most prominent because of its size and its relationship with General Electric, selling at 96 times earnings per share." Following the collapse of the bubble, utilities were "held up by Franklin Roosevelt and others as spectacles of financial abuse and confusion."

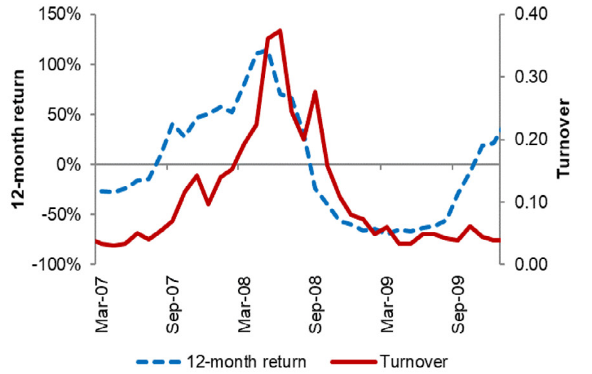


Panel E: Housing bubble of 2004–2006, price index

Panel F: Housing bubble of 2004–2006, 12-month returns and volume



Panel G: 2007–2008 Commodity bubble, USO price



Panel H: 2007–2008 Commodity bubble, 12-month returns and turnover of USO

Fig. 6. Continued

posite Home Price Index. This index, based on repeat transactions, seeks to measure the value of residential real estate in the 20 largest U.S. metropolitan areas.

The Case–Shiller Index rises from a base value of 100 in January 2000 to a peak of 206.61 in April 2006. In Panel F, we show the relationship between 12-month past returns and volume for the U.S. housing market; we use existing-home sales by month as a measure of volume.[22] The figure shows that, as for the two stock market bubbles, volume closely tracks the 12-month past return; their time-series correlation in monthly data from January 2003 to December 2008 is 0.96. This is higher than the correlation in the two-year post-bubble period from January 2009 to December 2010, namely, 0.2.

Whether the run-up in commodity prices in 2007 and 2008 can be easily explained by fundamentals or was instead a bubble is subject to debate, with some authors suggesting that the "financialization" of derivatives markets instigated demand from institutional investors (Irwin and Sanders, 2010; Cheng and Xiong, 2014; Hong et al., 2015). Panel G of Fig. 6 shows the run-up in oil prices as reflected in the share price of United States Oil (USO) Fund, the largest exchange-traded fund (ETF) with exposure to oil. USO more than doubled between December 2006 and June 2008. In Panel H of Fig. 6, we plot the monthly turnover and 12-month past return of this ETF, both obtained from CRSP. As in our other examples, the turnover of USO closely tracks the past return; their time-series correlation between April 2007 and December 2009 is 0.83. During the two-year post-bubble period, the correlation is 0.15.

7.2. *The source of trading volume in a bubble*

Another prediction of our model is that, during a bubble, a larger fraction of trading volume will be due to extrapolative investors. We test this prediction for the technology stock bubble of the late 1990s.

We do not have data on the trading of *all* investors during this period, but we do have data on the trading of mutual funds. We therefore test whether, as the bubble develops, a larger fraction of trading volume is due

[22] Existing-home sales are based on closing transactions of single-family homes, townhomes, condominiums, and cooperative homes and are provided by the National Association of Realtors.

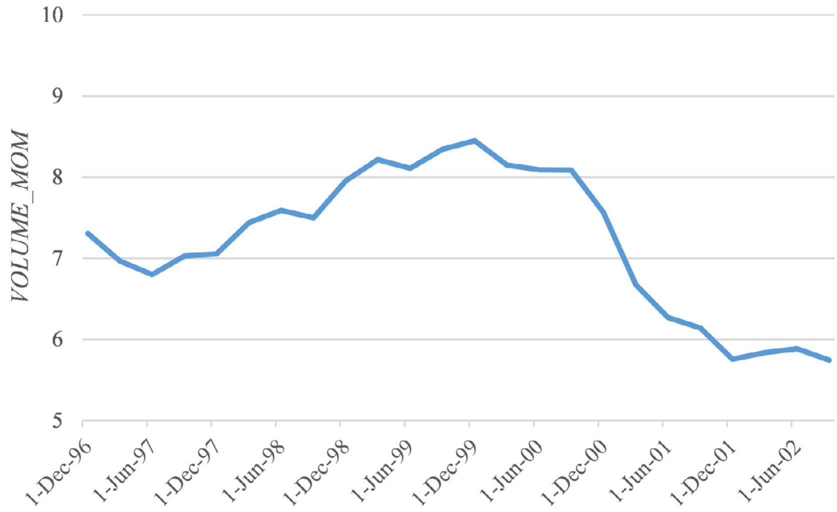


Fig. 7. Composition of volume during a bubble. In each quarter t from 1997 to 2002, and for each technology stock i, we compute $\text{Volume\_Mom}_{i,t}$, a measure of the extent to which trading in stock i in quarter t is due to mutual funds with extrapolator-like characteristics. To determine whether a fund is extrapolator-like, we look at its holdings six months earlier, in quarter $t-2$, and check whether these holdings earned high returns over the previous year. For each quarter, the figure plots the equal-weighted average of $\text{Volume\_Mom}_{i,t}$ across all technology stocks in that quarter.

to extrapolator-like mutual funds. We again identify .com stocks using the list of securities provided by Ofek and Richardson (2003) and then match this list to quarterly mutual fund holdings from the Thomson Reuters database. For technology stock i in quarter t, we compute a measure of extrapolator-weighted trading volume, namely,

$$\text{Volume\_Mom}_{i,t} = \frac{\sum_j \text{Buys}_{i,j,t}\text{Fundmom}_{j,t-2} + \sum_j \text{Sells}_{i,j,t}\text{Fundmom}_{j,t-2}}{\sum_j \text{Buys}_{i,j,t} + \sum_j \text{Sells}_{i,j,t}}, \tag{20}$$

where j indexes the mutual funds trading stock i in quarter t, so that $\text{Buys}_{i,j,t}$ and $\text{Sells}_{i,j,t}$ are the dollar buys and sells, respectively, of stock i by fund j in quarter t. In expression (20), the Buys and Sells in the numerator are weighted by the extent to which they were executed by extrapolator-like mutual funds. To determine how extrapolator-like fund j is, we look at its holdings six months prior to quarter t—in other words, in quarter $t-2$—and compute the “growthiness” of these holdings. Specifically, at time $t-2$, we sort all stocks—not only .com stocks—into NYSE deciles based on their past 12-month return. The growthiness of the fund’s portfolio at time $t-2$, $\text{Fundmom}_{j,t-2}$, is then measured as the position-weighted past-return decile of the stocks in the portfolio:

$$\text{Fundmom}_{j,t-2} = \sum_i w_{i,j,t-2} \ \text{stockmom}_{i,t-2}, \tag{21}$$

where stockmom takes an integer value between 1 and 10. A portfolio with a Fundmom score of 1, for example, contains only stocks that performed poorly over the trailing 12 months, while a portfolio with a score of 10 contains only stocks with high past returns.

