Nonlinear dynamics in asset pricing: the role of a sentiment index

Nonlinear Dynamics manuscript No.
(will be inserted by the editor)

Nonlinear dynamics in asset pricing: the role of a sentiment index

Giovanni Campisi · Silvia Muzzioli · Alberto Zaffaroni

Received: date / Accepted: date

Abstract This paper aims to contribute to the literature on the role of sentiment indices in heterogeneous asset pricing models. A new sentiment index in financial markets is proposed in which transactions take place between two groups of fundamentalists with divergent perceptions of fundamental value. It is assumed that the proportion of fundamentalists in the two groups depends on the sentiment index. After examining the analytical properties of the deterministic discrete dynamical system, stochastic components are added to the expectations of fundamentalists. First, the study measures the performance of the model in reproducing the stylized facts of financial data relying on the S&P 500 index. Second, the forecasting power of the model to predict the daily prices of the S&P 500 index is examined. For this purpose, the forecasting accuracy of the proposed dynamical model, where the sentiment index is explicitly modelled, is compared with a model where the sentiment index is not taken into account. In this case, the predictions are obtained by means of a machine learning technique (lasso regression). The results show that the sentiment index is important in explaining the stylized facts of financial returns and in forecasting prices.

Keywords Sentiment index · Nonlinear dynamical systems · Asset pricing · Stability and bifurcation analysis · Forecasting · Machine learning

An earlier version of this paper entitled "Fundamentalists heterogeneity and the role of the sentiment indicator" was published in March 2020 as Working Paper no. 167 of the Marco Biagi Department of Economics at the University of Modena and Reggio Emilia (Italy).

G. Campisi (corresponding author)
Marco Biagi Department of Economics, University of Modena and Reggio Emilia, Viale Jacopo Berengario 51, 41121, Modena, Italy
E-mail: giovannicampisi6@gmail.com

S. Muzzioli
Marco Biagi Department of Economics, University of Modena and Reggio Emilia, Viale Jacopo Berengario 51, 41121, Modena, Italy
E-mail: silvia.muzzioli@unimore.it

A. Zaffaroni
Marco Biagi Department of Economics, University of Modena and Reggio Emilia, Viale Jacopo Berengario 51, 41121, Modena, Italy
E-mail: alberto.zaffaroni@unimore.it

1 Introduction

Several studies have shown that stock returns in financial markets exhibit certain empirical regularities such as volatility clustering, asymmetry and excess of kurtosis (see e.g. Cont (2001), Lux and Marchesi (2000)). In order to replicate these stylized facts, many financial models have been developed, in particular, Heterogeneous Agent Models (HAMs) (see e.g. Brock and Hommes (1997), Tramontana et al. (2013), Anufriev et al. (2020), Schmitt et al. (2020)). Recently, evidence has been presented of the importance of investor sen-

timent in trading decisions and in predicting stock returns (Baker and Wurgler (2006), Corredor et al. (2013)), as it represents an opinion market indicator that traders can rely on in whole or in part in decision-making.

In line with these studies on the role of sentiment indices, the present study proposes a new sentiment index in a financial market model with two groups of fundamentalist traders adopting a homogeneous trading strategy, while displaying heterogeneous beliefs about the fundamental value of the asset, and a market maker engaging in transactions out of equilibrium. In addition to fundamentalists, HAMs usually consider chartists or noisy traders, showing that they are the main cause of complex dynamics in the model. In contrast with these studies, the present model considers only fundamentalists and assumes that type-1 fundamentalists believe that the fundamental value (F_1) is greater than the fundamental value (F_2) assumed by type-2 fundamentalists. As a result, when the stock price is above the two fundamentals, both types of fundamentalists tend to sell, though type-2 fundamentalists sell more aggressively than type-1 fundamentalists. Conversely, when the price falls below the two fundamental values, type-1 fundamentalists buy more aggressively than type-2 fundamentalists.

A number of studies consider models with only fundamentalists, such as Naimzada and Ricchiuti (2009), De Grauwe and Kaltwasser (2012), Cavalli et al. (2018), and Campisi and Muzzioli (2020). They assume that one type of fundamentalist underestimates the true fundamental value (the pessimist), whereas the other type overestimates it (the optimist). Based on the above considerations, the present model assumes that agents act on the basis of bounded rationality and may not be aware of the true fundamental value of the asset but still attempt to estimate it. Unlike previous studies, it is assumed only that the fundamental value estimated by type-2 fundamentalists is always lower than the fundamental value estimated by type-1 fundamentalists. This assumption differs from the one adopted in e.g. Cavalli et al. (2018), where fundamentalists are either optimists or pessimists. In particular, these fundamentalists are characterized by biased beliefs, expecting a constant price above (optimists) or below (pessimists) fundamental value. The present model generalizes the assumption by allowing both types of fundamentalists to underestimate or overestimate the true fundamental price. As a result, if fundamentalists heavily underestimate (overestimate) the true fundamental value, the estimated fundamental values may well be below (above) the true fundamental value. As explained in Section 3, this assumption allows for different scenarios.

This paper explores the role of investor sentiment in the dynamic of generic asset pricing in financial markets. For this purpose, the analysis consists of four main steps. First, the sentiment index responsible for the switching mechanism of agents is introduced. The sentiment index is meant to measure average sentiment in the market and to capture the difference between the average perception of the fundamental value and the market price.
By measuring in relative terms the distance between the price P_t and the two fundamental prices F_1 and F_2, the paper focuses on a switching mechanism based on average price perception.
Second, it analyses the deterministic skeleton of the resulting one-dimensional discrete dynamical system, highlighting the main characteristics of the model both analytically and numerically. Third, it considers a stochastic version of the present model adding a stochastic component to the expectations of both types of fundamentalists in order to reproduce the stylized facts and to identify the best fits to real data. For this purpose, the daily data of the S&P 500 index are considered as the benchmark for analysis. In particular, the paper follows Westerhoff and Franke (2012), Schmitt et al. (2020) and Anufriev et al. (2020) focusing mainly on the statistical properties of the returns, the analysis of volatility clustering and the leverage effect. Finally, the model is used to forecast the daily prices of the S&P 500 index, comparing its performance against a model that does not explicitly consider the sentiment index. In this case, the predictions rely on a machine learning technique (lasso regression) and compare the results with the model including the sentiment index.

This study is intended to contribute to the literature in three ways. First, it introduces a new sentiment index taking into account the difference between the average perception of fundamental value and the market price. Unlike De Grauwe and Kaltwasser (2012), the sentiment index considers the average perception of the fundamental value of the market, and is thus independent of the individual gains of traders. Moreover, the mathematical expression of the proposed index makes it possible to carry out an in-depth investigation of the local stability of the model both analytically and numerically.
Second, it demonstrates the ability of the model to capture certain features of financial markets. In particular, the sentiment index outlined in this study, that is based on the average perception of the fundamental price, captures the excess of kurtosis, volatility clustering, long-range dependence and the asymmetric relation between stock index returns and change in volatility in S&P 500.
Third, based on the available knowledge, most parts of the nonlinear dynamical models in discrete time have focused their attention on the analytical aspects, replicating the stylized facts and the parameter calibration with real data. Only a few authors have explored the forecasting performance of their nonlinear dynamical models (see Chiarella et al. (2012), Boswijk et al. (2007)). In particular, none of these studies has employed machine learning techniques in comparing the model's performance. The present findings show that the model with a sentiment index outperforms the model without a sentiment index, confirming the importance of sentiment indices in financial markets for investor decision-making.

The remainder of the paper is organized as follows. Section 2 reviews the related literature and highlights the contribution of this paper. Section 3 describes the model and outlines the new sentiment index. Section 4 examines the analytical and numerical results of the deterministic component of the model. In Section 5 outlines a statistical analysis of the stochastic model, in order to replicate the stylized facts of the S&P 500 index. Section 6 analyses the forecasting power of the dynamical model where the sentiment index is explicitly modelled against a model where the sentiment index is indirectly estimated by a machine learning technique (lasso regression). Section 7 concludes.

