

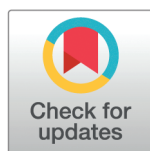
RESEARCH ARTICLE

Revisiting the excess volatility puzzle through the lens of the Chiarella model

Jutta G. Kurth^{1,2*}, Adam A. Majewski³, Jean-Philippe Bouchaud^{3,4}

1 CFM Chair of Econophysics and Complex Systems, École polytechnique, Palaiseau Cedex, France, **2** LadHyX UMR CNRS 7646, École polytechnique, Palaiseau Cedex, France, **3** Capital Fund Management, Rue de l'Université, Paris, France, **4** Académie des Sciences, Quai de Conti, Paris, France

* jutta.kurth@polytechnique.edu



OPEN ACCESS

Citation: Kurth JG, Majewski AA, Bouchaud J-P (2026) Revisiting the excess volatility puzzle through the lens of the Chiarella model. PLoS One 21(1): e0340409. <https://doi.org/10.1371/journal.pone.0340409>

Editor: Nicolò Pecora, Università Cattolica del Sacro Cuore Sede di Piacenza e Cremona Facoltà di Economia: Università Cattolica del Sacro Cuore Facoltà di Economia e Giurisprudenza, ITALY

Received: June 5, 2025

Accepted: December 19, 2025

Published: January 23, 2026

Copyright: © 2026 Kurth et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: The historical spot data used here and supplied by CFM that support the findings of this study can be purchased from

Abstract

We amend and extend the Chiarella model of financial markets to deal with arbitrary drift in long-term value in a consistent way. This allows us to improve upon existing calibration schemes, opening the possibility of calibrating individual monthly time series instead of classes of time series. The technique is employed on spot prices of four asset classes from ca. 1800 onward (stock indices, bonds, commodities, currencies). The so-called fundamental value is a direct output of the calibration, which allows us to (a) quantify the amount of excess volatility in these markets, which we find to be large (e.g. a factor ≈ 4 for stock indices) and consistent with previous estimates; and (b) determine the distribution of mispricings (i.e. the log-difference between market price and value), which we find in many cases to be bimodal. Both findings are strongly at odds with the Efficient Market Hypothesis. We also study in detail the ‘sloppiness’ of the calibration, that is, the directions in parameter space that are weakly constrained by data. The main conclusions of our study are remarkably consistent across different asset classes, and reinforce the hypothesis that the medium-term fate of financial markets is determined by a tug-of-war between trend followers and fundamentalists.

1 Introduction

The Efficient Market Hypothesis (EMH) maintains that market prices closely follow fundamental values at all times. Although still a cornerstone of Financial Economics and fiercely defended by some scholars – see, e.g., this [interview](#) for E. Fama’s latest quips on this topic – contradicting evidence has accumulated since the early eighties. Chiarella’s model can be seen as an explicit alternative to the EMH, opposing Fama’s peeve: *Now the problem is that behavioral finance doesn’t have any models of its own. It’s just a criticism of other models.* Among the most inconvenient facts are (i) Shiller’s excess volatility puzzle [1], i.e. that market volatility appears to be much too high to be explained by the volatility of fundamental values and (ii) the well-documented trend following anomaly, that is, the statistically significant, persistent

Finæon (formerly: Global Financial Data) (<https://www.finæon.com/data/>). The data cannot be made available by the authors as it was used under license. The same data may, however, be obtained and used by any entity holding a license.

Funding: The author(s) received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

and profitable positive correlation between past trends and future trends, across all asset classes [2–4] and refs. therein. Such correlations should not exist if markets were efficient.

A competing theory, that has gained momentum (no pun intended) over the past decades, is the order-driven view of market prices [5–8], a.k.a. the inelastic market hypothesis. In such a picture, prices are “mechanically” impacted by order flow, independently of fundamental value. Excess buy (resp. sell) pressure, even uninformed, makes prices go up (resp. down), and such an impact persists over the medium to long term. Therefore, accounting for price movements mostly means understanding order flows. Of course, the reasons why people buy or sell are multifarious and based on an infinite variety of incentives and trading signals. In order to model such a complex ecology of market participants [9], Carl Chiarella, and several authors after him, have proposed to retain only three main categories of traders [10–14]: “trend followers” (who buy/sell when the price has gone up/down), “fundamentalists” (who sell/buy when the price is above/below their perceived fair value), and “noise traders” (who buy or sell for random reasons, i.e. all other reasons not captured by trend or value). Chiarella’s model and its generalisations offer the simplest Heterogenous Agent-based framework (HABM) that captures important stylized facts of financial markets, including the excess volatility puzzle, volatility clustering, Black’s “factor 2” of persistent mispricings [15,16], as well as the long-term ecological coexistence of such strategies [14].

Using Bayesian filtering techniques, a full-fledged calibration of the Chiarella model was undertaken for a variety of assets in [13], and clearly supported the co-existence of trend-following and value mean-reversion in most markets, including the possibility to a bimodal distribution of mispricings – meaning that markets have a higher probability of being over- or under-valued than correctly valued. However, our attempt to extend such a calibration scheme to single stocks revealed problems and inconsistencies. The aim of the present study is to provide a new, more consistent specification of the Chiarella model, and its calibration of the same universe of assets as in [13]. This allows us to discuss in more detail the excess volatility puzzle and the issue of bimodality. We also apply the “sloppiness analysis” proposed by Sethna et al. [17] to our calibration procedure, allowing us to identify the most important features of the model that the data is able to identify. Nevertheless, the work in Ref. [13] provides the theoretical and empirical groundwork, on which our contribution is based.

In particular, our model differs from the one in [13] by an amendment of the dynamical system that renders the price dynamics stationary with respect to the fundamental value process. In [13] the non-stationarity of the two quantities implies that an asset price is not anchored by its fundamental value, which (unreasonably and without real supporting economic evidence) yields processes that diverge over long time scales. Further, a time-varying drift is introduced here instead of the constant drift in [13], which is not supported by the data and, in fact, leaves their calibration output imprecise. This requires a modification of the calibration scheme but allows for the calibration of *individual* asset prices, whereas in [13] most model parameters could only be determined for an asset class as a whole and not per asset.

This new version of the model and its calibration opens the possibility of not only using Chiarella-type models as a proof-of-concept, but in fact for a quantitative assessment of known stylised facts, such as excess volatility, or mispricing distributions. The extension of our method to single stocks, as well as possible further generalisations of the Chiarella model, will be detailed in forthcoming publications.

Related works in the literature involving behavioural economics HABMs with latent variables are [18], concluding that financial instability cannot be explained by exogenous fluctuations alone but only with endogenous fluctuations due to the speculative behaviour of different types of investors. Ref. [19] focuses on the emergence of cycles where different agents' beliefs are the unobserved latent states. Like us, both deploy Kalman filtering to infer latent variables. Similar HABMs have alternatively been successfully calibrated via Markov chain Monte Carlo (see, e.g., the works by Lux [20,21]), or via simulated maximum likelihood [22]. For a survey on empirical validations of (H)ABMs see [23]. Ter Ellen et al. [24] compare behavioural agent heterogeneity between asset classes and investigate their impact on market stability. For a general overview of financial ABMs, see the survey by Hommes [25].

The outline of this paper is as follows: in Sect 2 the model is introduced and its possible dynamical phases are analytically derived. Sect 3 describes the time series data used in this study on which the model is calibrated, detailed in Sect 4. The calibration results allow an investigation of excess volatility. To what extent prices typically depart from values is detailed in Sect 5 through the mispricing distribution, in particular its bimodality. The “sloppy” character of our calibration is discussed in Sect 6. Finally, a conclusion with an outlook is provided in Sect 7. More technical material and supplementary empirical analyses are provided in Appendices.

2 A modified Chiarella model

2.1 Model specification

We assume the evolution of (log-)prices P_t is governed by linear price impact as suggested in Kyle's seminal work [26]. Although a linear price impact is ruled out on daily or intraday time scales (see, e.g [5]), it is thought to be appropriate on longer, monthly time scales [6,7] which are of interest in the present work. This means that a price change in a (long enough) time interval $[t, t + \Delta t]$ is proportional to the total signed volume traded in that interval, where the total signed volume is represented by a cumulative demand imbalance $D(t, t + \Delta t)$, i.e.

$$P_{t+\Delta t} - P_t = \lambda D(t, t + \Delta t). \quad (1)$$

Here λ is Kyle's lambda, which is inversely proportional to the liquidity of the traded asset. Thus, a product is considered liquid if its price change resulting from a certain traded volume or demand imbalance is relatively small.

The aggregate demand of all investors is of course diverse and abundant. However, studies reveal that the two types of market participants accounting for a large share of demand imbalances are (a) fundamental value investors and (b) chartists or trend followers (TFs), as done in [10,11,27] and empirically confirmed in [28].

For this reason, a model inspired by [10] was proposed in [13], but whose analytical shortcomings we seek to alleviate in this new study. Like in the latter paper, the HABM-type model studied here contains three groups of investors or *agents* (for other HABM specifications, and their calibration on the S&P500, see the mini-review of T. Lux [29]). Those are:

1. **Fundamentalists:** investors who believe in a rational fundamental (log-)value V_t of a financial asset. They tend to only step-in when the (log-)price P_t is far away from its value V_t : they buy when assets are underpriced ($P_t < V_t$) and sell when they are overpriced ($P_t > V_t$). Fundamentalists' cumulative demand is proportional to the mispricing or price distortion $\delta_t := V_t - P_t$ with a factor $\tilde{\kappa}$ quantifying their weight in the market:

$$D^F(t, t + \Delta t) = \tilde{\kappa} \int_t^{t+\Delta t} (V_s - P_s) ds.$$

We will use below the quantity $\kappa := \lambda \tilde{\kappa}$.

Usually, value traders resort to fundamental analysis for their valuation and we will model value dynamics with an arbitrary (non-stationary) drift, g_t , plus random changes $\sigma_V dW_t^V$:

$$dV_t = g_t dt + \sigma_V dW_t^V.$$

The drift g_t describes the long-term evolution of fundamental value and one should expect its variations to be slow; any faster changes in value being captured by the random term dW_t^V .

2. **Trend Followers:** TFs' trading behaviour is independent of a notion fundamental value. Instead, their investment choices rely solely on past price dynamics. They buy, if the price moved up (relative to its long time drift) in the recent past, such that the trend signal $M_t > 0$ and sell if the price went down, $M_t < 0$. Popular choices for such trend signals are exponentially weighted moving averages (EWMA) of past returns. The demand is given by an increasing function of the signal M_t that saturates for strong signals $|M|$ (due to e.g. budget constraints or risk aversion [3]). With $\tilde{\beta}$ the weight of TFs and γ the signal saturation sensitivity, we posit that the demand of TFs reads

$$D^{\text{TF}}(t, t + \Delta t) = \tilde{\beta} \int_t^{t+\Delta t} \tanh(\gamma M_s) ds.$$

We will use below the quantity $\beta := \lambda \tilde{\beta}$.

3. **Noise Traders:** NTs subsume all those traders that follow strategies uncorrelated to the previous two. One may think of retail investors, or investors who trade on other signals or time horizons. Their cumulative demand is modelled as a Brownian motion $\sigma_N W_t^N$, where the standard deviation σ_N describes their impact in the asset, i.e.

$$D^{\text{NT}}(t, t + \Delta t) = \tilde{\sigma}_N \int_t^{t+\Delta t} dW_s^N,$$

where $\lambda \tilde{\sigma}_N =: \sigma_N$. The standard Brownian motion W_t^N is independent from random, fast changes of fundamental value described by W_t^V .

Thus, the overall cumulative demand leading to a price change is

$$D(t, t + \Delta t) = \sum_{i \in I} D^i(t, t + \Delta t), \quad (2)$$

where I is the set of investor types ($I = \{F, TF, NT\}$).

In the following we assume that the trend signal M_t is computed as an exponential moving average with a forget rate α . Consequently, for $\Delta t \rightarrow 0$, the price dynamics is described by the following set of stochastic differential equations

$$\begin{aligned} dP_t &= \kappa(V_t - P_t)dt + \beta \tanh(\gamma M_t)dt + g_t dt + \sigma_N dW_t^N \\ dM_t &= -\alpha M_t dt + \alpha(dP_t - g_t dt) \\ dV_t &= g_t dt + \sigma_V dW_t^V. \end{aligned} \quad (3)$$

The parameters $\alpha, \kappa, \beta, \gamma$ are all non-negative.

The quantity M_t denotes the trend signal, which is an EWMA of past *drift-adjusted* log-returns (which is similar to trending on mispricings as in [19]). In other words: we assume that trend followers react only to the excess returns of the asset,

and do not trend on the secular drift. This would lead to absurd instabilities in the model we want to avoid. As such, this specification is an improvement over the model proposed in [13], where a *time-independent* drift g appears only in the fundamental value V_t and not in the definition of M_t , nor in the dynamics of P_t . As will be shown next, the stability and dynamical phases of the system now become independent of the drift g_t , whereas in Ref. [13] the dynamical analysis was only valid for $g = 0$. This renders the dynamics and calibration improper for assets whose value can strongly drift upwards or downwards, like stocks. In these cases, the specification of Ref. [13] may lead to a divergence between P_t and V_t entirely caused by g_t and not as a consequence of demand imbalances.

