



# Simulation of Social Media-Driven Bubble Formation in Financial Markets using an Agent-Based Model with Hierarchical Influence Network

Gonzalo Bohorquez<sup>1</sup> and John Cartlidge<sup>2,\*</sup>

<sup>1</sup>School of Computer Science, University of Bristol, Bristol, UK

<sup>2</sup>School of Engineering Mathematics and Technology, University of Bristol, Bristol, UK

\*Corresponding author. Email address: john.cartlidge@bristol.ac.uk

## Abstract

We propose that a tree-like hierarchical structure represents a simple and effective way to model the emergent behaviour of financial markets, especially markets where there exists a pronounced intersection between social media influences and investor behaviour. To explore this hypothesis, we introduce an agent-based model of financial markets, where trading agents are embedded in a hierarchical network of communities, and communities influence the strategies and opinions of traders. Empirical analysis of the model shows that its behaviour conforms to several stylized facts observed in real financial markets; and the model is able to realistically simulate the effects that social media-driven phenomena, such as echo chambers and pump-and-dump schemes, have on financial markets.

**Keywords:** Agent-Based Model; Echo Chambers; Financial Market Simulation; Market Sentiment; Pump-and-Dump

## 1. Introduction

In recent years, there has been a dramatic shift in the accessibility of financial markets. New *fintech* applications such as *Robinhood* and *eToro* have enabled retail investors to trade directly on traditional markets, without the need for expensive brokerage services; and the unprecedented disruption of new cryptocurrency assets, with their overall valuation rising from zero to more than one trillion dollars in less than 15 years, has driven crypto-speculation by millions of novice investors. At the same time, social media has become a ubiquitous presence in the lives of billions of people across the globe, which has influenced behaviours for both good and bad.

This shift in financial market participants and networks of influence has resulted in new causal drivers of market dynamics, leading to previously rare events becoming relatively commonplace (e.g., *GameStop short squeeze*,

*Klein (2022)*). Additionally, the democratisation of markets has dramatically increased the opportunities for malicious actors to take advantage of the naïve through illicit practices such as *pump-and-dump schemes* (Li et al., 2021).

**Contribution:** In this paper, we introduce a new agent-based model of financial markets to capture this shift in dynamics. In particular, we extend the well-established model of Lux-Marchesi (Lux and Marchesi, 1999) by embedding agents on a hierarchical network to form *communities of influence*. These communities represent social media influence on behaviours. We are able to show that the model generates realistic behaviours under a variety of scenarios, and we are able to investigate the effects of social media influence, the effects of echo chambers, and markets most likely to be susceptible to pump-and-dump schemes. All code is available open source: <https://github.com/gonzalo-bo/Lux-Hierarchy>.



The rest of this paper is organised as follows. In Section 2, we present a detailed background, review related works, and identify a research gap. In Section 3, we detail our model implementation and describe a set of metrics that we use for evaluation. Section 4 presents model results, demonstrating realistic general behaviours (Section 4.1) and realistic responses to particular scenarios (Section 4.2). We then discuss findings and their significance (Section 5), and conclude by briefly illustrating limitations and opportunities for further work (Section 6).

## 2. State of the art

### 2.1. Social media and financial markets

Wang et al. (2022) established a causal link between social media sentiment and same-day stock returns by analysing online messages from stock investment forum *EastMoney Guba*. This finding is supported by Müller et al. (2023), who discovered the sentiment expressed in *X* (formerly *Twitter*) messages has a similar effect, and the influence on returns is stronger for “risky, volatile” assets with a market price that tends to deviate from the underlying fundamental value.

Sentiment expressed through social media has also been shown to significantly impact market volatility, such that positive news on social platforms can lead to rapid price increases, while negative news can cause sudden declines. For instance, Gilbert and Karahalios (2010) demonstrated a direct link between the sentiment on social media platforms and subsequent movements in the stock market, suggesting that increases in subjective expressions of anxiety could predict an increase in volatility for the S&P 500 index. Additionally, Jiao et al. (2020) found that, whilst influence from traditional news media predicts a decrease in an asset’s volatility, influence from social media has the opposite effect. They present evidence suggesting that this distinction can most accurately be explained by an “echo chamber” model, where traders in social media are disproportionately affected by information that gets repeated (i.e., echoes) within the network.

The *Echo Chamber Effect* refers to the idea that social media platforms can create environments where investors are exposed predominantly to opinions and information that reinforce their existing beliefs (Cinelli et al., 2021). Such environments can lead to overconfidence and exacerbate market anomalies, as investors might ignore contrary evidence or broader market signals. Barber and Odean (2008) found that individual investors tend to trade more aggressively under the influence of overconfidence, which could be heightened by echo chambers in social media settings, leading to suboptimal trading decisions.

Cookson et al. (2022) provide evidence for the existence of the *Echo Chamber Effect* by analysing the data of more than 400,000 users on social network *StockTwits*. They reveal that users of this forum have a strong tendency to selectively expose themselves to information that reinforces their pre-existing beliefs, which leads to persistent

disagreement within the forum and may help explain why assets subject to social media chatter display more price volatility. Furthermore, studies such as Jiao et al. (2020) and Cookson et al. (2022) suggest that the *Echo Chamber Effect* is more prevalent for positive opinions than for negative opinions; which leads to social media communities displaying disproportionately bullish tendencies (i.e., expectations that future prices will increase) regarding their most frequently discussed assets or asset classes.