2 Contribution to the literature

The present study is intended to contribute to the literature on sentiment indices and heterogeneous agent models. A number of studies have focused on how investor sentiment affects the financial markets. In general, the role of the sentiment index in determining stock prices is analysed from two points of view: a theoretical and an empirical one.
The theoretical strand of research provides evidence that stock prices and investor trading decisions are affected by noisy traders relying on unpredictable changes in sentiment (see De Long et al. (1990), Barberis et al. (1998)). In particular, De Long et al. (1990) stress that the effects of investor sentiment vary according to the behaviour of the investors dominating the market. Barberis et al. (1998) present a Markov-switching model with homogeneous traders, analysing the effect of investor sentiment on asset pricing.
On the other hand, the empirical strand of research has examined sentiment indices in relation to their ability to predict returns, as in the case of Corredor et al. (2013). In order to investigate the role of sentiment indices, both linear (Baker and Wurgler (2006)) and non-linear models (Boswijk et al. (2007)) have been employed. With regard to non-linear models, certain studies have analysed the role of investor sentiment in asset pricing, providing evidence that investor sentiment is one of the main determinants of asymmetry in stock returns (Jawadi et al. (2018)).

This paper is closely related to current HAMs with respect to heterogeneity and the switching behaviour of agents. From the seminal work of Brock and Hommes (1997, 1998), a number of other researchers have successfully modelled the price dynamics of a generic asset using an HAM framework to replicate the stylized facts observed in financial markets. The main idea behind these models is that asset pricing is the result of the interaction of various investors operating under different trading rules. In the Adaptive Belief System outlined by

Brock and Hommes (1997, 1998), investors adapt their beliefs over time by choosing from different predictors of future values of endogenous variables.
In addition to heterogeneity, a further assumption of HAMs is that traders act on the basis of bounded rationality in the sense that they do not possess all the relevant information to determine fundamental market value. Instead, they tend to rely on a simple rule of thumb (see for example Chiarella et al. (2009), Tramontana et al. (2015), Brianzoni and Campisi (2020)).

Combining these two strands of research, the present study analyses the effects of a non-linear sentiment index in an HAM when the number of traders varies according to this sentiment index. In particular, the model incorporates an endogenous switching mechanism generated by the sentiment index able to generate greed and fear scenarios.

This is believed to be the first paper comparing the effect of a sentiment index on the analytical properties of the model and on the empirical calibration with market data. The mathematical expression of the proposed index makes it possible to carry out an in-depth analytical and numerical investigation of the local stability of the model. In particular, all the fixed points of the model are computed, and the stability conditions are established mathematically.

The model is also able to capture certain features of financial markets. In particular, the sentiment index presented in the study is able to capture excess of kurtosis, volatility clustering and the asymmetric volatility returns relationship in S&P 500. Finally, the sentiment index plays a key role in the future decisions of investors. In comparing the present model with one without a sentiment index, the present model is more accurate in forecasting the future prices of S&P 500.

Campisi and Muzzioli (2020) have proposed an alternative sentiment index that relies on the RAX index of Elyasiani et al. (2018) constructed with option prices. In their study, the authors focus on adapting the index in an agent-based model showing that their index behaves like the RAX index and can capture the most important features of financial markets. In contrast, in the present study the focus is on the role of the sentiment index in investor decisions and its power in predicting future prices. In particular, only a few studies dealing with heterogeneous agents have explored the forecasting performance of their models (see Chiarella et al. (2012), Boswijk et al. (2007)). Moreover, none of these studies has considered only fundamentalists in their models and employed machine learning techniques in comparing the performance of the model.

3 The model

The model outlined consists of two types of fundamentalists and a market maker. It is assumed that fundamentalists are not aware of the underlying fundamental value and rely on their beliefs to place orders in the market, whereas the market maker adjusts price imbalances with respect to excess demand. Naimzada and Ricchiuti (2009) and De Grauwe and Kaltwasser (2012) analyse financial models in which agents adopt trading strategies according to a rule based on different fundamental values. In particular, they assume that fundamentalists are not aware of the true fundamental value and attempt to estimate it. Moreover, they assume that one type of fundamentalist always underestimates fundamental value (the pessimist) while the other always overestimates it (the optimist). The present model incorporates both types of fundamentalist, and assumes that type-2 fundamentalists only underestimate fundamental value (F_2) with respect to type-1 fundamentalists (F_1), i.e. $F_2 < F_1$. Based on this assumption, the model allows for cases in which both types of fundamentalist may be optimists or pessimists. This implies that if $P_t > F_1 > F_2$, then both types of fundamentalist sell, but type-2 fundamentalists sell more than type-1 (Fear predominance region). On the other hand, if $F_2 < P_t < F_1$, type-2 fundamentalists sell, whereas type-1 fundamentalists buy. This is similar to the bull-and-bear regime described in Day and Huang (1990) or Tramontana et al. (2009) (Fear and Greed mixed predominance region). Finally, if $P_t < F_2 < F_1$, both types of fundamentalist buy, but type-2 fundamentalists buy less than type-1 fundamentalists (Greed predominance region). The excess demand function of both types of

fundamentalists is assumed to be equal to:

$$D_t^{f_i} = \lambda(F_i - P_t) \tag{1}$$

where $D_t^{f_i}$ is the excess demand function for type-i fundamentalist; $i = 1, 2$, λ is a positive parameter and indicates how fast fundamentalist responds to the misalignment of the price from the fundamental value (F_i, with $i = 1, 2$).

As a result, when the price is below F_2, type-2 fundamentalists buy less than type-1 fundamentalists, whereas when the price is above F_1, type-2 fundamentalists sell more than type-1 fundamentalists.

Price adjustments are operated by market-makers. The price-setting rule of the market-maker is given by:

$$P_{t+1} = P_t + \left(w_1 D_t^{f_1} + w_2 D_t^{f_2}\right) \tag{2}$$

where w_i is the proportion of fundamentalists of type i, $i = 1, 2$ and $w_1 + w_2 = 1$ and D^{f_i}, with $i = 1, 2$, is the excess demand of type-i fundamentalist.

The proportion of the two types of fundamentalist has a strong impact on the dynamics of Eq. (2). In particular, it is assumed that the number of fundamentalists n_i, with $i = 1, 2$, is fixed[1] and equal to $2N$; i.e.:

$$n_1 + n_2 = 2N \tag{3}$$

As a result, the proportions of type-1 and type-2 agents are equal to:

$$w_1 = \frac{n_1}{2N} \qquad \text{and} \qquad w_2 = \frac{n_2}{2N} \tag{4}$$

To highlight the role of the sentiment index in the model, the numbers n_1, n_2 of type-1, and type-2 fundamentalists are allowed to vary according to a market sentiment index. In this respect, fundamentalists make assumptions about the fundamental value by relying on the following sentiment index:

$$\eta = \frac{(\frac{F_1+F_2}{2}) - P_t}{(F_1 - P_t)^2 + (F_2 - P_t)^2} \tag{5}$$

The sentiment index is meant to measure in relative terms the average perception of fundamental value. In particular, the numerator measures the proximity of the price to average fundamental value $\frac{F_1+F_2}{2}$. This difference is normalized by the sum of the two distances (the denominator). The range of the index depends on the distance between F_1 and F_2.

The dynamic of the sentiment index is incorporated in the number of the two types of fundamentalists. Following Lux (1995), let the number of type-1 fundamentalists be given by:

$$n_1 = n_2 - 2N\eta \tag{6}$$

Accordingly, the number of type-2 fundamentalists is:

$$n_2 = n_1 + 2N\eta \tag{7}$$

The value of η is responsible for switching between the two groups of fundamentalists. In particular, the number of type-1 (type-2) fundamentalists is greater than the number of type-2 (type-1) fundamentalists when $\eta < 0$ ($\eta > 0$). The two groups pair up when $\eta = 0$.

The closer the price to the fundamental value F_1, the greater the increase of the proportion of type-1 fundamentalists, since their performance was better than that of the other group in forecasting the equilibrium price (and in the market, fear predominates since investors expect the price to fall). On the other hand, the closer the price to the fundamental value F_2, the greater the increase of the proportion of type-2 fundamentalists, since they performed better than the other group in forecasting the equilibrium price (and greed predominates in the market since investors expect the price to increase).

Eqs. (3) and (6) result in $n_1 = N(1-\eta)$ and $n_2 = N(1+\eta)$. In order to have $n_1, n_2 \geq 0$, it is necessary to impose $-1 \leq \eta \leq 1$, which entails $(F_1 - F_2) \geq 0.5$[2]. In the case

[1] We highlight the fact that the sum of the number of fundamentalists, $n_1 + n_2$, is fixed and equal to $2N$. The total number of fundamentalists, $2N$, is constant and equal to 1 (it is assumed that $N = 0.5$), while the number of fundamentalists n_i, with $i = 1, 2$, varies with time according to the sentiment index η.