A way to mitigate such an effect and prevent the long-term divergence between price and value is to consider, as was done in Ref. [13], a non-linear anchoring term of the form $\kappa_3(V_t - P_t)^3$. They were the first to introduce non-linear fundamentalist demand functions within the Chiarella framework, which is also one of their most relevant contributions. The calibration of such a non-linear model however comes at a higher computational cost and it is important to have a linear model that makes sense and can be meaningfully calibrated before considering a non-linear extension (which we will have to do anyway for reasons explained below).

2.2 Non-linear model

Schmitt and Westerhoff [12] and Majewski et al. [13] further introduced a model with a non-linear demand function for fundamentalists. They argue and show that the linear fundamentalists' demand is not able to capture the complex nature of value investing, most importantly the uncertainty of investors about fundamental value, which cannot be directly observed but only estimated. It seems reasonable that the reaction of investors is not proportional to mispricing $\delta = P - V$, but much weaker for small mispricings (in view of the uncertainty, leading to an almost flat curve for small δ) and much stronger when mispricing becomes conspicuous [13].

In order to accommodate these departures from linearity, a cubic demand term can be added to the linear term. Within our adjustments the model then reads

$$\begin{aligned} dP_t &= f(V_t - P_t)dt + \beta \tanh(\gamma M_t)dt + g_t dt + \sigma_N dW_t^N \\ dM_t &= -\alpha M_t dt + \alpha(dP_t - g_t dt) \\ dV_t &= g_t dt + \sigma_V dW_t^V, \end{aligned} \quad (4)$$

where $f(x) = \kappa x + \kappa_3 x^3$ describes the modified demand. For the model to be compatible with strong mean reversion for large mispricing, one must impose $\kappa_3 > 0$.

2.3 Linear stability and bifurcation analysis

The deterministic equivalent to system (3), which can be analytically studied using methods from dynamical systems theory is obtained by letting $\sigma_N = \sigma_V = 0$ [30–32]. Since the mispricing $\delta = P - V$ associated with the system (3) is independent of the drift g_t , it is mathematically convenient to study price *relative to* value in δ - M -space, effectively reducing the dimensionality by one. The linear system ($\kappa_3 = 0$) then reads

$$\begin{aligned} \dot{\delta}_t &= -\kappa \delta_t + \beta \tanh(\gamma M_t) \\ \dot{M}_t &= -\alpha M_t + \alpha \delta \\ &= -\alpha M_t + \alpha(-\kappa \delta + \beta \tanh(\gamma M_t)). \end{aligned} \quad (5)$$

The δ -nullcline is

$$\delta_t = \frac{\beta}{\kappa} \tanh(\gamma M_t) \quad (6)$$

and the M -nullcline is

$$\delta_t = \frac{\beta}{\kappa} \tanh(\gamma M_t) - \frac{M_t}{\kappa}, \quad (7)$$

which is a sigmoidal function for small $|M_t|$ thanks to the tanh but for large $\pm M_t$ the function diverges to $\mp\infty$.

2.3.1 Spiral fixed point. From the two nullclines, which intersect exactly once, it follows that there is a single fixed point (FP) at $(M^*, \delta^*) = (0, 0)$, the origin. The FP's stability type can be inferred by inspecting the Jacobian

$$J = \begin{pmatrix} -\kappa & \beta\gamma(1 - \tanh^2(\gamma M_t)) \\ -\alpha\kappa & -\alpha + \alpha\beta\gamma(1 - \tanh^2(\gamma M_t)) \end{pmatrix} \quad (8)$$

of system (5) at the FP:

$$J|_{M=0, \delta=0} = J^* = \begin{pmatrix} -\kappa & \beta\gamma \\ -\alpha\kappa & \alpha(\beta\gamma - 1) \end{pmatrix}. \quad (9)$$

$\det(J^*) = \alpha\kappa > 0$ as the time scale $\alpha > 0$, and mean reversion strength $\kappa > 0$ for the dynamics not to diverge. This means that there are no saddles but only asymptotically (un)stable FPs. Further, $\text{tr}(J^*) = -\kappa + \alpha(\beta\gamma - 1)$. The fixed point is stable when $\text{tr}(J^*) < 0$, i.e. when $\kappa > \alpha(\beta\gamma - 1)$. Consequently, $\text{tr}(J^*) - 4 \det(J^*) < 0$, from which it follows that the FP is a *spiral* (because the eigenvalues of J^* then have non-zero imaginary part). Hence, the bifurcation point at which the FP becomes unstable and the flow in the δ - M -plane changes qualitatively is

$$\alpha^* = \frac{\kappa}{\beta\gamma - 1}. \quad (10)$$

That the fixed point is only stable when $\kappa > \alpha(\beta\gamma - 1)$ shows that when value investors dominate trading, the deterministic part of the price dynamics converges to the fundamental value, where it remains forever. When chartists dominate trading, the FP becomes unstable and, in fact, a stable limit cycle emerges, such that there is a periodic motion of price around value. The emergence of the limit cycle is proven in the next section.

The condition $\kappa > \alpha(\beta\gamma - 1)$ coincides with the result found in [13], however, for their model the condition was only true for $g_t = 0$ (which is not compatible with empirical results), while in our model the condition holds true generically.

An example of such a stable spiral dynamics in the mispricing $\delta = P - V$, meaning that price converges to value, is depicted in Fig 1 as a phase portrait in the δ - M -plane alongside its price, value and trend signal trajectories in the deterministic case. Its stochastic analogue using the same parameters is given in Fig 2.

2.3.2 Hopf-bifurcation: Emergence of a limit cycle. In this section it will be shown that a stable limit cycle emerges when the FP loses its stability. This qualitative change of dynamics – the loss of a FP's stability coinciding with the emergence of a periodic motion – is known as a *Hopf-bifurcation*. It occurs when a pair of complex conjugate eigenvalues of the Jacobian from the linearisation of the system around the FP crosses the imaginary axis in the complex plane as a parameter crosses its bifurcation point.

According to the Hopf-Bifurcation Theorem [31,32], two conditions on the eigenvalue pair of J^* , which in this case are

$$\lambda_{1/2} = \frac{1}{2} \left(\alpha\beta\gamma - \alpha - \kappa \pm \sqrt{(-\alpha\beta\gamma + \alpha + \kappa)^2 - 4\alpha\kappa} \right), \quad (11)$$

have to be fulfilled:

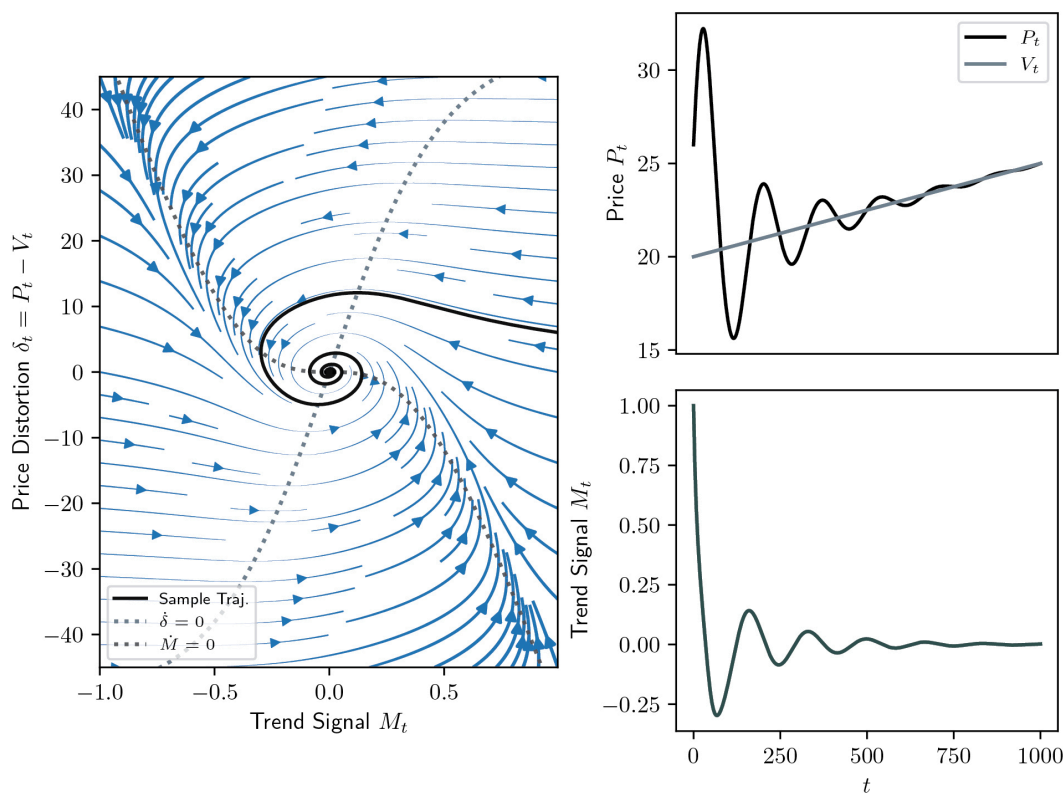


Fig 1. Typical dynamics of system (3) in the case where its limit set is a spiral, $\kappa > \alpha(\beta\gamma - 1)$, and without noise ($\sigma_N = \sigma_V = 0$). The parameters are $(\kappa, \alpha, \beta, \gamma) = (0.01, 1/7, 0.5, 2)$, and the system is initialised with $(P_0, V_0, M_0) = (26, 20, 1)$; the drift g is constant. Left: Phase portrait of the mispricing δ and the trend signal M together with its nullclines and a sample trajectory. The streamlines' (blue) width and density encode the magnitude of the velocity field. Right: evolution of the price P , value V and trend signal M .

<https://doi.org/10.1371/journal.pone.0340409.g001>

1. The eigenvalue pair becomes purely imaginary at the bifurcation point:

$$\begin{aligned} \lambda_{1/2}(\alpha^*) &= \frac{1}{2} \left(\underbrace{\frac{\kappa\beta\gamma}{\beta\gamma - 1} - \frac{\kappa}{\beta\gamma - 1}}_{=0} - \kappa \pm \sqrt{-4 \frac{\kappa^2}{\beta\gamma - 1}} \right) \\ &= i\kappa \sqrt{\frac{1}{\beta\gamma - 1}}, \end{aligned}$$

which is fulfilled since the condition involving $\text{tr}(J^*)$ in Sect 2.3.1 as $\beta\gamma > 1$ is a necessary condition for the FP to become unstable.

2. The real part of the derivative of the eigenvalues with respect to the bifurcation parameter, evaluated at the bifurcation point is non-zero:

$$\frac{\partial \text{Re}(\lambda)}{\partial \alpha}(\alpha^*) = \beta\gamma - 1 \neq 0,$$

which is true for the same reason.

The Hopf bifurcation may further be classified as *supercritical*.

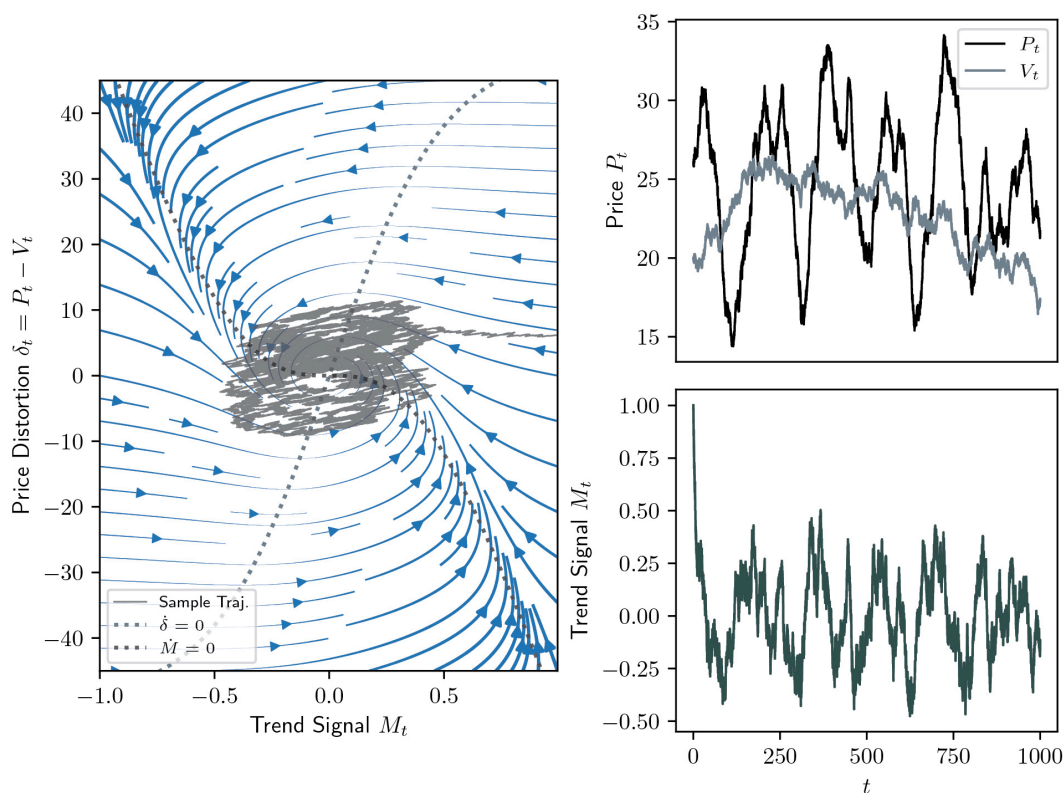


Fig 2. Same as Fig 1 but in the presence of noise ($\sigma_N = 0.35$ and $\sigma_V = 0.2$).

