A simple mechanism for financial bubbles: time-varying momentum horizon

Journal Article**Author(s):**

Lin, Li; Schatz, Michael; Sornette, Didier

Publication date:

2019

Permanent link:

<https://doi.org/10.3929/ethz-b-000317879>

Rights / license:

In Copyright - Non-Commercial Use Permitted

Originally published in:

Quantitative Finance 19(6), <https://doi.org/10.1080/14697688.2018.1540881>

A simple mechanism for financial bubbles: time-varying momentum horizon*

L. Lin^{1, 4 †}, M. Schatz² and D. Sornette^{2, 3}

October 3rd, 2018

1. School of Business, East China University of Technology and Science, 200237 Shanghai, China
2. ETH Zurich, Department of Management, Technology and Economics, Scheuchzerstrasse 7, CH-8092 Zurich, Switzerland
3. Swiss Finance Institute, c/o University of Geneva, 40 blvd. Du Pont d'Arve, CH 1211 Geneva 4, Switzerland
4. Research Institute of Financial Engineering, East China University of Technology and Science, 200237 Shanghai, China

Abstract

Building on the notion that bubbles are transient self-fulfilling prophecies created by positive feedback mechanisms, we construct the simplest continuous price process whose expected returns and volatility are functions of momentum only. The momentum itself is measured by a simple continuous moving average of past prices over a given time horizon. We introduce a simple dynamics of the time horizon used by the representative investor, which is motivated by the race of trend following agents to forerun their competitors. We provide the full set of solutions, which includes an explosive regime where the price and momentum explodes stochastically in finite time to infinity, transient price dynamics escaping to infinity and recurrent behaviors, where the momentum remains either strictly positive or undergoes instantaneous reflections at the origin. The proposed price generating process produces price dynamics that are in agreement with the main qualitative properties of empirical financial time series. Moreover, it produces realistic regime shifts between non bubble and bubble regimes. We construct a quasi-likelihood methodology to calibrate the model to empirical financial time series, which is applied to an Internet index and a “brick and mortar” index, over the period of the dotcom bubble and its subsequent crash, from Jan. 1998 to Dec. 2002. The Wilks test of nested hypotheses shows a very strong skill in diagnosing the bubble of the Internet index and in disqualifying a bubble in the “brick and mortar” index.

Keywords: financial bubbles, momentum, positive feedback, time-horizon, quasi-likelihood, finite-time-singularity

JEL classification: C52, G01, G17

*This is an Accepted Manuscript of an article published by Taylor & Francis in Quantitative Finance on 14 Nov 2018, available at DOI:10.1080/14697688.2018.1540881.

†Email: llin@ecust.edu.cn

1 Introduction

We provide a simple model that builds on the fundamental assumption that financial markets continuously exhibit transitions between phases of growth, exuberance (conventionally called bubbles) and crises [Sornette, 2004]. In this framework, most crises are actually endogenous and the consequence of procyclical (also called positive) feedbacks [Sornette and Cauwels, 2014, 2015a,b]. In our stylized framework, the excess return and volatility of the price generating process are assumed to be functions of the momentum only, which is estimated over a time horizon that may vary as a function of the momentum itself. The intermittent positive feedbacks between momentum and the time horizon over which momentum is estimated leads to regime shifts from non-bubble to transient bubble phases. As a result, the obtained price process exhibits the major known stylized facts of empirical financial time series. In particular, the model provides a straight-forward mechanism for the formation and evolution of bubbles, which combines the concept that momentum trading is ubiquitous amongst investors and the pervasive occurrence of bubbles.

Indeed, momentum or trend following strategies constitute one of the most popular technical investment techniques used by hedge-funds, mutual funds and individual investors in general [Antonacci, 2013, 2014]. Momentum, the tendency for rising asset prices to rise further, and falling prices to keep falling, enjoys strong empirical support [Jegadeesh and Titman, 1993; Grinblatt et al., 1995; Jegadeesh and Titman, 2001; Grinblatt and Moskowitz, 2004] and provides an improved explanatory power of risk premia in factor model regressions [Carhart, 1997; Fama and French, 2012]. Momentum effects have been documented in the US, Europe and Asia Pacific markets [Fama and French, 2012], and across different asset classes [Asness et al., 2013]. In addition, momentum is not only present in cross-sectional investigations for individual firms but also detected from the aggregated time series like stock indices and futures contracts [He and Li, 2015; Moskowitz et al., 2012].

Momentum is often attributed to investors' behavioral characteristics represented by their beliefs, such as over-confidence, self-attribution and confirmation biases [Daniel et al., 1998], and under-reaction/over-reaction [Barberis et al., 1998], or to preferences, such as the disposition effect derived from prospect theory and mental accounting operations [Grinblatt and Han, 2005]. Yet another explanation for momentum effects is based on herding [Hoitash and Krishnan, 2008; Demirel et al., 2015]. Orléan [1995] promoted the concept of convention to emphasise that prices reflect investors's beliefs. In self-fulfilling prophecies, investors make these beliefs come true through. Building on this understanding, Wyart and Bouchaud [2007] formulated a simple model for self-referential behavior in financial markets where agents build strategies based on their belief about the existence of correlation between some flow of information and price changes. Through their market impact, their strategies ensure those beliefs materialize, leading to excess volatility and regime shifts. Farmer [2002] argued that trend-following strategies cause both short-term trends in prices, as well as longer-term oscillations and boom-bust cycles.

Previous models developed by our group implementing the concept of positive feedbacks to characterize bubbles have emphasized mostly the influence of past prices level on future returns [Sornette and Andersen, 2002; Lin and Sornette, 2013; Hüsler et al., 2013; Corsi and Sornette, 2014; Lin et al., 2014; Sornette and Cauwels, 2015a] and the reinforcing role of social influence of the price formation process [Kaizoji et al., 2015]. The role of a nonlinear response to momentum [Ide and Sornette, 2002] has been suggested to be at the origin of the super-exponential price growth characterising bubbles [Sornette and Cauwels, 2015a; Leiss et al., 2015] and also as a model of hyperinflation [Sornette et al., 2003]. Here, we take a different route by constructing the simplest continuous price process whose expected returns and volatility are functions of momentum only. The momentum is measured as a simple continuous moving average of past prices over a given time horizon. This momentum is then used to surmise future returns by the investors. We formulate a simple self-consistent framework to embody the market impact of momentum strategies. The key idea is that investors use momentum strategies because they believe that these strategies have value. Due to their market impact, the momentum strategies in turn modify the price structure, leading to a reinforcement of the price momentum. In addition, taking into account the fact that competition between momentum traders make them decrease their time horizons as the price momentum develops, this leads to transient self-fulfilling bubbles

via the positive feedback mechanism between the triplet of price, momentum and the time horizon at which momentum is estimated.

The organization of this article is the following. Section 2 presents a derivation of our model, introducing and discussing the essential underlying assumptions. Section 3 presents a unified precise setting, which summarizes the derivation of section 2, and investigates the properties of the solutions to our model. Section 4 describes a quasi-likelihood methodology to calibrate the model to synthetic and empirical financial time series data and a hypothesis test for bubbles. Section 5 concludes. The proofs are given in the appendices.

2 Derivation of the momentum driven price model

2.1 General formulation of momentum dynamics

Consider a price series $\{S_t, t = 0, 1, 2, \dots\}$ in discrete time and its associated returns $\{r_t := \ln(S_t/S_{t-1}), t = 1, 2, \dots\}$. The exponential moving average momentum X_t over a typical time horizon $\tau = 1/\theta$, with $\theta \in (0, 1)$, is defined as

$$X_t = (1 - \theta)X_{t-1} + \theta r_t. \quad (1)$$

Let L be the lag operator defined by $L[x(t)] = x(t - 1)$. Then, eq.(1) can be reformulated as

$$[1 - (1 - \theta)L]X_t = \theta r_t = \theta \ln \left(\frac{S_t}{LS_t} \right) \quad (2)$$

Its solution reads

$$X_t = [1 - (1 - \theta)L]^{-1} \left(\theta \ln \frac{S_t}{LS_t} \right) = \theta \sum_{k=0}^{\infty} (1 - \theta)^k \ln \frac{S_{t-k}}{S_{t-k-1}}. \quad (3)$$

This shows that the average momentum X_t is nothing but the geometric moving average of the historical log price increments.

Eq.(1) can be reformulated in terms of the increments $\Delta X_t := X_t - X_{t-1}$ and $\Delta \ln S_t := \ln(S_t/S_{t-1})$ to obtain

$$\Delta X_t = -\theta X_{t-1} + \theta \Delta \ln S_t. \quad (4)$$

This suggests the following extension to continuous times of the average price momentum, obtained as the solution of the following stochastic differential equation

$$dX_t = -\theta X_t dt + \theta \frac{dS_t}{S_t}, \quad (5)$$

which can be interpreted as the continuous limit of the difference equation (4).

Example. Let us first consider the simplest example of a constant time horizon $\tau = 1/\theta$ and an asset price S that follows a geometrical Brownian motion (GBM),

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t. \quad (6)$$

Replacing this process in (5) shows that the average price momentum X must be the Ornstein-Uhlenbeck process

$$dX_t = -\theta(X_t - \mu)dt + \theta\sigma dW_t, \quad (7)$$

whose solution reads

$$X_t = \mu + (x_0 - \mu)e^{-\theta t} + \theta\sigma \int_0^t e^{-\theta(t-s)} dW_s. \quad (8)$$

This solution can be transformed into the more convenient form

$$X_t = \mu + \sigma \sqrt{\frac{\theta}{2}} e^{-\theta t} \cdot \widetilde{W}(e^{2\theta t}), \quad (9)$$

for a Brownian motion $\widetilde{W}$. This shows that, while the GBM price process (6) is non-stationary, its moving average X is a stationary process with a stationary distribution given by

$$X_t \sim N\left(\mu, \frac{\theta}{2}\sigma^2\right). \quad (10)$$

The longer the time horizon $\tau(\theta) = 1/\theta$ over which the average price momentum is estimated, the smaller is its variance $\theta\sigma^2/2$. This is nothing else but the statement that the average price momentum is a consistent estimator of the price drift μ , which converges with a standard deviation scaling as the inverse square root $1/\sqrt{\tau(\theta)}$ of the time horizon $\tau(\theta)$ used for its estimation.

2.2 Preliminary setting

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, \infty)}, \mathbb{P})$ be a filtered probability space with a right-continuous and $\mathbb{P}$ -complete filtration, let $(W_t)_{t \in [0, \infty)}$ be a real valued $(\mathcal{F}_t)_{t \in [0, \infty)}$ -adapted Brownian motion, let $f : [0, \infty) \times \Omega \rightarrow \mathbb{R}$ be adapted and $g : [0, \infty) \times \Omega \rightarrow [0, \infty)$ be predictable be *suitably regular*¹ stochastic processes, let $(S_t)_{t \in [0, \infty)}$ be an Itô process with the property that

$$\frac{dS_t}{S_t} = f_t dt + g_t dW_t. \quad (11)$$

Further, following the reasoning in section 2.1 above, let $\theta : [0, \infty) \times \Omega \rightarrow \mathbb{R}$ be a stochastic process describing the inverse time horizon and $(X_t)_{t \in [0, \infty)}$ be the average price momentum of S given by

$$dX_t = -\theta_t X_t dt + \theta_t \frac{dS_t}{S_t}. \quad (12)$$

To impose further structure on our model, we will make two major additional assumptions that are discussed below, a structural assumption on the impact of momentum on the price and an explicit mechanism allowing for positive feedback of increasing momentum on itself.