Social media is also associated with the illicit practice of *pump-and-dump*, which is described by Li et al. (2021) as the act of artificially inflating (i.e., “pump”) the price of an asset through exaggerated or entirely fabricated statements. Then, once the asset has been bought by unsuspecting buyers (drawn in by the hype), the schemers sell (i.e., “dump”) their holdings at the elevated price, causing excess supply which precipitates a sudden collapse in market price and significant losses for the unwitting investors left holding the (essentially worthless) asset. Lund (2022) suggests that pump-and-dump schemes are often facilitated by the reach and anonymity of social media, where it is easy to spread misleading information quickly to large and geographically dislocated audiences.

### 2.2. Lux-Marchesi model of financial markets (LM99)

Lux-Marchesi—hereafter referred to as LM99—is a seminal agent-based model of financial markets (Lux and Marchesi, 1999). Despite being 25 years old, its remarkable simplicity and versatility results in LM99 frequently being used as a foundation of contemporary agent-based models of financial markets (e.g., Alfarano et al., 2011; Meine and Vvedensky, 2023).

LM99 contains a single financial asset with a fundamental value,  $p_f$ , representing the “true” price of an asset in an efficient market containing rational agents with perfect information. Two types of trading agent strategy populate the model: *fundamentalists* and *chartists*. Fundamentalists make trading decisions based on the difference between the current market price and the current fundamental value, such that the asset is bought if market price is below the fundamental (i.e., when the asset is under-priced) and sold if the market price is above the fundamental (i.e., when the asset is over-priced). In this way, fundamentalists tend to drive the market price toward the fundamental value. In comparison, chartists make speculative trading decisions based on sentiment. Chartists fall into two categories: *optimists*, who are *bullish* and will always buy on the assumption that the market price will rise; and *pessimists*, who are *bearish* and will always sell on the assumption that the market price will fall. In this way, chartists tend to drive the market away from the fundamental value.

Importantly, agents in LM99 are adaptive and can switch between chartist and fundamentalist trading strategies based on the success, or profitability, of their current approach. LM99 also incorporates a mechanism for

the diffusion of opinions among chartists, such that pessimists (optimists) are more likely to become optimists (pessimists) if lots of optimists (pessimists) are observed in the market. This simple mechanism enables the model to capture the psychological aspects of trading and the impact of collective sentiment. Together, these evolutionary mechanisms, where agents adapt strategy over time, allow the relatively simple model of LM99 to generate complex dynamics that realistically reproduce key aspects of real-world financial markets. For full details, see Lux and Marchesi (1999, 2000).

### 2.3. Hierarchical models of social media networks

When attempting to model the interplay between social networks and financial markets, hierarchies repeatedly emerge as a simple yet effective way to represent several interesting phenomena. Modeling social network dynamics via hierarchies is a well-established practice that has proven to be quite powerful through several studies (Watts et al., 2002; Clauset et al., 2008). Furthermore, papers such as Zhang et al. (2017) and Dang et al. (2019) showcase the substantial improvements that hierarchical designs provide when it comes to modeling social media networks specifically.

### 2.4. Hierarchical ABMs of financial markets

While the apparent usefulness of hierarchical structures to model social networks is clear, there are very few papers exploring their use as a way to model communication between financial trading agents in an ABM. A systematic *Google Scholar* search for “(abm OR “agent based model”) AND “financial market” AND (intitle:hierarchy OR intitle:hierarchical)” returns only ten unique results, and only two of these are pertinent to the current study. These are: Alfarano et al. (2011), who show that an ABM with a hierarchical-like structure leads to increased volatility in simulated markets; and Meine and Vvedensky (2023), who reveal that a hierarchical structure leads to agents in simulated markets displaying a high degree of speculative behaviour. We therefore present the use of hierarchical networks to model social media influence in agent-based models of financial markets as a relatively underexplored gap in the research literature.

The model introduced by Meine and Vvedensky (2023)—hereafter referred to as MV23—extends the model of LM99 by embedding trading agents into a hierarchical network structure, such that each level in the network contains communities of *children* traders with mutual connection to a single *parent*. These communities represent traders interacting via online forums or social media groups, and the opinions (i.e., optimism/pessimism) and behaviours (i.e., chartist/fundamentalist) of community members affect the opinions and behaviours of others. Adding this network community structure enables the model of MV23 to demonstrate a greater variety of speculative *herding* behaviour than the underlying LM99 model, which it ex-

tends.

### 2.5. Research gap: social media-driven financial ABM

We have seen that social media networks have a significant effect on the behaviour of financial markets (Section 2.1), and hierarchical network models can successfully model aspects of social media dynamics (Section 2.3). Combined, the two aforementioned ideas highlight the potential of applying hierarchical structures to models of social media communication between participants trading in a financial market. While some models, such as MV23, explore hierarchical structures as a way to model communication dynamics within financial markets (Section 2.4), we identified no models in the literature that are explicitly tailored to the context of social media. Therefore, we present hierarchical agent-based models of social-media influence on financial markets as a research gap in the literature.

In this paper, we address this research gap by introducing and exploring a hierarchical extension of LM99 (Section 2.2). We initially considered generalizing MV23 to explore social media effects under different scenarios. However, we found that the implementation details of MV23 are difficult to interpret and the model’s parameter configurations are not presented in Meine and Vvedensky (2023) (for detailed analysis, see Bohorquez, 2024). Therefore, while we acknowledge MV23 as a strong inspiration for the present work, we present an alternative model.

## 3. Materials and methods

In this section, we introduce a new agent-based model of financial markets, where trading agent opinions and behaviours are influenced by a hierarchical network of communities. We begin by detailing LM99, which forms the base of the model (Section 3.1), before describing the community network structure that we introduce as an extension (Section 3.2). Finally, we define a series of metrics that will be used to evaluate the model (Section 3.3).

### 3.1. Model core: LM99

The core of the agent-based model closely follows LM99, introduced in Lux and Marchesi (1999). We refer the reader to Table 1 for a description of model parameters and configuration values.