<https://doi.org/10.1371/journal.pone.0340409.g002>

Note that the linear stability analysis of the center fixed point (0, 0) as well as the implications of the Hopf-bifurcation theorem continue to hold for the model with the cubic fundamental demand (system (4)) as the Jacobian of the *linearised* system evaluated at this fixed point is the same.

An example of the deterministic dynamics (analog. to Fig 1) including the phase portrait with the limit cycle in the δ - M -plane, as well as the evolutions of price, value and trend signal is provided with Fig 3. In Fig 4 is a full, stochastic example.

Thus, according to the model, price converges to value on long time scales if fundamentalists dominate trading, while it may oscillate when the presence of TFs is strong enough.

3 Data

All data used for the subsequent calibration of the model are *monthly* spot prices. In order to show the generality of the results, we study four different asset classes: Indices (IDX), Commodities (CMD), Bonds (BND) & Currencies (FXR). In a subsequent paper, single name stocks will be further considered, which requires a model adaptation and modified calibration scheme, wherefore they are not included here.

To show that our model is an improvement over the one in [13] and to demonstrate the generality of the model, which can be applied to many asset classes, the model is calibrated on the same data set they used. The provider is *Global Financial Data* (now named *Finaneon*) and the data set available covers the period 1791 to 2015. As in [13], we restrict the asset pool to only those products with a long enough history, which in the case of indices, bonds, and currencies means Australia, Canada, Germany, Japan, Switzerland, the UK and the US. Further, only exchange rates of the named

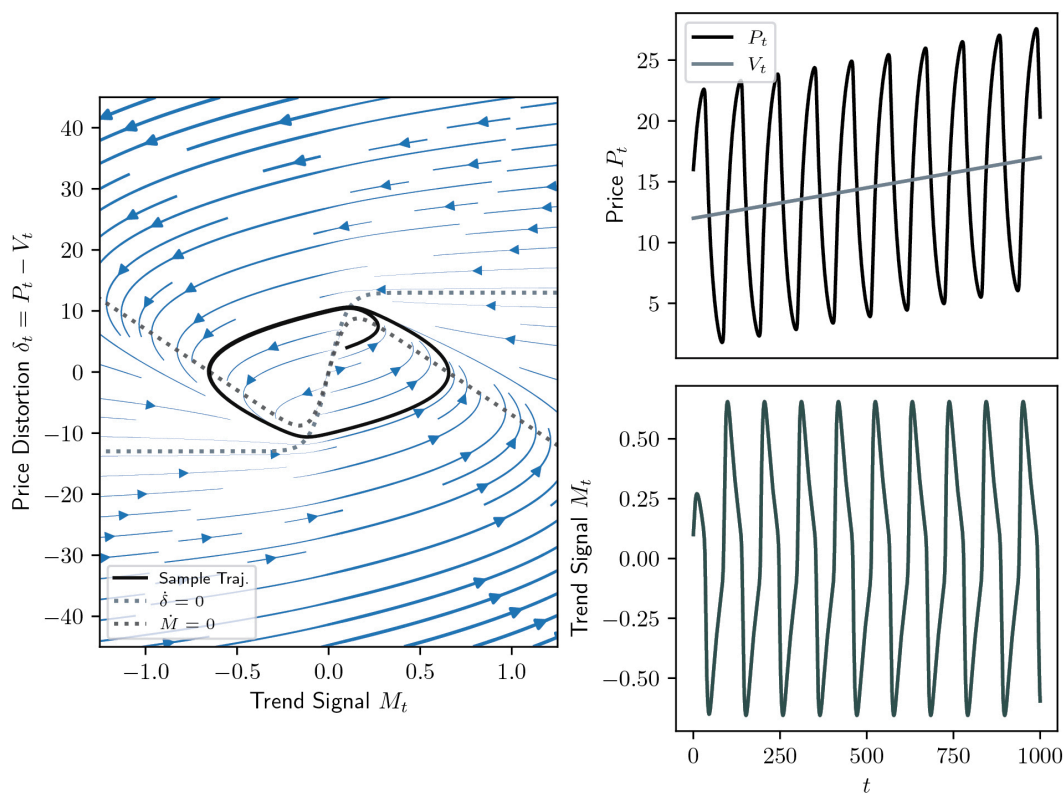


Fig 3. Same as Fig 1 but in the case where the limit set is a limit cycle, $\kappa < \alpha(\beta\gamma - 1)$. The parameters are $(\kappa, \alpha, \beta, \gamma) = (0.05, 1/7, 0.65, 10)$, and the system is initialised with $(P_0, V_0, M_0) = (16, 12, 0.1)$.

<https://doi.org/10.1371/journal.pone.0340409.g003>

countries' currencies against the US dollar are regarded. The considered commodities are copper, corn, crude oil, Henry Hub natural gas, live cattle, sugar, and wheat.

Furthermore, as in [13], the time series of each asset is restricted to when it was publicly traded with high liquidity. Thus, the used exchange rate series start in 1973 as from 1945 to 1973 all considered currencies were pegged against the US dollar in the Bretton Woods system. Government bond prices are used only after 1920 for they were not liquid before. For all commodities the prices during World War II are excluded and additionally the period of 1939-1985 for crude oil. Before 1986 oil prices were largely administered/regulated rather than discovered in a deep spot/futures market, so the data are not comparable to modern oil prices (e.g. due to OPEC setting 'official' prices via long-term contracts and Saudi Arabia acting as a swing producer). Equity index prices in particular display occasional extreme events as they are strongly impacted by political events, wherefore World War II is removed from the German and Japanese index' price series and World War I from the German and British ones. For Germany, the period around the so-called 'hyperinflation' of the Weimar Republic, which concerns the post WW I period is removed. Finally, the years 1973 and 1974, marked by the fall of the Bretton Woods system, are excluded from the British index. Whenever a time series is discontinued, the left end of the gap (with all the data preceding it) will be brought to the same level as the right end of the gap to avoid price jumps for which the model is not designed.

In addition to that, index prices are inflation adjusted by multiplying their nominal price with the respective Consumer Price Index (CPI) value belonging to that time stamp, normalised by (i.e. divided by) the final observed CPI value (such that the last price has a CPI multiplier of 1). Commodities are inflation adjusted with the US CPI in the same fashion.

Further details on the data set are provided in [3].

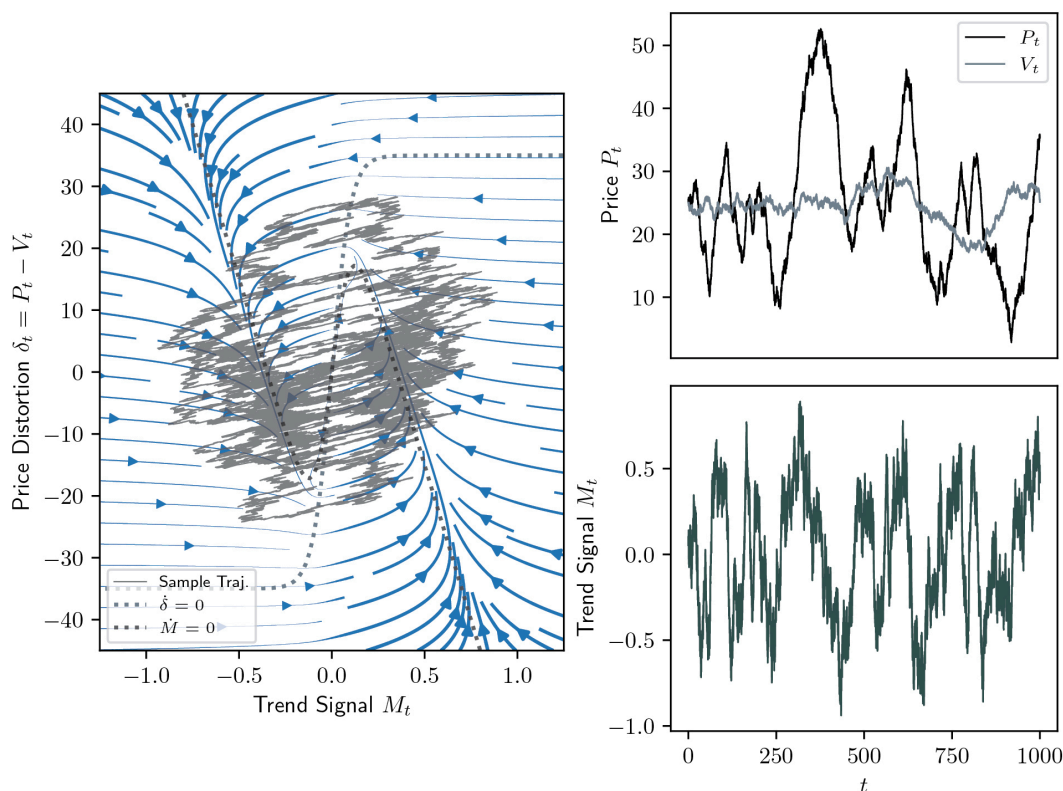


Fig 4. Same as Fig 3 but in the presence of noise ($\sigma_N = 0.6$ and $\sigma_V = 0.2$). The parameters are $(\kappa, \alpha, \beta, \gamma) = (0.01, 1/7, 0.35, 10)$, and the system is initialised with $(P_0, V_0, M_0) = (26.5, 26, 0.1)$.

<https://doi.org/10.1371/journal.pone.0340409.g004>

4 Calibration

The two key steps in the calibration of a dynamical system of the type of (3) are a combination of the Expectation-Maximisation Algorithm and Kalman filtering, as done in Ref. [13], where the full calibration scheme on which we rely is derived. For further resources on Bayesian filtering and the EM algorithm, see, e.g., [33–36].

A discrete time version of system (3), where a time increment $dt = \Delta t = 1$ corresponds to one month, is given by

$$\begin{aligned} p_{t+1} &= p_t + \kappa(v_t - p_t) + \beta \tanh(\gamma m_t) + g_t + \eta_{t+1}^N \\ m_{t+1} &= (1 - \alpha)m_t + \alpha(p_t - p_{t-1} - g_t) \\ v_{t+1} &= v_t + g_t + \eta_{t+1}^V, \end{aligned} \quad (12)$$

where we use small cap notation for discrete time variables, and $\eta^{N/V}$ are Gaussian white noise processes with variance $\sigma_{N/V}^2$. Due to the model set-up in Majewski et al. [13] and its implications on the Kalman relations and EM-algorithm the drift g_t had to be fixed to a time independent value g . In our case, the (integrated) drift impacts price and value alike, our model becomes more canonical and $G_t = \int_0^t g_s ds$ may be removed from log-price series ex ante, allowing us to consider any time-dependent g_t . This improves the model as we do not generally find empirical evidence supporting the choice of a constant drift, and deem it too restrictive. Fig 5 substantiates this claim for the US stock index whose evolution does not justify the assumption of a constant drift.

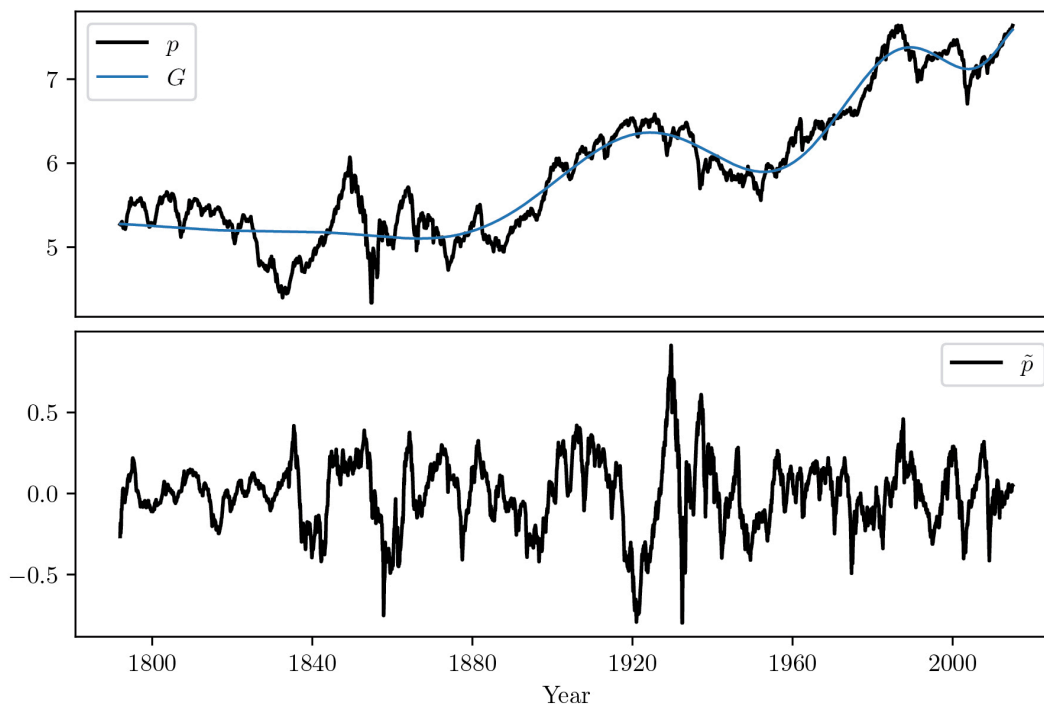


Fig 5. Evolution of the log-price p , the integrated drift G , and the de-drifted log-'price' \tilde{p} of the US index. G is estimated as a polynomial with one order per decade of data, i.e. a 22nd order polynomial here for data ranging from 1791-12 to 2014-12.