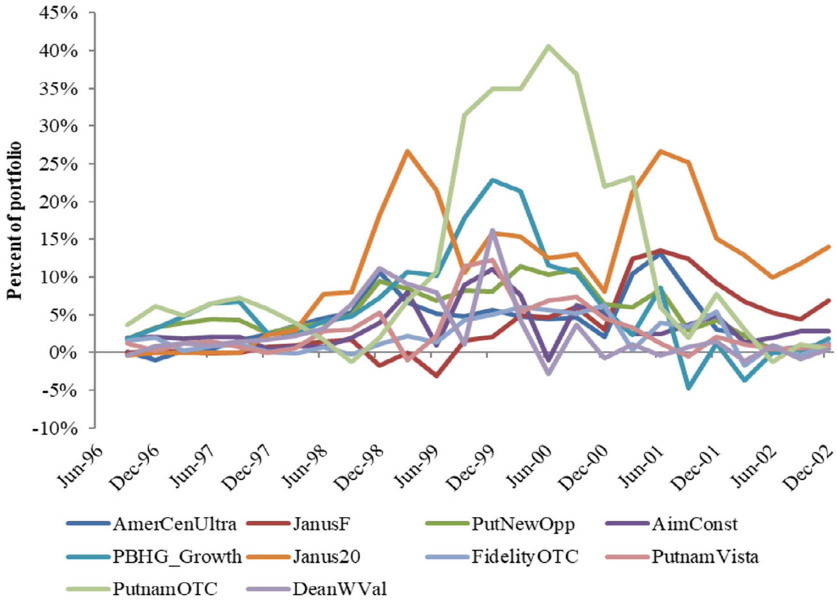
In summary, $\text{Volume\_Mom}_{i,t}$ measures the extent to which a dollar of trading volume in stock i in quarter t is associated with extrapolative investors. When it takes a high value, the stock is being bought and sold primarily by extrapolators; when it takes a low value, contrarian investors play a larger role in the trading.

For each quarter t, we compute an equal-weighted average of $\text{Volume\_Mom}_{i,t}$ across all technology stocks at that date. Fig. 7 plots the resulting time series between 1997 and 2002. The figure shows that, even at the beginning of the bubble in 1997 and early 1998, there is substantial trading by extrapolative investors. As prices rise in 1998 and 1999, volume becomes increasingly dominated by extrapolators. Then, as the bubble collapses in 2000 and 2001, the average Volume_Mom drops substantially, from a peak of 8.45 in December 1999 to 5.75 in December 2001. Overall, the figure provides support for our prediction that, as the bubble grows, a larger fraction of volume is due to trading by investors with extrapolator-like characteristics.

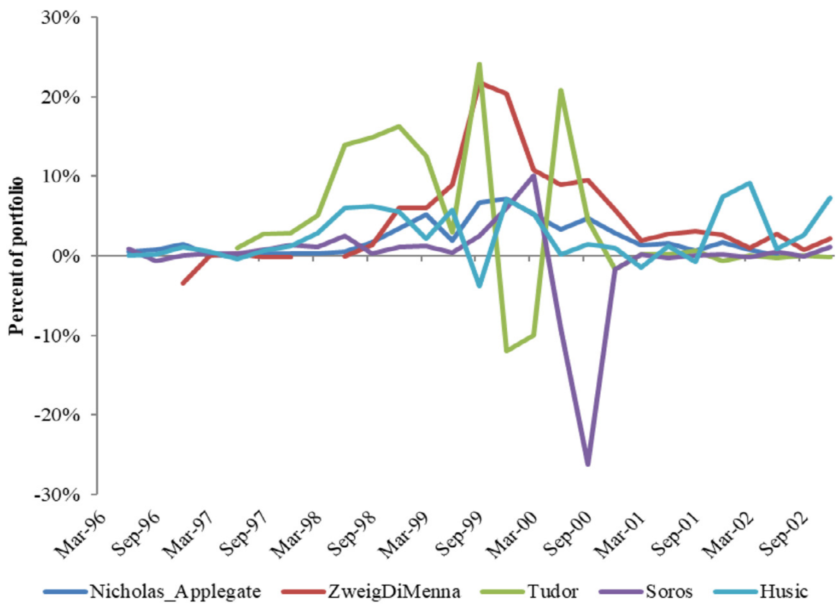
7.3. Evidence of wavering

In the previous section, we saw that, as the technology stock bubble grew, a larger fraction of trading volume was due to extrapolative mutual funds. We now show that these funds also exhibited signs of wavering.

To provide evidence of wavering, we look at *fund*-level holdings: if, as our model assumes, wavering is independent across funds, it will be undetectable in aggregate mutual fund holdings. For each mutual fund in our sample, we compute its maximum dollar exposure to the Internet sector between 1996 and 2000, and focus on the ten funds with the highest maximum exposure. For each of these ten funds, we measure the *active* changes in its exposure to the Internet sector during the bubble period. Specifically, for each of the fund’s positions in each quarter, we compute the change in the value of the position due to trading as a percentage of the total value of the fund’s long posi-



Panel A: Mutual fund net purchases



Panel B: Hedge fund net purchases

Fig. 8. Wavering by mutual funds and hedge funds. The figure shows the active change in weight in Internet stocks, defined as the net purchases of Internet stocks in any quarter as a percentage of the total portfolio value. Panel A shows net purchases for the ten largest mutual fund holders of Internet stocks during the bubble. Panel B shows net purchases for the five largest hedge fund holders. Purchases and sales are based on split-adjusted holdings from the Thomson Reuters mutual fund and institutional ownership database.

tions that quarter; we denote this by Δw. For a newly established position, Δw is simply the value of the position scaled by the total value of all long positions. For positions already held at the end of quarter $t-1$, Δw at time t is the change in shares held at time t multiplied by the time t price, again scaled by the total value of all long positions. In each quarter, we sum Δw up over all Internet stocks held by the fund, with positive values indicating an active increase in exposure to these stocks and negative values an active decrease in exposure. As before, we identify Internet stocks using the Ofek and Richardson (2003) list.

The top chart in Fig. 8 shows the time series of Δw for the ten funds. Several of the funds exhibit trading behavior that is consistent with wavering: as the bubble forms, they display substantial shifts in their enthusiasm for the Internet sector. Moreover, and also consistent with our model, the wavering appears largely uncorrelated across funds: their enthusiasm waxes and wanes at different times.

We conduct an analogous exercise for hedge funds. Specifically, we take the five hedge funds in the Brunnermeier and Nagel (2004) sample with the highest maximum dollar exposure to the Internet sector between 1996 and 2000. For each of the five funds, we compute the time series of Δw. The lower chart in Fig. 8 presents the results. Zweig DiMenna's trading behavior is best described as extrapolative with little wavering. However, Tudor exhibits behavior that is more consistent with wavering: the fund moves in, out, and back into bubble stocks as the bubble grows. Nicholas Applegate, Husic, and Soros' trading is similar to Tudor's: their enthusiasm for the Internet sector also fluctuates substantially over time, albeit less dramatically. While hedge fund behavior varies across funds, the standard deviation of each fund's time series of Δw during the January 1998–March 2000 bubble period, a rough measure of wavering-induced trading, greatly exceeds the standard deviation of this quantity both before and after the bubble period.[23]

Other models may also be able to explain investors' alternately increasing and decreasing their exposure to an overvalued asset. Abreu and Brunnermeier (2003) present a model in which one hedge fund after another becomes aware that an asset is overpriced. Each fund is unsure how many other funds know about the overpricing, and this creates a dilemma: if relatively few other funds know about the mispricing, then it is better for the fund to ride the bubble because there is unlikely to be enough selling pressure to burst it. However, if many other funds are aware of the bubble, it is better for the fund to sell in order to avoid a crash caused by other funds exiting.