[2] This is a conservative assumption since 0.5 may represent a multiple of the minimum tick size (e.g. for S&P 500 the

$F_1 - F_2 = 0.5$, the range of η is the interval $[-1, 1]$. If $F_1 - F_2 > 0.5$, then η ranges in an interval $[-\alpha, \alpha]$ with $0 < \alpha = \frac{1}{2(F_1 - F_2)} < 1$. Assuming $w_1 = \frac{n_1}{2N}$ and $w_2 = \frac{n_2}{2N}$ the proportions of type-1 and type-2 agents, which are given by:

$$w_1 = \frac{(1-\eta)}{2} \qquad \text{and} \qquad w_2 = \frac{(1+\eta)}{2} \tag{8}$$

the result is that $0 < w_i < 1$ with $i = 1, 2$. As a result, the switching mechanism is not only due to the information included in the index, but also depends on the distance between the two fundamental values. The farther apart the fundamental values F_1 and F_2 are spaced, the lower the number of agents who change their mind about fundamental value.

After this outline of the main characteristics of the new index, it is now necessary to briefly examine its behaviour. The sentiment index is positive for $P_t < \overline{F} = \frac{F_1+F_2}{2}$, and attains its maximum when P_t is equal to F_2, since for prices lower than F_2 both types of investors buy, and for prices $P_t < \frac{F_1+F_2}{2}$ and higher than F_2 type-1 investors buy with a higher intensity compared to the selling behaviour of type-2 investors, giving rise to an overall buy signal.

On the other hand, the index is negative for $P_t > \overline{F}$ and attains its minimum when $P_t = F_1$ since for prices higher than F_1 both types of investors sell, and for prices $P_t > \frac{F_1+F_2}{2}$ and lower than F_1, type-1 investors buy with a lower intensity compared to the selling behaviour of type-2 investors, giving rise to an overall sell signal.

Finally, if $P_t = \frac{F_1+F_2}{2}$ the sentiment index is equal to zero, indicating neither buy nor a sell signal.

To better appreciate the role of the sentiment index outlined in this study and how it affects the dynamics of the model, in the following a comparison is provided with models with only fundamentalists based on different switching mechanisms. In particular, two models are considered based on different switching mechanisms: the models of De Grauwe and Kaltwasser (2012), and Campisi and Muzzioli (2020).

minimum tick is equal to 0.25 index points), in order to have a sufficient difference between the two fundamental values F_1 and F_2.

Whereas De Grauwe and Kaltwasser (2012) use the well-known Brock and Hommes (1998) discrete choice model, the present study adopts a switching mechanism relying on the methodology of Lux (1995). Unlike the sentiment index in the present study, their model includes the intensity of choice in the mathematical formula of the sentiment index. In De Grauwe and Kaltwasser (2012) fundamentalists switch to a more profitable strategy according to personal gains. On the other hand, the sentiment index in the present study considers the average perception of the fundamental value of the market, and is thus independent of the individual gains of market traders.

Moreover, the sentiment index in the present study essentially differs from the one proposed by Campisi and Muzzioli (2020), a study that concentrates on the risk asymmetry in the market and models a fear index (i.e. periods of declining prices impact volatility more than periods of rising prices). The index does not depend on the distance between the two fundamental values. Moreover, their model, that is highly non-linear, is more suitable for a numerical study than for analytical results. On the other hand, the sentiment index in the present study is more general, behaving symmetrically for fear and greed scenarios. It depends on the distance between the two fundamental values. Finally, it is possible to carry out an in-depth investigation of the model both analytically and numerically.

3.1 Properties of the sentiment index

A mathematical description of the new sentiment index is provided in the following. Note that the analysis of the sentiment index performed in this section is intended to give details of its behavior at time t, as shown in Figure 1. In particular, the price P_t and P_{t+1} are assumed to be continuous variables and that the sentiment index is continuous.

Given $F_1 > F_2 > 0$, consider the sentiment index

$$\eta = \frac{\overline{F} - P_t}{(F_1 - P_t)^2 + (F_2 - P_t)^2}$$

where $\overline{F} = (F_1+F_2)/2$. The result is that η is increasing for $P_t < F_2$ and $P_t > F_1$, and decreasing for $F_2 < P_t < F_1$. It tends to 0 when P_t tends to $\pm\infty$.

Moreover, the following holds:

$$\eta = \frac{1}{2}\frac{F_1 + F_2 - 2P_t}{(F_1 - P_t)^2 + (F_2 - P_t)^2}$$

and

$$\frac{d\eta}{dP_t}(P_t) = \frac{1}{2}\frac{-2(F_1 - P_t)^2 - 2(F_2 - P_t)^2 + }{((F_1 - P_t)^2 + (F_2 - P_t)^2)^2}$$
$$\frac{+(F_1 + F_2 - 2P_t)(2F_1 + 2F_2 - 4P_t)}{((F_1 - P_t)^2 + (F_2 - P_t)^2)^2} =$$
$$= \frac{(F_1 + F_2 - 2P_t)^2 - (F_1 - P_t)^2 - (F_2 - P_t)^2}{((F_1 - P_t)^2 + (F_2 - P_t)^2)^2}$$

the numerator of which boils down to

$$2P_t^2 - 2(F_1 + F_2)P_t + 2F_1F_2 = 2(P_t - F_1)(P_t - F_2),$$

a quadratic expression which vanishes at $P = F_1$ and $P = F_2$, being, respectively, the global maximum and global minimum point of η. The maximum and the minimum values of η depend on F_2 and F_1, with

$$\eta(F_1) = \frac{1}{2(F_2 - F_1)} < 0 \quad \text{and} \quad \eta(F_2) = \frac{1}{2(F_1 - F_2)} > 0$$

which are unbounded as F_1 moves closer and closer to F_2. The index assumes its values in the interval $(-1, 1)$, which is true[3] for $F_1 - F_2 > 0.5$. Fig.(1) shows three different shapes of the sentiment index function depending on the distance of the two fundamentals. In panel (a) of Fig.(1) the sentiment index attains values in the interval $[-1, 1]$. In panels (b) and (c) of Fig.(1) $-1 < \eta < 1$ and the interval of values attained by the sentiment index shrinks as $F_1 - F_2$ increases.
From Fig.(1) two main features of the index are emphasized:

1. the fear and greed region (i.e. the region between the maximum and the minimum points of the sentiment index depicted in panels (a)-(c) of Fig.(1)) expands with the distance of the two fundamentals. If $F_1 - F_2 = 0.5$ the range of η is the interval $[-1, 1]$. The farther apart the fundamental values F_1 and F_2 are spaced, the smaller the number of agents who change their mind about fundamental value;

2. for large values of $F_1 - F_2$ the role of the sentiment index in the switching mechanism decreases significantly, implying a smaller number of agents switching from one group to the other.

[3] In the next section the restriction on the difference $F_1 - F_2$ will be made clear also from a mathematical point of view when the local stability of fixed points is analysed.

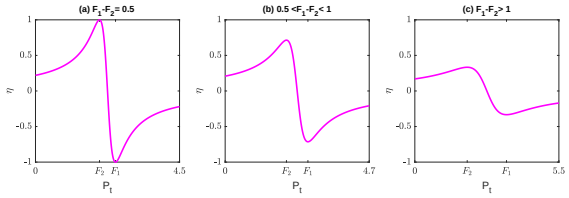

Figure 1 The sentiment index function. This figure demonstrates the role of the sentiment index when the distance $F_1 - F_2$ increases. The greater the distance $F_1 - F_2$, the lower its role as indicator of market sentiment.

The next section analyses the role of the sentiment index in the present model.