2.3 Assumption I: Impact of momentum on absolute price level

Assuming the preliminary setting 2.2 and based on the intuition that investors focus in aggregate on the momentum as the leading indicator of pricing, we introduce

Assumption (A). There exists a C^2 -function $\lambda : [0, \infty) \rightarrow [0, \infty)$ with strictly positive first and second derivative such that the price can be written as a function of the momentum X , in the sense that

$$S_t = \lambda(X_t), \quad t \in [0, \infty). \quad (13)$$

Lemma 1 in the appendix shows that assumption (A) naturally implies that the time-varying horizon θ_t and the volatility represented by g_t^2 satisfy

$$\theta_t = \frac{\lambda(X_t)}{\lambda'(X_t)} \quad \text{and} \quad g_t^2 = \frac{2\lambda(X_t)X_t}{\theta_t^2 \lambda''(X_t)}, \quad \text{for all } t \in [0, \infty). \quad (14)$$

¹Below we will allow for processes with explosion and thus deal with solutions S and coefficient functions f, g being defined only on stochastic intervals $[0, \zeta)$ for a random time $\zeta : \Omega \rightarrow [0, \infty)$.

Comments. The impact of momentum on price has been considered in the literature by letting its drift and volatility functions depend on X to arrive at a specification of the form

$$\frac{dS_t}{S_t} = f(X_t) dt + g(X_t) dW_t. \quad (15)$$

This kind of structure has been employed in the stochastic volatility model of [Hobson and Rogers, 1998], which builds the price process as depending on an exponentially weighted moving average of some past price reference level, similar to equation (15). However, in their model, X_t is the price change surprise, defined as the difference between the current log-price and an exponentially weighted average of past log-prices, rather than the price momentum itself used in our model. Kim and Omberg [1996] and Garleanu and Pedersen [2013] develop models of the impact of trading strategies and portfolio allocation on prices that have similar forms as (15).

Note that the above assumption (A) is a somewhat stronger assumption than (15) and imposes a lot of structure on our model. To see this, recall the example of Brownian motion discussed in section 2.1 above: S is a geometric Brownian motion while its exponential moving average X is an Ornstein-Uhlenbeck process. Consequently, there can be no function λ with the property that $S_t = \lambda(X_t)$ for $t \in [0, \infty)$, indicating that assumption (A) a priori restricts feasible asset price dynamics. Viewed from a different perspective, we start from the inherently non-Markovian structure given by equations (11) and (12) (the price increments depend on the momentum, which itself depends on the price history) and look for a Markovian special case that still satisfies the momentum equations. Also note that assumption (A) does not impose a specific form of the drift f_t of S .

2.4 Assumption II : Positive feedback of momentum on momentum time-horizon

Assuming again the preliminary setting 2.2 and based on the possible effect of positive feedback of momentum on its own time horizon $\tau = 1/\theta$ during a bubble, we introduce

Assumption (B). There exist constants $\theta^* \in (0, \infty)$ and $\eta \in [0, 1)$ such that the random time horizon $(\theta_t)_{t \in [0, \infty)}$ can be written as

$$\theta_t = \theta(X_t) = \theta^* + \eta X_t. \quad (16)$$

Together with assumption (A) (in the form of equation (14)), this implies that there exists a constant $\underline{S} \in (0, \infty)$ such that

$$S_t = \underline{S} \left(1 + \frac{\eta}{\theta^*} X_t\right)^{\frac{1}{\eta}} = \underline{S} \left(\frac{\theta_t}{\theta^*}\right)^{\frac{1}{\eta}}. \quad (17)$$

and

$$g_t = g(X_t) = \sqrt{\frac{2}{1-\eta}} X_t. \quad (18)$$

Note that the case $\eta = 0$ is embedded in this formulation via its pointwise limit

$$S_t = \lim_{\eta \rightarrow 0} \left(1 + \frac{\eta}{\theta^*} X_t\right)^{\frac{1}{\eta}} = \exp\left(\frac{X_t}{\theta^*}\right). \quad (19)$$

Comments. Statistical tests of the performance of momentum or trend-following strategies suggest that they have delivered significant positive performance through very different economic and financial environments and over very long time horizons [Hurst et al., 2014; Lempérière et al., 2014]. This is reflected in their broad use by hedge-funds and professional investors. However, their implementation in general involves time-varying exposures to manage the highly variable risk of momentum [Taylor, 2014; Barroso and Santa-Clara, 2015; Elaut and Erdős, 2016]. Moreover, they exhibit periods with weak performance [Hwang and Rubesam,

2015], as for instance since 2011 for most developed financial markets. Pressured by the need to provide positive risk-adjusted return, many professional investors periodically tinker the parameters of their trend-following strategies to adapt to the recent past. For this purpose machine learning algorithms are often employed, with the rational to approximate best the unknown price generating process by adapting the parameters of the momentum strategies. This tends to create destabilising feedback loops in the form of bubble-corrections cycles [Farmer, 2002; Bouchaud and Challet, 2016].

Thus, during a typical financial cycle, the time horizon used in momentum strategies is progressively reduced as an adaptation to the increasing price levels and to the accelerating volatility dynamics. In a *reflexive* race to forerun competitors, trend-followers attempt to adjust their position before the herd. They believe that, by decreasing the time horizon of their momentum indicators, they will be able to forerun their peers². In general, this culminates in a bubble peak, followed by a high volatility regime, possibly a crash and the subsequent detuning of trend-following strategies. And it is only after a recovery period that they become again usable when a new trend appears.

The goal of assumption (B) is to include this effect in the momentum driven price model. To represent the attempt by trend-followers not to be left behind by their peers, we introduce this second-order trend following in the sense that the parameter θ , introduced in section 2.1 to estimate the price momentum, is itself a function of the observed momentum X . Specifically, we assume that, for every t , θ_t is a linearly increasing function of X_t . The representative investor thus becomes more and more myopic for increasing momentum, as his time horizon $\tau = 1/\theta$ is shrinking.

2.5 Complementary Assumption: linear impact of momentum on the drift

Finally, to complement assumptions (A) and (B) of sections 2.3 and 2.4 and for practical ease, we add the following assumption on the drift f_t in the preliminary setting 2.2. In analogy with earlier assumptions, we assume that the drift can be written as a function of the momentum, specifically

Assumption (C). There exist constants $c \in \mathbb{R}$, $d \in (0, \infty)$ such that the drift f of S can be written as the following linear function

$$f_t = f(X_t) = cX_t + d. \quad (20)$$

While all arguments and classification obtained below apply to both $c < 0$ and $c \geq 0$, we will focus on the later case $c \geq 0$, corresponding to the more natural positive relationship between momentum and average growth rate.

Collecting assumptions (A), (B) and (C) leads us to the following final model setting for the momentum driven price model.

3 Precise model setting and solutions

3.1 Setting

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, \infty)}, \mathbb{P})$ be a filtered probability space with a right-continuous and $\mathbb{P}$ -complete filtration, let $(W_t)_{t \in [0, \infty)}$ be a real valued $(\mathcal{F}_t)_{t \in [0, \infty)}$ -adapted Brownian motion, let $\eta \in [0, 1)$, $\theta^* \in (0, \infty)$, $c \in [0, \infty)$, $d \in (0, \infty)$ be constants, let $f : [0, \infty) \rightarrow \mathbb{R}$, $g : [0, \infty) \rightarrow [0, \infty)$ and $\theta : [0, \infty) \rightarrow (0, \infty)$ be functions given

²This description reflects information provided by professionals through many private communications.

by

$$f(x) = cx + d, \quad (21)$$

$$g(x) = \sqrt{\frac{2x}{1-\eta}}, \quad (22)$$

$$\theta(x) = \theta^* + \eta x, \quad (23)$$

and $X_0, \underline{S} \in (0, \infty)$. Then, following the discussion in section 2 above, there exists³ a unique two-dimensional diffusion $(S_t, X_t)_{t \in [0, \infty)}$, describing a stock price S and its momentum X , with the property that

$$dX_t = -\theta(X_t)X_t dt + \theta(X_t) \frac{dS_t}{S_t} \quad \text{and} \quad X(t=0) = X_0, \quad (24)$$

$$\frac{dS_t}{S_t} = f(X_t)dt + g(X_t)dW_t, \quad (25)$$

$$S_t = \underline{S} \left(1 + \frac{\eta}{\theta^*} X_t\right)^{\frac{1}{\eta}}. \quad (26)$$

3.2 Classification based on boundary behavior

The setting above allows for diverse behavior of the stock price S and its momentum X . To further differentiate this, we need some terminology.

Definition 1. Let the setting in section 3.1 be fulfilled and $X : [0, \infty) \times \Omega \rightarrow [0, \infty)$ be a strong solution of the SDE (24) with initial value $X_0 \in (0, \infty)$ up to its explosion time $\tau = \inf\{t \in (0, \infty] | S_t = \infty\}$. Then X is called

- (a) *explosive*, if $\mathbb{P}[\tau < \infty] > 0$,
- (b) *diffusive-transient*,⁴ if $\mathbb{P}[\tau = \infty] = 1$ and $\lim_{t \rightarrow \infty} X_t = \infty$,
- (c) *recurrent*, if $\forall a \in (0, X_0)$ it holds that $\mathbb{P}[X_t = a \text{ for some } t \in (0, \infty)] = 1$,
- (d) *strictly positive*, if $\mathbb{P}[X_t > 0 \text{ for all } t \in (0, \infty)] = 1$,
- (e) *instantaneously reflected at 0*, if $\mathbb{P}[X_t = 0 \text{ for some } t \in (0, \infty)] > 0$ and for all $t, \epsilon \in (0, \infty)$ we have $\mathbb{P}[X \mathbb{1}_{[t, t+\epsilon]} \equiv 0] = 0$.

The notions (a)-(c) describe the behavior of X at the upper boundary infinity, while the notions (d)-(e) describe the behavior at the lower boundary 0. We will see below that the full parameter range of section 3.1 can be covered with these descriptions.

3.2.1 The case $\eta = 0$.

In this special case, the characteristic time-horizon used by the representative investor to gauge momentum is constant, $\theta_t = \theta^*$, and the governing equation (24) for the momentum X reads

$$dX_t = \theta^* (X_t (c - 1) + d) dt + \theta^* \sqrt{2X_t} dW_t, \quad (27)$$

and

$$S_t = \lim_{\eta \rightarrow 0} \underline{S} \left(1 + \frac{\eta}{\theta^*} X_t\right)^{\frac{1}{\eta}} = \underline{S} \exp\left(\frac{X_t}{\theta^*}\right), \quad (28)$$

Using theorems 5.1, 5.5 (for the case $c = 1$) and theorems 5.3, 5.7 (for the case $c \neq 1$) in Cherny and Engelbert [2005], the behavior of X can be distinguished by

³Up to a random time, at which the processes may explode. For (pathwise) uniqueness and (strong) existence of solutions see, e.g., theorem 3.1 and 3.2 in [Ikeda and Watanabe, 1981]. For details on explosion see comments below.

⁴To avoid confusion with the term *transient* as used in economics for temporary phenomena, we explicitly add *diffusive*.

1. X is diffusive-transient for $c \in [1, \infty)$ and $c = 1 \wedge d > \theta^*$
2. X is recurrent for $c \in [0, 1)$ and $c = 1 \wedge d \in (0, \theta^*]$
3. X is strictly positive for $d \geq \theta^*$ and instantaneously reflected at 0 if $d \in (0, \theta^*)$.

Comments and classification

- In the recurrent case, as (i) there is no positive feedback of increasing momentum on the momentum time-horizon, and (ii) the momentum is roughly mean-reversal with shrinking volatility, we will refer to this case as the **non-bubble regime**. Indeed, here X is equal to the well known CIR model [Cox et al., 1985], which can thus be motivated as the dynamics of a variable X that averages over a fixed memory length the past realizations of an underlying process S .
- In the diffusive-transient case, eq. (27) suggests that X has a stochastic exponential growth rate bounded from below by $\theta^*(c - 1)$ and with shrinking volatility. In particular, ignoring the stochastic term, for large S we have $S_t \sim \exp(\exp(\theta^*(c - 1) t))$, and S grows double-exponentially. As S shows super-exponential growth but is not explosive, we will refer to this case as **mild bubble regime**.