In the basic version of the Abreu and Brunnermeier (2003) model, there is nothing that resembles wavering: each fund sells the asset some time after learning that it is overpriced, and then stays out of the market. However, in an extension of the model that incorporates uninformative synchronizing events, something reminiscent of wavering emerges: funds sell when a synchronizing event occurs, and, if the selling pressure fails to burst the bubble, re-enter the market until the next such event, and so on.

The Abreu and Brunnermeier (2003) predictions for investor behavior differ from ours in at least two important ways. First, the Abreu and Brunnermeier model predicts that the high volume we observe during bubble periods will manifest itself in the form of occasional large spikes in volume. In our model, however, volume is high during bubbles on a more "continuous" basis. To our eyes, the evidence on volume during bubble periods is more consistent with the latter view.

Second, in Abreu and Brunnermeier (2003), hedge fund trading is highly correlated: funds exit and re-enter the market at the same time, namely, on the dates of the synchronizing events. In our model, however, wavering is assumed to be uncorrelated across investors: at each date during the bubble period, some extrapolators increase their exposure to the risky asset while others decrease it. This assumption is important: if extrapolators' trades are too highly correlated, our model will not generate high volume at the peak of the bubble.

To distinguish the two models empirically, we compute the average pairwise correlation of the times series of Δw—the quarterly change in a fund's exposure to the Internet sector—for the five hedge funds represented in the lower panel of Fig. 8. We find that this average is a modest 0.19. In other words, consistent with our framework, but less consistent with the Abreu and Brunnermeier (2003) model, changes in hedge fund exposure to the Internet sector are relatively uncorrelated.

8. Conclusion

Although historical accounts of price bubbles typically emphasize extrapolative expectations (Kindleberger, 1978; Shiller, 2000), recent models of bubbles have moved away from this feature. In this paper, we embrace it. In our model, some investors hold extrapolative expectations, but also waver in their convictions in that they worry more or less about the possible overvaluation of the asset. The model generates occasional bubbles in asset prices. Such bubbles occur in response to particular patterns of good news, a phenomenon Kindleberger (1978) called displacement. They are characterized by very high trading volume documented in earlier literature, which to a significant extent comes from trading between the wavering extrapolators. The model generates a new prediction that trading volume is driven by high past returns which distinguishes it from some popular recent models and appears to be consistent with some historical evidence.

Our analysis has left several important issues to future work. First, we have not addressed the controversy of whether bubbles actually exist, and whether investors can tell in the middle of a rapid price increase of an asset that it is actually overpriced. Second, our model assumes a simple and stabilizing form of arbitrage, and thus does not consider the possibility of destabilizing arbitrage, whereby rational investors buy an overpriced asset in the hope of selling at an even higher price to extrapolators

[23] Mutual funds and hedge funds as a group do not correspond neatly to either the extrapolators or the fundamental traders in our model: some of these funds follow extrapolative strategies, while others behave more like fundamental traders. Given their large average exposure to the Internet sector in the late 1990s, the 15 mutual and hedge funds we study in Fig. 8 resemble extrapolators more than they do fundamental traders.

(De Long et al., 1990; Brunnermeier and Nagel, 2004). But we also have not considered other stabilizing forces, such as arbitrage by security issuers themselves through greater issuance or asset creation (Galbraith, 1954). Perhaps most important, we have adopted a standard but ad hoc formulation of extrapolative beliefs by some investors. The fundamental psychological mechanisms of extrapolation remain to be understood.

Appendix A. Proofs

A.1. A micro-foundation for fundamental trader demand in Eq. (3)

Consider an economy with the timing and asset structure described at the start of Section 2. There are two types of traders: one type, which makes up a fraction μ^X of the population, has time t per capita demand N_t^X for shares of the risky asset; the other type, which makes up a fraction $\mu^F = 1 - \mu^X$ of the population, is a fundamental trader who, at time t, chooses his per capita share demand N_t^F by maximizing a utility function with constant absolute risk aversion γ and defined over next period's wealth. Specifically, his objective is

$$\max_{N_t^F} \mathbb{E}_t^F[-e^{-\gamma(W_t + N_t^F(\widetilde{P}_{t+1} - P_t))}]. \tag{A.1}$$

This trader is boundedly rational, in a way that we make precise in what follows.

To determine his time t demand for the risky asset, the fundamental trader reasons as follows. At the final date, date T, the price of the risky asset P_T must equal the cash flow realized on that date, so that $P_T = D_T$. At time $T-1$, the fundamental trader's first-order condition implies that his share demand is

$$N_{T-1}^F = \frac{\mathbb{E}_{T-1}^F(\widetilde{P}_T) - P_{T-1}}{\gamma \mathrm{Var}_{T-1}^F(\widetilde{P}_T - P_{T-1})} = \frac{D_{T-1} - P_{T-1}}{\gamma \sigma_\varepsilon^2}, \tag{A.2}$$

where we have used the fact that $\mathbb{E}_{T-1}^F(\widetilde{P}_T) = D_{T-1}$, and have assumed that the fundamental trader takes the conditional distribution of the change in price to be Normal—an assumption that is confirmed in equilibrium—and that, for simplicity, he sets the conditional variance of price changes equal to the variance of cash-flow shocks. Market clearing implies

$$\mu^F\left(\frac{D_{T-1} - P_{T-1}}{\gamma\sigma_\varepsilon^2}\right) + \mu^X N_{T-1}^X = Q, \tag{A.3}$$

which, in turn, implies

$$P_{T-1} = D_{T-1} - \frac{\gamma\sigma_\varepsilon^2}{\mu^F}(Q - \mu^X N_{T-1}^X). \tag{A.4}$$

At time $T-2$, the fundamental trader's demand is

$$N_{T-2}^F = \frac{\mathbb{E}_{T-2}^F(\widetilde{P}_{T-1}) - P_{T-2}}{\gamma\sigma_\varepsilon^2}. \tag{A.5}$$

It is here that his bounded rationality comes into play. To compute $\mathbb{E}_{T-2}^F(\widetilde{P}_{T-1})$, in other words, to compute the expectation of the quantity in (A.4), he needs a value for $\mathbb{E}_{T-2}^F(N_{T-1}^X)$. We assume that the fundamental trader does not try to forecast the evolution of the other traders' demand N^X, but instead sets $\mathbb{E}_{T-2}^F(N_{T-1}^X) = Q$; in other words, he assumes that the other traders will simply hold an amount of the risky asset that corresponds to their weight in the population. Under this assumption, $\mathbb{E}_{T-2}^F(\widetilde{P}_{T-1}) = D_{T-2} - \gamma\sigma_\varepsilon^2 Q$, so that

$$N_{T-2}^F = \frac{D_{T-2} - \gamma\sigma_\varepsilon^2 Q - P_{T-2}}{\gamma\sigma_\varepsilon^2}. \tag{A.6}$$

We assume that the fundamental trader continues to reason in this way, working back from time T to the current time t, and, at each time, forecasting that the other traders' per capita demand at the next date will simply equal Q. Under these assumptions,

$$N_t^F = \frac{D_t - \gamma\sigma_\varepsilon^2(T-t-1)Q - P_t}{\gamma\sigma_\varepsilon^2}, \tag{A.7}$$

which is Eq. (3).