4 Dynamics of the deterministic model

The final map: Based on the considerations in Section 3, the first-order non-linear discrete dynamical equation is obtained that describes the price evolution over time. It takes the following form:

$$P_{t+1} = P_t + \Big\{\Big[\frac{n_2}{2N} - \frac{(\frac{F_1+F_2}{2}) - P_t}{(F_1 - P_t)^2 + (F_2 - P_t)^2}\Big]\lambda(F_1 - P_t) + \\ + \Big[\frac{n_1}{2N} + \frac{(\frac{F_1+F_2}{2}) - P_t}{(F_1 - P_t)^2 + (F_2 - P_t)^2}\Big]\lambda(F_2 - P_t)\Big\} \tag{9}$$

In this section, a stability analysis of Map (9) is performed in order to describe the local properties of the fixed points of the model under investigation. By means of a straightforward computation, the following may be stated:

Lemma 1 *(**Existence of fixed points**): Let $P_t > 0$ $\forall t$ and $F_1 > F_2$. The Map (9) allows for up to three*

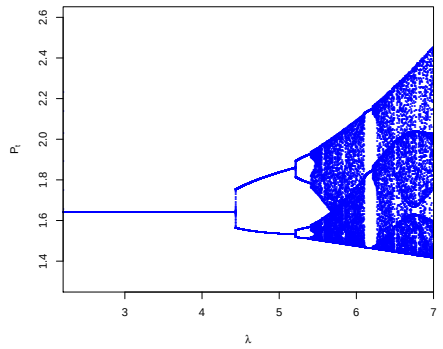

Figure 2 Stability of fixed point P_1^* on varying the intensity of trading parameter λ in the interval $[2.2, 7]$ and other parameters fixed as follows: $n_2 = 0.47$, $F_1 = 3$, $F_2 = 1.2$, $N = 0.5$, i.c. $P_0 = 1.65$.

fixed points

$$P_1^* = \frac{F_1 + F_2}{2}$$
$$P_2^* = \frac{F_1 + F_2}{2} - \frac{\sqrt{(F_1 - F_2)(F_2 - F_1 + 1)}}{2}$$
$$P_3^* = \frac{F_1 + F_2}{2} + \frac{\sqrt{(F_1 - F_2)(F_2 - F_1 + 1)}}{2}$$

Moreover, the fixed points belong to the interval $[F_2, F_1]$. *The fixed points* (P_2^*, P_3^*) *exist provided that* $0 < F_1 - F_2 < 1$.

If $F_1 - F_2 = 1$ *the three fixed points collapse to the fixed point* $P_1^* = \dfrac{F_1 + F_2}{2}$.

It should be noted that the fixed points (P_2^*, P_3^*) exist when the fundamental values of the two traders are not too far apart. To establish the local stability of the fixed point P_1^*, reference may be made to Proposition 1 below:

Proposition 1 *Let* $F_1 - F_2 > 1$ *then there exists only one fixed point for the Map (9),* P_1^*, *which is locally asymptotically stable provided that*

$$\lambda < \frac{2(F_1 - F_2)}{F_1 - F_2 - 1}$$

For $\lambda = \frac{2(F_1 - F_2)}{F_1 - F_2 - 1}$, P_1 *loses stability via flip bifurcation and a stable two-cycle arises.*

Proof In order to analyse the local stability of the fixed point P_1^*, it is necessary to calculate the derivative of price evaluated in the equilibrium[4]. Consider Map (9). Rewriting n_1, n_2 in terms of η and then solving for P_t the result is:

$$P_{t+1} = g(P_t) = P_t + \lambda \frac{\bar{F} - P_t}{(F_1 - P_t)^2 + (F_2 - P_t)^2} [2P_t^2 + - 2(F_1 + F_2)P_t + F_1^2 + F_2^2 + \frac{1}{2}(F_2 - F_1)] \tag{10}$$

The expression inside square brackets, denoted by $M(P_t)$, is of second order in P_t and is always positive under our assumptions, since it holds $\Delta = (F_1 - F_2)(1 - F_1 + F_2) < 0$. Thus the sole fixed point is $P_1^* = (F_1 + F_2)/2$.

Calculating the first derivative of (10), the result is:

$$\frac{dg}{dP_t}(P_t) = 1 + \frac{[-\lambda M(P_t) + \lambda(\bar{F} - P_t)4(P_t - \bar{F})]((F_1 - P_t)^2 + (F_2 - P_t)^2)^2 - \lambda(\bar{F} - P_t)M(P_t)4(P_t - \bar{F})}{[(F_1 - P_t)^2 + (F_2 - P_t)^2]^2} \tag{11}$$

In the present case, the derivative at the fixed point P_1^* is equal to

$$\frac{dP_{t+1}}{dP_t}(P_1^*) = 1 - \lambda M(\overline{F}) = 1 - \lambda \frac{F_1 - F_2 - 1}{F_1 - F_2}$$

and applying the condition for local stability

$$-1 < \frac{dP_{t+1}}{dP_t} < 1$$

the result of Proposition (1) is found. □

As long as the reactivity of the fundamentalists is lower than the critical value found in Proposition (1), that is $\lambda < \frac{2(F_1 - F_2)}{F_1 - F_2 - 1}$, the asset price converges on the locally stable fixed point P_1^*. When $\lambda = \frac{2(F_1 - F_2)}{F_1 - F_2 - 1}$, the fixed point P_1^* is locally unstable and the price converges on a stable two-cycle. In this case, the price goes through alternate periods when it is overvalued and periods when it is undervalued. By further increasing the value of λ, the two-cycle also becomes unstable, giving rise to a chaotic attractor (see Fig.(2)).

[4] see Gandolfo (2010) and Medio and Lines (2001).

In addition, mention should be made of three main features of the model. First, in the present model coexistence is achieved only under certain conditions, i.e. a low degree of heterogeneity among fundamentalists. Second, under certain conditions concerning the reactivity parameter of fundamentalists (λ), the central fixed point (P_1^*) is stable. Finally, for all combinations of parameters, the central fixed point, $P_1^* = \frac{F_1+F_2}{2}$ is symmetric with respect to the two fixed points (P_2^*, P_3^*).

In the case of a low degree of heterogeneity ($F_1 - F_2 < 1$), the following dynamics apply:

Property 1 Assume $F_1 \neq F_2$. Let c_{min}, c_{max} be the local minimum and the local maximum of the Map (9); C_{min} and C_{max} are their iterates respectively, that is $C_{min} = P_{t+1}(c_{min})$ and $C_{max} = P_{t+1}(c_{max})$, then there exist two disjoint invariant[5] intervals $I = [P_{t+1}(c_{min}), P_{t+1}(C_{min})]$ and $J = [P_{t+1}(c_{max}), P_{t+1}(C_{max})]$ such that:

1. **Fear or Greed scenario:** if $P_{t+1}(C_{min}) < P_1^*$ and $P_{t+1}(C_{max}) > P_1^*$, then Map (9) has two coexisting attractors, P_2^* and P_3^*;
2. **Fear and Greed scenario:** for $P_{t+1}(C_{min}) = P_{t+1}(C_{max}) = P_1^*$ the two attractors merge and a contact bifurcation occurs.

The present analysis makes it possible to summarize the most interesting cases with regard to the values of the parameters in the model. In particular, all the possible cases are resumed in the following:

Proposition 2 *Consider Map (9), moreover, let $F_1 > F_2$, then:*

1. *If $F_1 - F_2 > 1$, hence there exists one positive fixed point, $P_1^* = \frac{F_1+F_2}{2}$ and it is globally asymptotically stable provided that $\lambda < \dfrac{2(F_1 - F_2)}{F_1 - F_2 - 1}$,*
2. *Let $0 < F_1 - F_2 < 1$ there exist three positive fixed points mentioned in Proposition 1. The fixed point P_1^* is always unstable, while P_2^*, P_3^* are stable provided that $\lambda < \dfrac{1}{1 - F_1 + F_2}$.*

[5] A set $I \subseteq \mathbb{R}_+$ is positively (negatively) invariant if $T^i(I) \subseteq I$ ($T^i(I) \supseteq I$) $\forall i \in \mathbb{Z}_+$. Moreover, I is invariant when it is both positively and negatively invariant. Finally, a closed and positively invariant region is called trapping.

Proof In order to study the local stability of P_2^* and P_3^* the same steps are followed as in Proof of Proposition 1.

The quadratic expression $M(P_t)$ in (10) vanishes at the points found in Proposition 1: $P_2^*, P_3^* = \bar{F} \pm \frac{\sqrt{\Delta}}{2}$, since $\Delta = (F_1 - F_2)(1 - F_1 + F_2)$ is positive by assumption.