<https://doi.org/10.1371/journal.pone.0340409.g005>

Denoting as \tilde{p} , \tilde{v} , \tilde{m} the de-drifted versions of p , v , and m , the formulation simplifies to

$$\begin{aligned}\tilde{p}_{t+1} &= \tilde{p}_t + \kappa(\tilde{v}_t - \tilde{p}_t) + \beta \tanh(\gamma \tilde{m}_t) + \eta_{t+1}^N \\ \tilde{m}_{t+1} &= (1 - \alpha)\tilde{m}_t + \alpha(\tilde{p}_t - \tilde{p}_{t-1}) \\ \tilde{v}_{t+1} &= \tilde{v}_t + \eta_{t+1}^V.\end{aligned}\tag{13}$$

The question how the drift g_t is to be chosen is quite important, since one can obviously find a perfect fit to the data by choosing $g_t \equiv p_{t+1} - p_t$, $\sigma_{N/V} = 0$, $\beta = 0$, such that $p_t = v_t$ at all times. This corresponds, in a sense, to the Efficient Market limit, where the price evolution is fully explained by changes of value. However, the well-documented presence of auto-correlation in the return time series (i.e. trend following on medium time scales and mean-reversion on long times scales [2,3,15,16]) would mean that prices are *not* properly anticipated future values.

Following the main tenet of the Chiarella model, we rather assume that trend-followers and noise traders have a non-zero impact on prices, i.e. $\beta > 0$ and $\sigma_N > 0$. We also assume that the long-term drift of value g_t changes smoothly over time, higher frequency changes being captured by the noise term η^V . Assuming a business cycle of ten years, we propose to extract the long-term drift by fitting the price time series with a polynomial of order $k = \lfloor T/10 \rfloor$, where T is the total length of the series in years. However, we do not want g_t to capture fluctuations on scales shorter than 5 years or so, since these should emerge from the dynamics of Chiarella's model itself. In fact, one should choose a time scale T/k longer than κ^{-1} in order to self-consistently assume that prices follow value on long enough time scales. Calibration will indeed suggest that κ^{-1} is typically in the range 2–5 years, see Table 1. An example of log-price p , integrated drift G and de-drifted price \tilde{p} is shown in Fig 5, corresponding the US stock index, with an order 22 polynomial fit over circa 220 years. We have checked that our results are very robust against changing k in a reasonable range. A case study

Table 1. Calibration results for the fit of the de-drifted linear model (13) on the four asset classes (calibration step 3. in Sect 4.3). First group: stock indices, second: commodities, third: bonds, fourth: currencies. The data for some of the assets, esp. commodities, goes back 200 years. $\bar{\mathcal{L}}$ is the predictive log-likelihood normalised by the series length. Std. deviations for all values are listed in Table IV in Appendix B in S1 Appendices. $\beta\gamma < 1$ for all assets.

	κ	β	γ	σ_N	σ_V	v_0	$\bar{\mathcal{L}}$
US	0.027	0.076	4.168	0.043	0.011	0.017	1.723
UK	0.026	0.100	4.101	0.036	0.009	0.019	1.907
AU	0.028	0.079	4.238	0.039	0.010	0.012	1.823
CH	0.034	0.100	4.232	0.043	0.011	0.024	1.708
JP	0.028	0.095	4.147	0.058	0.015	0.128	1.398
CA	0.045	0.101	4.138	0.045	0.012	0.027	1.678
DE	0.036	0.125	4.279	0.044	0.011	0.025	1.702
SUGAR	0.063	0.146	2.246	0.073	0.020	-0.011	1.186
CORN	0.096	-0.027	2.324	0.119	0.033	0.103	0.699
LCATTLE	0.270	0.381	2.198	0.045	0.013	0.004	1.276
WHEAT	0.081	-0.002	2.518	0.092	0.025	-0.066	0.960
COPPER	0.055	0.059	2.100	0.061	0.017	0.044	1.365
NATGAS	0.195	-0.111	2.321	0.174	0.048	0.104	0.300
CRUDE	0.089	0.276	2.766	0.106	0.029	-0.362	0.812
USBND	0.069	0.048	5.316	0.046	0.003	-0.002	1.644
UKBND	0.084	0.068	5.382	0.048	0.003	0.006	1.599
CHBND	0.062	0.089	4.802	0.051	0.004	-0.004	1.546
JPBND	0.038	-0.008	6.066	0.087	0.006	-0.035	1.009
AUBND	0.066	0.067	5.849	0.045	0.003	-0.001	1.645
CABND	0.068	0.064	4.901	0.036	0.003	-0.000	1.896
DEBND	0.066	0.103	4.574	0.045	0.003	-0.005	1.668
CHFUSD	0.035	0.032	6.455	0.035	0.006	-0.007	1.923
JPYUSD	0.033	0.057	6.451	0.032	0.005	-0.017	1.996
AUDUSD	0.040	0.028	7.011	0.033	0.005	-0.002	1.980
GBPUSD	0.044	0.051	6.701	0.029	0.005	-0.013	2.102
CADUSD	0.019	0.004	6.764	0.019	0.003	-0.013	2.510
EURUSD	0.025	0.030	6.403	0.032	0.005	-0.014	2.000

<https://doi.org/10.1371/journal.pone.0340409.t001>

demonstrating this is given in Sect 3 of Appendix B in S1 Appendices, where the values of calibrated parameters are given for $k = 14, 22$ and 30 , on the example of the US stock index. Importantly, we further tested that our results are robust when determining g_t as a causal, backward-looking EWMA over 10 years. Such a determination of the drift allows trend followers to compare medium term trend to long term trend and trade the difference, as is done in practice.

Finally, let us comment on the philosophy of our approach, where we let the data determine the “fundamental value” of assets, taking stock of the fact that the true fundamental value of an asset is probably unknowable. Some theoretical models, often based on questionable equilibrium assumptions, have been proposed for stocks, bonds and currencies. But there are so many moving parts in these models that it is probably better to accept that the fundamental value must itself be determined empirically – admittedly again using a theoretical framework to extract it from price time series, in our case the Chiarella model.

The merit of this approach is that we leave maximum flexibility on the definition of fundamental value, allowing one to apply our framework universally, including asset classes for which there are no obvious pricing models, like commodities or even Bitcoin. Of course, as noted above, the EMH solution would be to say that price equals value at all times, end of story. If one believes that this is not the case and that trend followers/value traders impact prices (following [5–8]), then our framework provides quantitative information about mispricing and excess volatility, while leaving the best possible chance to “value” to explain the price path.

4.1 EM-algorithm

The estimation of the parameters and the de-drifted log-value methodology overlaps with the algorithm proposed in [13]. Thus, we only briefly describe the algorithm here as it coincides with the one of Ref. [13], but now with $g = 0$ since \tilde{p} , \tilde{m} , \tilde{v} in Eqs (13) are already de-drifted. Further, we propose methods different from those in [13] for the ex-ante estimation of the trend parameters γ and α (where in Ref. [13] they were rather motivated than calibrated) in Sect 4.2, which are then fixed in our version of the EM-algorithm to reduce the number of parameter estimates from six to four and because there exist simpler and more canonical ways of estimating them than through the EM-algorithm.

Note that were the parameters known, the unobservable fundamental value could be inferred as what is known in the control theory literature as a hidden or latent variable via Bayesian filtering techniques [33]. Indeed, system (3) is linear in v and the noise is assumed to be Gaussian, hence the optimal filter is a Kalman filter. If this was not the case, particle filters could be used.

Having obtained a first value proxy by initialising the calibration, the EM-algorithm is used to obtain a set of optimal parameters based on the current fundamental value by maximising an otherwise difficult to compute marginal log-likelihood by instead maximising a joint log-likelihood. Each iteration follows a two step procedure:

1. **E-step:** calculation of a conditional expectation of the joint log-likelihood of the posterior distribution over the hidden variable v , given past prices and the current best guess of the parameters.
2. **M-step:** calculation of the parameters by optimising the joint log-likelihood.

After each iteration there is a new estimate of the fundamental value v together with the set of parameters $\theta = (\kappa, \beta, \sigma_N, \sigma_V, v_0)$, until the algorithm terminates when the increase in likelihood falls below a tolerance of $\epsilon = 10^{-5}$.

For the model with a non-linear demand function of the fundamentalists, system (4), we also use the adapted version ($g = 0$) of the algorithm presented in [13], utilising unscented Kalman filtering to treat the cubic fundamentalists' demand.

The implementations of both the linear and non-linear model are provided in [S2 Code](#).

4.2 Estimation of α and γ

The trend time scale α is chosen as the time scale that maximises the Sharpe ratio of the assets' *de-drifted* trend signal \tilde{m} . Choosing \tilde{m} over m also undercuts the appropriate criticism of trend signals often used in the literature that are defined on past returns directly, thus reflecting mostly the long-only bias (buying and holding an asset while its price and value tend to increase over long horizons due to overall market growth), rather than the actual short-term to medium-term trend. The Sharpe ratio is the expected return from a strategy in excess of a benchmark return (here: the return from the long-only strategy) divided by the standard deviation of that excess return:

$$SR = \frac{\mathbb{E}[\tilde{r}]}{\sqrt{\text{Var}[\tilde{r}]}} \quad (14)$$

where \tilde{r} is the excess (log-)return of the signal. This quantifies the expected performance of an investment after adjusting for its involved risk. To wit,

$$\alpha = \underset{\alpha'}{\operatorname{argmax}} SR(\tilde{m}(\alpha')). \quad (15)$$

The typical EWMA time scale of the trend signals \tilde{m} that maximises the Sharpe ratio is $\alpha \approx 1/5$. α could be dissected further for each individual time series but we refrain from doing so as it varies as much over the centuries as over the products. Further, we find (as in [13]) that results are almost invariant for $\alpha \in \{1/4, 1/5, 1/6, 1/7\}$. The sloppiness analysis in Sect 6 will formalise this justification.

Next, it is shown why the fundamentalists' demand imbalances are chosen to be a hyperbolic tangent of the trend signal. Note, albeit, that the hyperbolic tangent is just an example of a function that saturates for relatively large values of the signal. The necessary conditions that such a function should obey have been derived for the original Chiarella model in Ref. [10].

This functional relationship relating the (normalised) returns \tilde{r}_n and the trend signal \tilde{m}_n calculated from those returns is depicted in Fig 6. Its shape is common to all regarded asset classes, serving an ex-post justification of Chiarella's assumptions. Fig 6 shows a slight departure from the hyperbolic tangent for large positive trend signals beyond two standard deviations. This phenomenon has been reported in the literature as trend-reversion [16,37]. However, practitioners usually clip their signals at $\pm 2\sigma$, eliminating this effect. Because such strong trend signals are rare anyway, we deem the tanh an appropriate parametric choice. The returns and trend signals had to be normalised in Fig 6 in order to make them comparable among different assets that may show different levels of volatility and thus differently sized returns.

In order to estimate the parameter γ per instrument, the function

$$h(x) = a + b \tanh(\tilde{\gamma}x + c) \quad (16)$$

is fitted to data from the sets of assets belonging to one asset class. For each product the slope γ is then brought back into its natural units via

$$\gamma = \frac{\tilde{\gamma}}{\sqrt{\text{Var}[\tilde{m}]}} \quad (17)$$

$\text{Var}[\tilde{m}]$ is the variance of the trend signal. This is done for each product.

The γ -values for each product are reported together with all calibration output in Table 1 and the respective standard deviations in Table IV in Appendix A in S1 Appendices.

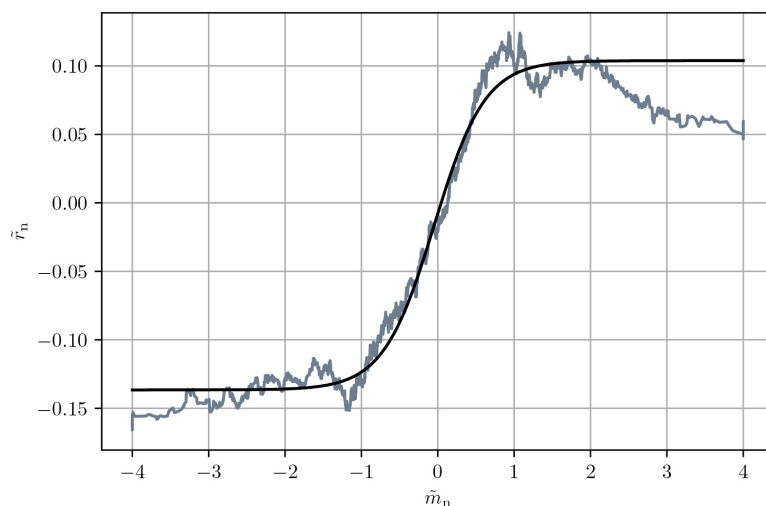


Fig 6. De-drifted and normalised future returns \tilde{r}_n as a function of the trend signal \tilde{m}_n calculated from all returns from currencies against the US dollar. The grey curve is a rolling average over 1000 consecutive points along the abscissa of the real data aggregated over all six FXR pairs, while the solid black line is a parametric fit of Eq (16) over all data.