3.2.2 The case $\eta \in (0, 1)$.

Here, the characteristic time-horizon of the representative investor is an decreasing function of the momentum and equation (24) reads

$$dX_t = (\theta^* + \eta X_t) (X_t (c - 1) + d) dt + (\theta^* + \eta X_t) \sqrt{\frac{2X_t}{1 - \eta}} dW_t. \quad (29)$$

Lemma 2 in the appendix shows that

1. X is explosive for $c \in \left(\frac{1}{1 - \eta}, \infty\right)$ and recurrent for $c \in \left[0, \frac{1}{1 - \eta}\right]$,
2. X is strictly positive for $d \in \left[\frac{\theta^*}{1 - \eta}, \infty\right)$ and instantaneously reflected at 0 if $d \in \left(0, \frac{\theta^*}{1 - \eta}\right)$.

As the self-referential influence of the momentum on its own time horizon creates the possibility of large hikes of X (and thus S), we will refer to both these cases as the **bubble regimes**.

We can make the following further distinction for two kinds of bubbles, given $\eta \in (0, 1)$.

Comments and classification

- In the recurrent case $c \in \left[0, \frac{1}{1 - \eta}\right]$, X will experience periodical rallies with rapid growth followed by sharp declines. S , the exponential of X , will thus experience similar but steeper rallies that mimic the pattern of bubbles. For details see section 3.3.2 below. We refer to this case as **recurrent bubble regime**.
- In the explosive case $c \in \left(\frac{1}{1 - \eta}, \infty\right)$, X and S both experience super-exponential growth, faster than double exponential, with a stochastic finite-time-singularity. We refer to this case as the **wild bubble regime**.

3.2.3 Interpretation of model parameters

Based on the model derivation in section 2, let us discuss the model parameters (θ^*, η, c, d) and their interpretation, before we tackle the calibration of the model.

1. The parameter $1/\theta^*$ determines the characteristic time-horizon of the exponential moving average momentum in the regime $\eta = 0$, cf. section 2.1. For $\eta \in (0, 1)$, θ^* sets the scale $\sim \theta^*/\eta$ for the minimum momentum that can influence the time horizon with which momentum is estimated by investors.

2. The parameter η represents the impact per unit of momentum on the time-horizon that is used to measure momentum. The term $\frac{\theta^*}{\eta}$ is the characteristic momentum necessary to influence the time horizon θ in the bubble regime: ranging from $\sim \infty$ for $\eta \sim 0$ (no feedback of momentum on the time horizon) to θ^* for $\eta \sim 1$ (small momentum creates considerable feedback on time horizon). Consistent with our view of bubbles being characterised by the influence of momentum on its on time-scales, η can be understood as a *bubblieness* parameter.
3. The parameter c quantifies the linear dependence of stock returns on unit momentum. For given $\eta \in [0, 1)$, it allows one to distinguish the four main regimes discussed in sections 3.2.1 and 3.2.2 above, *non-* and *mild*, *recurrent*, or *wild* bubble. As such, c , together with η , determines the bubble type.
4. The parameter d represents an intrinsic return independent of momentum and is thus not related to any notion of bubblieness. However, it determines the nature of price or momentum dynamics near zero, in particular whether the momentum can vanish or not. As such, it has a considerable influence on the stylized dynamics of the process.

Figure 1: Different types of bubbles distinguished by the relationship between η and c

3.3 Nondimensionalization

Below we investigate the role of different parameters in our model and propose a reparameterization yielding a standardized set of parameters, which will simplify calibration and further discussion.

3.3.1 Change of time horizon

Here, we conduct a time change using the characteristic time horizon of the momentum, $1/\theta^*$. This shows how X can be derived from a *time*-dimensionless process Y and reduces the characteristic parameters (θ^*, η, c, d) to (η, c, d) .

Let the setting in section 3.1 be fulfilled and (S, X) the corresponding diffusion, additionally let $d^Y \in (0, \infty)$ be a constant and let $f^Y : [0, \infty) \rightarrow \mathbb{R}$ and $\theta^Y : [0, \infty) \rightarrow (0, \infty)$ be functions given by

$$f^Y(y) = cy + d^Y, \quad (30)$$

$$\theta^Y(y) = 1 + \eta y. \quad (31)$$

As above, there exists a two-dimensional Itô process $(S_t^Y, Y_t)_{t \in [0, \infty)}$ with initial value (S_0, X_0) , describing a stock price S^Y and its momentum Y , with the property that

$$dY_t = -\theta^Y(Y_t)Y_t dt + \theta^Y(Y_t) \frac{dS_t^Y}{S_t^Y}, \quad (32)$$

$$\frac{dS_t^Y}{S_t^Y} = f^Y(Y_t)dt + g(Y_t)dW_t, \quad (33)$$

$$S_t^Y = \underline{S} (1 + \eta Y_t)^{\frac{1}{\eta}}. \quad (34)$$

Then we have the following result.

Proposition 1. *Let the setting above be fulfilled and assume $d = d^Y \theta^*$. Then it holds (in the sense of pathwise uniqueness) that*

$$\theta^* Y_{t\theta^*} = X_t \quad \text{and} \quad S_{t\theta^*}^Y = S_t. \quad (35)$$

The proposition follows immediately from corollary 3 in Øksendal [1990], which shows equivalence in distribution, and the fact that the involved stochastic differential equations have strong solutions. We can conclude that the parameter θ^* acts as a time change in our model and the behavior of X can be fully characterized in the reduced parameter space (η, c, d) .

3.3.2 Reparameterization

Following our classification of bubbles based on boundary behavior in section 3.2 above, we introduce two reduced parameters $C \in [0, \infty)$, $D \in (0, \infty)$ as

$$C = c(1 - \eta), \quad D = \frac{d(1 - \eta)}{\theta^*}. \quad (36)$$

As mentioned in subsection 2.5, we focus on the case $c \in [0, \infty)$ and thus $C \in [0, \infty)$ corresponding to the normal relationship between momentum and average growth rate. By plugging (21), (22) into (24) and (25), the dynamics of X and S can be rewritten as follows. For initial values $X(t = 0) = X_0$ and $S(t = 0) = \underline{S} \left(1 + \frac{\eta}{\theta^*} X_0\right)^{\frac{1}{\eta}}$, we have

- $\eta = 0$:

$$dX_t = [(C - 1)\theta^* X_t + D\theta^{*2}] dt + \theta^* \sqrt{2X_t} dW_t, \quad (37)$$

$$dS_t = S_t \theta^* (C \ln S_t + D) dt + S_t \sqrt{2\theta^* \ln S_t} dW_t. \quad (38)$$

- $\eta \in (0, 1)$:

$$dX_t = (\theta^* + \eta X_t) \left(\frac{C - 1 + \eta}{1 - \eta} X_t + \frac{D\theta^*}{1 - \eta} \right) dt + (\theta^* + \eta X_t) \sqrt{\frac{2X_t}{1 - \eta}} dW_t \quad (39)$$

$$dS_t = \frac{\theta^*}{\eta(1 - \eta)} [C S^{\eta+1} + (D\eta - C)S_t] dt + \sqrt{\frac{2\theta^*}{(1 - \eta)\eta}} \cdot \sqrt{S_t^{\eta+2} - S_t^2} dW_t. \quad (40)$$

This reparameterization allows for a standardized categorization of bubble types in the parameter space (C, D) across different values of η . Fig. 2 illustrates these categories.

Figure 2: Bubble classification using the standardized parameters C and D . *FTS* refers to a finite-time singularity (explosive). The term *Ghost FTS* refers to transient explosive behavior, see section 3.3.3 below for details.

Note that, for fixed parameters θ^* , $C \in [0, \infty)$, $D \in (0, \infty)$, the growth rate of a bubble increases with $\eta \in [0, 1)$. This can be seen directly from the drift terms of eq.(38) and eq.(40), noting that the function $h(\eta) = \frac{S^\eta - 1}{\eta(1-\eta)}$ is monotonically increasing for $S > 0$. This justifies the use of η as a measure of *bubbliness* and confirms the terminology introduced in section 3.2,

- *wild* bubbles as opposed to *mild* bubbles for $C \in (1, \infty)$ and
- *recurrent* bubbles as opposed to *non*-bubbles for $C \in (0, 1)$.

Below we will show that recurrent bubbles ($\eta \in (0, 1)$, $C \in [0, 1]$) experience volatility-driven so-called *ghost-finite time singularities* followed by sharp declines. This means that both bubble types with $\eta \in (0, 1)$ show a certain type of explosive behaviour, which motivates us to test for the hypothesis of $\eta = 0$ in bubble detection, see section 4.4.

3.3.3 Recurrent bubbles and ghost-finite time singularities

The recurrent bubble case, $\eta \in (0, 1)$ and $C \in [0, 1]$, is characterised by solutions that are reminiscent of the so-called *ghost finite-time-singularity* introduced in Smug et al. [2018]. Indeed, the explosive volatility term leads to transiently explosive growth phases followed by decline phases, these two regimes alternating and repeating themselves endlessly. Appendix B presents a detailed explanation and description of the properties of the recurrent bubble case in the light of the ghost finite-time-singularity.

3.4 Simulation of price trajectories

In order to illustrate the behavior of solutions for different parameters as discussed above, we present several simulations. Each set of simulations is organized in four subfigures, constructed by fixing two of the parameters (η, C, D) and plotting trajectories varying the third parameter. Each set uses a single random seed for all paths, identical observation length of 1000, same initial values $X_0 = 0.001$ and $\underline{S} = 1$ and the same

discrete time increment $\Delta t = 0.01$ in the simulation algorithm. Based on the discussion in section 3.3.1, without loss of generality we fix the time-horizon θ^* to be 0.002.

Fig. 3 shows non-bubble ($\eta = 0, C = 0.5$) and mild-bubble ($\eta = 0, C = 1.5$) regimes with positive momentum ($D = 1.5$) and non-negative momentum ($D = 0.5$) for each regime. The trajectories in the top graphs experience recurrent fluctuations while the two trajectories in the bottom graphs show persistent super-exponential growth.

Figure 3: Simulated price time series for S with fixed C, D and $\eta = 0$. Note the different vertical scales in the top compared with bottom graphs.

Fig. 4 demonstrate the four types of bubbles classified in fig. 2 (b) with nine simulated trajectories corresponding to different sizes of η . As shown, larger values of η generate stronger bubble for both type of recurrent bubble and wild bubble, indicated by the higher volatility and shorter life for the bubble to be fully pumped up to infinity.

Figure 4: Simulated price time series for S with fixed C, D and $\eta \in [0.1, 0.9]$. Note that (increasing) η represents (increasing) *bubbliness* strength, as θ^*/η is the characteristic momentum necessary to influence the time horizon $1/\theta$. Note also the different vertical scales in the top compared with bottom graphs.

Fig. 5 shows how the shape of bubbles changes for different values of C . The top subfigures show simulated trajectories of bubbles with non-negative momentum, while the bottom subfigures show simulated trajectories of bubbles with strictly positive momentum. The simulated bubbles on the right are stronger (more explosive) than the ones in the left, due to the larger size of η . A distinct phase transition happens for C less than 1 to larger than 1, moving from recurrent to explosive behavior. For larger values of D , the bubble characteristics are more pronounced. Larger values of η , besides earlier explosion times, seems to increase the number of rallies on a given time interval.

Figure 5: Simulated price time series for S with fixed D, η and $C \in [0, 2]$. Note that X (and thus S) is explosive for $C \in (1, \infty)$. Note the different vertical scales in the top compared with bottom graphs.

Fig. 6 shows how the shape of bubble changes with different values of D . The top subfigures show simulated trajectories of recurrent bubbles, while the bottom subfigures show simulated trajectories of wild bubbles. The simulated bubbles on the right are stronger (more explosive) than ones in the left, due to the larger size of η . There is no distinct phase transition of bubble shape around $D = 1$. However, larger D allows for larger relative price levels of S . We can also confirm the observation that larger values of η lead to shorter time intervals between distinct rallies, as mentioned for fig. 5.