Evaluating the derivative of (10) at $P_{2,3}^*$, and recalling that $M(P_{2,3}^*) = 0$, we obtain:

$$\begin{aligned}\frac{dg}{dP_t}(P_{2,3}^*) &= 1 + \lambda - \frac{4(\bar{F} - P_{2,3}^*)^2}{(F_1 - P_{2,3}^*)^2 + (F_2 - P_{2,3}^*)^2} \\ &= 1 - \lambda \frac{4\frac{\Delta}{4}}{(F_1 - P_{2,3}^*)^2 + (F_2 - P_{2,3}^*)^2}\end{aligned} \tag{12}$$

from which we see that $\frac{dg}{dP_t}(P_{2,3}^*) < 1$. In order to have $\frac{dg}{dP_t}(P_{2,3}^*) > -1$, it must hold:

$$\begin{aligned}\lambda\Delta &< 2((F_1 - P_{2,3}^*)^2 + 2(F_2 - P_{2,3}^*)^2) \\ &= 2(2(P_{2,3}^*)^2 - 2P_{2,3}^*(F_1 + F_2) + F_1^2 + F_2^2 \\ &= 2[2(\bar{F}^2 + \frac{\Delta}{4} \pm \bar{F}\sqrt{\Delta}) - 4\bar{F}(\bar{F} \pm \frac{\sqrt{\Delta}}{2})F_1^2 + F_2^2] \\ &= \Delta - 4\bar{F}^2 + 2F_1^2 + 2F_2^2\end{aligned} \tag{13}$$

from which we derive:

$$\begin{aligned}\frac{\Delta}{2}(\lambda - 1) &< F_1^2 + F_2^2 - 2\bar{F}^2 \\ &= F_1^2 + F_2^2 - \frac{(F_1 + F_2)^2}{2} \\ &= \frac{1}{2}(F_1^2 + F_2^2 - 2F_1F_2) \\ &= \frac{1}{2}(F_1 - F_2)^2\end{aligned} \tag{14}$$

and finally:

$$\begin{aligned}\lambda - 1 &< \frac{(F_1 - F_2)^2}{(F_1 - F_2)(1 - F_1 + F_2)} = \frac{F_1 - F_2}{1 - F_1 + F_2} \\ \lambda &< \frac{1}{1 - F_1 + F_2}\end{aligned} \tag{15}$$

This concludes the proof. □

Moreover, the bifurcation conditions of the three fixed points of the Map (9) are provided in the following:

Proposition 3 (Bifurcation conditions of fixed points)

Let P_1^*, P_2^*, P_3^* the fixed points of the Map (9) as defined in Lemma 1.

1. *Assume $F_1 - F_2 > 1$ and let $\lambda^* = \frac{2(F_1-F_2)}{F_1-F_2-1}$ the flip bifurcation value. Then the unique fixed point P_1^* is stable for $\lambda < \lambda^*$ and unstable for $\lambda > \lambda^*$. For $\lambda = \lambda^*$, P_1^* exhibits a flip bifurcation.*
2. *Assume $0 < F_1 - F_2 < 1$ and let $\lambda^{**} = \frac{1}{1-F_1+F_2}$ the flip bifurcation value. Then the fixed points P_2^*, P_3^* are stable for $\lambda < \lambda^{**}$ and unstable for $\lambda > \lambda^{**}$. For $\lambda = \lambda^{**}$, P_2^*, P_3^* exhibit a flip bifurcation. The fixed point P_1^* is always unstable.*

In order to complete the analysis, in the following a graphical analysis is shown to shed light on the dynamics occurring in a neighbour of the fixed points reported in Proposition 2.

Figures (3) and (4) present two scenarios generated by the model showing the coexistence of attractors. In this case, depending on the initial condition, P_0, the price trajectory converges on the greed or fear attractor (Fig. 3(a-b) and (c-d), respectively). In Fig. (4) a homoclinic bifurcation is found. It should be noted that the complex dynamics generated by the model when the two attractors merge give rise to a single complex attractor.

An interesting situation is described in Fig. (5), delineating the different bifurcation scenarios. Both the fear and greed scenarios generate symmetric dynamics with respect to the number of fundamentalists. The greater the proportion of one type of fundamentalist, the more complex the price dynamics (see Figure (5) greed scenario and fear scenario). With regard to the greed scenario shown in Fig. (5), it should be noted that the model converges on a stable fixed point only if the proportion of one of the two fundamentalists is sufficiently large: otherwise, complex dynamics emerge. It is interesting to compare the dynamics in the fear scenario with those in the greed scenario in Fig. (5).

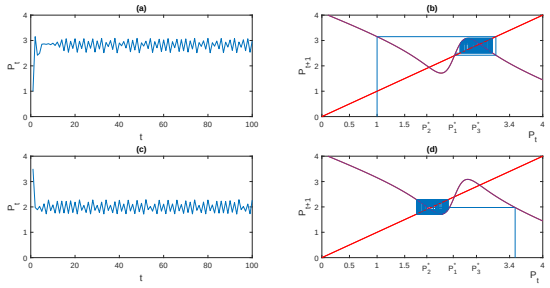

Figure 3 Coexistence of attractors for parameters $\lambda = 1.8$, $F_1 = 2.7$, $F_2 = 2.1$, $N = 0.5$, $n_2 = 0.5$. In (a) and (b) (greed scenario) an i.c. $P_0 = 1$ generates a trajectory converging to the attractor P_3^*. While in (c) and (d) (fear scenario) an i.c. $P_0 = 3.5$ generates a trajectory converging to the stable attractor P_2^*.

The dynamics in the fear scenario are similar to those in the greed scenario if $n_2 \in [0.2, 0.8]$, while for values of n_2 outside this interval more complicated dynamics arise in the fear scenario. Finally, the fear-and-greed scenario, it emerges when the distance between the two fundamental values is greater than the distance of the fundamental values in both the fear and the greed scenario.

To sum up, the model is able to generate different intricate price dynamics. First, the map can exhibit the coexistence of attractors. In addition, this model is able to reproduce the scenario in which there is one attractor that can be globally asymptotically stable provided that the conditions relating to the parameters established in Proposition (3) hold. Second, the price trajectory is quite different in the three scenarios, the fear-and-greed scenario, the fear scenario, and the greed scenario. In particular, both the fear-and-greed scenario and the fear scenario exhibit symmetric price dynamics. Moreover, for extreme values of n_2 (with a slight variation in the value of the parameter of reaction of fundamentalists, λ) more intricate dynamics emerge. However, the greed scenario allows for stability only when the proportion of one of the two fundamentalists is sufficiently large.

In the next section a stochastic version of the model is outlined in order to test its ability to reproduce the stylized features of the S&P 500 index.

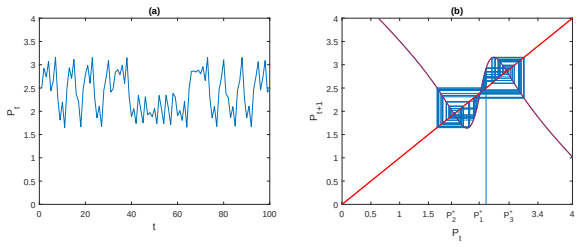

Figure 4 Homoclinic bifurcation scenario. In (a) and (b) two complex attractors merge for $n_2 = 0.5$, $N = 0.5$, $F_1 = 2.7$, $F_2 = 2.1$, $\lambda = 2.1$, i.c. $P_0 = 2.5$.

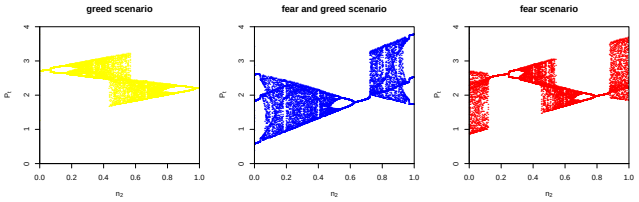

Figure 5 The three scenarios. Greed scenario (yellow diagram) with parameters $\lambda = 1.8$, $F_1 = 2.7$, $F_2 = 2.2$, $N = 0.5$, i.c. $P_0 = 1.2$. Fear and greed scenario (blue diagram) with parameters $\lambda = 2.1$, $F_1 = 2.6$, $F_2 = 1.75$, $N = 0.5$, i.c. $P_0 = 1.9$. Fear scenario (red diagram) with parameters $\lambda = 2.2$, $F_1 = 2.6$, $F_2 = 1.95$, $N = 0.5$, i.c. $P_0 = 2.7$.