<https://doi.org/10.1371/journal.pone.0340409.g006>

4.3 Calibration results and excess volatility

In order to improve the robustness of the calibration and because the algorithm generally exhibits difficulties in pinning down σ_V , a multi-step calibration is used. The three step calibration can be summarised as follows:

1. Calibration of model (13) on the de-drifted log-price series \tilde{p} of each asset to get a first estimate of the parameters (after ex-ante estimations of G , α , γ).
2. Calibration of the factor $\Sigma = \langle \frac{\sigma_N}{\sigma_V} \rangle_{a \in \mathcal{A}}$ per asset class \mathcal{A} by maximising the cumulated likelihood of all the assets a in one class (keeping all other parameters fixed).
3. Recalibration of model (13) (or non-linear model: Eqs (4)) per asset but with $\sigma_V = \frac{1}{\Sigma} \sigma_N$ fixed.

Table 1 lists the calibrated parameters from the calibration of the model with the linear fundamentalist demand, system (3); the standard deviations of those parameters are listed in Table IV in S1 Appendices.

The results of Table 1 reveal that indices show the highest level of trend following beyond the long-only bias as the corresponding trend following parameter, β , averages the highest. This is also supported by paper-trading Sharpe ratios that may be calculated for the different asset classes over different time horizons (see Appendix A in S1 Appendices), where the Sharpe on indices clearly dominates. However, it is noteworthy that certain commodities, such as Crude, which trends the most according to our calibration, or sugar, exceed the level of TF by index traders.

One of the highest level of value investing (next to commodities) may be found for bonds whose mean reversion or value parameters consistently score high. This implies that for bonds large mispricings persist shorter than for other asset classes (as the inverse of κ has units of time) for investors realise returns from mispricings comparatively quicker. This is supported by the Sharpe ratios from value investing that are higher on bonds than on the other asset classes. Such swift reactions and mean-reverting corrections may be explained by a very low volatility in fundamental value (further demonstrated by a very small σ_V) due to a low uncertainty – or high consensus – among investors about what the fundamental value is (as there are, e.g., widely used formulas for bond valuation). Perhaps surprisingly, commodities also show very high levels of mean reversion. This simultaneous existence of high values of κ and β suggests a strong presence of both TFs and value investors, possibly leading to frequent alternations between excursions due to the TFs and subsequent mean-reversion; this variability of the price is also reflected in the large price volatility of commodities.

From Table 2 (center column), which lists the calibrated factor Σ between the two noise sources per asset class, we infer that the contribution from noise traders is crucial and much exceeds the noise of the fundamental (log-)value process v .

This may be interpreted as a quantification of the famous excess volatility puzzle, first formalized by Shiller in 1981 [1]; see, e.g., also [38]. Within our framework, this excess volatility is mostly due to excess trading from noise traders, which has been reported on all asset classes studied here [1,39–41]. Trend-following activity, on the other hand, does not contribute much to short-term volatility because the signal is computed over rather long time scales. However, trend-following is responsible for further long term decoupling between price and value.

Table 2. Estimated ratio $\Sigma = \frac{\sigma_N}{\sigma_V}$ between the two noise sources per asset class. The error is given by the standard deviation within one class. Left: linear model, right: non-linear. The value of Σ is found to be nearly identical for the two models. Furthermore, replacing σ_N by σ_P (price volatility) gives almost indistinguishable results.

Asset Class	Σ (linear)	Σ (non-linear)
Indices	3.87 ± 0.61	3.81 ± 0.65
Commodities	3.62 ± 1.62	3.35 ± 1.13
Currencies	5.93 ± 1.10	5.89 ± 1.09
Bonds	13.94 ± 4.46	13.82 ± 4.35

<https://doi.org/10.1371/journal.pone.0340409.t002>

More precisely, we find (comp. Table 2) that the volatility of the price due to noise traders is significantly larger (by a factor 4 to 14) than the volatility of the fundamental value in all cases. This is one of the central results of our study, and justifies the title of this paper. In general, σ_N is indistinguishable from σ_p , the price volatility calculated via the dedrifted log-returns, within the error margins, justifying its usage in determining the excess volatility. This also demonstrates that the noise traders are responsible for the largest fraction of the price volatility.

For indices and commodities, σ_N is typically around four times as large as σ_V , suggesting that prices depart from value due to strong excess trading. Such an amplification factor is compatible with other estimates from the literature, see e.g. [42].

For exchange rates, the ratio between the noise trader volatility and fundamental volatility is higher, with a value around 6, twice as large as the value reported in [42]. The largest ratio is found for bonds, for which we report a ratio of almost 12. However, such a high ratio for bonds is not due to an extreme amount of excess trading on that asset class, which would correspond to a large σ_N that we do not observe, see Table 1. Instead, it stems from the fundamental volatility being particularly small for bonds, as expected since the fundamental value of bonds is expected to be much more stable than the fundamental value of indices. Foreign exchange rates are in this sense intermediate.

As noted above, an important feature of this model and calibration is that it directly outputs a model-implied notion fundamental value, V . The Kalman relations allow for two different notions of fundamental value: The first is the filtered value, which is determined through the dynamical system and the information of *past* prices and values. The second is the smoothed value, which takes both *past and future* information into account. As a result, the filtered value is the value that could have been known to the fundamentalist at the time of trading, while the smoothed value is an ex-post best estimate of what the true fundamental value really was.

As an illustration, the log-price p together with its filtered and smoothed calibrated fundamental values v are shown in Fig 7 (top plot). The blue shaded area indicates one standard deviation of the smoothed value according to the Kalman smoother relations. The bottom plot of Fig 7 provides the same insight but on the de-drifted (log-)price \tilde{p} and value \tilde{v} .

In addition to that, the calibrated parameters of the linear model in Table 1 are such that none of the assets satisfies the bifurcation condition for oscillations as for all assets $\beta\gamma < 1$, ensuring that $\kappa > \alpha(\beta\gamma - 1)$. Similar results are reported in [13]. It was further shown in Chiarella et al. [43] that the limit set being a cycle is a necessary condition for the distribution of trend signals to be bimodal. Since we *do* find that for some assets the distribution of mispricing is bimodal (even with $\beta\gamma < 1$, see Table 3) we are compelled to reject the linear specification of the Chiarella model and turn to the non-linear version, see Eqs (4), which has a much richer phase diagram that is in fact not yet fully explored analytically.

The calibration results for the non-linear model using the unscented Kalman filter with the cubic fundamentalist demand are detailed in Table V in S1 Appendices. As for the linear model, the results are illustrated on the US stock index, which is illustrated in Fig 8 – see Appendix A in S1 Appendices, in which the estimated filtered and smoothed fundamental values from the cubic model are shown with an error bar alongside the de-drifted log-price and the linear comparison. As was also noted in [13], we find a small, often negative value of κ (enhancing trend-like behaviour for small mispricing) and a staunchly positive value of κ_3 , confirming that mean-reverting behaviour only becomes appreciable for large mispricings. The noise ratios Σ are listed in the right column of Tab. 2; they are very similar to those for the linear model, which means that σ_N is estimated similarly in both models (note that for the non-linear model σ_V is fixed to the value obtained from the linear model).

In conclusion of this section, we have shown that it is now possible to calibrate meaningfully our modified versions of Chiarella's model, both linear and non-linear, on *individual* time series, while it was previously only possible to jointly calibrate classes of similar time series, i.e. conduct a joint calibration of an entire asset class instead of an individual contract [13].

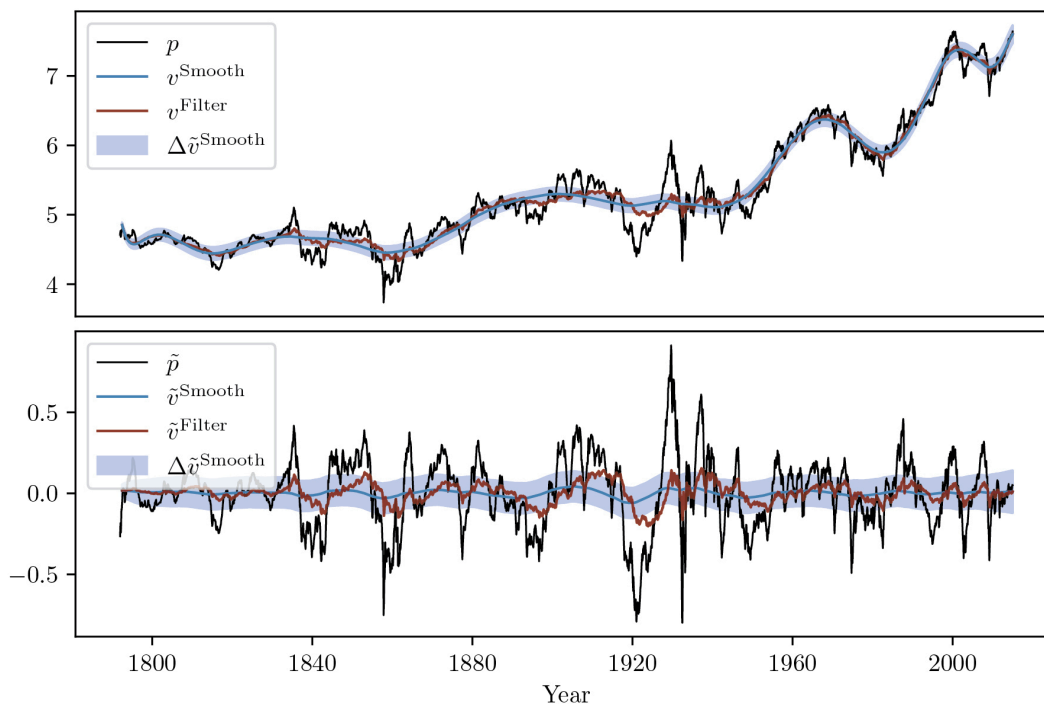


Fig 7. Top: Evolution of the log-price p together with the filtered and smoothed fundamental values v^{Filter} and v^{Smooth} for the US stock index. Δv^{Smooth} denotes the confidence interval obtained as one standard deviation of the value from the Kalman relations. Bottom: same for the de-drifted log-price \tilde{p} and values.

<https://doi.org/10.1371/journal.pone.0340409.g007>

5 Mispricing distribution and bimodality

The asset price dynamics described by Eqs (3) may be further classified by the shape of the resulting mispricing distribution $\rho(\delta)$ (where again $\delta = p - v$), in particular with respect to its uni- or bimodality.

For the classical Chiarella model (without drift) this was numerically investigated by Chiarella et al. in [43], where the conditions for ρ to be bimodal were established. Empirically, bimodal distributions of price distortions independent of Chiarella-type models were reported in [12]. Majewski et al. confirmed this finding for a modified Chiarella model with a constant drift in the fundamental value v for a wider range of asset classes [13]; see also [29] for further discussions.

As alluded to above, the Chiarella model can only generate bimodal mispricing distributions if $\kappa < \alpha(\beta\gamma - 1)$, that is when the limit set is a stable limit cycle [43]. It was then shown in Majewski et al. that in the latter case, the distribution of the trend signal is also bimodal [13]. Still empirical results are at odds with these predictions: not only one finds bimodality of mispricing even when the calibrated parameters are such that $\beta\gamma < 1$, but one does not necessarily find a corresponding bimodality in the trend distribution. The non-linear version of the model does not suffer from these limitations. The mispricing distribution results presented in this section are restricted to the model (4) with a cubic demand function, where bimodality can occur without a bimodal trend signal distribution when $\kappa_3 > 0$ and $\kappa \leq 0$.

For the analytical stationary probability density is unknown, we probe bimodality via Silverman's test for multimodality [44], which tests for a distribution having a minimum of $k + 1$ modes, while the null hypothesis is a distribution with at most k modes. Consequently, we perform the test with $k = 1$, such that a rejection of the null hypothesis is tantamount to rejecting a unimodal distribution, hence suggesting bimodality. More than two modes is not possible within the models investigated here, and there is no such empirical evidence either. A significance level of 0.02 is chosen; this means the null-hypothesis of unimodality is rejected when the p-value is below 0.02.

Table 3. p-Values of the Silverman test for bimodality (column 2 and 3). The null hypothesis is unimodality with a significance level of 0.02. The test is on the mispricing δ , once with the empirical filtered fundamental value and once with the smoothed. \times : acceptance of the unimodal null hypothesis for both empirical time series, \checkmark : twofold rejection (bimodality), \sim : inconclusiveness. Center Right (Numerical): modality test results of simulated time series of the non-linear model, using parameters from Table V in S1 Appendices. Right: Jensen-Shannon (J-S) distances between the empirical and numerical mispricing distributions. A J-S distance of 0.2 corresponds to a Kolmogorov-Smirnov distance of 8% (i.e. the maximum distance between two cumulative distributions functions). For comparison, the J-S distance between LCATTLE and AU is 0.19.