A.2. A micro-foundation for extrapolator demand in Eq. (5)

Consider an economy with the timing and asset structure described at the start of Section 2. Now consider a trader who, at time t, maximizes a utility function with constant absolute risk aversion γ and defined over next period's wealth. Specifically, his objective is

$$\max_{N_t^X} \mathbb{E}_t^X[-e^{-\gamma(W_t + N_t^X(\widetilde{P}_{t+1} - P_t))}]. \tag{A.8}$$

From the first-order condition, optimal demand is

$$N_t^X = \frac{\mathbb{E}_t^X(\widetilde{P}_{t+1}) - P_t}{\gamma \mathrm{Var}_t^X(\widetilde{P}_{t+1} - P_t)}, \tag{A.9}$$

where, for simplicity, the investor takes the conditional distribution of the change in price to be Normal. Suppose that this investor forms beliefs about future price changes by extrapolating past price changes, so that

$$\mathbb{E}_t^X(\widetilde{P}_{t+1} - P_t) = (1-\theta)\sum_{k=1}^{\infty}\theta^{k-1}(P_{t-k} - P_{t-k-1}) \equiv X_t, \tag{A.10}$$

which, for an economy that starts at time 0, can be written as

$$\mathbb{E}_t^X(\widetilde{P}_{t+1} - P_t) = (1-\theta)\sum_{k=1}^{t-1}\theta^{k-1}(P_{t-k} - P_{t-k-1}) + \theta^{t-1}X_1. \tag{A.11}$$

Suppose also, for simplicity, that he sets the conditional variance of price changes equal to the variance of cash-flow shocks, namely, σ_ε^2. His demand then becomes

$$N_t^X = \frac{1}{\gamma\sigma_\varepsilon^2}\left((1-\theta)\sum_{k=1}^{t-1}\theta^{k-1}(P_{t-k} - P_{t-k-1}) + \theta^{t-1}X_1\right), \tag{A.12}$$

as in (5).

Proposition 1. In the economy described in Section 2, a unique market-clearing price always exists and is determined *as follows. Let $\overline{P}_i$, $i \in \{0, 1, \ldots, I\}$, be the risky asset price at which trader i's short-sale constraint starts to bind. Let $N_{\overline{P}_i}$ be the aggregate risky asset share demand across all traders when the price equals $\overline{P}_i$. If $\max_{i\in\{0,1,\ldots,I\}} N_{\overline{P}_i} < Q$, then, in equilibrium, all traders have strictly positive demand for the risky asset and the asset's price equals*

$$P_t = D_t + \frac{\sum_{i=1}^{I}\mu_i(1-w_{i,t})}{\mu_0 + \sum_{i=1}^{I}\mu_i w_{i,t}}X_t - \gamma\sigma_\varepsilon^2 Q\frac{(\mu_0 + \sum_{i=1}^{I}\mu_i w_{i,t})(T-t-1)+1}{\mu_0+\sum_{i=1}^{I}\mu_i w_{i,t}}. \quad \text{(A.13)}$$

Otherwise, let i^* be the value of $i \in \{0, 1, \ldots, I\}$ for which $N_{\overline{P}_i}$ exceeds Q by the smallest amount, and let I^* be the set of $i \in \{0, 1, \ldots, I\}$ such that trader i has strictly positive demand for the risky asset at price $\overline{P}_{i^*}$. In this case, in equilibrium, only the traders in I^* have strictly positive demand for the risky asset and the asset's price equals

$$P_t = D_t + \frac{\sum_{i\in I^*}\mu_i(1-w_{i,t})}{\sum_{i\in I^*}\mu_i w_{i,t}}X_t - \gamma\sigma_\varepsilon^2 Q\frac{(\sum_{i\in I^*}\mu_i w_{i,t})(T-t-1)+1}{\sum_{i\in I^*}\mu_i w_{i,t}}. \quad \text{(A.14)}$$

Proof of Proposition 1. From expressions (9) and (10), we see that aggregate demand for the risky asset, $\mu_0 N_t^F + \sum_{i=1}^{I}\mu_i N_t^{E,i}$, can take an arbitrarily high value if the price P_t is sufficiently low, and a value as low as zero if the price is sufficiently high. Moreover, it is a continuous function of P_t and is strictly decreasing in P_t until it falls to zero. Taken together, these observations imply that there is a unique price P_t at which aggregate demand at time t equals the risky asset supply Q.

We find the market-clearing price in the following way. As noted in the statement of the proposition, we define $\overline{P}_i$ to be the price at which trader i's short-sale constraint binds, namely,

$$\overline{P}_0 = D_t - \gamma\sigma_\varepsilon^2(T-t-1)Q$$
$$\overline{P}_i = D_t - \gamma\sigma_\varepsilon^2(T-t-1)Q + \frac{1-w_{i,t}}{w_{i,t}}X_t, \quad i \in \{1,\ldots,I\}. \quad \text{(A.15)}$$

We now order these $I+1$ "cut-off" prices, so that

$$\overline{P}_{i(0)} \geq \overline{P}_{i(1)} \geq \ldots \geq \overline{P}_{i(I)},$$

where $i(l)$ indexes the trader $i \in \{0, 1, \ldots, I\}$ with the $(l+1)$th highest cut-off price. If $N_{\overline{P}_{i(l)}}$ is aggregate demand at price $\overline{P}_{i(l)}$, we have

$$0 = N_{\overline{P}_{i(0)}} \leq N_{\overline{P}_{i(1)}} \leq \cdots \leq N_{\overline{P}_{i(I)}}.$$

Finally, let $I(l)$ be the set of traders i who have strictly positive demand at price $\overline{P}_{i(l)}$. Note that $I(0)$ is the empty set and that $I(l)$ is a subset of $I(l+1)$.

We consider two cases. Suppose that $N_{\overline{P}_{i(I)}} < Q$. This indicates that the market-clearing price is below $\overline{P}_{i(I)}$, and that, in equilibrium, all traders in the economy have strictly positive demand. Aggregate demand at the market-clearing price P_t therefore equals

$$\sum_{i=0}^{I}\mu_i\left[w_{i,t}\left(\frac{D_t - \gamma\sigma_\varepsilon^2(T-t-1)Q - P_t}{\gamma\sigma_\varepsilon^2}\right) + (1-w_{i,t})\frac{X_t}{\gamma\sigma_\varepsilon^2}\right],$$

where $w_{0,t} \equiv 1$, indicating that fundamental traders put a weight of 1 on the value signal. Setting this aggregate demand equal to supply Q leads to the equilibrium price in (A.13).

We now turn to the other case. Suppose that $N_{\overline{P}_{i(l)}} \leq Q \leq N_{\overline{P}_{i(l+1)}}$. We then know that the market-clearing price is somewhere between $\overline{P}_{i(l+1)}$ and $\overline{P}_{i(l)}$, and that, in equilibrium, only the traders in the set $I(l+1)$, denoted I^* in the statement of the proposition, have strictly positive demand for the risky asset. Aggregate demand at the market-clearing price P_t therefore equals

$$\sum_{i\in I(l+1)}\mu_i\left[w_{i,t}\left(\frac{D_t - \gamma\sigma_\varepsilon^2(T-t-1)Q - P_t}{\gamma\sigma_\varepsilon^2}\right) + (1-w_{i,t})\frac{X_t}{\gamma\sigma_\varepsilon^2}\right].$$

Setting this equal to supply Q, we obtain the equilibrium price in (A.14). □

Proposition 2. Suppose that there is a continuum of extrapolators and that each extrapolator draws an independent weight $w_{i,t}$ at time t from a bounded and continuous density $g(w)$, $w \in [w_l, w_h]$, with mean $\overline{w}$ and with $0 < w_l < w_h < 1$. Suppose that the economy has been in its steady state up to time $l-1$ and that there is then a sequence of positive shocks $\varepsilon_l, \varepsilon_{l+1}, \ldots, \varepsilon_n$ that move the economy from the first stage of the bubble to the second stage at some intermediate date j with $l<j<n$. Also suppose that the economy remains in its second stage through at least date $N>n$.