5 Statistical properties of the stochastic model

After this outline of the new sentiment index, its performance in reproducing stylized facts is tested. The analysis carried out in this section is inspired by the work of Schmitt et al. (2020), Anufriev et al. (2020), Westerhoff and Franke (2012). For this purpose, noise is added to each of the demand components and examine the dynamics of the stochastic model presented in this study ($SM1$). In detail, the demand of type-1 fundamentalists becomes:

$$D_t^{f_1} = \lambda(F_1 - P_t) + \epsilon_t^{f_1} \qquad \epsilon_t^{f_1} \sim N(0, (\sigma)^2) \tag{16}$$

and, similarly, the demand of type-2 fundamentalists is given by:

$$D_t^{f_2} = \lambda(F_2 - P_t) + \epsilon_t^{f_2} \qquad \epsilon_t^{f_2} \sim N(0, (\sigma)^2) \tag{17}$$

where σ is a positive parameter representing the standard deviation of the normal random variables that are assumed to be equal for both types of fundamentalists. To assess how well the simulated data match the real data in terms of statistical and qualitative properties, the S&P 500 index (S&P 500) is used as a benchmark for comparison. For the present analysis 5464 daily observations from 1 June 1998 to 14 February 2020 were used. Each simulation was run for 5464 time periods and the first 1000 were dropped to wash out the initial effect of the results. The parameters and noisy demand functions used are shown in Table(1).

Table 1 Parameter setting and initial values

F_1	F_2	n_2	N	λ	σ	P_0
2700	2000	0.5	0.5	0.1	0.1	2200

5.1 Properties of the returns

Using log-returns of the time series (defined as $r_t = ln(P_t) - ln(P_{t-1})$), it is intended to capture the main stylized features observed from the S&P 500 index: negative skewness, excess kurtosis, and non-normality of the returns. A closer look at Table 2 confirms that the model presents the same sign of the S&P 500 skewness, that is, the distribution of the returns of both time series is more left-skewed than the normal distribution. This implies that negative returns have a greater impact than positive returns. The kurtosis level is greater than 3 for both SM and S&P 500 models, indicating that extreme events are more frequent than those occurring in the normal distribution.

As a further complement, to confirm the non-normality of the returns the Jarque-Bera test (see Table (2))is performed. In particular, the null hypothesis that returns follow a normal distribution at the 1% significant level is tested. The value of $J - B = 1$ indicates that the Jarque-Bera test rejects the null hypothesis at the 1% significance level, since the test statistic, J-B statistic, is greater than the critical value, which is 5.8461.

5.2 Volatility clustering and long-range dependence

This section analyses the properties of the returns and shows that the model can reproduce two other critical stylized features of the S&P 500 index: volatil-

Table 2 Summary statistics of returns. The table reports the summary statistics including mean, standard deviation (sd), skewness (Skew), kurtosis (Kurt), minimum and maximum value, J-B test (1 if normality is rejected, 0 otherwise) and J-B statistic of S&P 500 and SM.

	Mean	sd	Min	Max	Skew	Kurt	J-B	J-B statistic
S&P500	$207 * 10^{-6}$	0.012	-0.095	0.109	-0.237	11.085	1	14932
SM	$36 * 10^{-7}$	0.059	-0.324	0.285	-0.076	5.063	1	974.53

ity clustering and long-range dependence. With volatility clustering, the market goes through high volatility periods alternating with low volatility periods. In Figs.(6)(a)-(b), the autocorrelation functions (ACFs) of returns (blue line), absolute returns (red line), and squared returns (yellow line) against the lags q are plotted. Volatility clustering is evident for S&P 500 and SM, respectively. In fact, the persistence of the ACFs of returns for both absolute and squared returns is evident (both ACFs display a positive and significant sign even after 50 lags), whereas returns do not exhibit significant autocorrelation.

Moreover, to confirm this stylized feature, an estimate of the power component in the ACFs of absolute returns is provided. The literature converges on the insight that the empirical autocorrelations are similar to a power law. Then, in addition the speed of the ACF of absolute returns decays is measured, following He and Zheng (2016) and Cont (2001):

$$corr(|r_t|, |r_{t+q}|) \simeq \frac{\zeta}{q^d} \tag{18}$$

with $corr(|r_t|, |r_{t+q}|)$ indicating the ACF of absolute returns, q is the number of lags, ζ is the ACF of absolute returns with lag one, and d is the parameter indicating the decay of the ACFs. The parameter d is estimated by employing non-linear least squares. The empirical evidence agrees with a value of d ranging in the interval $[0.2, 0.4]$. Table (3), shows that the value of d of the SM model is in line with S&P 500.

As stressed by He and Zheng (2016), volatility clustering is an indicator of long memory, but it does not necessarily lead to long memory. Long-range dependence or long memory may be defined for stationary processes as the significant correlation of far distant observations (Hauser (1997)). To account for this evidence, the current analysis tests the hypothesis of long-range dependence in the volatility measured by the time series of

Table 3 Persistence of ACFs of absolute returns. For each return series (S&P500, SM) we estimate $corr(|r_{t+q}|, |r_t| \simeq \zeta/q^d$ with nonlinear least squares and report d, ζ, q and R^2.

	S&P500 $corr(\lvert r_{t+q}\rvert, \lvert r_t\rvert)$	SM $corr(\lvert r_{t+q}\rvert, \lvert r_t\rvert)$
d	0.3073^{***}	0.3868^{***}
ζ	0.4700^{***}	0.3426^{***}
q	200	200
R^2	0.7578	0.7009

*** significant at 1%.

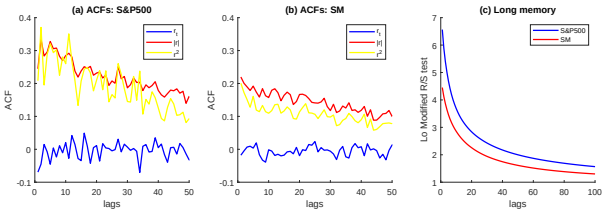

Figure 6 Volatility clustering and long-range dependence. Panels (a)-(b) plot the ACFs of returns (blue line), absolute returns (red line) and squared returns (yellow line) for S&P500 and SM respectively. Panel (c) plots the Lo modified R/S statistic of the absolute returns.

$|r_t|$. To this end, the Lo-modified range over standard deviation or R/S statistic (also called re-scaled range) (see Lo (1991)) is computed given by:

$$\mathcal{Q}_n \equiv \frac{1}{\hat{\sigma}_n(q)} \left[\max_{1 \le k \le n} \sum_{j=1}^{j=k} (r_j - \bar{r}_n) - \min_{1 \le k \le n} \sum_{j=1}^{j=k} (r_j - \bar{r}_n) \right] \tag{19}$$

where $\hat{\sigma}_n(q)$ is the sample variance:

$$\hat{\sigma}_n(q) = \frac{1}{n} \sum_{j=1}^{j=n} (r_j - \bar{r}_n) + \frac{2}{n} \sum_{j=1}^{j=q} \left(1 - \frac{j}{q+1}\right) \sum_{i=j+1}^{i=n} (r_i - \bar{r}_n)(r_{i-j} - \bar{r}_n)$$

In the square brackets of Eq. (19), the maximum and the minimum of the partial sums of the first k differences between r_t and their sample mean $\bar{r}_t$, are shown respectively. Fig. 6(c) shows the R/S statistic of the absolute returns for the first 100 lags. Given that for lags $q \le 50$, the Lo modified R/S statistic is outside the 95% critical interval $[0.809, 1.862]$, it is found that long memory exists for SM and S&P 500 models.

5.3 The asymmetric relation between stock index returns and changes in volatility

The empirical evidence converged on the insight of an asymmetric relation between stock index returns and

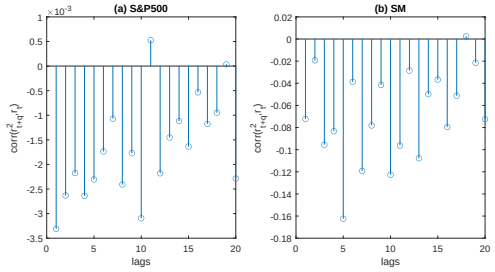

Figure 7 The leverage effect. This figure plots the degree of leverage effect measured by $corr(r^2_{t+q}, r_t)$.