	Filtered	Smoothed	Bimodality (Empirical)	Bimodality (Numerical)	J-S Distance
US	0.364	0.353	\times	\checkmark	0.128
UK	0.011	0.010	\checkmark	\checkmark	0.136
AU	0.001	0.001	\checkmark	\checkmark	0.164
CH	0.031	0.098	\times	\checkmark	0.152
JP	0.001	0.001	\checkmark	\checkmark	0.202
CA	0.067	0.182	\times	\checkmark	0.148
DE	0.019	0.047	\sim	\checkmark	0.200
SUGAR	0.007	0.001	\checkmark	\checkmark	0.304
CORN	0.002	0.001	\checkmark	\checkmark	0.172
LCATTLE	0.639	0.898	\times	\times	0.115
WHEAT	0.706	0.258	\times	\checkmark	0.166
COPPER	0.099	0.040	\times	\checkmark	0.257
NATGAS	0.074	0.313	\times	\checkmark	0.245
CRUDE	0.001	0.785	\sim	\times	0.214
USBND	0.362	0.346	\times	\times	0.149
UKBND	0.023	0.014	\sim	\times	0.136
CHBND	0.233	0.026	\times	\checkmark	0.224
JPBND	0.002	0.001	\checkmark	\checkmark	0.203
AUBND	0.331	0.306	\times	\times	0.158
CABND	0.322	0.341	\times	\times	0.123
DEBND	0.004	0.073	\sim	\checkmark	0.148
CHFUSD	0.077	0.277	\times	\checkmark	0.247
JPYUSD	0.175	0.275	\times	\times	0.138
AUDUSD	0.738	0.117	\times	\checkmark	0.168
GBPUSD	0.890	0.643	\times	\times	0.151
CADUSD	0.855	0.452	\times	\times	0.133
EURUSD	0.024	0.165	\times	\checkmark	0.270

<https://doi.org/10.1371/journal.pone.0340409.t003>

Silverman's test is performed on two different kinds of empirical mispricing series δ for the simple reason that the Kalman relations allow for two different notions of fundamental value v . The first kind of mispricing is derived from the asset log-prices and the filtered fundamental values, and the second kind from the smoothed ones.

Those results are summarised in the first three columns of Table 3, which focus on the empirical mispricing distributions, i.e. those from the real log-prices p and the two types of calibrated log-values v . If both types of empirical mispricings suggest bimodality within the given significance, we mark the series as empirically bimodal (\checkmark). If none of the two reject the null-hypothesis of unimodality, the asset is classified as empirically unimodal (\times). And if one accepts and one rejects the null-hypothesis, the test is marked as inconclusive (\sim).

Table 3 shows that there is in many cases clear empirical evidence for a rejection of unimodality. At the 2% significance level, the empirical mispricing distributions of stock indices are bimodal in nearly half of the cases. For commodities there is further empirical evidence for bimodality in the cases of sugar and corn. In addition to that, the test is inconclusive for Crude Oil, UK Bonds and German Bonds.

It is however known [45] that the Silverman test suffers from a conservatism bias – especially for small samples like we have here – where the null-hypothesis is falsely not rejected. Therefore, we have repeated the test on data generated from numerical simulations of the non-linear model (4) using the calibrated optimal parameters listed in Table V in S1 Appendices. Simulating the model for a total duration of $T = 10^5$ with time increments of $dt = 0.01$, yields $N = 10^7$ data

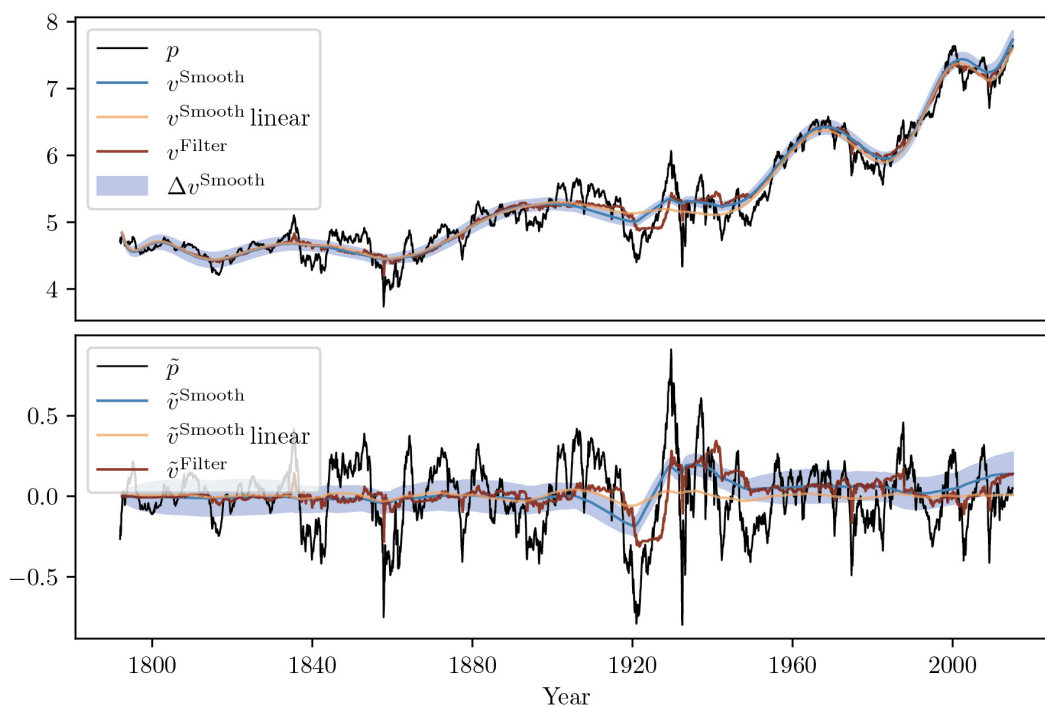


Fig 8. Same as Fig 7 but for the non-linear (cubic) model. The smoothed value from the linear model is given for comparison (orange).

<https://doi.org/10.1371/journal.pone.0340409.g008>

points per product and alleviates the data scarcity problem. Indeed, the test on simulated data yields very accurate results even for weak bimodalities, i.e. when the system is close to the critical bifurcation point. The results are given in the second rightmost column of Table 3. It shows that in all the cases where the test suggests bimodality based on the empirical data, it finds bimodality based on the numerical data, confirming the results as well as the success of the calibration from a different angle. Furthermore, in the cases where the test on empirical data is inconclusive, the test on numerical data leans towards bimodality (except for one asset, crude oil). The numerical study also finds bimodality in some cases where the empirical one did not. Strikingly, the numerical test suggests bimodality for all indices, for most commodities and around half of the bonds and currencies. Note that our flexible definition of the long-term drift, which allows for low-frequency oscillations akin to “business cycles”, tends to lessen any sign of bimodality.

Hence, bimodality appears to be the rule rather than the exception. This is interesting for two reasons: first, it suggests that in many cases assets are more often under- or overpriced than correctly priced. According to numerical data this holds true for all considered indices – a stunning albeit not necessarily surprising result, and in line with the result reported in [12] for the S&P 500, and in [13] for the US and Canadian stock markets. Second, it provides evidence that in such cases prices perform noisy oscillations around value, at least according to the studied model, and even after having accounted for business cycles.

One example of an empirical and its respective numerical mispricing distribution in the case where the distribution is unimodal (live cattle) and another one in the bimodal case (Australian stock index) is given in Fig 9 alongside one of the cases (US stock index) where the empirical and numerical distributions suggest a different type of modality (even though the bimodality in the numerical data is extremely weak). Schmitt et al. [12] report a more pronounced bimodality based on monthly historic S&P 500 prices. This may be due to differences in estimation of fundamental value where they apply Shiller’s method based on discounting the index’ real dividend payments [46], instead of our flexible definition of a time varying drift with low-frequency oscillations. It might also be due to their overestimation the trend signal, which is defined

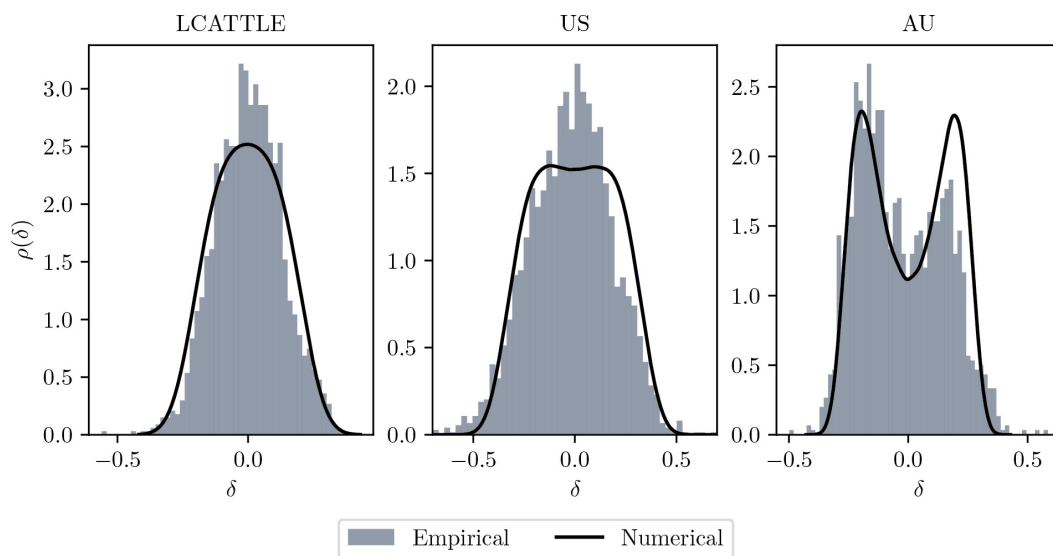


Fig 9. Empirical: mispricing distribution of the log-price and the smoothed value derived from the Kalman filter with parameters from Table V in S1 Appendices for the commodity live cattle (LCATTLE) and the US and Australian (AU) stock index. Numerical: mispricing distribution from a simulation of system (4) with $T = 10^5$, $dt = 0.01$ and the same parameters.

<https://doi.org/10.1371/journal.pone.0340409.g009>

on log-returns directly, instead of returns in excess of some benchmark, such as the long-term drift (as done in our study), or the sectorial or market drifts as often done in practice, esp. for stocks. The extreme bimodalities obtained from simulating several models based on the trend signals of [12] may be a symptom of this.

The comparison of the empirical and numerical distributions show that the mispricings can be captured rather well, considering that non-linear model (4) is a highly schematic model with only three investor types. More importantly, the calibration relies on price trajectories and is unaware of the empirical mispricing distribution, so a good match with the predicted mispricing distribution can be seen as an independent validation of the model. Jensen-Shannon (J-S) distances between the empirical and numerical distributions are given in the rightmost column of Table 3 to quantify the similarity between numerical and empirical distributions for all assets [47].

The J-S distance is the square root of the J-S divergence, which is a symmetrised and smoothed version of the well-known Kullback-Leibler divergence. It has the benefit of being a metric that can be interpreted as a distance measure. Another merit is that it remains well-defined if there are bins with observations for only one of the two discrete distributions. The distance's lower bound of zero means identical distributions, while the upper bound indicating maximal difference is one [47]. As the J-S distance requires the same domain support and binning for both distributions, the domain is chosen to be between the minimum and maximum of the empirical and numerical distribution with the number of bins being the square root of the length of the empirical time series. This is important as the length difference between both series is many orders of magnitude, such that using the length of the shorter series is critical to have an empirical distribution that is not zero for many intervals in the domain.

Before calculating the J-S distances, we have made sure that the mean and the variance of the empirical distributions exactly match those of the numerical simulation. This is only approximately true with the parameters obtained from the calibration, but a small shift of these parameters in the direction of the gradient of the variance allows us to fix this issue with a minimal change of the log-likelihood of the calibration (less than 5% in most cases).

6 Sloppiness analysis

The so-called *sloppiness analysis* is based on the Fisher information matrix, with the aim of gaining some insight about the hierarchy of parameter importance in a model or in a dynamical system, see [17] and [48] for more recent developments. In this context, a parameter (or combination of parameters) is termed *sloppy* if a perturbation in its direction does not change the model output significantly: the model is insensitive to the exact value of that parameter, further implying that it is hard to estimate empirically – parameter uncertainty is high, as encapsulated by the Cramer-Rao bound. Likewise, a *stiff* parameter (or combination of parameters) leads to significant changes in a model's output upon perturbation, implying that it can be estimated well on empirical data.