No price spiral. *If all the extrapolators are in the market at all dates—we specify the condition for this below—the overpricing generated at time t by the cash-flow shocks $\varepsilon_l, \varepsilon_{l+1}, \ldots, \varepsilon_n$ is*

$$\mathcal{O}_t \equiv P_t - P_t^F = \begin{cases} \sum_{m=l}^{t-1}\mathcal{L}_1(t-m)\varepsilon_m & l \leq t < j \\ \sum_{m=j}^{t-1}\mathcal{L}_2(t-m)\varepsilon_m + \mathcal{O}_t^1 & j \leq t \leq N \end{cases}, \quad \text{(A.16)}$$

where $\mathcal{O}_t^1$ is the component of the time t overpricing generated by the shocks $\{\varepsilon_i\}_{i=l}^{j-1}$ that occurred during the first stage of the bubble and is given by

$$\mathcal{O}_t^1 = \begin{cases} (\alpha_2 + \theta\alpha_2\alpha_1^{-1})\mathcal{O}_{j-1} - \alpha_2\mathcal{O}_{j-2} + \alpha_2\varepsilon_{j-1} - \frac{\mu_0}{\overline{w}(1-\mu_0)}\gamma\sigma_\varepsilon^2 Q & t = j \\ (\alpha_2+\theta)\mathcal{O}_{t-1}^1 - \alpha_2\mathcal{O}_{t-2}^1 - \frac{\mu_0(1-\theta)}{\overline{w}(1-\mu_0)}\gamma\sigma_\varepsilon^2 Q & j < t \leq N \end{cases}, \quad \text{(A.17)}$$

where $\alpha_1 \equiv (1-\theta)(1-\mu_0)(1-\overline{w})/(\mu_0+(1-\mu_0)\overline{w})$ *and* $\alpha_2 \equiv (1-\theta)(1-\overline{w})/\overline{w}$. *The quantities* $\{\mathcal{L}_i(j)\}_{j\ge 0}$ *are determined as follows. If* $\alpha_i < 2-\theta-2\sqrt{1-\theta}$ *or* $\alpha_i > 2-\theta+2\sqrt{1-\theta}$, *then*

$$\mathcal{L}_i(j) = 2^{-j}\alpha_i((\alpha_i+\theta)^2-4\alpha_i)^{-0.5} \times[(\alpha_i+\theta+\sqrt{(\alpha_i+\theta)^2-4\alpha_i})^j -(\alpha_i+\theta-\sqrt{(\alpha_i+\theta)^2-4\alpha_i})^j]. \quad \text{(A.18)}$$

If $2-\theta-2\sqrt{1-\theta} < \alpha_i < 2-\theta+2\sqrt{1-\theta}$, *then*

$$\mathcal{L}_i(j) = 2\alpha_i^{0.5j+1}\left(4\alpha_i-(\alpha_i+\theta)^2\right)^{-0.5}\sin(j\beta), \quad \text{(A.19)}$$

where $\beta = \cos^{-1}(0.5(\alpha_i+\theta)\alpha_i^{-0.5})$. *If* $\alpha_i = 2-\theta+2\sqrt{1-\theta}$ *or* $\alpha_i = 2-\theta-2\sqrt{1-\theta}$, *then*

$$\mathcal{L}_i(j) = j\alpha_i^{0.5(j+1)}. \quad \text{(A.20)}$$

Price spiral. *If, at some date* j', $j \le j' \le N$, *the overpricing* $\mathcal{O}_{j'}$ *computed using (A.16) is greater than* $\bar{\mathcal{O}} \equiv ((1-w_h)+(1-\mu_0)(w_h-\overline{w}))\gamma\sigma_\varepsilon^2 Q/((1-\mu_0)(w_h-\overline{w}))$, *then a price spiral begins at* j'. *During the spiral, the time t overvaluation is*

$$\mathcal{O}_t = \frac{1-\overline{w}(X_t)}{\overline{w}(X_t)}X_t + \gamma\sigma_\varepsilon^2 Q - \frac{\gamma\sigma_\varepsilon^2 Q}{(1-\mu_0)\eta(X_t)\overline{w}(X_t)}, \quad t \ge j', \quad \text{(A.21)}$$

where

$$\overline{w}(X_t) = \int_{w_l}^{w^*} wg(w)dw \Big/ \int_{w_l}^{w^*} g(w)dw, \quad \eta(X_t) = \int_{w_l}^{w^*} g(w)dw, \quad \text{(A.22)}$$

$w^*(X_t)$ *is determined by*

$$w^*\gamma\sigma_\varepsilon^2 Q = X_t(1-\mu_0)\int_{w_l}^{w^*}(w^*-w)g(w)dw, \quad \text{(A.23)}$$

and the growth signal X_t *evolves as*

$$X_t = \begin{cases} \dfrac{\theta(\mu_0+(1-\mu_0)\overline{w})(\mathcal{O}_{t-1}-\gamma\sigma_\varepsilon^2 Q)+\theta\gamma\sigma_\varepsilon^2 Q}{(1-\mu_0)(1-\overline{w})} \\ \quad +(1-\theta)(\mathcal{O}_{t-1}-\mathcal{O}_{t-2}+\varepsilon_{t-1}+\gamma\sigma_\varepsilon^2 Q) \\ \text{for} \quad t = j' \\ \theta X_{t-1}+(1-\theta)(\mathcal{O}_{t-1}-\mathcal{O}_{t-2}+\varepsilon_{t-1}+\gamma\sigma_\varepsilon^2 Q) \\ \text{for} \quad t > j' \end{cases} \quad \text{(A.24)}$$

At each time t, extrapolators with $w_{i,t} < w^*(X_t)$ *stay in the market, while those with* $w_{i,t} \ge w^*(X_t)$ *stay out of the market. If the price spiral ends before time N, Eqs. (A.21), (A.22), and (A.24) still apply but with* w^* *set to* w_h.

Before we prove the proposition, we explain it in more detail. Eq. (A.16) gives the magnitude of overvaluation in the absence of a price spiral. To understand it, suppose that, up until time $l-1$, the economy has been in its steady state, and that, at time l, there is a unit cash-flow shock $\varepsilon_l = 1$, after which the cash-flow shocks revert to zero forever. The quantities $\mathcal{L}_1(1)$, $\mathcal{L}_1(2)$, $\mathcal{L}_1(3),\dots$ are equal to the overvaluation of the risky asset 1, 2, 3,… periods after the shock, in other words, at dates $l+1$, $l+2$, $l+3,\dots$, conditional on the bubble staying in the first stage, so that fundamental traders and all extrapolators are in the market. The first row of Eq. (A.16) shows that our model has a linear structure, in the sense that, during the first stage of the bubble, the total overvaluation at time t caused by a *sequence* of shocks $\varepsilon_l, \varepsilon_{l+1},\dots,\varepsilon_{t-1}$ is given by

$$\mathcal{L}_1(1)\varepsilon_{t-1}+\mathcal{L}_1(2)\varepsilon_{t-2}+\cdots+\mathcal{L}_1(t-l)\varepsilon_l.$$