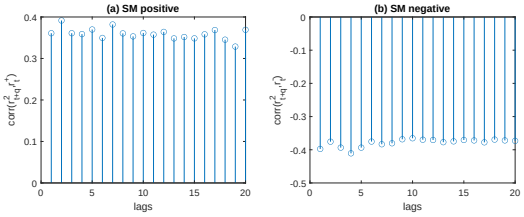

Figure 8 Volatility asymmetry. Panels (a)-(b) plot the degree of leverage effect measured by $corr(r^2_{t+q}, r^+_t)$ and $corr(r^2_{t+q}, r^-_t)$ respectively for SM.

changes in volatility, that is, negative returns have a greater impact on the volatility index than positive returns. In particular, two main hypotheses may explain this asymmetric relation: the leverage effect (Black (1976) and Christie (1982)), and feedback effect hypotheses (French et al. (1987), Campbell and Hentschel (1992), Bekaert and Wu (2000)). According to the leverage effect hypothesis, when returns are negative, the leverage ratio of the firm increases, and the firm's debt exceeds total equity. As the firm's equity is more exposed to the overall risk, the volatility of the equity should increase in turn. On the other hand, the feedback hypothesis states that the asymmetric relation is based on changes in conditional volatility, implying a change in the stock market price (for further details see Badshah (2013) and Bekiros et al. (2017)). For this purpose, the asymmetric returns-volatility relation is assessed following Cont (2001), and He and Zheng (2016) graphing the $corr(r^2_{t+q}, r_t)$ as a function of lag q in Fig. (7). In this case, volatility is found to be negatively correlated with stock index returns for the S&P 500 (Fig. 7(a)), and SM (Fig. 7(b)), and in fact $corr(r^2_{t+q}, r_t)$ starts from negative values in all cases, suggesting that negative returns are associated with higher contemporary volatility and followed by higher short-term volatility.

Finally, in order to evaluate the performance of the model, returns are divided into positive (r^+_t) and negative (r^-_t) and $corr(r^2_{t+q}, r^+_t)$ and $corr(r^2_{t+q}, r^-_t)$ are shown in the graph as a function of lag q for SM (Fig. 8(a-b)). As shown in Fig. (8), SM is able to detect the asymmetric volatility-returns relationship.

6 Forecasting prices: a comparison between the present model and a model without the sentiment index

To evaluate the model specification under the sentiment index, data from the S&P 500 are used. The out-of-sample predictability of the model incorporating the sentiment index is compared with that of the same model without a sentiment index. In the model without the sentiment index, the parameters in Table 1 are used to compute the excess demand functions of both types of fundamentalists $D^{f_1}_t$ and $D^{f_2}_t$ as defined in (16) and (17), with P_t being the S&P 500 price series. In Eq. (2), w_1 and w_2 are estimated by using lasso regression. Instead, in the model with the sentiment index, w_1 and w_2 are computed by means of Eq. (8).

In order to compare the forecasts at time t+1 of the two models (with and without the sentiment index) with the market prices at time t+1 of the S&P 500, two metrics are employed: the test error rate and the accuracy. The test error (or equivalently, the mean squared error (MSE)) is defined as:

$$test\ error = \frac{1}{m}\sum_{j=0}^{m-1}[P_{t+1+j} - P_{t+j}]^2$$

where $j = 0, ..., m-1$ and $m = 5464$ observations. The accuracy of the model is given by:

$$accuracy = 1 - test\ error$$

Before comparing the predictive power, the lasso regression is reviewed in detail. This method, also called the penalized method, appends a penalty to the original loss function. This regularization approach has the

advantage of reducing variance. Moreover, with lasso regression, some of the coefficients may be estimated to be exactly zero, and for this reason the lasso regression belongs to the category of shrinkage methods.
To provide a description of the lasso regression method, the conditional expectation takes the following form:

$$g(P_t; w) = P_t + w' D_t \tag{20}$$

where $w = (w_1, w_2)$ is the vector containing the proportion of fundamentalists as defined in (4) and $D_t = (D_t^{f_1}, D_t^{f_2})$ is the vector of the excess demand functions of both types of fundamentalists as defined in (16) and (17).

The estimation procedure consists of minimizing the following objective function:

$$\mathcal{L}(w; \rho) = \mathcal{L}(w) + \phi(w; \rho) \tag{21}$$

where

$$\mathcal{L}(w) = \frac{1}{T} \sum_{t=1}^{T} (P_{t+1} - g(P_t; w))^2 \tag{22}$$

is the objective function of the traditional least squares, where P_t and P_{t+1} are the prices at time t and at time $t+1$ of the S&P 500. $\phi(w; \rho)$ is the penalty term which makes it possible to regularize the model, i.e. to control for its complexity. In particular, in the case of the lasso regression, the following functional form of $\phi(w; \rho)$ is considered:

$$\phi(w; \rho) = \rho \|w\|_k \tag{23}$$

where $\rho > 0$ is the tuning parameter. This parameter serves to control the relative impact of the regression coefficient estimates. When $\rho = 0$, the penalty function (23) has no effect in the objective function (21) and the same results of the least squares estimates are found. When $\rho \to \infty$, the coefficient estimates approach zero. The notation $\|w\|_k$ denotes the l_k norm of a vector, and is defined as $\|w\|_k = \sum_{i=1}^{2} |w_i|^k$. The $k = 1$ corresponds to the lasso regression.

The tuning parameters are adaptively optimized using k-fold cross-validation. In particular, the dataset is randomly divided into k folds of approximately equal size. First, the model is trained on $k-1$ subsets, then tested on the remaining one subset. Finally, the test error (i.e., the mean squared error) is calculated. This procedure is repeated k times, obtaining k estimates of the test error. The k-fold cross-validation estimate is computed by averaging these values. Following the literature on machine learning applications to finance, it was decided to set $k = 10$ (see James et al. (2013) or Gu et al. (2020) for further details). The forecasting results for the two models are shown in Table 4.

Table 4 Test error rate and accuracy for the model with and without the sentiment index

Model	test error	accuracy
Model without the sentiment index	0.4173	0.5827
Model with the sentiment index	0.3894	0.6106

Note that the smaller (larger) the value of test error (accuracy), the better the forecasting accuracy. The results shown in Table 4 suggest that the model incorporating the sentiment index has a greater forecasting accuracy than the model without the sentiment index (based on lasso regression). Thus, the sentiment index is effective in forecasting prices of the S&P 500.

7 Conclusions

This study examined both analytically and numerically a financial market model consisting of two types of fundamentalist. Although the model takes rational traders into account, investors are assumed to be unable to forecast the fundamental value exactly, that is, uncertainty about true fundamental value is assumed to be the case. All the fixed points of the model are computed, and the stability conditions determined mathematically. Moreover, by means of simulations, the model with the sentiment index demonstrated an ability to generate interesting dynamics.

In addition, the model was checked to see whether it replicates the statistical properties of the returns of the S&P 500. In particular, the presence of excess kurtosis, volatility clustering, long-range dependence and the asymmetric relation between stock index returns and

changes in volatility documented in the empirical literature were tested. The model matches the stylized facts observed on the S&P 500 index.

More importantly, the role of the sentiment index in forecasting stock prices is highlighted. To this end, the forecast prices of the model, based on a sentiment index, are compared with those of a model without a sentiment index that uses a machine learning technique to forecast prices.

The model based on a sentiment index was found to be superior in terms of accuracy to the model without the sentiment index.

In conclusion, the sentiment index is important in capturing the global sentiment of the market and characterizing trading decisions by investors.

This study lends itself to further extensions. First, it can be considered to be closely related to the switching mechanism of Brock and Hommes (1998) relying on their discrete choice model with multinomial logit probabilities. Second, following Hommes and in 't Veld (2017), two different fundamental benchmark price models can be proposed: the Gordon model and the Campbell-Cochrane consumption-habit model with a time-varying risk premium. Finally, it is relevant to the continuous-time set-up of He and Li (2012) to consider the fullness of uncertainty and randomness at any time.

Acknowledgements The authors gratefully acknowledge financial support from the University of Modena and Reggio Emilia for the FAR2017 and FAR2019 projects. They also wish to thank William Bromwich for his painstaking attention to the copy-editing of this paper, and the Editor and the anonymous reviewers for their helpful and valuable comments.

Declarations

Funding

This study was supported by the University of Modena and Reggio Emilia as part of the FAR2017 and FAR2019 projects

Conflict of interest

The authors declare that they have no conflict of interest.

Availability of data and materials

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Code availability

Not applicable

Appendix A

In order to generalize the model, in this appendix, the two fundamental values are allowed to change over time. The robustness of the model to a more general set-up is demonstrated. Following Anufriev et al. (2020) and Schmitt et al. (2020), the fundamental values, $F_{i,t}$, $i = 1, 2$, follow a random walk:

$$F_{i,t+1} = F_{i,t} + \epsilon_{t+1} \tag{24}$$

where ϵ_{t+1} is a Gaussian discrete white noise with constant standard deviation σ_F. In the numerical simulations, the value of σ_F is set equal to 0.05 (in line with Anufriev et al. (2020)). In particular, Map (9) is used to numerically reconsider the stability of fixed point P_1^* (former shown in Fig. (2)), and of the fixed points P_1^*, P_2^* and P_3^* (former shown in Fig. (5)). The results are shown in Fig. (9) and (10). A comparison with Fig. (2)) and Fig. (5)) shows that the results of the model are robust to exogenous shocks in the two fundamental values.