As described in Section 2.2, LM99 contains two agent types: *fundamentalists* and *chartists*, with chartists taking one of two roles: *optimist* or *pessimist*. The model contains a single asset and agents can perform one of two actions: buy the asset; or sell the asset. Excess profit is used to measure the opportunity cost inherent to these two actions; i.e., if an agent bought the asset, its excess profit is how much it made from the purchase minus how much it would have made from not buying the asset, and vice-versa for the excess profit if it decided to sell. We calculate the excess profits for fundamentalists,  $EP_f$ , excess profits

**Table 1.** Parameter settings. Top: Values taken from Lux and Marchesi (2000, p.692). Bottom: Parameter settings of hierarchical influence network.

	II	III	IV	Description
$\alpha_2$	0.25	0.25	0.2	How much chartists are influenced by changes in the asset price
$\alpha_3$	1	0.75	1	How much traders are influenced by a role's profit
$v_1$	4	0.5	2	How often pessimists try to become optimists & vice versa
$v_2$	1	0.5	0.6	How often fundamentalists try to become chartists & vice versa
$\beta$	4	2	4	How often the market price changes
$r$	0.004	0.004	0.004	Dividends paid by the asset
$R$	0.0004	0.0004	0.0004	Returns from alternative investments
$s$	0.75	0.75	0.75	Factor by which a fundamentalist's profit is reduced (i.e., discount factor)
$p_f$	10	10	10	Fundamental value
$\sigma$	0	0	0	Magnitude of fundamental value fluctuations
$\mu$	0.1	0.1	0.05	Noise when price changes due to excess demand/supply
$\gamma$	0.01	0.02	0.01	How strongly fundamentalists react to deviations from the fundamental price
$t_c$	0.015	0.02	0.01	How much of the asset is bought or sold (by optimists or pessimists respectively)
$\delta t$	0.01	0.01	0.01	Time-step/interval. The simulation runs for $\frac{1}{\delta t}$ time-steps per unit time
$\delta t'$	0.002	0.002	0.002	Compare the current price and the price from $\delta t'$ time-steps ago to determine how fast the price is changing
$L$	5	5	5	Number of levels in the hierarchy
$k$	5	5	5	Number of children per community
$b$	1.8	2.25	2.4	Strength of hierarchy (community) influence on optimist $\leftrightarrow$ pessimist transition
$\phi$	0.5	0.5	0.5	Efficiency of information diffusion within the hierarchy
$\omega$	1	1	1	Optimist child's influence on its parent
$\nu$	1	1	1	Pessimist child's influence on its parent

\*Lux and Marchesi (2000, p.692) parameters  $N$  (number of traders) and  $\alpha_1$  (influence of chartists on chartists) are replaced by network hierarchy; see equations (9) - (13).

for optimists,  $EP_+$ , and excess profits for pessimists,  $EP_-$ , as:

$$EP_f = s \left| \frac{p_f - p}{p} \right|, \quad EP_+ = \left( r + \frac{\dot{p}}{v_2} \right) / p - R, \quad (1)$$

$$EP_- = R - \left( r + \frac{\dot{p}}{v_2} \right) / p$$

where  $p$  is the current market price,  $p_f$  is the current fundamental price, and  $\dot{p}$  is the price trend, calculated by comparing the current price with the price from  $\delta t'$  time-steps ago:

$$\dot{p} = (p_t - p_{t-\delta t'}) / \delta t' \quad (2)$$

At any given time-step  $t$ , traders have a probability of switching from one role to another. This probability depends on a set of transition pressures. As defined in equation (3),  $U_{21}$  represents the pressure exerted by the *optimist*  $\leftrightarrow$  *fundamentalist* transition, with positive values of  $U_{21}$  meaning optimists feel pressure to become fundamentalists (vice-versa for negative values). Similarly,  $U_{22}$  and  $U_1$  represent the pressure exerted by the *pessimist*  $\leftrightarrow$  *fundamentalist* and the *optimist*  $\leftrightarrow$  *pessimist* transitions, respectively:

$$U_{21} = \alpha_3 (EP_f - EP_+), \quad U_{22} = \alpha_3 (EP_f - EP_-), \quad (3)$$

$$U_1 = \alpha_1 \left( \frac{n_p - n_o}{n_p + n_o} \right) + \alpha_2 \frac{\dot{p}}{v_1}$$

where  $n_o$  and  $n_p$  represent the number of optimists and

pessimists in the network, respectively.

Furthermore, as in real markets, the trader's actions affect the market price of the asset being traded. At any given time-step, the probability that the price rises,  $\pi_{\uparrow p}$ , or falls,  $\pi_{\downarrow p}$ , is determined by the total demand in the market:

$$\pi_{\uparrow p} = \max[0, \beta(ED_c + ED_f + \mu)] \quad (4)$$

$$\pi_{\downarrow p} = \min[0, -\beta(ED_c + ED_f + \mu)]$$

$ED_c$  represents the total demand from chartists and is determined by the total number of optimists,  $n_o$ , and pessimists,  $n_p$ .  $ED_f$  represents the total demand exerted by fundamentalists and is determined by the number of fundamentalists,  $n_f$ , as well as whether fundamentalists are buying or selling the asset.  $ED_c$  and  $ED_f$  are defined as:

$$ED_c = (n_o - n_p) * t_c \quad (5)$$

$$ED_f = n_f * \gamma * (p_f - p)$$

Finally, equations (6), (7) and (8) outline the transition probabilities between trader states:

$$\pi_{o \rightarrow p} = v_1 \cdot \frac{n_c}{N} \cdot e^{-U_1} \cdot \delta t, \quad \pi_{p \rightarrow o} = v_1 \cdot \frac{n_c}{N} \cdot e^{U_1} \cdot \delta t \quad (6)$$