Now suppose that the bubble is in the second stage, but with no price spiral, so that the fundamental traders are not in the market but all extrapolators are. Suppose that there is a unit cash-flow shock at time j, $\varepsilon_j = 1$, after which the shocks equal zero forever. The quantities $\mathcal{L}_2(1)$, $\mathcal{L}_2(2)$, $\mathcal{L}_2(3),\dots$ measure how much *additional* overvaluation this shock creates 1, 2, 3,… periods later, in other words, at dates $j+1$, $j+2$, $j+3,\dots$, relative to the case in which $\varepsilon_j = 0$, and conditional on all extrapolators staying in the market. The second row of Eq. (A.16) shows that, in this second stage of the bubble, the total overvaluation at time t caused by a *sequence* of shocks $\varepsilon_l, \varepsilon_{l+1},\dots,\varepsilon_n$ has two components. The first is the overvaluation created by the cash-flow shocks that arise during the second stage of the bubble. This is again linear in structure and equals

$$\mathcal{L}_2(1)\varepsilon_{t-1}+\mathcal{L}_2(2)\varepsilon_{t-2}+\cdots+\mathcal{L}_2(t-j)\varepsilon_j.$$

The second component of the overvaluation, $\mathcal{O}_t^1$, is typically much smaller in magnitude. It is the overvaluation at time t caused by the lingering effect of the cash-flow shocks that occurred during the *first* stage of the bubble.

Eqs. (A.18)–(A.20) provide explicit expressions for $\mathcal{L}_1(\cdot)$ and $\mathcal{L}_2(\cdot)$. They show that $\mathcal{L}_i(j)$ can take one of four shapes when plotted for $j = 1, 2, \dots$. The two most common shapes are a curve that rises and then falls monotonically and a curve that oscillates with decreasing amplitude. The other possibilities are a curve that oscillates with increasing amplitude and a curve that increases monotonically.

Proof of Proposition 2. Given the assumptions about extrapolators in the statement of Proposition 2 and the results from Proposition 1, the equilibrium price of the risky asset is

$$P_t = D_t + \alpha_1\sum_{k=1}^{\infty}\theta^{k-1}(P_{t-k}-P_{t-k-1}) - (T-t-1)\gamma\sigma_\varepsilon^2 Q - \frac{\gamma\sigma_\varepsilon^2 Q}{\mu_0+(1-\mu_0)\overline{w}} \quad \text{(A.25)}$$

in the first stage of the bubble, where $\alpha_1 \equiv \frac{(1-\theta)(1-\mu_0)(1-\overline{w})}{\mu_0+(1-\mu_0)\overline{w}}$. In the second stage of the bubble, so long as all the extrapolators are in the market, the equilibrium price is

$$P_t = D_t + \alpha_2\sum_{k=1}^{\infty}\theta^{k-1}(P_{t-k}-P_{t-k-1}) - (T-t-1)\gamma\sigma_\varepsilon^2 Q - \frac{\gamma\sigma_\varepsilon^2 Q}{(1-\mu_0)\overline{w}}, \quad \text{(A.26)}$$

where $\alpha_2 \equiv \frac{(1-\theta)(1-\overline{w})}{\overline{w}} > \alpha_1$.

From (4), (A.25), and (A.26), the level of overpricing $\mathcal{O}_t$, defined as the difference between the price of the risky asset and its fundamental value, is

$$\mathcal{O}_t = \begin{cases} \dfrac{1-\overline{w}}{\overline{w}}X_t - \dfrac{\gamma\sigma_\varepsilon^2 Q}{(1-\mu_0)\overline{w}} + \gamma\sigma_\varepsilon^2 Q \\ \text{for} \quad X_t > \dfrac{\gamma\sigma_\varepsilon^2 Q}{(1-\mu_0)(1-\overline{w})} \\ \dfrac{(1-\mu_0)(1-\overline{w})}{\mu_0+(1-\mu_0)\overline{w}}X_t - \dfrac{\gamma\sigma_\varepsilon^2 Q}{\mu_0+(1-\mu_0)\overline{w}} + \gamma\sigma_\varepsilon^2 Q \\ \text{for} \quad X_t \le \dfrac{\gamma\sigma_\varepsilon^2 Q}{(1-\mu_0)(1-\overline{w})} \end{cases}. \tag{A.27}$$

Note that $\mathcal{O}_t$ is continuous at the switching point between the first and second stages of the bubble; at this point, $\mathcal{O}_t = \gamma\sigma_\varepsilon^2 Q$. Also note that, when the value of X_t equals its steady-state level of $\gamma\sigma_\varepsilon^2 Q$, the overpricing is zero; in this case, the per capita demand of both extrapolators and fundamental traders for the risky asset equals the supply Q.

From (A.25) and (A.26) it is apparent that, if the economy stays within stage one or within stage two with all the extrapolators in the market, the model has a linear structure: in stage i, a fundamental shock of ε_{t_1} at t_1 and a fundamental shock of ε_{t_2} at t_2 generate, at a later time t, a total overvaluation of $\mathcal{L}_i(t-t_1)\varepsilon_{t_1} + \mathcal{L}_i(t-t_2)\varepsilon_{t_2}$. It is also straightforward to check that $\mathcal{L}_1(\cdot)$ and $\mathcal{L}_2(\cdot)$ can be defined recursively as

$$\begin{aligned} &\mathcal{L}_i(0) = 0, \quad \mathcal{L}_i(1) = \alpha_i, \\ &\mathcal{L}_i(l) = (\alpha_i+\theta)\mathcal{L}_i(l-1) - \alpha_i\mathcal{L}_i(l-2), \\ &\qquad l \ge 2 \quad \text{for } i = 1, 2. \end{aligned} \tag{A.28}$$

This is a standard difference equation with the general solution

$$\mathcal{L}_i(j) = A_1(K_1)^j + A_2(K_2)^j, \tag{A.29}$$

where K_1 and K_2 are the roots of the quadratic equation

$$K^2 - (\alpha_i+\theta)K + \alpha_i = 0 \tag{A.30}$$

and where A_1 and A_2 can be obtained from the boundary conditions $\mathcal{L}_i(0) = 0$ and $\mathcal{L}_i(1) = \alpha_i$. When $(\alpha_i+\theta)^2 > 4\alpha_i$, (A.30) has two real roots; matching (A.29) with the boundary conditions gives (A.18). When $(\alpha_i+\theta)^2 < 4\alpha_i$, (A.30) has two complex roots with nonzero imaginary components; matching (A.29) with the boundary conditions gives (A.19). When $(\alpha_i+\theta)^2 = 4\alpha_i$, applying L'Hôpital's rule to either (A.18) or (A.19) gives (A.20).

The linear structure implies that, at time t with $l \le t < j$, the overpricing generated by $\{\varepsilon_i\}_{i=l}^{t-1}$ is $\sum_{m=l}^{t-1}\mathcal{L}_1(t-m)\varepsilon_m$; and that, at time t with $j \le t \le N$, the additional overpricing generated by $\{\varepsilon_i\}_{i=j}^{t-1}$ is $\sum_{m=j}^{t-1}\mathcal{L}_2(t-m)\varepsilon_m$.