References

Anufriev M, Gardini L, Radi D (2020) Chaos, border collisions and stylized empirical facts in an asset pricing model with heterogeneous agents. Nonlinear Dynamics 102:993–1017

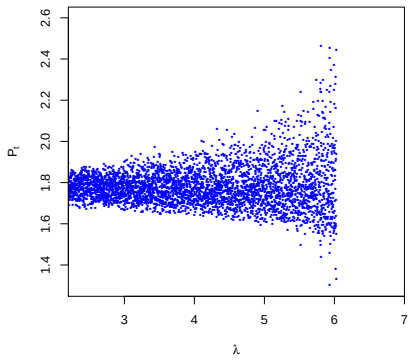

Figure 9 Bifurcation diagram showing the stability of fixed point P_1^* on varying the intensity of trading parameter λ under constellation of parameter values as in Fig. 2, but F_1 and F_2 representing random walk fundamental processes

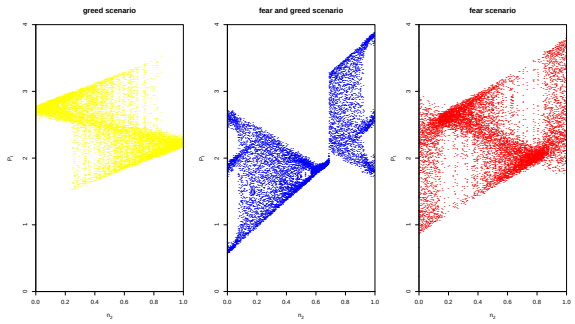

Figure 10 Bifurcation diagrams of the three scenarios, under constellation of parameter values as in Fig. 5, but F_1 and F_2 representing random walk fundamental processes

Badshah IU (2013) Quantile regression analysis of the asymmetric return-volatility relation. Journal of Futures Markets 33(3):235–265

Baker M, Wurgler J (2006) Investor sentiment and the cross-section of stock returns. The Journal of Finance 61(4):1645–1680

Barberis N, Shleifer A, Vishny R (1998) A model of investor sentiment. Journal of Financial Economics 49(3):307–343

Bekaert G, Wu G (2000) Asymmetric volatility and risk in equity markets. The Review of Financial Studies 13(1):1–42

Bekiros S, Jlassi M, Naoui K, Uddin GS (2017) The asymmetric relationship between returns and implied volatility: Evidence from global stock markets. Journal of Financial Stability 30:156–174

Black F (1976) Studies of stock market volatility changes. 1976 Proceedings of the American Statistical Association Bisiness and Economic Statistics Section

Boswijk HP, Hommes CH, Manzan S (2007) Behavioral heterogeneity in stock prices. Journal of Economic Dynamics and Control 31(6):1938–1970

Brianzoni S, Campisi G (2020) Dynamical analysis of a financial market with fundamentalists, chartists, and imitators. Chaos, Solitons & Fractals 130:109434

Brock WA, Hommes CH (1997) A rational route to randomness. Econometrica: Journal of the Econometric Society pp 1059–1095

Brock WA, Hommes CH (1998) Heterogeneous beliefs and routes to chaos in a simple asset pricing model. Journal of Economic Dynamics and Control 22(8-9):1235–1274

Campbell JY, Hentschel L (1992) No news is good news: An asymmetric model of changing volatility in stock returns. Journal of Financial Economics 31(3):281–318

Campisi G, Muzzioli S (2020) Investor sentiment and trading behavior. Chaos: An Interdisciplinary Journal of Nonlinear Science 30(9):093103, doi: 10.1063/5.0011636

Cavalli F, Naimzada AK, Pecora N, Pireddu M (2018) Agents' beliefs and economic regimes polarization in interacting markets. Chaos: An Interdisciplinary Journal of Nonlinear Science 28(5):055911

Chiarella C, Dieci R, He X (2009) Heterogeneity, market mechanisms and asset price dynamics. Handbook of Financial Markets: Dynamics and Evolution pp 277–344

Chiarella C, He XZ, Huang W, Zheng H (2012) Estimating behavioural heterogeneity under regime switching. Journal of Economic Behavior & Organization 83(3):446–460

Christie AA (1982) The stochastic behavior of common stock variances: Value, leverage and interest rate effects. Journal of Financial Economics 10(4):407–432

Cont R (2001) Empirical properties of asset returns: stylized facts and statistical issues. Quantitative Finance 1:223–236

Corredor P, Ferrer E, Santamaria R (2013) Investor sentiment effect in stock markets: Stock characteristics or country-specific factors? International Review of Economics & Finance 27:572–591

Day RH, Huang W (1990) Bulls, bears and market sheep. Journal of Economic Behavior & Organization 14(3):299–329

De Grauwe P, Kaltwasser PR (2012) Animal spirits in the foreign exchange market. Journal of Economic Dynamics and Control 36(8):1176–1192

De Long JB, Shleifer A, Summers LH, Waldmann RJ (1990) Noise trader risk in financial markets. Journal of Political Economy 98(4):703–738

Elyasiani E, Gambarelli L, Muzzioli S, et al. (2018) The risk-asymmetry index as a new measure of risk. Multinational Finance Journal 22(3-4):173–210

French KR, Schwert GW, Stambaugh RF (1987) Expected stock returns and volatility. Journal of Financial Economics 19(1):3

Gandolfo G (2010) Economic Dynamics (Forth ed.). Springer: Heidelberg

Gu S, Kelly B, Xiu D (2020) Empirical Asset Pricing via Machine Learning. The Review of Financial Studies 33(5):2223–2273, DOI 10.1093/rfs/hhaa009

Hauser MA (1997) Semiparametric and nonparametric testing for long memory: A monte carlo study. Empirical Economics 22(2):247–271

He XZ, Li K (2012) Heterogeneous beliefs and adaptive behaviour in a continuous-time asset price model. Journal of Economic Dynamics and Control 36(7):973–987

He XZ, Zheng H (2016) Trading heterogeneity under information uncertainty. Journal of Economic Behavior & Organization 130:64–80

Hommes C, in 't Veld D (2017) Booms, busts and behavioural heterogeneity in stock prices. Journal of Economic Dynamics and Control 80:101–124, DOI https://doi.org/10.1016/j.jedc.2017.05.006

James G, Witten D, Hastie T, Tibshirani R (2013) An Introduction to Statistical Learning, vol 112. Springer

Jawadi F, Namouri H, Ftiti Z (2018) An analysis of the effect of investor sentiment in a heterogeneous switching transition model for G7 stock markets. Journal of Economic Dynamics and Control 91:469–484

Lo W Andrew (1991) Long-term memory in stock market prices. Econometrica 59(5):1279–1313

Lux T (1995) Herd behaviour, bubbles and crashes. The Economic Journal 105(431):881–896

Lux T, Marchesi M (2000) Volatility clustering in financial markets: a microsimulation of interacting agents. International Journal of Theoretical and Applied Finance 3(04):675–702

Medio A, Lines M (2001) Nonlinear Dynamics: A Primer. Cambridge University Press

Naimzada AK, Ricchiuti G (2009) Dynamic effects of increasing heterogeneity in financial markets. Chaos, Solitons & Fractals 41(4):1764–1772

Schmitt N, Tramontana F, Westerhoff F (2020) Nonlinear asset-price dynamics and stabilization policies. Nonlinear Dynamics 102(2):1045–1070

Tramontana F, Gardini L, Dieci R, Westerhoff F (2009) The emergence of bull and bear dynamics in a nonlinear model of interacting markets. Discrete Dynamics in Nature and Society 2009

Tramontana F, Westerhoff F, Gardini L (2013) The bull and bear market model of Huang and Day: Some extensions and new results. Journal of Economic Dynamics and Control 37(11):2351–2370

Tramontana F, Westerhoff F, Gardini L (2015) A simple financial market model with chartists and fundamentalists: market entry levels and discontinuities. Mathematics and Computers in Simulation 108:16–40

Westerhoff F, Franke R (2012) Converse trading strategies, intrinsic noise and the stylized facts of financial markets. Quantitative Finance 12(3):425–436