Figure 6: Simulated price time series for S with fixed C, η and $D \in [0, 2]$. Note that X is strictly positive (and thus S strictly larger than its lower bound $\underline{S}$) for $D \in [1, \infty)$.

4 Statistical inference and calibration for the bubble model

4.1 Formulation of the quasi-likelihood calibration method

In the present study, we adopt the maximum quasi-likelihood method to calibrate the model, partly because there is no available closed-form solution for the transition density of our bubble model. As the dynamics specification for $\eta = 0$ is different from that for $\eta \in (0, 1)$, these two cases are treated separately.

- Conditional on $\eta = 0$, the dynamics of S_t follows eq.(38). The Euler discretisation for such SDE reads

$$S_{t+\Delta t} - S_t = S_t \theta^* (C \ln S_t + D) \Delta t + S_t \sqrt{2 \theta^* \ln S_t} (W_{t+\Delta t} - W_t). \quad (41)$$

As $W_{t+\Delta t} - W_t \sim \mathcal{N}(0, \Delta t)$, the approximate transition densities in this case is

$$S_{t+\Delta t} | S_t \sim \mathcal{N}(S_t [1 + \theta^* (C \ln S_t + D)] \Delta t, 2 \theta^* S_t^2 \ln S_t \Delta t). \quad (42)$$

Therefore, the conditional log quasi-likelihood for a series of observations $\{S_t\}_{t=0}^T$ can be expressed as

$$\begin{aligned} \ln \mathcal{L}_\Theta &= \sum_{t=0}^{T-1} \ln \Pr^\Theta (S_{t+1} | S_t) \\ &= \sum_{t=0}^{T-1} -\frac{1}{2} \ln 2\pi - \ln \sqrt{2 \theta^* S_t^2 \ln S_t} - \frac{[S_{t+1} - S_t - S_t \theta^* (C \ln S_t + D)]^2}{4 \theta^* S_t^2 \ln S_t}. \end{aligned} \quad (43)$$

Thereafter, we denote $\Theta^{ne} = \{\theta^*, C, D\}$ and re-express eq.(43) as $\ln \mathcal{L}_{\Theta^{ne}}^{ne}$. The superscript ne indicate the case where the dynamics of S_t and X_t is of a non-explosive nature ($\eta = 0$).

- Conditional on $\eta \in (0, 1)$, the dynamics of S_t follows eq.(40). In this case, the Euler discretisation for such SDE can be written as

$$S_{t+\Delta t} - S_t = \frac{\theta^*}{\eta(1-\eta)} [C S^{\eta+1} + (D\eta - C)S_t] \Delta t + \sqrt{\frac{2\theta^*}{(1-\eta)\eta}} \cdot \sqrt{S_t^{\eta+2} - S_t^2} (W_{t+\Delta t} - W_t). \quad (44)$$

As $W_{t+\Delta t} - W_t \sim \mathcal{N}(0, \Delta t)$, the approximate transition densities in this case is given by

$$S_{t+\Delta t} | S_t \sim \mathcal{N}\left(S_t + \frac{\theta^*}{\eta(1-\eta)} [C S^{\eta+1} + (D\eta - C)S_t] \Delta t, \frac{2\theta^*(S_t^{\eta+2} - S_t^2)}{\eta(1-\eta)} \Delta t\right). \quad (45)$$

Therefore, the conditional log quasi-likelihood for a series of observations $\{S_t\}_{t=0}^T$ can be expressed as ($\Delta t = 1$)

$$\begin{aligned} \ln \mathcal{L}_\Theta &= \sum_{t=0}^{T-1} \ln \Pr^\Theta(S_{t+1} | S_t) \\ &= \sum_{t=0}^{T-1} -\frac{1}{2} \ln 2\pi - \frac{1}{2} \ln \frac{2\theta^*(S_t^{\eta+2} - S_t^2)}{\eta(1-\eta)} \\ &\quad - \frac{[\eta(1-\eta)(S_{t+1} - S_t) - \theta^*(C S^{\eta+1} + (D\eta - C)S_t)]^2}{4\theta^*\eta(1-\eta)(S_t^{\eta+2} - S_t^2)}. \end{aligned} \quad (46)$$

To distinguish from the case $\eta = 0$, we use $\Theta^e = \{\theta^*, \eta, C, D\}$ for the set of parameters and re-express eq.(46) as $\ln \mathcal{L}_{\Theta^e}^e$. The superscript e refers to the possibility that the dynamics of S_t and X_t can be explosive with the existence of a (ghost) finite-time-singularity.

The conditional quasi-likelihood maximization problem to determine the optimal parameters in each case can be written respectively

$$\max_{\Theta^{ne}} \ln \mathcal{L}_{\Theta^{ne}}^{ne} = \max_{\theta^*, C, D} \ln \mathcal{L}^{ne}(\{S_t\}_{t=0}^{t=T}, \theta^*, C, D) \quad \text{conditional on } \eta = 0 \quad (47)$$

$$\max_{\Theta^e} \ln \mathcal{L}_{\Theta^e}^e = \max_{\theta^*, \eta, C, D} \ln \mathcal{L}^e(\{S_t\}_{t=0}^{t=T}, \theta^*, \eta, C, D) \quad \text{conditional on } \eta \in (0, 1). \quad (48)$$

The corresponding variance for each estimated parameter can be obtained from the diagonal terms of the following Fisher information matrix

$$\text{Var}(\hat{\Theta}^{ne}) = \left(\frac{\partial^2 \mathcal{L}^{ne}}{\partial \Theta^{ne} \partial \Theta^{ne}} \right)^{-1} \Big|_{\Theta^{ne} = \hat{\Theta}^{ne}}, \quad \text{conditional on } \eta = 0 \quad (49)$$

$$\text{Var}(\hat{\Theta}^e) = \left(\frac{\partial^2 \mathcal{L}^e}{\partial \Theta^e \partial \Theta^e} \right)^{-1} \Big|_{\Theta^e = \hat{\Theta}^e}, \quad \text{conditional on } \eta \in (0, 1). \quad (50)$$

Given a financial time series of daily close prices $\{S_t\}_{t=0}^{t=T}$, we use the L-BFGS-B method to optimize the log-likelihood. The limited-memory BFGS optimization algorithm belongs to the family of quasi-Newton methods that approximates the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm using a limited amount of computer memory. The L-BFGS-B extends L-BFGS to handle simple bound constraints in the search of the optimal parameters [Byrd et al., 1995].

4.2 Calibration of synthetic price time series

In order to employ the L-BFGS-B optimization algorithm and taking into account computation efficiency, we explicitly assign a prior searching space for all four parameters $\{\theta^*, \eta, C, D\}$ before performing their estimation.

- For parameter θ^* , we assign its search to be in the interval $[10^{-8}, 0.004]$. Recall that θ^* has two roles. In the case $\eta = 0$, it is the inverse time horizon over which the momentum is supposed to be evaluated by traders. For $\eta \in (0, 1)$, it is the characteristic momentum value above which a given realised momentum has an impact on the time horizon over which it is itself estimated (see expression (23) with (24)). Thus, the lower bound encompasses the cases where there is essentially an almost infinite memory (compared to the longest existing historical financial time series) of how past price changes influence momentum. 10^{-8} corresponds approximately to 10 million days $\approx$ 400 thousand years. It also allows for a full feedback of momentum on its time horizon. For the upper bound, we are informed by a large body of literature mentioning that momentum trading typically make profit when using times scales from one to five years (See [He and Li, 2015; Moskowitz et al., 2012] and the literature therein). We thus assign the upper search bound for θ^* to be $\frac{1}{250} = 0.004$, where one time unit corresponds to one trading day.
- For η , we search in what is essentially the whole possible interval $[10^{-8}, 1 - 10^{-8}]$, with the exclusion of tiny portions at the two boundaries for the sake of computational precision. We indeed want to avoid the divergence of the log likelihood function $\ln \mathcal{L}^e$ at the lower boundary $\eta = 0$.
- The search interval for parameter C is chosen to be $[0, 5]$. First, the condition that C is non-negative just means that a positive momentum tends to self-perpetuate itself (contrarian behaviors are excluded from the model in the present work). Since C has the meaning of an expected return per unit of momentum adjusted by the bubblieness η , the upper bound 5 is likely to be more than large enough: given some momentum, which reflects the belief by investors on the average price trend over a given time horizon modified by the inverse of the bubblieness $1 - \eta$, the coefficient C transforms this momentum into an actual average price that is $C \frac{1}{1 - \eta}$ times larger over the same time horizon; it is thus reasonable that C should be larger than a few units. Thus the upper bound of 5 is largely sufficient.
- Last, the search interval for parameter D is chosen to be $[0, 500]$. Given that $D = \frac{d(1 - \eta)}{\theta^*}$, the upper bound D is amply sufficient as it would correspond to more than 100% return in one time unit (one day in the subsequent synthetic time series and in the real life applications below) for momentum trading strategy with a typical time-horizon memory of $\frac{1}{0.002} = 200$ days during the bubble with only small bubblieness.

We test our calibration algorithm on synthetic simulated price time series. For $\eta = 0$ and $\eta \in (0, 1)$, tables 1 and 2 respectively list the true parameters used to generate the price time series, the values of calibrated parameters, as well as the values of the logarithmic quasi-likelihood calculated for each simulated series with those parameters. Table 1 gives the results for 4 synthetic time series shown in fig. 3. Table 2 gives the results for 8 synthetic time series with positive bubblieness ($\eta \in (0, 1)$) shown in fig.4, which cover the different regimes for C and D and contrasts the cases $\eta = 0$ and $\eta \in (0, 1)$.

C	$\hat{C}$	D	$\hat{D}$	θ^*	$\hat{\theta}^*$	$-2 \ln \mathcal{L}^{ne}$	$-2 \ln \widehat{\mathcal{L}}^{ne}$
0.5	1.003 (0.905)	0.5	1.038 (0.702)	0.002	0.0021 (<0.0001)	358.6	355.9
0.5	1.068 (0.723)	1.5	1.058 (1.297)	0.002	0.0021 (<0.0001)	3842	3840
1.5	1.284 (0.587)	0.5	1.119 (0.917)	0.002	0.0022 (<0.0001)	4696	4693
1.5	1.864 (0.428)	1.5	1.372 (1.282)	0.002	0.0024 (<0.0001)	10640	10639

Table 1: Estimation of the 4 synthetic series shown in figure 3 with $\eta = 0$. The true parameter values used to generate the synthetic time series are indicated without the hat $\hat{\cdot}$. The estimated parameters are shown with the $\hat{\cdot}$. The values in brackets below the estimated values give the corresponding standard errors.

η	$\hat{\eta}$	C	$\hat{C}$	D	$\hat{D}$	θ^*	$\hat{\theta}^*$	$-2 \ln \mathcal{L}^{ne}$	$-2 \ln \widehat{\mathcal{L}}^{ne}$
0.9	0.936 (0.005)	0.5	0.194 (0.168)	0.5	0.127 (0.008)	0.002	0.004 (0.0002)	769.5	355.3
0.9	0.873 (0.010)	0.5	0.319 (0.171)	1.5	0.923 (0.186)	0.002	0.004 (0.0003)	2094	1996
0.9	0.942 (0.009)	1.5	0.694 (0.470)	0.5	0.084 (0.013)	0.002	0.004 (0.0005)	247.3	187.9
0.9	0.893 (0.019)	1.5	0.499 (0.470)	1.5	0.750 (0.358)	0.002	0.004 (0.0007)	608.1	564.7
0.5	0.224 (0.05)	0.5	0.294 (0.849)	0.5	0.140 (0.018)	0.002	0.004 (0.0003)	2760	2472
0.5	0.464 (0.019)	0.5	0.910 (0.543)	1.5	1.109 (0.547)	0.002	0.002 (<0.0001)	479.5	480.8
0.5	0.384 (0.022)	1.5	0.989 (0.228)	0.5	0.098 (0.029)	0.002	0.004 (0.0002)	3357	3243
0.5	0.588 (0.020)	1.5	1.703 (0.417)	1.5	1.224 (0.581)	0.002	0.002 (<0.0001)	229.0	109.0

Table 2: Estimation of 8 synthetic series shown in fig. 4. The true parameter values used to generate the synthetic time series are indicated without the hat $\hat{\cdot}$. The estimated parameters are shown with the $\hat{\cdot}$. The values in brackets below the estimated values give the corresponding standard errors.