The crux of the method lies in calculating the eigen-decomposition of the Hessian matrix of a loss function \mathcal{L} , where the loss is a quantifier of the change in model output upon perturbation of the set of model parameters θ to $\theta' = \theta(1 + \Delta)$, where Δ is small. Here, the loss function is defined as the (normalized) L_2 distance between the realizations of observable y measured before and after a perturbation is applied:

$$\mathcal{L}(\theta, \theta') = \frac{1}{T} \sum_{t=1}^T \left(\frac{y_t(\theta) - y_t(\theta')}{\sigma} \right)^2, \quad (18)$$

where t are increments of the simulation time T . Note that the random seed must be fixed in this analysis to only measure the loss due to the parameter perturbation and not due to noise. Further, the beginning of each time series may want to be dropped due to stationarity. σ is the standard deviation over time of $y_t(\theta)$. In the following we will choose the mispricing δ as the observable y .

The model sensitivity to parameter variations may then be regarded through the Hessian of the loss, a.k.a. the Fisher information matrix:

$$\mathcal{H}_{ij} = \frac{d^2 \mathcal{L}(\theta, \theta')}{d \log(\theta_i) d \log(\theta_j)} \Big|_{\theta'=\theta}, \quad (19)$$

where it is standard to take log-derivatives to regard relative parameter changes as parameters usually have inconsistent units and their magnitudes may span multiple decades. Subsequent analysis of \mathcal{H} via its eigendecomposition corresponds to an approximation of the surfaces of constant model deviations as N -dimensional ellipsoids [17]. Numerically, \mathcal{H}_{ij} can be computed using only first derivatives of $y_t(\theta)$, see e.g. [49].

A model is termed *sloppy* if its sensitivity eigenvalue spectrum spans multiple decades in a rather consistent manner, in other words that the eigenvalues decay very quickly with rank, meaning that only very few parameters (or linear combinations thereof) can be identified. Applying this rationale to the mispricing $\delta = p - v$ in the linear model 13 over $N = 6$ parameters $\theta = (\kappa, \beta, \gamma, \alpha, \sigma_N, \sigma_V)$ with optimal parameters from Table 1 and $\Delta = 10^{-2}$, strong evidence for sloppiness is reported as the sensitivity spectra span from five to nine decades (comp. Fig 18 in S1 Appendices). The same is reported for the non-linear model with the additional parameter κ_3 (Fig 21 in S1 Appendices). For the linear model and the asset class commodities, this is depicted in Fig 10 exemplarily.

Next, the eigenvectors of the Hessian in Eq (19) for the linear model are analysed. Since it is not sensible to regard each time series' eigenvectors standalone and because financial time series within one asset class share many key characteristics, we analyse the *average* Hessian within one asset class.

The eigenvectors of the Hessian point in the eigendirections in parameter space. Thus, one can infer from those whether the eigendirections coincide with individual parameter axes, or whether parameter *combinations* determine a system's dynamics and how it reacts to perturbations. Since the Hessian of the loss of the system perturbed from its optimal parameters θ is determined, those eigendirections (descendingly with the magnitude of the corresponding eigenvalue) denote the direction in which the fit is degraded the quickest, i.e. it orders the directions towards which the model dynamics are most sensitive.

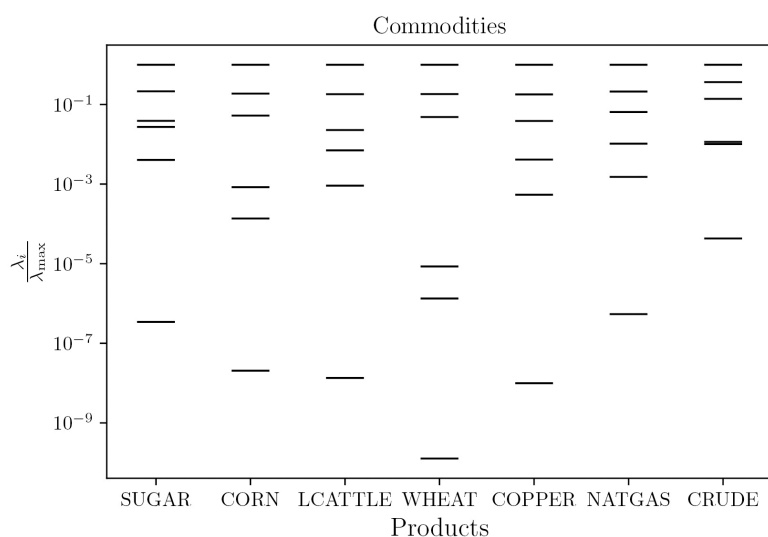


Fig 10. Parameter sensitivity spectrum due to the eigenvalues of Eq (19) for the linear model (13) and the asset class commodities. Per asset the spectrum is rescaled by its maximum eigenvalue for visualisation.

<https://doi.org/10.1371/journal.pone.0340409.g010>

For stock indices, the normalised eigenvectors of the Hessian \mathcal{H} of the linear model perturbed around its optimal parameters (comp. Table 1) and ranked by their eigenvalues are given in Fig 11. Interestingly, the eigenvectors for other asset classes (commodities, currencies, bonds) are qualitatively very similar compared to Fig 11. Overall it is clear at first glance that the eigendirections in parameter space do not generally align with the parameter directions because most

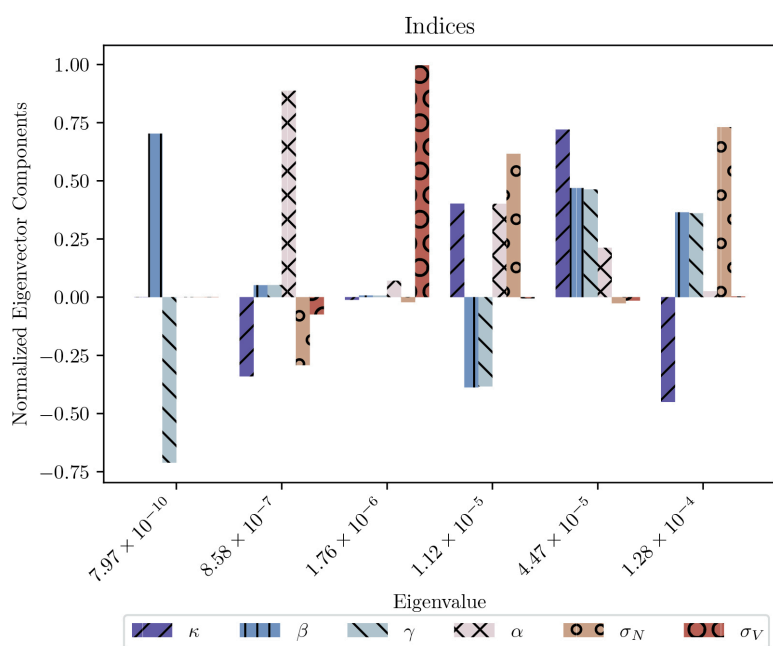


Fig 11. The 6 normalised eigenvectors for stock indices, ranked by eigenvalue magnitude (in increasing order) of the average Hessian for the linear model. The observable in the loss \mathcal{L} (comp. Eq 18) is the price distortion $\delta = p - v$. The total number of observations is $N = 10^4$.

<https://doi.org/10.1371/journal.pone.0340409.g011>

modes are mixed, except for σ_V , which has its own isolated mode for all four asset classes (mode four, corresponding to the fourth largest eigenvalue, in all classes except currencies, comp. Fig 11). Moreover, the eigenvectors of all asset classes show that the contribution of β and γ to the eigendirections is always the same. This is no surprise considering that the local linearisation of the model as well as the bifurcation condition depend solely on the product $\beta\gamma$ and not on the parameters individually (comp. Sect 2.3). More generally, the different eigenvectors (or *modes*) can be interpreted in the context of the model.

- The first mode reflects the quick degradation in quality of fit when $(\sigma_N, \beta, \gamma)$ are perturbed in one direction and κ in the other. The first three parameters control the dispersion between price and value. In particular, when they are increased (decreased), the dispersion increases (decreases) as σ_N is the largest influence on the variance of the price and both β and γ are parameters associated with trend following, also increasing the departure of price from value when increased. Meanwhile, κ has the opposite effect, describing mean reversion towards value. Hence, decreasing (increasing) κ also results in larger dispersion from value. The first mode may thus be termed the *variance mode*. In total, it can be concluded that this mode governs the variance of the price distortion; it is also the stiffest parameter direction in the system to which it is most sensitive and which can be calibrated most reliably (which is confirmed, e.g., by the small errors on σ_N and κ in Appendix A in S1 Appendices). In all other asset classes the first mode looks qualitatively similar. Quantitatively it can additionally be deduced through comparison of the first modes in Fig 18 in S1 Appendices that for currencies the contributions of the parameters associated with trend following, β and γ , are almost negligible, which suggests relatively weak trend following in currency markets, compatible with the results of [3] but not really born out by the analysis of [4]. For commodities the opposite is visible: β and γ – and thus trend following in general – are the most pronounced, a further known fact for after all almost the entire CTA industry was built on this strategy [50].
- The second mode in Fig 11 may be termed the *critical* or *bifurcation mode* as it depends on the parameters occurring in the bifurcation condition (comp. Sect 2.3). The mode shows that an increase (decrease) in mean reversion strength (by κ) accompanied by a simultaneous increase (decrease) in trend following (through $\beta\gamma$ and α) leads to a deterioration of the fit even though there is no direct implication on the level of the price distortion as the effects counterbalance each other. The bifurcation condition – and therewith the overall dynamical state – is however sensitive to such perturbations as the different parameters do not enter the condition equally but in a non-linear way.
- The third mode is similar to the first mode in composition, except for a stronger contribution of α . As in the first mode, σ_N is the dominant contributor. κ seems to be a larger contributor in this mode in all asset classes than on mode one. The relative sign of κ and $\beta\gamma$ are inverted compared to the first mode, as a consequence of the orthogonality condition acting on the different eigenvectors.
- The fourth mode, the *value noise mode*, is the only case in which a parameter direction coincides precisely with an eigendirection. It describes the response of the price distortions to perturbations in σ_V . Its eigenvalue is however two orders of magnitude lower than that of the first mode. This implies that one needs to perturb the system $\sqrt{\lambda_1/\lambda_4} \approx 10$ times as hard in that direction to achieve a comparable variation in model output. This, of course, implies that σ_V is relatively loosely constrained, which makes its estimation ten times harder. At the same time it means that the exact value of σ_V is less relevant compared to σ_N , κ , β and γ .
- The fifth mode can be interpreted as the *trend speed mode* as it is mostly determined by α . Its small associated eigenvalue confirms our previous statement that changes in α (that we have hard-coded to 1/5) do not change the results significantly, neither qualitatively, nor quantitatively (comp. the footnote in Sect 4).
- The sixth and last mode is a pure trend following mode as it only depends on the trend parameters β and γ . It is the only mode in which the two have different orientations. This mode can be interpreted as the *trend saturation mode*. It is the consequence of the breaking of the linear dependence of the trend signal on $\beta\gamma$ as higher order terms of the tanh function become relevant. This mode has the smallest impact as the sloppiest direction for only returns beyond two standard deviations fall in this saturation regime, which is rare, such that the overall influence is comparatively small. Principally,

it shows how an increase (decrease) in β accompanied by a decrease (increase) in γ reduces the quality of fit, which is an immediate consequence of the tanh and its contributions that are $O(\beta\gamma^3)$ and ensures that β and γ can both be estimated and not just their product – albeit not with very high precision.

The non-linear model, Eqs (4), whose sloppy analysis is detailed in Appendix B in [S1 Appendices](#), naturally has one more eigendirection as it has an additional parameter, κ_3 . As it turns out, the mode interpretation is very similar to the linear model. The additional seventh eigendirection can be interpreted as a *value mode*, quantifying the response to perturbations in (κ, κ_3) , and ranks as the sixth mode, whereas the sixth from the linear model becomes the final seventh mode.

7 Conclusion

In this paper the dynamical interplay between trend and value anomalies that pervade (almost) all financial markets was revisited. Specifically, the generalized Chiarella model proposed by Majewski et al. [13] was corrected for its analytical shortcomings, which also impacted its calibration. We proposed a new self-consistent model, in which the stability conditions for the dynamically possible phases exactly match the model dynamics – even for non-zero, arbitrary long term value drift. This was achieved by letting the trend signal and price dynamics, and not only the fundamental value, be drift-dependent. The idea is to define the trend signal on *mispricing* returns rather than on standard returns, removing any long term bias, which should not be considered as part of the trend.

Our model is therefore able to accommodate arbitrary time-dependent drifts, whereas previously only linear drifts were allowed, which we deem unsatisfactory. A calibration scheme adapted to this new model was proposed and implemented. This leads to a notable estimation improvement as it enables one to calibrate the model on individual price time series, whereas previously only asset-class-wide calibrations were possible. As in the literature, this was performed on a model that is linear in the fundamentalist's demand as well as one that is non-linear (cubic). We find that only the non-linear model is consistent with many of the stylised facts, including the bimodality of the mispricing distribution while the trend signal remains unimodal.

One important output of the calibration is the fundamental value of an asset. By proposing a new price vs. value variance estimation technique, the ratio between the noise trader induced volatility and the value volatility could be estimated per asset class. This ratio sheds light on the long-standing excess volatility puzzle as it confirms and quantifies by how much price volatilities are amplified over volatilities in value and that changes in value do not justify the amplitude of price changes whatsoever, putting the rationality and efficiency of prices into question. Our estimate of an amplification of a factor 4 for stock indices is comparable to other estimates from the literature, including Shiller's original paper [1]. Differences in excess volatility per asset class could be qualitatively accounted for. It was possible to separate the variance contribution of noise traders and of fundamental value for each asset individually.