$$\pi_{o \rightarrow f} = v_2 \cdot \frac{n_o}{N} \cdot e^{-U_{21}} \cdot \delta t, \quad \pi_{f \rightarrow o} = v_2 \cdot \frac{n_f}{N} \cdot e^{U_{21}} \cdot \delta t \quad (7)$$

$$\pi_{p \rightarrow f} = v_2 \cdot \frac{n_p}{N} \cdot e^{-U_{22}} \cdot \delta t, \quad \pi_{f \rightarrow p} = v_2 \cdot \frac{n_f}{N} \cdot e^{U_{22}} \cdot \delta t \quad (8)$$

where  $\pi_{a \rightarrow b}$  represents the probability that a trader of type  $a$  becomes a trader of type  $b$  at any given time-step, and

$n_c$  represents the total number of chartists in the market.

### 3.2. Model extension: hierarchical network

We extend LM99 by embedding agents within a hierarchical network structure, which allows us to model a trading agent's opinions and behaviours being influenced by other members of their local community and wider community-of-communities. The network contains two types of node: *community* and *trading agent*. Trading agents exist within the bottom level (leaf nodes) of the hierarchy, while all higher-level nodes are designated as communities.

We assume that the network hierarchy has  $L$  levels and every non-leaf node has exactly  $k$  children. Then the total number of trading agents,  $N_t$ , is equal to the number of leaf nodes, and the total number of communities,  $N_c$ , is equal to the number of non-leaf nodes, i.e.:

$$N_t = k^{L-1}, \quad N_c = \frac{k^L - 1}{L - 1} - N_t \quad (9)$$

Unlike the base model of LM99, where agents have perfect knowledge of the distribution of opinions across the whole market, in our extended model, chartists only have access to information that is transmitted through the network structure. As such, the *optimist*  $\leftrightarrow$  *pessimist* transition equation (3) is replaced by equation (10):

$$U_1 = b \left( \frac{C_o - C_p}{C_o + C_p} \right) + \alpha_2 \frac{\dot{p}}{v_1} \quad (10)$$

where  $b$  represents *hierarchy strength*, and  $C_o$ ,  $C_p$  represent the proportion of optimists and pessimists in the local community, respectively.

Community nodes do not store a discrete state but rather a vector:  $[o, p, f]$ , where  $o$ ,  $p$ , and  $f$  are positive, continuous variables representing the optimist, pessimist, and fundamentalist states respectively.

Additionally, by letting traders behave like a homogeneous community (i.e., a community only populated by one child type), equation (11) simplifies modelling the interaction between communities and traders:

$$\begin{aligned} \text{Optimist} &= [\omega, 0, 0] \\ \text{Pessimist} &= [0, \nu, 0] \\ \text{Fundamentalist} &= [0, 0, 1] \end{aligned} \quad (11)$$

where  $\omega$  and  $\nu$  represent the influence of optimist and pessimist children, respectively.

Community nodes update their state in two distinct phases:

1. **Backward pass:** for each child,  $u$ , a community's state is updated according to equation (12). In other words, each community state,  $C$ , is determined by the average of its children's states:

$$C_{[o, p, f]} = \frac{1}{k} \sum u_{[o, p, f]} \quad (12)$$

2. **Forward pass:** following equation (13), each community state,  $C$ , is updated to account for the state of their respective parent community,  $Q$ :

$$C'_{[o, p, f]} = \frac{1}{2} C_{[o, p, f]} + \phi \cdot Q_{[o, p, f]} \quad (13)$$

where  $\phi$  represents the *network efficiency of information diffusion*. As  $\phi \rightarrow \infty$ , each trader is equally affected by the averaged opinions of all other traders in the network, while lower values of  $\phi$  imply that traders are more strongly affected by opinions in the local community, and  $\phi = 0$  represents insular traders that are not affected by the opinions of any other trader, regardless of network distance.

At each time-step, the backward pass is followed by the forward pass, which is then followed by traders switching their state according to the model's transition pressures. Finally, following the LM99 paradigm defined in equation (4), the market price is updated through the interaction of trading agents buying and selling.

The full set of model parameters and their descriptions are listed in Table 1. For the base LM99 parameters, we follow the three standard parameter sets presented in Lux and Marchesi (2000, p.692). Note that number of traders,  $N$ , is replaced by  $N_t$  from equation (9) and influence of chartists,  $\alpha_1$ , is replaced by  $b$  as described in equation (10).

Network parameters are presented at the bottom of Table 1. We fix hierarchy levels  $L = 5$  and children per community  $k = 5$ , giving  $N_t = 625$  trading agents in total as per equation (9). To preserve model symmetry, the influence of optimists  $\omega = 1$  and pessimists  $\nu = 1$  are fixed. Efficiency of information diffusion  $\phi = 0.5$  is fixed to balance the relative strength of influence of child and parent communities in equation (13). Finally,  $b$  controls the strength of hierarchy influence on the *optimist*  $\leftrightarrow$  *pessimist* transition, as defined in equation (10).

Note that, while there is no toggle to quickly reduce the hierarchical model into the base LM99 model, some specific parameter arrangements will achieve this purpose, e.g., setting  $b = 0$  in equation (10) is equivalent to LM99 with  $\alpha_1 = 0$ , and as  $\phi \rightarrow \infty$  in equation (13) the model approaches LM99 with  $\alpha_1 = b$ .