From the above two tables, we can draw the following conclusions concerning the performance of our conditional quasi-likelihood estimation method.

- The estimation of θ^* conditional on $\eta = 0$ (non-bubble or mild-bubble cases) is quite accurate. The estimation of θ^* conditional on $\eta \in (0, 1)$ (recurrent or wild bubbles) is less precise but still in a reasonable range. In particular, for moderate bubblieness ($\eta = 0.5$) and strictly positive momenta ($D \in [1, \infty)$), the estimation of θ^* is good. For other bubble types, η is overestimated. It compensates the underestimation

of both C and D , at the cost of an overestimation of the volatility (recall eq.(38) and eq. (40)), as we will discuss soon.

- The estimation of η looks quite reliable for both recurrent or wild bubbles, except for the cases with $\eta = 0.5$ and $D \in (0, 1)$, for which η are underestimated. Nonetheless, the estimator is able to identify correctly the range of the bubblieness, as the estimations of true values $\eta = 0.5$ are all well separated from those with true values $\eta = 0.9$.
- The estimations of C are quite far from their true values for all bubble types. Moreover, the standard deviation of these estimations are quite large, so that the estimations of C fall within two standard deviations of the true value. In other words, given the rather large standard deviations, the estimated values can be deemed compatible with the true values. It is encouraging to observe that the larger the true value of C , the largest its estimator, inter alia. The source of the discrepancy between estimated and true values of C is likely to be found in the Euler discretization scheme, which has used a time step $\Delta t = 1$ to match the daily observation scale. With this choice $\Delta t = 1$, we are quite far from the consistency condition that the Euler discretization converges in the limit $\Delta t \rightarrow 0$. For the stronger bubbles ($\eta \in (0, 1)$), the parameters are underestimated (except for $\eta = 0.5$ and $D \in (0, 1)$ for which the parameters are slightly overestimated), which leads to an underestimation of the real strength of the positive feedbacks between momentum and time scale. Actually, as the drift and diffusion coefficients have θ^* as a common parameter, the quasi-likelihood approach tend to give more weight on such a shared parameter, which thus leads to an underestimation of C .
- For the non-bubble and mild-bubble cases, the estimations of D fall within two standard deviations of the true value. As for C , the larger the true value of C , the largest its estimator, inter alia. The origin of the discrepancy between estimated and true values of D is also likely to be found in the Euler discretization scheme, as discussed for C . For the stronger bubbles, the estimation biases are approximately the same, around -0.4 for all cases.

In order to investigate in detail the performance of this estimation method based on Euler quasi-likelihood approximation, the empirical density of the estimated value of each parameter is created by repeating the estimation procedure on an ensemble of 200 synthetic time series generated with fixed specific model parameters but different random seeds. The following figures 7 -10 summarise the results. The solid line shows the empirical density function, and the vertical dashed line gives the empirical first moment (statistical mean value).

Figure 7 confirms the relatively good performance of the estimation of η , with the empirical density function of $\hat{\eta}$ peaking close the true value. However, the empirical density function is quite broad, exhibiting a significant half-width of about $0.1 - 0.2$, which is much larger than the standard error $0.01 \sim 0.03$ estimated on each calibration using the Fisher information matrix, as reported in figure 8. Moreover, one can observe a significant skewness and a fat-tailed structure of the density functions. This reflects the strong sensitivity of the estimation of η to the specific realisation of the stochastic innovations associated with a given price path trajectory, such that the ensemble standard deviation is much larger than the single realisation standard deviation. This phenomenon is typical of time series without ergodicity, which is the case here given the relatively small sizes of time windows and the nature of the bubble dynamics in our model.

(a) $C = 0.5, \eta = 0.5$

(b) $C = 0.5, \eta = 0.9$

(c) $C = 1.5, \eta = 0.5$

(d) $C = 1.5, \eta = 0.9$

Figure 7: Empirical density functions of estimated η on an ensemble of 200 synthetic time series with different random seeds, generated with fixed specific model parameters $D = 1.5$, $\theta^* = 0.002$ and C indicated below each graph. The vertical dashed line gives the empirical first moment (statistical mean value).

Figure 8: Empirical density functions of estimated standard deviations of η obtained from the Fisher information matrix for each fit within an ensemble of 200 synthetic time series with different random seeds, generated with fixed specific model parameters $D = 1.5$, $\theta^* = 0.002$ and C indicated below each graph. The vertical dashed line gives the empirical first moment (statistical mean value) of the estimated standard deviations of η .

Figure 9 shows that smaller (resp. larger) values of bubblieness tend to yield an over-estimation (resp. under-estimation) of C . The shown empirical density functions of C over an ensemble of 200 synthetic time series are quite broad with half-widths comparable to the Fisher matrix based estimation of the standard deviation for each individual time series, as illustrated by the empirical density functions of C , corrected, respectively, by $+$ and $-$ two times the standard deviation of C obtained from the estimated Fisher matrix for each calibration. Figure 10 shows that the modes and means of the empirical density functions of D are smaller than the true value for all shown values of C and η parameters. The dispersion of the estimated D is also quite broad and similar across the ensemble of synthetic time series, as estimated for each time series with the Fisher information matrix. In all cases, the estimated values of both C and D fall within two standard deviations of their true values.

4.3 Binary classification of recurrent ($C < 1$) versus wild ($C > 1$) bubbles

Given the biases and broad statistical distributions of estimated parameters C for $\eta \in (0, 1)$, we seek to reliably characterise a given time series as being either in the recurrent ($C \in (0, 1)$) or wild ($C \in (1, \infty)$) bubble class, using some combination of estimated parameters.

To achieve this, we construct a classifier metric defined by $\hat{C} + h \cdot \hat{\sigma}_C$, where h is a parameter to be determined to ensure maximum sensitivity and specificity of the resulting classification, as defined below. The proposed classification rule is shown in figure 11.

Figure 9: Empirical density functions of C over an ensemble of 200 synthetic time series with different random seeds, generated with fixed model parameters $D = 1.5$, $\theta^* = 0.002$ and C and η as indicated below each graph. As indicated in the graphs, the two other curves are the empirical density functions of C corrected respectively by $+$ and $-$ two times the standard deviation of C obtained from the estimated Fisher matrix for each calibration. For each of the three obtained density functions, the vertical dashed line gives its statistical mean value.

Figure 10: Empirical density functions of D over an ensemble of 200 synthetic time series with different random seeds, generated with the same fixed specific model parameters $D = 1.5$, $\theta^* = 0.002$ and C and η as indicated below each graph. As indicated in the graphs, the two other curves are the empirical density functions of D corrected respectively by $+$ and $-$ two times the standard deviation of D obtained from the estimated Fisher matrix for each calibration. For each of the three obtained density functions, the vertical dashed line gives its statistical mean value.

Figure 11: Classification scheme to decide if a given time series is in the recurrent or wild bubble regime (conditional of $\eta \in (0, 1)$). The parameter h is chosen to optimise sensitivity and specificity properties (see text).

We consider two populations of 200 time series of 1000 time steps each, with 200 different random seeds. Both populations share the same initial momentum and parameters $\theta^* = 0.002, \eta = 0.9, D = 1.5$. The two populations are distinguished by one having $C = 0.5$ (recurrent bubbles) and the other having $C = 1.5$ (wild bubbles).

For each value of h , we estimate the probability $\Pr_h(\text{No}|R)$ that a time series will be qualified as a recurrent bubble (answer ‘No’ to the question in figure 11), given that the time series has been generated in the recurrent bubble regime $C = 0.5 < 1$. $\Pr_h(\text{No}|R)$ is determined as the fraction among the 200 time series generated with $C = 0.5$ that answer ‘No’ to the question in figure 11 for that given h value. Similarly, for each value of h , we estimate the probability $\Pr_h(\text{No}|W)$ that a time series will be qualified as a recurrent bubble (answer ‘No’ to the question in figure 11), given that the time series has been generated in the wild bubble regime $C = 1.5 > 1$. $\Pr_h(\text{No}|W)$ is determined as the fraction among the 200 time series generated with $C = 1.5$ that answer ‘No’ to the question in figure 11 for that given h value.

Since the classification proposed in figure 11 is binary, we have

$$\Pr_h(\text{Yes}|R) = 1 - \Pr_h(\text{No}|R) \quad \text{and} \quad \Pr_h(\text{Yes}|W) = 1 - \Pr_h(\text{No}|W). \quad (51)$$

Knowing $\Pr_h(\text{No}|R)$ and $\Pr_h(\text{No}|W)$, for an unknown time series, the probability $\Pr(R|\text{No})$ that it is a recurrent bubble, given that the answer to the question in figure 11 is ‘No’ is obtained as

$$\Pr_h(R|\text{No}) = \frac{\Pr_h(\text{No}|R) p(R)}{\Pr_h(\text{No}|R) p(R) + \Pr_h(\text{No}|W) p(W)}, \quad (52)$$

where $p(R)$ (resp. $p(W)$) is the unconditional probability that the time series is a recurrent (resp. wild) bubble. The probability $\Pr_h(W|\text{No})$ that it is a wild bubble, given that the answer to the question in figure 11 is ‘No’ is then

$$\Pr_h(W|\text{No}) = 1 - \Pr_h(R|\text{No}), \quad (53)$$

since, for a given answer ‘No’, the state of the time series is binary, either recurrent or wild.

Similarly, using $\Pr_h(\text{Yes}|R)$ and $\Pr_h(\text{Yes}|W)$ given in (51), we can derive $\Pr_h(R|\text{Yes})$, the probability that the time series is in a recurrent bubble, given that the answer to the question in figure 11 is ‘Yes’:

$$\Pr_h(R|\text{Yes}) = \frac{\Pr_h(\text{Yes}|R) p(R)}{\Pr_h(\text{Yes}|R) p(R) + \Pr_h(\text{Yes}|W) p(W)}, \quad (54)$$

where $\Pr_h(\text{Yes}|R)$ and $\Pr_h(\text{Yes}|W)$ are known from (51). The probability $\Pr(W|\text{Yes})$ of being a wild bubble, given that the answer to the question in figure 11 is ‘Yes’, then reads

$$\Pr_h(W|\text{Yes}) = 1 - \Pr_h(R|\text{Yes}), \quad (55)$$

since, for a given answer ‘Yes’, the state of the time series is binary, either recurrent or wild.

As an illustration, assuming that we have no prior information on the relative frequencies of recurrent vs wild bubbles, we assign the no-informative priors $p(R) = p(W) = 1/2$ and expressions (52) and (54) reduce to

$$\Pr_h(R|\text{No}) = 1 - \Pr_h(W|\text{No}) = \frac{\Pr_h(\text{No}|R)}{\Pr_h(\text{No}|R) + \Pr_h(\text{No}|W)}, \quad (56)$$

and

$$\Pr_h(R|\text{Yes}) = 1 - \Pr_h(W|\text{Yes}) = \frac{\Pr_h(\text{Yes}|R)}{\Pr_h(\text{Yes}|R) + \Pr_h(\text{Yes}|W)} . \quad (57)$$

For instance, let us assume $\Pr_h(\text{No}|R) = 0.7$ (true positive rate of 70% for recurrent bubbles) and $\Pr_h(\text{Yes}|R) = 0.3$ (false negative rate of 30% for recurrent bubbles) and $\Pr_h(\text{No}|W) = 0.2$ (false negative rate of 20% for wild bubbles) and $\Pr_h(\text{Yes}|W) = 0.8$ (true positive rate of 80% for wild bubbles). Expressions (56) and (57) yields

$$\Pr_h(R|\text{No}) = 78\% ; \Pr_h(R|\text{Yes}) = 27\% ; \Pr_h(W|\text{No}) = 22\% ; \Pr_h(W|\text{Yes}) = 73\% . \quad (58)$$

Thus, if the answer is ‘No’ and this is the only information we have, we will attribute a probability of 78% that the time series is recurrent and 22% that it is wild. If the answer is ‘Yes’ and this is the only information we have, we will attribute a probability of 73% that the time series is wild and 27% that it is recurrent.