Besides, the distribution of instantaneous mispricings was empirically and numerically analysed for the non-linear model. Statistical tests confirm the existence of bimodality, which have been previously reported [12,13]. In fact much stronger evidence was found than in Majewski et al. [13], especially for stock indices and commodities, while for bonds and currencies less so. This finding shakes the Efficient Market Hypothesis to its core for it suggests that assets are more likely to be mispriced than correctly priced – but it also pinpoints markets (bonds and currencies) that are closer to efficiency.

A Hessian or sloppy analysis allowed for a systematic multi-parameter sensitivity study of the model to small changes in parameter combinations, defining a strong hierarchy in the eigenvalues of the Fisher information matrix. This analysis, on the one hand, justifies why certain parameters are difficult to estimate, but on the other hand also suggests that their exact value may not be crucial to the model's dynamical signature. It may also be interesting from a regulatory standpoint as it can quantify which market contributors and effects can affect the price dynamics most notably. We were able to explain these parameter directions in the context of market perturbation modes.

This research opens the path to several follow-up questions and topics: first, it would be interesting to see how the parameters have evolved over centuries. This may be a difficult endeavour as the dynamical time-scales are on the order of decades, making sequential calibrations difficult in many cases due to data scarcity. However, it could elucidate whether prices and markets have become more efficient with time and whether levels of mispricing have grown or shrunk, see also [37]. Second, an extension of the model towards single stocks would be worthwhile in order to understand their level of mispricing for different economic sectors, and to study the excess volatility puzzle further through our lens and in its original context. We will tackle this question in a subsequent publication. It would also be interesting to repeat this analysis on crypto assets, which, as noted above, do not have a tangible notion of fundamental value. Finally, a more micro-founded model of demand leads to an enhanced version of the Chiarella model that we are currently investigating.

Supporting information

S1 Appendices. Supporting Information to the Paper. The supporting information contains appendices with further analyses, verifications and results that support and underpin the findings of this work.

S2 Code. Computer Algorithm Implementations. This supporting information contains the algorithms used in this study, implemented in Python3.

Acknowledgments

We would like to thank Frank Westerhoff, Julius Bonart, Damien Challet and Doyne Farmer for their insights and useful discussions, as well as Fabian Aguirre-Lopez for mathematical remarks. We also thank the referees for very valuable suggestions that helped us improve the manuscript. This research was conducted within the Econophysics & Complex Systems Research Chair, under the aegis of the Fondation du Risque, the Fondation de l'Ecole polytechnique, the Ecole polytechnique and Capital Fund Management.

Author contributions

Conceptualization: Jutta Kurth, Jean-Philippe Bouchaud, Adam A. Majewski.

Data curation: Adam A. Majewski.

Formal analysis: Jutta Kurth.

Investigation: Jutta Kurth, Jean-Philippe Bouchaud, Adam A. Majewski.

Methodology: Jutta Kurth, Jean-Philippe Bouchaud, Adam A. Majewski.

Project administration: Jean-Philippe Bouchaud.

Supervision: Jean-Philippe Bouchaud, Adam A. Majewski.

Validation: Jutta Kurth.

Visualization: Jutta Kurth.

Writing – original draft: Jutta Kurth, Jean-Philippe Bouchaud.

Writing – review & editing: Jutta Kurth, Jean-Philippe Bouchaud, Adam A. Majewski.

References

1. Shiller R. Do stock prices move too much to be justified by subsequent changes in dividends?. *Am Econ Rev*. 1981;71:421–36.
2. Asness CS, Moskowitz TJ, Pedersen LH. Value and momentum everywhere. *J Finance*. 2013;68(3):929–85. <https://doi.org/10.1111/jofi.12021>

3. Lempérière Y, Deremble C, Seager P, Potters M, Bouchaud J-P. Two centuries of trend following. *JOIS*. 2014;3(3):41–61. <https://doi.org/10.21314/jois.2014.043>
4. Hurst B, Ooi YH, Pedersen LH. A century of evidence on trend-following investing. *JPM*. 2017;44(1):15–29. <https://doi.org/10.3905/jpm.2017.44.1.015>
5. Bouchaud JP, Bonart J, Donier J, Gould M. *Trades, quotes and prices: Financial markets under the microscope*. Cambridge University Press; 2018.
6. Gabaix X, Koijen R. In search of the origins of financial fluctuations: The inelastic markets hypothesis. National Bureau of Economic Research, Inc.; 2021.
7. Bouchaud J-P. The inelastic market hypothesis: A microstructural interpretation. *Quant Finance*. 2022;22(10):1785–95. <https://doi.org/10.1080/14697688.2022.2068052>
8. van der Beck P, Bouchaud JP, Villamaina D. Ponzi funds. *arXiv preprint arXiv:240512768*. 2024.
9. Farmer JD. Market force, ecology and evolution. *Ind Corp Change*. 2002;11(5):895–953. <https://doi.org/10.1093/icc/11.5.895>
10. Chiarella C. The dynamics of speculative behaviour. *Ann Oper Res*. 1992;37(1):101–23. <https://doi.org/10.1007/bf02071051>
11. Lux T, Marchesi M. Scaling and criticality in a stochastic multi-agent model of a financial market. *Nature*. 1999;397(6719):498–500. <https://doi.org/10.1038/17290>
12. Schmitt N, Westerhoff F. On the bimodality of the distribution of the S&P 500's distortion: Empirical evidence and theoretical explanations. *J Econ Dyn Control*. 2017;80:34–53. <https://doi.org/10.1016/j.jedc.2017.05.002>
13. Majewski AA, Ciliberti S, Bouchaud J-P. Co-existence of trend and value in financial markets: Estimating an extended Chiarella model. *J Econ Dyn Control*. 2020;112:103791. <https://doi.org/10.1016/j.jedc.2019.103791>
14. Scholl MP, Calinescu A, Farmer JD. How market ecology explains market malfunction. *Proc Natl Acad Sci U S A*. 2021;118(26):e2015574118. <https://doi.org/10.1073/pnas.2015574118> PMID: 34172576
15. Black F. Noise. *J Finance*. 1986;41(3):528–43.
16. Bouchaud JP, Ciliberti S, Lempérière Y, Majewski A, Seager P, Ronia KS. Black was right: Price is within a factor 2 of value. *arXiv q-finPM preprint arXiv:171104717*. 2018.
17. Gutenkunst RN, Waterfall JJ, Casey FP, Brown KS, Myers CR, Sethna JP. Universally sloppy parameter sensitivities in systems biology models. *PLoS Comput Biol*. 2007;3(10):1871–8. <https://doi.org/10.1371/journal.pcbi.0030189> PMID: 17922568
18. Delli Gatti D, Gusella F, Ricchiuti G. Endogenous vs exogenous fluctuations: Unveiling the impact of heterogeneous expectations. *Macroecon Dyn*. 2025;29. <https://doi.org/10.1017/s1365100525100345>
19. Gusella F, Ricchiuti G. Endogenous cycles in heterogeneous agent models: A state-space approach. *J Evol Econ*. 2024;34(4–5):739–82. <https://doi.org/10.1007/s00191-024-00870-w>
20. Lux T. Estimation of agent-based models using sequential Monte Carlo methods. *J Econ Dyn Control*. 2018;91:391–408. <https://doi.org/10.1016/j.jedc.2018.01.021>
21. Lux T. Bayesian estimation of agent-based models via adaptive particle Markov Chain Monte Carlo. *Comput Econ*. 2021;60(2):451–77. <https://doi.org/10.1007/s10614-021-10155-0>
22. Kukacka J, Barunik J. Estimation of financial agent-based models with simulated maximum likelihood. *J Econ Dyn Control*. 2017;85:21–45. <https://doi.org/10.1016/j.jedc.2017.09.006>
23. Lux T, Zwickels RC. Empirical validation of agent-based models. *Handbook of computational economics*. Elsevier; 2018. p. 437–88.
24. ter Ellen S, Hommes CH, Zwickels RCJ. Comparing behavioural heterogeneity across asset classes. *J Econ Behav Org*. 2021;185:747–69. <https://doi.org/10.1016/j.jebo.2019.12.013>
25. Hommes C. Behavioral and experimental macroeconomics and policy analysis: A complex systems approach. *J Econ Lit*. 2021;59(1):149–219. <https://doi.org/10.1257/jel.20191434>
26. Kyle AS. Continuous auctions and insider trading. *Econometrica: Journal of the Econometric Society*; 1985. p. 1315–35.
27. Lux T. The socio-economic dynamics of speculative markets: Interacting agents, chaos, and the fat tails of return distributions. *J Econ Behav Org*. 1998;33(2):143–65. [https://doi.org/10.1016/s0167-2681\(97\)00088-7](https://doi.org/10.1016/s0167-2681(97)00088-7)
28. Boswijk HP, Hommes CH, Manzan S. Behavioral heterogeneity in stock prices. *J Econ Dyn Control*. 2007;31(6):1938–70. <https://doi.org/10.1016/j.jedc.2007.01.001>
29. Lux T. Can heterogeneous agent models explain the alleged mispricing of the S&P 500?. *Quant Finance*. 2021;21(9):1413–33. <https://doi.org/10.1080/14697688.2021.1909744>
30. Strogatz SH. *Nonlinear dynamics and chaos: With applications to physics, biology, chemistry, and engineering*. CRC Press; 2018.
31. Guckenheimer J, Holmes P. *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*. Springer Science & Business Media; 2013.
32. Lorenz HW. *Nonlinear dynamical economics and chaotic motion*. vol. 334. Springer; 1993.
33. Särkkä S, Svensson L. *Bayesian filtering and smoothing*. Cambridge University Press; 2023.

34. Baum LE, Petrie T, Soules G, Weiss N. A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Ann Math Statist.* 1970;41(1):164–71. <https://doi.org/10.1214/aoms/1177697196>
35. Dempster AP. Maximum likelihood from incomplete data via the EM algorithm. *J R Stat Soc Ser B.* 1977;39:1–38.
36. Chen C-F. The EM approach to the multiple indicators and multiple causes model via the estimation of the latent variable. *J Am Stat Assoc.* 1981;76(375):704–8. <https://doi.org/10.1080/01621459.1981.10477708>
37. Schmidhuber C. Trends, reversion, and critical phenomena in financial markets. *Phys A: Stat Mech Applic.* 2021;566:125642. <https://doi.org/10.1016/j.physa.2020.125642>
38. LeRoy SF. Excess volatility. *The new Palgrave dictionary of economics*, 2nd ed. Palgrave Macmillan; 2006. p. 13.
39. Deaton A, Laroque G. On the behaviour of commodity prices. *Rev Econ Stud.* 1992;59(1):1. <https://doi.org/10.2307/2297923>
40. Meese RA, Rogoff K. Empirical exchange rate models of the seventies: Do they fit out of sample?. *J Int Econ.* 1983;14(1–2):3–24.
41. Campbell JY, Shiller RJ. Yield spreads and interest rate movements: A bird's eye view. *Rev Econ Stud.* 1991;58(3):495. <https://doi.org/10.2307/2298008>
42. Wehrli A, Sornette D. The excess volatility puzzle explained by financial noise amplification from endogenous feedbacks. *Sci Rep.* 2022;12(1):18895. <https://doi.org/10.1038/s41598-022-20879-0> PMID: 36344614
43. Chiarella C, He X-Z, Wang D, Zheng M. The stochastic bifurcation behaviour of speculative financial markets. *Phys A: Stat Mech Applic.* 2008;387(15):3837–46. <https://doi.org/10.1016/j.physa.2008.01.078>
44. Silverman BW. Using Kernel density estimates to investigate multimodality. *J R Stat Soc Ser B: Stat Methodol.* 1981;43(1):97–9. <https://doi.org/10.1111/j.2517-6161.1981.tb01155.x>
45. Mammen E, Marron JS, Fisher NI. Some asymptotics for multimodality tests based on kernel density estimates. *Probab Th Rel Fields.* 1992;91(1):115–32. <https://doi.org/10.1007/bf01194493>
46. Shiller RJ. *Irrational exuberance*. Princeton: Princeton University Press; 2000.
47. Endres DM, Schindelin JE. A new metric for probability distributions. *IEEE Trans Inform Theory.* 2003;49(7):1858–60. <https://doi.org/10.1109/tit.2003.813506>
48. Quinn KN, Abbott MC, Transtrum MK, Machta BB, Sethna JP. Information geometry for multiparameter models: New perspectives on the origin of simplicity. *Rep Prog Phys.* 2022;86(3):10.1088/1361-6633/aca6f8. <https://doi.org/10.1088/1361-6633/aca6f8> PMID: 36576176
49. Naumann-Woleske K, Knicker MS, Benzaquen M, Bouchaud JP. Exploration of the parameter space in macroeconomic models. *Routledge international handbook of complexity economics*. Routledge; 2024. p. 190–207.
50. Elaut G, Erdős P. Trends' signal strength and the performance of CTAs. *Financ Anal J.* 2019;75(1):64–83. <https://doi.org/10.1080/0015198x.2018.1547052>