### 3.3. Model evaluation: stylized facts

To validate the *realism* of financial market simulations, it is common to compare the behaviour of the simulation model with the behaviour of real financial markets. Although real financial markets exhibit complex dynamics, there are broad patterns that consistently emerge. These so-called *stylized facts* can be used as a benchmark to validate simulation models, i.e., if the simulation reproduces the stylized facts of real markets, then we can consider that the simulation behaves in a somewhat realistic fashion. In this work, we consider four stylized facts of financial markets, and we describe a metric for detecting financial bubbles.

### 3.3.1. Fat tails of returns

The *fat tails* phenomenon is a stylized fact observed for an asset's (logarithmic) returns, indicating that large deviations from the mean are more common than would be expected under a normal distribution. Fat tails are usually represented through leptokurtic distributions; however, as explained in Lux and Marchesi (2000), kurtosis is a somewhat ambiguous concept, and it is not entirely clear how to compare the kurtosis statistics obtained for various time series. For this reason, rather than measuring kurtosis, it is standard practice to measure fat-tailedness by assuming that the tails of the returns decay according to a Pareto distribution:  $1 - ax^{-\alpha}$ .

Equation (14) details how to estimate  $\alpha$  for a tail of size  $m$  out of a total of  $n$  observations (Hill, 1975):

$$\alpha_H = \frac{1}{\frac{1}{m} \sum_{i=1}^m [\ln(R_{n-i-1}) - \ln(R_{n-m})]} \quad (14)$$

Following standard practice, we will present values for the 2.5%, 5%, and 10% right-tails of the absolute returns.

**Fact:** *The absolute returns of most financial assets present a tail decay such that  $2 \leq \alpha \leq 6$ , with lower values of  $\alpha$  representing fatter tails (Cont, 2001).*

### 3.3.2. Volatility clustering

The volatility of an asset,  $\sigma$ , is generally measured as the standard deviation of market price, defined as:

$$\sigma = \sqrt{\frac{\sum (p_i - \mu)^2}{N}} \quad (15)$$

where  $\mu$  is mean market price and  $p_i$  is market price at time-step  $i$ .

In simulation, but not in real markets, we can also calculate the deviation of market price from the fundamental value,  $F_\sigma$ , defined as:

$$F_\sigma = \sqrt{\frac{\sum (p_i - f_i)^2}{N}} - \sqrt{\frac{\sum (p_i - \mu)^2}{N}} \quad (16)$$

where  $f_i$  is the fundamental value at time-step  $i$ .

*Volatility clustering* is the observation that asset returns have a tendency to cluster based on their magnitude/significance. That is, large changes tend to be followed by large changes; while small changes tend to be followed by small changes. To evaluate volatility clustering, we measure the auto-correlation function (ACF) for the absolute and squared returns when  $\tau = 10$  and  $T = 70$ . We compute ACF by importing the `acf` function from `statsmodels.api.tsf`.

**Fact:** *We expect ACF value to be larger than 0, indicating that, for all time-steps, there is a correlation between the current returns and returns from 10 time-steps ago.*

**Table 2.** SADF and GSADF critical values at significance levels 90–100%.

T	Asymptotic SADF			Finite-sample GSADF*		
	90%	95%	100%	90%	95%	100%
100	1.10	1.37	1.88	1.65	2.00	2.57
200	1.12	1.41	2.03	1.84	2.08	2.70
400	1.20	1.49	2.07	1.92	2.20	2.80
800	1.21	1.51	2.06	2.10	2.34	2.79
1600	1.23	1.51	2.06	2.19	2.41	2.87

\*GSADF ( $T, r_0$ ) values: (100,0.190), (200,0.137), (400,0.100), (800,0.074), (1600,0.055).

### 3.3.3. Return to gaussianity

The *return to gaussianity* or *aggregational gaussianity* phenomenon refers to the expectation that the distribution of the asset's (logarithmic) returns should resemble a normal distribution for large values of  $T$ . To evaluate this, we measure the kurtosis of returns with  $T = 1$ ,  $T = 10$ , and  $T = 50$ .

To compute the excess kurtosis of the returns, we use the `norm.pdf` function from `scipy.stats` to fit the returns to a normal distribution, after which we use the `pandas.kurt` function to compute the distribution's kurtosis.

**Fact:** *We expect excess kurtosis to be large when  $T = 0$ , and approach zero as  $T$  becomes large.*

### 3.3.4. Slow decay of autocorrelation

This stylized fact is related to the observation that the autocorrelation in the absolute returns of an asset tends to decay slowly, following a power law such that  $f(x) = at^{-\beta}$ .

To evaluate the slow decay of the ACF, we measure the decay of absolute returns with  $T = 70$  as a function of  $\tau$ . This value is calculated by using the `curve_fit` function of `scipy.optimize` library to fit an exponential decay function  $f(x) = at^{-\beta}$  to the ACF of the absolute returns.

**Fact:** *When calculated through GARCH-type models, real markets exhibit  $\beta \in [0.2, 0.4]$  (Cont, 2001). However, using our methodology, we expect values of  $\beta \in [0.6, 1]$ .*

### 3.3.5. Financial bubbles & explosive behaviour

*Explosive behaviour* in an asset price describes a rapid and anomalous increase. A related concept is a *bubble*, which also describes a rapid and anomalous increase in market price. However, to be classified as a bubble, it is necessary for the price increase to be driven by speculation rather than an increase in fundamental value; whereas explosive behaviour covers both scenarios.

To detect financial market bubbles, we use the simple and well-documented GSADF test (Phillips et al., 2015). For the GSADF method with finite samples, we derive our significance level and minimum window size,  $r_0$ , values from Phillips et al. (2011); see Table 2. We also use the PWY procedure to detect *explosive periods* (Phillips et al., 2015), and follow the methodology of Phillips et al. (2012) to identify significant bubbles through the asymptotic SADF critical values (see Table 2).