In sum, the complete information of the binary classifier defined in figure 11 for a given h value is contained in just two conditional probabilities $\Pr_h(\text{No}|R)$ and $\Pr_h(\text{No}|W)$, which are plotted as a function of h in figure 12. These two probabilities are obtained as explained above from two populations of 200 time series of 1000 time steps each, with 200 different random seeds, one population being in the recurrent bubble regime, while the other one is in the wild bubble regime. One can observe that the true positive rate $\Pr_h(\text{No}|R)$ for recurrent bubbles is all the larger, the smaller h is. In contrast, the false negative rate $\Pr_h(\text{No}|W)$ for wild bubble decreases with h . This leads to conflicting requirements: to have the largest possible $\Pr_h(\text{No}|R)$, we need h small, while to have the smallest possible $\Pr_h(\text{No}|W)$, we need to have a large h . We can thus expect that the correct identification rates $\Pr_h(R|\text{No})$ and $\Pr_h(W|\text{Yes})$ depends nonlinearly and non-monotonously on h . Indeed, figure 13 shows that choosing $h = 0$ would lead to a perfect classification score $\Pr_h(W|\text{Yes}) = 1$ for the identification of wild bubbles at the cost of a modest performance $\Pr_h(R|\text{No}) = 0.57$ for the identification of recurrent bubbles. Choosing $h = 3.5$ gives the maximum possible score $\Pr_h(R|\text{No}) = 0.8$ for the identification of recurrent bubbles together with a quite honourable $\Pr_h(W|\text{Yes}) = 0.85$ for the identification of a wild bubble.

As h can be varied, the optimal classification scheme applied to an unknown time series is the following. One should first try the value $h = 0$ and check what is the answer to the question in figure 11 for $h = 0$. If it is ‘Yes’, then we know with 100% certainty that the time series is in the wild bubble regime. If the answer is ‘No’ for $h = 0$, one idea is to choose $h = 3.5$ to ask again the question in figure 11. If the answer is ‘No’ with $h = 3.5$, we conclude that the time series is recurrent with 80% probability. If the answer is ‘Yes’ with $h = 3.5$, we deduce that the time series is wild with 85% probability. We can actually do better. If the answer is ‘No’ for $h = 0$, we can scan all the values of h from 0 to 3.5 to check if there is a value $h < 3.5$ for which the answer is ‘Yes’. Since $\Pr_h(W|\text{Yes})$ is monotonically decreasing from 1 for $h = 0$ to 0.85 for $h = 3.5$, the first time we get a ‘Yes’ for the smallest $h < 3.5$ gives us a success probability $\Pr_h(W|\text{Yes})$ larger than 85%. If there is no ‘Yes’ found for $h < 3.5$, the best conclusion is that the time series is recurrent with a probability of 80%.

Figure 12: The two conditional probabilities, $\Pr_h(\text{No}|W)$ (true positive) and $\Pr_h(\text{No}|R)$ (false negative), as a function of h .

Figure 13: The two derived posterior probabilities $\Pr_h(R|\text{No})$ and $\Pr_h(W|\text{Yes})$ given by expressions (52) and (55) as a function of h , with the assumption that both of the unknown priori probabilities $p(R) = p(W) = 1/2$. $\Pr_h(R|\text{No})$ (resp. $\Pr_h(W|\text{Yes})$) is the probability or success score for the correct identification of a recurrent (resp. wild) bubble, given that the answer to the question figure 11 is ‘No’ (resp. ‘Yes’) for a given h .

4.4 Bubble detection based on Wilks’ Test

Wilks’ Theorem [Wilks, 1938] provides a simple and efficient test to compare two nested models using their log-likelihood ratio. For the proposed time-varying momentum horizon bubble model, the bubble type of $\eta = 0$ with dynamics given by (38) is indeed nested into the model characterized by (40) with $\eta \in [0, 1]$. Given a price time series, we can thus subject it to the Wilks’ test with the following null and alternative hypotheses, based on the :

H_0 : $\eta = 0$, no bubble or mild-bubble (without finite-time-singularity).

H_1 : $\eta \in (0, 1)$, recurrent or wild bubble (with ghost- or true finite-time-singularity and likely decline/regime-change).

The Wilks statistics $T_{ne:e}$ is usually defined as

$$T_{ne:e} := -2 \left(\max_{\theta^*, C, D} \ln \mathcal{L}^{ne} - \max_{\eta, \theta^*, C, D} \ln \mathcal{L}^e \right). \quad (59)$$

Under the hypothesis that the null H_0 holds, $T_{ne:e}$ is distributed according to the chi-square distribution with a number of degrees of freedom equal to the number of parameters that needs to be fixed to go from H_1 to H_0 , here 1, since the parameter space for H_0 , i.e., $(\theta^*, C, D) \in \mathbb{R}^{+3}$ lies on a 3-dimensional hyperplane within the 4 dimensional space represented by $(\eta, \theta^*, C, D) \in [0, 1] \otimes \mathbb{R}^{+3}$. However, there is a correction needed to account for the fact that H_0 is recovered from H_1 by fixing η on the boundary 0 of the interval $\eta \in [0, 1]$. In

this case, rather than $T_{ne:e}$ being distributed according to $\chi^2(1)$, the distribution of $T_{ne:e}$ needs to be adjusted (see Chernoff [1954])

$$T_{ne:e} := -2 \left(\max_{\theta^*, C, D} \ln \mathcal{L}^{ne} - \max_{\eta, \theta^*, C, D} \ln \mathcal{L}^e \right) \sim \text{Mixed}(0, \chi^2(1); \frac{1}{2}). \quad (60)$$

where $\text{Mixed}(0, \chi^2(1); \frac{1}{2})$ denote the mixed law such that the statistic is 0 with one chance out of two and follows a $\chi^2(1)$ with the other one chance out of two. Under the null hypothesis H_0 , this statistic T should be distributed according to the above mixed distribution: the number of parameters to fix in the general model (40) to arrive at the particular model (38) in the nested hypothesis testing is 1 (parameter η), while $\eta = 0$ is on the border of $\eta \in [0, 1)$. Thus, the q -quantile for $\chi^2(1)$ is equivalent to the $\frac{1+q}{2}$ -quantile for the distribution of $T_{ne:e}$.

Fig. 14 and fig. 15 illustrate the results obtained by implementing the Wilks' test on synthetic price time series to examine the capability of differentiating the regimes $\eta = 0$ and $\eta \in (0, 1)$. As shown in fig. 14, for all 400 simulated price time series generated with $\eta \in (0, 1)$, the null H_0 ($\eta = 0$) can be rejected with significance level of 2.5%. By and large, for the time series with larger bubblieness (larger η), the null H_0 is easier to reject, regardless of whether they are recurrent or wild bubbles. For bubble with moderate bubblieness (left panel with $\eta = 0.5$), the null H_0 is easier to reject for the wild bubbles (triangle) than for the recurrent bubbles (small circle). For bubbles with strong bubblieness (right panel with $\eta = 0.9$), the easier rejection of the null for wild bubbles than for recurrent bubble can be attributed to the explosive behavior and short duration of some of the wild bubble time series.

As shown in fig. 15 for time series generated with $\eta = 0$, the Wilks' test performs well as there is only two false alarm at the 97.5% confidence level and, for all other time series, the null cannot be rejected.

Figure 14: Wilks' statistics (59) calculated for synthetic price time series generated for two types of strong bubbles, with $\eta \in (0, 1)$. There are 400 time series in total. All time series are generated with the fixed parameter values $\theta^* = 0.002$, $D = 1.5$. For each pair (η, C) , there are 100 synthetic time series to test. The number 3.84 correspond to the 95%-quantile for the $\chi^2(1)$ distribution and indicates the 97.5% confidence level to reject H_0 .

Figure 15: Wilks' statistics (59) calculated on synthetic price time series generated with $\eta = 0$. There are 200 time series in total. All time series are generated with fixed parameter values $\theta^* = 0.002$, $D = 1.5$. For each value of C , there are 100 synthetic timeseries to test. The number 3.84 corresponds to the 95%–quantile for the $\chi^2(1)$ distribution and indicates the 97.5% confidence level to reject H_0 .

We implement the test on two historical financial price trajectories. They are indices constructed with the 505 stock prices that entered in the composition of the S&P 500 stock index from the beginning of 1998 to the end of 2002. The first Internet stock index is an equally weighted portfolio of the firms related to the Internet sector according to GICS. The second “brick and mortar” index is an equally weighted portfolio of the remaining “brick and mortar” firms listed as S&P 500 components. Fig.16 shows that the “brick and mortar” index fluctuated by not more than $\pm 25\%$ over the 5 years. In contrast, the Internet stock index was multiplied by a factor more than 11 (corresponding to a return of more than $> 1100\%$ from 1998 to its peak in the first quarter of 2000). This index then shrunk with a great crash from a value above 150 to below 50 in March-April 2000, followed by a jumpy decay to achieve a level at the end of 2002 below its value in January 1998. The contrast between the behavior of these two indices over the same 5 years horizon cannot be more striking.

Table 3 demonstrates that the Wilks' test can cleanly classify the two indices into two different classes. For the Internet index, H_0 is very strongly rejected, while it is not for the “brick and mortar” stock index.

	$\eta = 0$				$\eta > 0$					$T_{ne:e}$	p -Value
	C	D	θ^*	$-\mathcal{L}^{ne}$	η	C	D	θ^*	$-\mathcal{L}^e$		
Internet	0.5	< 0.01	5.32×10^{-4}	1663.4	0.281	0.719	0.538	1.31×10^{-4}	1567.7	191.5	< 0.001
Non-internet	0.5	4.84	2.06×10^{-5}	235.08	< 0.01	0.442	5.98	2.06×10^{-5}	235.08	0.017	0.948

Table 3: Estimation of the two models $\eta = 0$ and $\eta \in (0, 1)$ and Wilk's test on two stock indices during the dot-com bubble and following crash from 1998 to 2002.

Table 4 is the same as Table 3, with the difference of restricting the time interval from the beginning of 1998 to the end of 1999, i.e. only during the run-up of the bubble and before the crash occurs and following bearish regime develops. This is important to disentangle the impact of the bubble regime that has not yet revealed the underlying finite time singularity from the crash regime in distinguishing the two classes of dynamics. Again,

Figure 16: Logarithm of the price level of the Internet (top panel) and of the “brick and mortar” (bottom panel) stock indices. The time periods goes from the beginning of 1998 to the end of 2002. Note the difference in vertical scales in the two plots.

we find that, for the Internet index, the null hypothesis H_0 is rejected at the 95% confidence level (with a p-value of 0.018). In contrast, the null is not rejected for the “brick and mortar” stock index.

	The class of $\eta = 0$				η	The class of $\eta > 0$				$T_{ne:e}$	p -Value
	C	D	θ^*	$-\mathcal{L}^{ne}$		C	D	θ^*	$-\mathcal{L}^e$		
Internet	0.007	10.54	3.7×10^{-4}	858.9	0.099	0.452	22.5	2.4×10^{-4}	856.7	4.42	0.018
Non-internet	5.210	12.33	1.51×10^{-5}	5.653	< 0.001	4.75	10.54	1.51×10^{-5}	5.648	0.01	0.960

Table 4: Same as table 3 from the beginning of 1998 to the end of 1999.