We now derive $\mathcal{O}_t^1$ at time $t \ge j$. For $t = j$,

$$X_j = (1-\theta)(P_{j-1} - P_{j-2}) + \theta X_{j-1}. \tag{A.31}$$

From (A.16) we know

$$P_{j-1} - P_{j-2} = \mathcal{O}_{j-1} - \mathcal{O}_{j-2} + \varepsilon_{j-1} + \gamma\sigma_\varepsilon^2 Q \tag{A.32}$$

and

$$X_{j-1} = \frac{\mu_0+(1-\mu_0)\overline{w}}{(1-\mu_0)(1-\overline{w})}\left(\mathcal{O}_{j-1} + \frac{\gamma\sigma_\varepsilon^2 Q}{\mu_0+(1-\mu_0)\overline{w}} - \gamma\sigma_\varepsilon^2 Q\right). \tag{A.33}$$

Substituting (A.32) and (A.33) into (A.31), and then substituting (A.31) back into (A.27) gives $\mathcal{O}_j^1$ in (A.17). For $j < t \le N$, similar steps lead to $\mathcal{O}_t^1$ in (A.17).

Substituting Eq. (A.26) into the extrapolator share demand in (10) shows that, whenever $X_t > w_h\gamma\sigma_\varepsilon^2 Q/((1-\mu_0)(w_h-\overline{w}))$, the extrapolator with $w_{i,t} = w_h$ exits the market and a price spiral occurs. Eq. (A.27) shows that this condition is equivalent to $\mathcal{O}_t > \bar{\mathcal{O}}$. Applying (A.27) at time $j'-1$ gives $X_{j'-1}$ as a function of $\mathcal{O}_{j'-1}$, and further applying (A.31) and (A.32) gives (A.24).

Assume that, at time t, extrapolators with $w_{i,t} \in [w_l, w^*)$ are in the market. Integrating the share demands of these extrapolators in (10) and equating the result to the aggregate per-extrapolator supply of $Q/(1-\mu_0)$ gives (A.21) and (A.22). Setting the share demand of the extrapolator with $w_{i,t} = w^*$ to zero then gives (A.23). Given that $X_t > w_h\gamma\sigma_\varepsilon^2 Q/((1-\mu_0)(w_h-\overline{w}))$, the left-hand side of (A.23) is smaller than the right-hand side when $w^* = w_h$; however, the left-hand side of (A.23) is greater than the right-hand side when $w^* = w_l$. As a result, there must exist a w^* that solves (A.23). □

Proof of Proposition 3. Substituting the equilibrium asset price in (A.25) and (A.26) into extrapolator i's share demand in (10) gives

$$N_t^{E,i} = \begin{cases} \max\left[\dfrac{\overline{w}-w_{i,t}}{\overline{w}\gamma\sigma_\varepsilon^2}X_t + \dfrac{w_{i,t}Q}{(1-\mu_0)\overline{w}}, 0\right] \\ \text{for} \quad X_t > \dfrac{\gamma\sigma_\varepsilon^2 Q}{(1-\mu_0)(1-\overline{w})} \\ \max\left[\dfrac{\mu_0(1-w_{i,t}) + (1-\mu_0)(\overline{w}-w_{i,t})}{(\mu_0+(1-\mu_0)\overline{w})\gamma\sigma_\varepsilon^2}X_t \right. \\ \qquad \left. + \dfrac{w_{i,t}Q}{\mu_0+(1-\mu_0)\overline{w}}, 0\right] \\ \text{for} \quad X_t \le \dfrac{\gamma\sigma_\varepsilon^2 Q}{(1-\mu_0)(1-\overline{w})} \end{cases} \tag{A.34}$$

Suppose that the fundamental traders and all extrapolators are in the market; from (A.34) we know that this is true when $-w_l\gamma\sigma_\varepsilon^2 Q/(\mu_0(1-w_l) + (1-\mu_0)(\overline{w}-w_l)) \le X_t < \gamma\sigma_\varepsilon^2 Q/((1-\mu_0)(1-\overline{w}))$. In this case, the component of the change in extrapolator i's share demand between time t and $t+1$ that is due to wavering is

$$\frac{(w_{i,t+1}-w_{i,t})(\gamma\sigma_\varepsilon^2 Q - X_{t+1})}{(\mu_0+(1-\mu_0)\overline{w})\gamma\sigma_\varepsilon^2}. \tag{A.35}$$

Taking the absolute value of this quantity—conditional, for simplicity, on $X_{t+1} = X_t$—and integrating over $w_{i,t+1}$ and $w_{i,t}$ shows that wavering-induced trading volume is equal to

$$\frac{|X_t - \gamma\sigma_\varepsilon^2 Q|\overline{\Delta}_0}{(\mu_0+(1-\mu_0)\overline{w})\gamma\sigma_\varepsilon^2}, \tag{A.36}$$

where $\overline{\Delta}_0$ is defined in (17). Now suppose that the fundamental traders are out of the market but all extrapolators are in; this is true when $\gamma\sigma_\varepsilon^2 Q/((1-\mu_0)(1-\overline{w})) \le X_t \le w_h\gamma\sigma_\varepsilon^2 Q/((w_h-\overline{w})(1-\mu_0))$. In this case, the component of the change in extrapolator i's share demand between t

and $t+1$ that is due to wavering is

$$\frac{(w_{i,t+1}-w_{i,t})(\gamma\sigma_\varepsilon^2 Q-(1-\mu_0)X_{t+1})}{(1-\mu_0)\overline{w}\gamma\sigma_\varepsilon^2}. \quad \text{(A.37)}$$

A similar calculation to that used to obtain (A.36) shows that, in this case, wavering-induced trading volume is given by

$$\frac{((1-\mu_0)X_t-\gamma\sigma_\varepsilon^2 Q)\overline{\Delta}_0}{(1-\mu_0)\overline{w}\gamma\sigma_\varepsilon^2}. \quad \text{(A.38)}$$

When $X_t > w_h\gamma\sigma_\varepsilon^2 Q/((w_h-\overline{w})(1-\mu_0))$, extrapolators with a sufficiently high level of w stay out of the market but may re-enter in the next period. For those extrapolators who stay in for both periods, replace $\overline{w}$, $1-\mu_0$, and $\overline{\Delta}_0$ in (A.38) by $\overline{w}(X_t)$, $(1-\mu_0)\eta(X_t)$, and $\overline{\Delta}(X_t)$, respectively, where

$$\overline{w}(X_t) \equiv (\eta(X_t))^{-1}\int_{w_l}^{w_\eta(X_t)} wg(w)dw$$

$$\eta(X_t) \equiv \int_{w_l}^{w_\eta(X_t)} g(w)dw$$

$$\overline{\Delta}(X_t) \equiv \int_{w_l}^{w_\eta(X_t)}\int_{w_l}^{w_\eta(X_t)} |w_1-w_2|g(w_1)g(w_2)dw_1dw_2,$$

and where $w_\eta(X_t)$ is the implicit solution to

$$(1-\mu_0)w_\eta\left(\int_{w_l}^{w_\eta} g(w)dw\right)X_t - w_\eta\gamma\sigma_\varepsilon^2 Q = (1-\mu_0)\left(\int_{w_l}^{w_\eta} wg(w)dw\right)X_t. \quad \text{(A.39)}$$