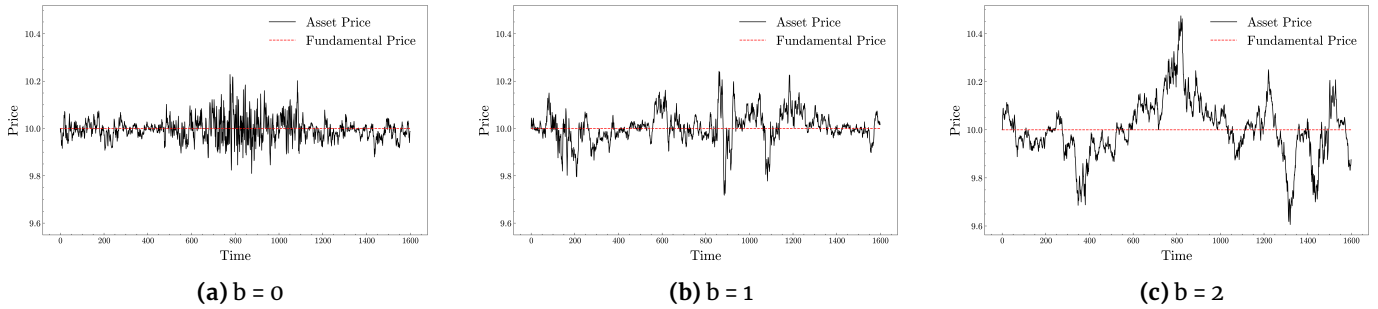


Figure 1. Sample runs showing how price volatility and fundamental deviation increases with hierarchy strength,  $b$ .

## 4. Results

In this section we first show that the model produces realistic dynamics (Section 4.1). We then use the model to explore real-world scenarios of social media influence, echo chambers, and pump-and-dump schemes (Section 4.2).

### 4.1. Reproducing stylized facts of financial markets

Fig. 1 shows sample runs with fixed fundamental value,  $p_f = 10$ , under increasing values of  $b$  (refer to equation (10)). We see that price movements increase in magnitude as the hierarchy's effect becomes more pronounced, with the market price displaying sustained and significant deviation from the fundamental price when  $b$  is large.

As shown Table 3, more rigorous analysis reveals that: for all parameter sets, an increase in  $b$  is associated with a significant increase in price volatility and a significant increase in the likelihood of explosive periods.

Furthermore, by looking at individual runs (e.g., see Fig. 2), we can see that instances of explosive behaviour are often sustained in time or occur in clustered bursts. This suggests that bubbles are driven by endogenous forces rather than occurring by random chance. Indeed, as shown in Fig. 3, we see that the model leads to *herding behaviour* of chartists, such that traders collectively swing from optimists to pessimists and back again over time. This behaviour is not present in the core model of LM99 and results directly from the hierarchical influence network.

Finally, Table 4 shows that all metrics fall within expected bounds, demonstrating that the simulation model is capable of reproducing all stylized facts of real markets introduced in Section 3.3.

### 4.2. Scenario testing

In the previous section, we validated the model by showing that it produces realistic financial market dynamics. In this section, we *scenario test* the model to see how it behaves under specific situations of interest.

#### 4.2.1. Social media effect on volatility and bubbles

Probably the most salient distinction between social media networks and their traditional counterparts is the speed

Table 3. Mean volatility and proportion of runs with explosive periods, showing both measures tend to increase with hierarchy strength,  $b$ .

$b$	Volatility			Explosive Periods		
	Set II	Set III	Set IV	Set II	Set III	Set IV
0.0	4.7	1.6	3.9	0%	4%	0%
0.1	5.9	1.7	3.6	2%	4%	0%
0.25	5.5	2.2	4.0	6%	8%	6%
0.5	5.6	2.1	4.5	4%	12%	6%
1.0	6.7	2.8	5.7	2%	14%	15%
2.0	9.1	4.4	8.9	44%	34%	22%

\* Cells show mean value of 50 i.i.d. simulation trials, with  $8 \times 10^4$  time-steps.

\* Volatility values are scaled by a factor of  $10^{-2}$ .

\* Explosive behaviour is determined via the GSADF test at the 90% significance level.

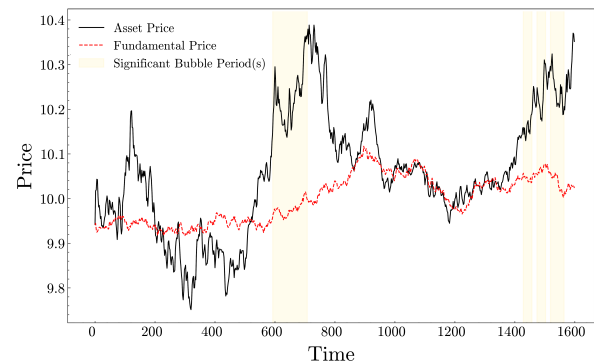


Figure 2. PWY procedure on hierarchical model identifies explosive periods (shaded yellow) that are prolonged (left), and clustered in time (right).

and efficiency with which information can be distributed across the former medium. Here, we note that higher values of  $\phi$  in equation (13) represent less network decay, equivalent to higher network efficiency. Thus, we take larger values of  $\phi$  to represent social media, while smaller values of  $\phi$  represent slower, more conventional modes of information diffusion.