5 Concluding remarks

In this work, we have introduced a reduced form model of financial bubbles, based on a simple specification of the stochastic dynamics of the price momentum of a financial asset. The model assumes that the price is a deterministic function of the instantaneous momentum, a quantity that constitutes one of the most popular technical indicators powering investment techniques used by hedge-funds, mutual funds and individual investors in general. In our model, the instantaneous momentum was constructed by a simple continuous moving average of past prices over a given time horizon. One key feature is to consider that this time horizon is not fixed but tends to shrink as a bullish market develops, reflecting the race of trend-following agents to forerun their competitors. The feedback of momentum onto its time horizon, namely the stronger the momentum, the shorter the time horizon, leads to several interesting market dynamics that can be classified in terms of an index η of bubblieness and of the dependence of stock returns on momentum. We have identified four main classes of market behavior: non-bubble regime (an exponential of a CIR process), a mild bubble regime characterised by a transient exponential of an exponential growth of the price, a recurrent bubble regime with transient stochastic ghost-singularity and a wild bubble regime with a genuine stochastic finite-time singularity. We have shown how to test for the presence of these different regimes through intensive tests on synthetic time series. Finally, we have presented a bubble detection test based on our model, which we have applied to the Internet bubble regime. This has allowed us to confirm the reasonable performance of our proposed calibration method to qualify the Internet stocks as being described by the recurrent bubble regime compared with the non-Internet stocks well characterised by the non-bubble class.

We end by pointing out two limits of our model, which should be addressed in future papers. First, since only positive momenta are allowed, the model does not describe negative bubbles characterised by downward accelerating prices spiraling towards a bottom followed by a rebound or rally. Second, the model assumes representative momentum investors with a single typical time scale. This should be improved in the future by accounting for the multiple time scales characterising the universe of trading strategies, from high-frequency trading, intra-day trading, to long term investment horizons. Re-introducing these multiple time scales can offer in principle the possibility to account for price jumps.

Acknowledgments We acknowledge financial support from the National Natural Science Funds of China (Grant No. 71771086) and (Grant No. 71301051).

A Proofs

Lemma 1. Assume the setting in section 2.2 and let $\lambda : [0, \infty) \rightarrow [0, \infty)$ be a C^2 -function with strictly positive first and second derivative such that $S_t = \lambda(X_t)$, $t \in [0, \infty)$. Then it holds that

$$\theta_t = \frac{\lambda(X_t)}{\lambda'(X_t)} \quad \text{and} \quad g_t^2 = \frac{2\lambda(X_t)X_t}{\theta_t^2 \lambda''(X_t)}, \quad \text{for } t \in [0, \infty). \quad (61)$$

Proof. Equations (11) and (12) imply that the quadratic variation of X is given by

$$[X]_t = \theta_t^2 g_t^2 dt. \quad (62)$$

Assume that S satisfies $S_t = \lambda(X_t)$ for $t \in [0, \infty)$. Then Itô's formula shows that

$$\begin{aligned} dS_t &= \lambda'(X_t) dX_t + \frac{1}{2} \lambda''(X_t) d[X]_t \\ &= \lambda'(X_t) dX_t + \frac{1}{2} \lambda''(X_t) \theta_t^2 g_t^2 dt \end{aligned} \quad (63)$$

Rearranging equation (12) we get

$$dS_t = \frac{S_t}{\theta_t} dX_t + S_t X_t dt, \quad (64)$$

and comparing the ($\mathbb{P}$ -a.s. unique) Itô-equation terms in (63) and (64), we conclude that

$$\theta_t = \frac{\lambda(X_t)}{\lambda'(X_t)} \quad \text{and} \quad g_t^2 = \frac{2\lambda(X_t)X_t}{\theta_t^2 \lambda''(X_t)}. \quad (65)$$

The proof of lemma 1 is thus completed.

Lemma 2. Assume the setting in section 3.1 and let $\beta < \infty$. Then the process X is given by the SDE

$$dX_t = \left(\theta^* + \frac{X_t}{\beta} \right) (X_t (c - 1) + d) dt + \left(\theta^* + \frac{X_t}{\beta} \right) \sqrt{\frac{2\beta X_t}{\beta - 1}} dW_t \quad (66)$$

and it holds that

- (a) X is explosive for $c \in \left(\frac{\beta}{\beta - 1}, \infty \right)$ and recurrent for $c \in \left[0, \frac{\beta}{\beta - 1} \right]$, and
- (b) X is strictly positive for $d \in \left[\frac{\theta^* \beta}{\beta - 1}, \infty \right)$ and instantaneously reflected at 0 if $d \in \left(0, \frac{\theta^* \beta}{\beta - 1} \right)$.

Proof. The form of X in (66) follows immediately from equations (21)-(25).

For some $a \in (0, \infty)$, define the functions $b : [0, \infty) \rightarrow (-\infty, \infty)$, $\sigma : [0, \infty) \rightarrow [0, \infty)$, $\rho : [a, \infty) \rightarrow [0, \infty]$ and $s : [a, \infty) \rightarrow [0, \infty]$ by

$$b(x) = \frac{c - 1}{\beta} x^2 + \left((c - 1)\theta^* + \frac{d}{\beta} \right) x + d\theta^*, \quad (67)$$

$$\sigma(x) = \sqrt{\frac{2}{\beta(\beta - 1)}} x^{\frac{3}{2}} + \theta^* \sqrt{\frac{2\beta}{\beta - 1}} x^{\frac{1}{2}}, \quad (68)$$

$$\rho(x) = \exp \left(- \int_a^x \frac{2b(y)}{\sigma(y)^2} dy \right), \quad (69)$$

$$s(x) = \int_x^\infty \rho(y) dy. \quad (70)$$

Theorem 4.1 in [Cherny and Engelbert, 2005] shows that X is recurrent if for some $a \in (0, \infty)$ it holds that $s(a) = \infty$. Observe that there exists a constant $K \in (0, \infty)$ such that, for $(c-1)(\beta-1) \leq 1$, we have

$$\int_a^\infty \rho(x) dx \geq K \int_a^\infty \exp\left(-\int_a^x \frac{(c-1)(\beta-1)}{y} dy\right) dx = K \int_a^\infty \left(\frac{a}{x}\right)^{(c-1)(\beta-1)} dx = \infty. \quad (71)$$

This proves that X is recurrent for $c \leq \beta/(\beta-1)$. Theorem 4.3 in [Cherny and Engelbert, 2005] shows that X is explosive if for some $a \in (0, \infty)$ it holds that

$$s(a) < \infty \quad \text{and} \quad \int_a^\infty \frac{s(x)}{\rho(x)\sigma^2(x)} dx < \infty. \quad (72)$$

To see this, we assume that $c > \beta/(\beta-1)$ and choose

$$\epsilon \in \left(0, \min\left\{\frac{2c}{\beta} - \frac{2}{\beta-1}, \frac{\left(\frac{2}{\beta(\beta-1)}\right)^2}{\frac{2(c-1)}{\beta} - \frac{2}{\beta(\beta-1)}}\right\}\right) \quad (73)$$

and $a = a(\epsilon) \in (0, \infty)$ such that

$$\left(\sqrt{\frac{2}{\beta(\beta-1)}} y^{\frac{1}{2}} + \theta^* \sqrt{\frac{2\beta}{\beta-1}}\right)^2 < \left(\epsilon + \frac{2}{\beta(\beta-1)}\right) y^2, \quad \text{for all } y \in [a, \infty). \quad (74)$$

This implies that, for an exponent

$$E = \frac{2(c-1)}{\beta\left(\epsilon + \frac{2}{\beta(\beta-1)}\right)} > 1, \quad (75)$$

it holds that

$$s(a) = \int_a^\infty \rho(x) dx < \int_a^\infty \exp\left(-\int_a^x \frac{2\frac{c-1}{\beta}y}{\left(\epsilon + \frac{2}{\beta(\beta-1)}\right)y^2} dy\right) dx = \int_a^\infty \left(\frac{a}{x}\right)^E dx < \infty. \quad (76)$$

To verify the second part of inequalities (72), we first note that, from equation (76), we get

$$s(x) \leq \int_x^\infty \left(\frac{a}{y}\right)^E dx = \frac{a^E}{E-1} \frac{1}{x^{E-1}}, \quad (77)$$

and thus, in combination with the inequality on $\rho(x)$ used in equation (71), we can conclude that there exists a constant $C' \in (0, \infty)$ such that

$$\int_a^\infty \frac{s(x)}{\rho(x)\sigma^2(x)} dx \leq C' \int_a^\infty \frac{1}{x^{E-1}} \frac{1}{x^3} x^{(c-1)(\beta-1)} dx, \quad (78)$$

which is finite, as by choice of ϵ in (73) we have

$$E + 2 - (c-1)(\beta-1) > 1. \quad (79)$$

The prove of claim (a) is thus completed.

To analyze the behavior of X at 0, define, for $a \in (0, \infty)$, the functions $\tilde{\rho}: [a, \infty) \rightarrow [0, \infty]$ and $\tilde{s}: [a, \infty) \rightarrow [0, \infty]$ by

$$\tilde{\rho}(x) = \exp\left(\int_x^a \frac{2b(y)}{\sigma(y)^2} dy\right), \quad (80)$$

$$\tilde{s}(x) = \int_x^a \tilde{\rho}(y) dy, \quad (81)$$

with b and σ as in equations (67)-(68). Note that there exist constants $K, K', K'' \in (0, \infty)$ and $a = a(K'') \in (0, \infty)$ such that

$$\tilde{\rho}(x) \leq K \left(\frac{a}{x} \right)^{\frac{(\beta-1)d}{\theta^* \beta}} \quad \text{and} \quad (82)$$

$$\tilde{\rho}(x) \geq K' \left(\frac{a \left(K'' x + (\theta^*)^2 \frac{2\beta}{\beta-1} \right)}{x} \right)^{\frac{(\beta-1)d}{\theta^* \beta}} \quad \text{for } x \in (0, a]. \quad (83)$$

Theorem 2.12 in [Cherny and Engelbert, 2005] shows that X is instantaneously reflected at 0 if there exists some $a \in (0, \infty)$ such that

$$\int_0^a \tilde{\rho}(x) dx < \infty, \quad \int_0^a \frac{1 + |b(x)|}{\tilde{\rho}(x) \sigma^2(x)} dx < \infty \quad \text{and} \quad \int_0^a \frac{|b(x)|}{\sigma^2(x)} dx = \infty. \quad (84)$$

Assume that $d < \theta^* \beta / (\beta - 1)$. Then, the first inequality of (84) follows immediately from equation (82). For the second inequality of (84), we use estimate (83) and note that there exists a constant $K \in (0, \infty)$ such that

$$\sigma^2(x) \geq (\theta^*)^2 \frac{2\beta}{\beta-1} x \quad \text{and} \quad |b(x)| \leq K d \theta^*, \quad x \in (0, a]. \quad (85)$$

To derive the third inequality of (84), it suffices to note that there exists a constant $K' \in (0, \infty)$ and some $a = a(K') \in (0, \infty)$ such that

$$\int_0^a \frac{|b(x)|}{\sigma^2(x)} dx \geq K' \int_0^a \frac{1}{x} dx = \infty. \quad (86)$$

Thus the sufficient condition of theorem 2.12 in [Cherny and Engelbert, 2005] is fulfilled.