For those extrapolators who are in at time t but out at time $t+1$, their change in share demand is

$$\frac{X_t}{\gamma\sigma_\varepsilon^2} - \frac{w_{i,t}((1-\mu_0)\eta(X_t)X_t-\gamma\sigma_\varepsilon^2 Q)}{(1-\mu_0)\eta(X_t)\gamma\sigma_\varepsilon^2\overline{w}(X_t)} \geq 0 \quad \text{(A.40)}$$

for $w_{i,t} \leq w_\eta(X_t)$. Integrating (A.40) over $w_{i,t}$ from w_l to $w_\eta(X_t)$ and then further integrating it over $w_{i,t+1}$ from $w_\eta(X_t)$ to w_h gives $(1-\eta(X_t))Q/(1-\mu_0)$. The trading volume generated by extrapolators who are out at time t but in at time $t+1$ can be computed in a similar way; it also equals $(1-\eta(X_t))Q/(1-\mu_0)$. Overall, wavering-induced trading volume in this case equals

$$\frac{(\eta(X_t)X_t-\gamma\sigma_\varepsilon^2 Q(1-\mu_0)^{-1})\overline{\Delta}(X_t)}{\eta(X_t)\overline{w}(X_t)\gamma\sigma_\varepsilon^2} + \frac{2(1-\eta(X_t))Q}{1-\mu_0}. \quad \text{(A.41)}$$

Taking the derivative of expressions (A.36), (A.38), and (A.41) gives the results in Proposition 3. □

References

Abreu, D., Brunnermeier, M., 2003. Bubbles and crashes. Econometrica 71, 173–204.
Allen, F., 1931. Only Yesterday: An Informal History of the 1920's. Harper and Brothers, New York.
Bagehot, W., 1873. Lombard Street: A Description of the Money Market. Wiley, New York.
Barberis, N., Greenwood, R., Jin, L., Shleifer, A., 2015. X-CAPM: an extrapolative capital asset pricing model. J. Financ. Econ. 115, 1–24.
Barberis, N., Shleifer, A., 2003. Style investing. J. Financ. Econ. 68, 161–199.
Blanchard, O., Watson, M., 1982. Bubbles, rational expectations, and financial markets. In: Wachtel, P. (Ed.), Crises in the Economic and Financial Structure. Lexington Books, Lexington, MA, pp. 296–316.
Brunnermeier, M., Nagel, S., 2004. Hedge funds and the technology bubble. J. Financ. 59, 2013–2040.
Brunnermeier, M., Oehmke, M., 2013. Bubbles, financial crises, and systemic risk. In: Constantinides, G., Harris, M., Stulz, R. (Eds.), Handbook of the Economics of Finance, Vol. 2b. North Holland, pp. 1221–1288.
Carlos, A., Neal, L., Wandschneider, K., 2006. Dissecting the Anatomy of Exchange Alley: The Dealings of Stockjobbers During and After the South Sea Bubble. Unpublished Working Paper. University of Illinois.
Case, K., Shiller, R., Thompson, A., 2012. What have they been thinking? Homebuyer behavior in hot and cold markets. Brook. Pap. Econ. Act. 45, 265–315.
Cheng, I.H., Xiong, W., 2014. The financialization of commodity markets. Ann. Rev. Financ. Econ. 6, 419–441.
Cutler, D., Poterba, J., Summers, L., 1990. Speculative dynamics and the role of feedback traders. Am. Econ. Rev. Pap. Proc. 80, 63–68.
De Long, B., Shleifer, A., Summers, L., Waldmann, R., 1990. Positive feedback investment strategies and destabilizing rational speculation. J. Financ. 45, 379–395.
Fama, E., 2014. Two pillars of asset pricing. Am. Econ. Rev. 104, 1467–1485.
Fehr, E., Rangel, A., 2011. Neuroeconomic foundations of economic choice—recent advances. J. Econ. Perspect. 25, 3–30.
Galbraith, J., 1954. The Great Crash: 1929. Houghton Mifflin Co., Boston, MA.
Giglio, S., Maggiori, M., Stroebel, J., 2016. No-bubble condition: model-free tests in housing markets. Econometrica 84, 1047–1091.
Glaeser, E., Nathanson, C., 2017. An extrapolative model of house price dynamics. J. Financ. Econ. 126, 147–170.
Greenwood, R., Shleifer, A., 2014. Expectations of returns and expected returns. Rev. Financ. Stud. 27, 714–746.
Griffin, J., Nardari, F., Stulz, R., 2007. Do investors trade more when stocks have performed well? Evidence from 46 countries. Rev. Financ. Stud. 20, 905–951.
Harrison, M., Kreps, D., 1978. Speculative investor behavior in a stock market with heterogeneous expectations. Q. J. Econ. 89, 323–336.
Haruvy, E., Lahav, Y., Noussair, C., 2007. Traders' expectations in asset markets: experimental evidence. Am. Econ. Rev. 97, 1901–1920.
Hong, H., de Paula, A., Singh, V., 2015. Hoard Behavior and Commodity Bubbles. NBER Working Paper No. 20974.
Hong, H., Sraer, D., 2013. Quiet bubbles. J. Financ. Econ. 110, 596–606.
Hong, H., Stein, J., 1999. A unified theory of underreaction, momentum trading, and overreaction in asset markets. J. Financ. 54, 2143–2184.
Hong, H., Stein, J., 2007. Disagreement and the stock market. J. Econ. Perspect. 21, 109–128.
Irwin, S., Sanders, D., 2010. The Impact of Index and Swap Funds on Commodity Futures Markets: Preliminary Results. OECD Food, Agriculture, and Fisheries Working Paper No. 27.
Jin, L., 2015. A speculative asset pricing model of financial instability. California Institute of Technology, Unpublished working paper.
Kindleberger, C., 1978. Manias, Panics, and Crashes: A History of Financial Crises. Basic Books, New York.
Mackay, C., 1841. Extraordinary Popular Delusions and the Madness of Crowds. Harmony Books, New York.
Ofek, E., Richardson, M., 2003. Dotcom mania: the rise and fall of internet stock prices. J. Financ. 58, 1113–1138.
Pastor, L., Veronesi, P., 2006. Was there a Nasdaq bubble in the 1990s? J. Financ. Econ. 81, 61–100.
Scheinkman, J., Xiong, W., 2003. Overconfidence and speculative bubbles. J. Polit. Econ. 111, 1183–1219.
Shiller, R., 2000. Irrational Exuberance. Princeton University Press, New Jersey.
Simsek, A., 2013. Belief disagreements and collateral constraints. Econometrica 81, 1–53.
Smith, V., Suchanek, G., Williams, A., 1988. Bubbles, crashes, and endogeneous expectations in experimental spot asset markets. Econometrica 56, 1119–1151.
Statman, M., Thorley, S., Vorkink, K., 2006. Investor overconfidence and trading volume. Rev. Financ. Stud. 19, 1531–1565.
Tirole, J., 1985. Asset bubbles and overlapping generations. Econometrica 53, 1071–1100.
Towal, R.B., Mormann, M., Koch, C., 2013. Simultaneous modeling of visual saliency and value computation improves predictions of economic choice. Proc. Natl. Acad. Sci. 110, E3858–3867.
White, E., 1990. The stock market boom and crash of 1929 revisited. J. Econ. Perspect. 4, 67–83.

Wigmore, B., 1985. The Crash and Its Aftermath: A History of Securities Markets in the United States, 1929–1933. Greenwood Press, Westport, CT.
Xiong, W., 2013. Bubbles, crises, and heterogeneous beliefs. In: Fouque, J.-P., Langsam, J. (Eds.), Handbook for Systemic Risk. Cambridge University Press, New York, pp. 664–713.
Xiong, W., Yu, J., 2011. The Chinese warrants bubble. Am. Econ. Rev. 101, 2723–2753.