Fig. 4 plots volatility and percentage of runs displaying explosive behaviour for different values of  $\phi$ . Clearly, there is a trend towards higher volatility and likelihood of explosive behaviour as  $\phi$  increases. In other words, as the network becomes more efficient, the market becomes more volatile and prone to bubble formation. This result matches observations of real-world markets, where social media

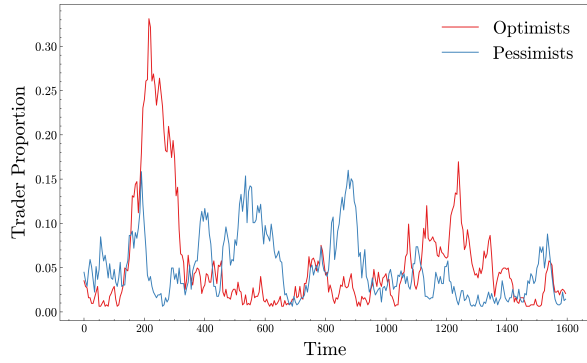


Figure 3. Opinion distributions within the chartist population reveals strong herding tendencies (at  $T=175-300$ ;  $450-700$ ;  $1175-1350$ ).

Table 4. Stylized facts reproduced by the model.

	Set II	Set III	Set IV
<i>Fat tails of (log) returns (<math>2 \leq x \leq 6</math>)</i>			
Tail Decay (2.5%)	4.72	5.22	5.16
Tail Decay (5%)	3.48	4.21	3.82
Tail Decay (10%)	2.72	3.24	2.83
<i>Return to Gaussianity (<math>\lim_{T \rightarrow \infty} x = 0</math>)</i>			
Excess Kurtosis ( $T = 1$ )	1.79	1.15	1.35
Excess Kurtosis ( $T = 10$ )	1.72	1.12	1.22
Excess Kurtosis ( $T = 50$ )	0.37	0.02	0.16
<i>Volatility clustering of returns (<math>x &gt; 0</math>)</i>			
Log Autocorrelation ( $\tau = 10, T = 70$ )	0.29	0.52	0.45
Squared Autocorrelation ( $\tau = 10, T = 70$ )	0.4	0.49	0.40
<i>Slow decay of returns autocorrelation (<math>x &lt; 1</math>)</i>			
Absolute Autocorrelation Decay	0.77	0.68	0.67

\* Cells show mean value of 50 i.i.d. simulation trials, with  $4 * 10^4$  time-steps.

can drive explosive behaviours of unprecedented scale in so-called *meme stocks* (e.g., see Klein, 2022). The result is also consistent with previous findings in the literature, which suggest that social media (i.e., an efficient network) is correlated with higher volatility, whilst traditional news media (i.e., a comparatively inefficient network) is related to reduced volatility levels (Jiao et al., 2020).

#### 4.2.2. Echo chamber effect

As described in Section 2.1, social media can act as an *echo chamber*, which can lead to a strengthening of pre-existing opinions. We can simulate this effect by modifying  $\omega$  and  $v$  in equation (11) via a third variable  $\mathcal{E}$ , such that the influence of a trader’s opinion is multiplied by  $\mathcal{E}$  if the trader’s opinion conforms to the majority opinion within its parent community. We consider two distinct scenarios:

- **Asymmetric Model:** Where only optimists are affected by  $\mathcal{E}$ , such that:

$$\omega = \begin{cases} \mathcal{E} & \text{if optimists} > \text{pessimists} \\ 1 & \text{otherwise} \end{cases} \quad (17)$$

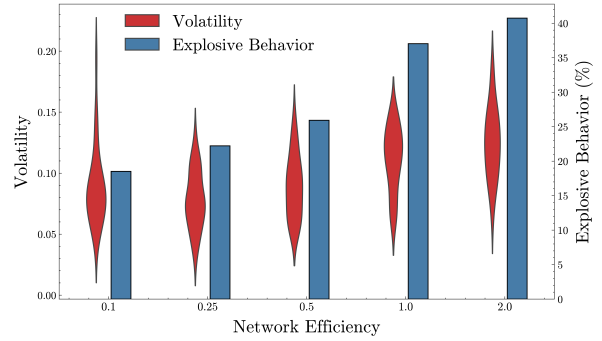


Figure 4. Parameter set IV. Volatility and explosive behaviour both increase with network efficiency,  $\phi$ .

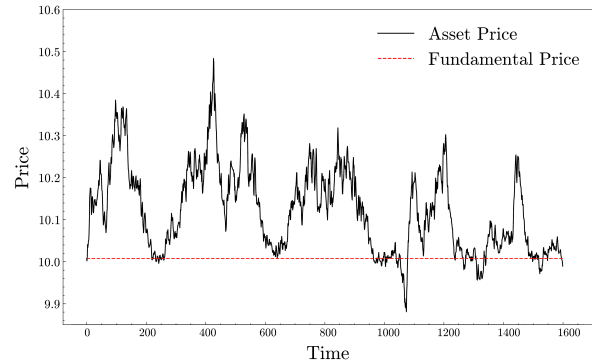


Figure 5. Sample run of the asymmetric echo chamber scenario, showing a volatile asset price that remains consistently above the fundamental.

- **Symmetric Model:** Where all chartists are affected by  $\mathcal{E}$ , such that, in addition to the effect from equation (17):

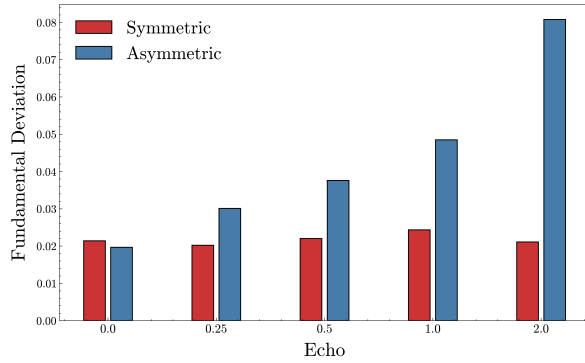
$$v = \begin{cases} \mathcal{E} & \text{if pessimists} > \text{optimists} \\ 1 & \text{otherwise} \end{cases} \quad (18)$$

For small (but non-zero) values of  $b$ , increasing  $\mathcal{E}$  in both the symmetric and asymmetric echo chamber scenarios has a similar effect to that of simply increasing the hierarchy strength. However, for large values of  $b$ , these two scenarios differ significantly. Namely, whilst the symmetric scenario’s effect on market behaviour becomes negligible, the asymmetric scenario is associated with large and persistent deviations between the asset’s market price and its fundamental value (see Figs. 5 and 6). This result is consistent with the finding that echo chambers lead to higher levels of speculation, and this effect is exacerbated when it is asymmetric in the positive direction (Cookson et al., 2022).