Theorem 2.16 in [Cherny and Engelbert, 2005] shows that X is strictly positive if there exists some $a \in (0, \infty)$ such that

$$\int_0^a \tilde{\rho}(x) dx = \infty \quad \text{and} \quad \int_0^a \frac{1 + |b(x)|}{\tilde{\rho}(x) \sigma^2(x)} \tilde{s}(x) dx < \infty. \quad (87)$$

Assume that $d \geq \theta^* \beta / (\beta - 1)$. Then, the first condition in (87) follows immediately from equation (83). To derive the second inequality of (87), we use estimates (82), (83), (86) and note that there exists a constant $K'' \in (0, \infty)$ and $a = a(K'') \in (0, \infty)$ such that

$$\tilde{s}(x) \leq \frac{K''}{x^{\frac{(\beta-1)d}{\theta^* \beta} - 1}}, \quad x \in (0, a]. \quad (88)$$

This completes the proof of claim (b) and lemma 2.

B Explanation of the notion of ghost finite time singularities applied to the recurrent bubble regime

Here, we make precise the concept of a ghost finite-time singularity mentioned in subsection 3.3.3 as applied to the recurrent bubble case $\eta \in (0, 1)$ and $C \in [0, 1]$. In a nutshell, the term ghost finite-time singularity describes processes that follow explosive dynamics temporarily, shadowing a true finite time singularity dynamics up to a time when it departs from it and saturates as a result of an inherent stabilizing mechanism that takes over and prevents the divergence.

Following section 3.3.1, we set $\theta^* \equiv 1$ without loss of generality. First, we introduce a bounded process

$$Z_t = \frac{1}{\gamma \eta} \tan^{-1} \left(\sqrt{S_t^\eta - 1} \right), \quad Z_t \in [\underline{Z}, \overline{Z}] = \left[0, \frac{\pi}{2\gamma \eta} \right], \quad (89)$$

where $\gamma = \sqrt{\frac{2}{\eta(1-\eta)}}$. Conversely,

$$S_t = [\tan^2(\gamma\eta Z_t) + 1]^{\frac{1}{\eta}} = \sec^{\frac{2}{\eta}}(\gamma\eta Z_t). \quad (90)$$

With Itô's lemma, we can check that Z is a unitary diffusion process with

$$dZ_t = \gamma \left[\frac{2(C-1)-\eta}{4} \tan(\gamma\eta Z_t) + \frac{\eta(2D-1)}{4} \cot(\gamma\eta Z_t) \right] dt + dW_t. \quad (91)$$

The drift coefficient of Z depends on the balance of power between the tangent and cotangent terms, being $\frac{2(C-1)-\eta}{4} \tan(\gamma\eta Z_t)$ and $\frac{\eta(2D-1)}{4} \cot(\gamma\eta Z_t)$, respectively. If $C \leq 1$ and $D > \frac{1}{2}$, the tangent term is always negative and the cotangent term is always positive. For this set of parameters, the evolution of S can be described in three consecutive stages.

- When S_t is close to 1^+ , i.e., $\gamma\eta Z_t \approx 0^+$, the positive cotangent term dominates the drift of Z_t and contribute a strong impetus to blow up the bubble.
- After this initial pumping up, at some point the effect from the cotangent term will be offset by the negative effect of the tangent term. Then the drift of Z may be neglected, and $Z_t \sim W_t$. Hence,

$$S_t \sim \sec^{\frac{2}{\eta}}(\gamma\eta W_t) \sim \left(\frac{\pi}{2} - \gamma\eta W_t \right)^{-\frac{2}{\eta}}, \quad (92)$$

which gives an hyperbolic (stochastic) growth for S_t . The above approximate dynamics for S indeed show a finite-time-singularity with a theoretical random critical time point $\tilde{T}_c$. If we solve for the first $\frac{\pi}{2}$ -passage time of the diffusion $\gamma\eta W$, we get a Levy distribution for $\tilde{T}_c$, in particular

$$\tilde{T}_c \sim \text{Levy}(0, \frac{\pi^2}{4\gamma^2\eta^2}). \quad (93)$$

- When S becomes large, the negative tangent term will eventually dominate the cotangent term and act as a strong repeller to prevent explosion of S . This mechanism drives the recurrent behavior, such that the finite time singularity only serves as an approximate, *ghost*-like reference.

The case $D \leq \frac{1}{2}$ evolves analogously, without the initially contributing positive drift term.

References

- Antonacci, G., 2013. Following the trend (diversified managed future trading). Wiley (West Sussex, UK).
- Antonacci, G., 2014. Dual Momentum Investing: An Innovative Strategy for Higher Returns with Lower Risk. McGraw-Hill Education.
- Asness, C. S., Moskowitz, T. J., Pedersen, L. H., 2013. Value and momentum everywhere. The Journal of Finance 68 (3), 929–985.
- Barberis, N., Shleifer, A., Vishny, R., 1998. A Model of Investor Sentiment. Journal of Financial Economics 49 (3), 307–343.
- Barroso, P., Santa-Clara, P., 2015. Momentum has its moments. Journal of Financial Economics 116, 111–120.
- Bouchaud, J.-P., Challet, D., 2016. Why have asset price properties changed so little in 200 years. Available at <https://arxiv.org/abs/1605.00634>.

- Byrd, R. H., Lu, P., Nocedal, J., Zhu, C., 1995. A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing* 16 (5), 1190–1208.
- Carhart, M. M., Mar. 1997. On Persistence in Mutual Fund Performance. *The Journal of Finance* 52 (1), 57–82.
- Chernoff, H., 1954. On the distribution of the likelihood ratio. *The Annals of Mathematical Statistics*, 573–578.
- Cherny, A. S., Engelbert, H.-J., 2005. *Singular Stochastic Differential Equations*.
- Corsi, F., Sornette, D., 2014. Follow the money: The monetary roots of bubbles and crashes. *International Review of Financial Analysis* 32, 47–59.
- Cox, J., Ingersoll, J., Ross, S., 1985. A theory of the term structure of interest rates. *Econometrica* 53, 385–407.
- Daniel, K. D., Hirshleifer, D., Subrahmanyam, A., 1998. Investor psychology and security market under- and overreactions. *Journal of Finance* 53 (6), 1839–1885.
- Demirer, R., Lien, D., Zhang, H., 2015. Industry herding and momentum strategies. *Pacific-Basin Finance Journal* 32, 95–110.
- Elaut, G., Erdős, P., 2016. Adaptive time series momentum – benchmark for trend-following funds. Available at SSRN: <http://ssrn.com/abstract=2772047>.
- Fama, E., French, K., 2012. Size, value, and momentum in international stock returns. *Journal of Financial Economics* 105, 457–472.
- Farmer, J. D., 2002. Market force, ecology and evolution. *Industrial and Corporate Change* 11 (5), 895–953.
- Garleanu, N., Pedersen, L. H., 2013. Dynamic trading with predictable returns and transaction costs. *Journal of Finance* 68 (6), 2309–2340.
- Grinblatt, M., Han, B., 2005. Prospect theory, mental accounting, and momentum. *Journal of financial economics* 78 (2), 311–339.
- Grinblatt, M., Moskowitz, T. J., 2004. Predicting stock price movements from past returns: the role of consistency and tax-loss selling. *Journal of Financial Economics* 71, 541–579.
- Grinblatt, M., Titman, S., Wermers, R., 1995. Momentum investment strategies, portfolio performance, and herding: a study of mutual fund behavior. *The American Economic Review* 85 (5), 1088–1105.
- He, X. Z., Li, K., Apr. 2015. Profitability of time series momentum. *Journal of Banking & Finance* 53, 140–157.
- Hobson, D. G., Rogers, L., 1998. Complete models with stochastic volatility. *Mathematical Finance* 8 (1), 27–48.
- Hoitash, R., Krishnan, M., 2008. Herding, momentum and investor over-reaction. *Review of Quantitative Finance and Accounting* 30 (1), 25–47.
- Hurst, B., Ooi, Y. H., Pedersen, L. H., 2014. A century of evidence on trend-following investing. *AQR Capital Management* (Fall), 1–12.
- Hüsler, A., Sornette, D., Hommes, C. H., 2013. Super-exponential bubbles in lab experiments: evidence for anchoring over-optimistic expectations on price. *Journal Economic Behavior and Organization* 92, 304–316.

- Hwang, S., Rubesam, A., 2015. The disappearance of momentum. *The European Journal of Finance* 21 (7), 584–607 (DOI:10.1080/1351847X.2013.865654).
- Ide, K., Sornette, D., 2002. Oscillatory Finite-Time Singularities in Finance, Population and Rupture. *Physica A* 307 (1-2), 63–106.
- Ikedda, N., Watanabe, S., 1981. Stochastic differential equations and diffusion processes. North-Holland Publishing Co., Amsterdam.
- Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of finance* 48 (1), 65–91.
- Jegadeesh, N., Titman, S., Apr. 2001. Profitability of Momentum Strategies: An Evaluation of Alternative Explanations. *The Journal of Finance* 56 (2), 699–720.
- Kaizoji, T., Leiss, M., Saichev, A., Sornette, D., 2015. Super-exponential endogenous bubbles in an equilibrium model of rational and noise traders. *Journal of Economic Behavior and Organization* 112, 289–310.
- Kim, T. S., Omberg, E., 1996. Dynamic nonmyopic portfolio behavior. *Review of Financial Studies* 9 (1), 141–161.
- Leiss, M., Nax, H. H., Sornette, D., 2015. Super-exponential growth expectations and the global financial crisis. *Journal of Economic Dynamics and Control* 55, 1–13.
- Lempérière, Y., Deremble, C., Seager, P., Potters, M., Bouchaud, J.-P., 2014. Two centuries of trend following. *The Journal of Investment Strategies* 3 (3), 41–61.
- Lin, L., Ren, R., Sornette, D., 2014. The Volatility-Confined LPPL Model: A Consistent Model of ‘Explosive’ Financial Bubbles With Mean-Reversing Residuals. *International Review of Financial Analysis* 33, 210–225.
- Lin, L., Sornette, D., 2013. Diagnostics of Rational Expectation Financial Bubbles with Stochastic Mean-Reverting Termination Times. *The European Journal of Finance* 19 (5-6), 344–365.
- Moskowitz, T. J., Ooi, Y. H., Pedersen, L. H., May 2012. Time series momentum. *Journal of Financial Economics* 104 (2), 228–250.
- Øksendal, B., 1990. When is a stochastic integral a time change of a diffusion? *Journal of Theoretical Probability* 3 (2), 207–226.
- Orléan, A., 1995. Bayesian interactions and collective dynamics of opinion: herd behavior and mimetic contagion. *Journal of Economic Behavior & Organization* 28, 257–274.
- Smug, D., Ashwin, P., Sornette, D., Mar. 2018. Predicting financial market crashes using ghost singularities. *PLOS One* 13 (3), e0195265.
- Sornette, D., 2004. *Why Stock Markets Crash*. Princeton University Press.
- Sornette, D., Andersen, J. V., 2002. A Nonlinear Super-Exponential Rational Model of Speculative Financial Bubbles. *Int. J. Mod. Phys. C* 13 (2), 171–188.
- Sornette, D., Cauwels, P., 2014. 1980-2008: The illusion of the perpetual money machine and what it bodes for the future. *Risks* 2, 103–131.

- Sornette, D., Cauwels, P., 2015a. Financial bubbles: mechanisms and diagnostics. *Review of Behavioral Economics* 2 (3), 279–305.
- Sornette, D., Cauwels, P., 2015b. Managing risks in a creepy world. *Journal of Risk Management in Financial Institutions (JRMFI)* 8 (1), 83–108.
- Sornette, D., Takayasu, H., Zhou, W.-X., 2003. Finite-Time Singularity Signature of Hyperinflation. *Physica A* 325, 492–506.
- Taylor, N., 2014. The rise and fall of technical trading rule success. *Journal of Banking & Finance* 40, 286–302.
- Wilks, S., 1938. The large-sample distribution of the likelihood ratio for testing composite hypotheses. *Ann. Math. Statist.* 9, 60–62.
- Wyart, M., Bouchaud, J.-P., 2007. Self-referential behavior, overreaction and conventions in financial markets. *Journal of Economic Behavior & Organization* 63, 1–24.