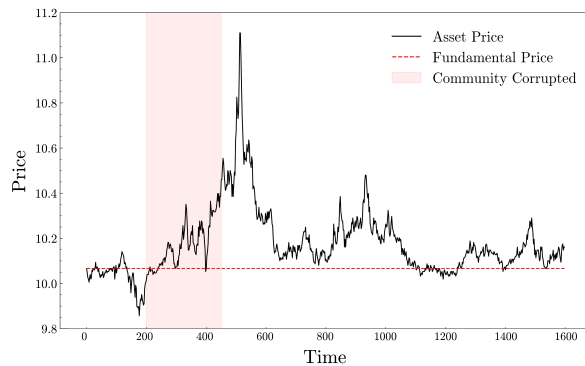
#### 4.2.3. Pump-and-dump

We model pump-and-dump schemes through the process of a community becoming “corrupted”: such that between time-steps  $T_0$  and  $T_1$ , a corrupted community’s (forward pass) signal conveys a deceptively high proportion of optimists to its children. Hence, if a community is subject





**Figure 6.** Symmetric vs asymmetric echo chambers. Note that the asymmetric model exhibits a strong correlation between hierarchy strength and fundamental deviation, a relation not present for the symmetric model.



**Figure 7.** Sample run for a successful pump-and-dump scheme where a community becoming corrupted (shaded red) results in a very pronounced peak in the asset price. Notice that the peak ( $T=550$ ) occurs after the community ceases to be corrupted ( $T=450$ ).

to becoming corrupted and the current time-step  $t$  is between  $T_0$  and  $T_1$ , its forward pass equation (13) is replaced by equation (19):

$$C'_{[o, p, f]} = \frac{1}{2} C_{[o, p, f]} + \phi \cdot Q_{[S*(o+p+f), p, f]} \quad (19)$$

where  $S$  represents the *signal strength* and  $S * (o + p + f)$  is the magnitude of the signal emitted by the corrupted community.

We define a pump-and-dump scheme as being “successful” if the maximum price reached during or after the community is corrupted is larger than the maximum price reached by 95% of 50 “uncorrupted” simulations using identical model parameter settings. Fig. 7 shows an example successful pump-and-dump scheme. However, for all standard parameter sets (see Table 1), the success rate of pump-and-dump schemes (with reasonable signal strengths s.t.  $S \leq 50$ ) is very low, and not significantly greater than random chance. This should not come as a surprise, since pump-and-dump schemes are relatively infrequent, especially in mature markets predominated by professional investors who carefully monitor the fundamental value of assets.

Notwithstanding, some parameter configurations do lead to more frequently successful pump-and-dump schemes. Broadly, there are four distinct classes:

- **Speculative markets:** where, by altering the transition pressures from equations (3) and (10), the importance that traders place on the actual price trend of the asset,  $\alpha_2$ , or the profit made by other traders,  $\alpha_3$ , becomes much lower than the importance placed on one another’s type,  $b$ .
- **Volatile markets:** where the price changes quickly. This effect can be further exacerbated through changes to the parameters in equation (5), such that the price becomes heavily influenced by the actions of speculative traders (i.e., large  $t_c$ ), or only mildly influenced by fundamentalist traders (i.e., small  $\gamma$ ).
- **Chartist-heavy markets:** where, through equation (1), small values of  $s$  make it unlikely for chartists to transition into fundamentalists.
- **Nascent markets:** pump-and-dump schemes are generally more successful when  $T_0$  is small, since the market is young and presumably still unstable.

In other words, markets that are speculative, volatile, nascent, and chartist-dominated are most susceptible to pump-and-dump. This fits with the common understanding of social media-driven pump-and-dump schemes, where targets are often naïve traders who are tricked into buying relatively obscure cryptocurrency assets after other users in a *WhatsApp* group claim to be making great profits.

## 5. Discussion

We have introduced a new agent-based model of financial markets, where agent behaviours and opinions are influenced by a hierarchical network of communities. Our model extends the well-regarded Lux-Marchesi model (LM99) of financial markets, with the addition of a hierarchical influence network to simulate social media effects.

We explored the general behaviour of the model and showed that it reproduces several stylized facts of real markets (Table 4; Section 4.1), which strongly evidences model *realism*. Of particular note are the much higher autocorrelation of returns than the base model (LM99 logarithmic autocorrelation at  $\tau = 10$ ,  $T = 70$  is smaller than 0.25 for all parameter sets, and squared autocorrelation is smaller than 0.2 (Bohorquez, 2024)). This higher degree of volatility clustering indicates a lower degree of market efficiency (Hameed et al., 2009), which is another sign pointing toward the model’s validity, since strong social media effects are commonly believed to reduce the efficiency of the market upon which they act (Al-Yahyaee et al., 2018; Bundi and Wildi, 2019). The finding that a stronger hierarchy results in higher levels of volatility and more frequent instances of explosive behaviour (Table 3) is also consistent with results presented by Alfarano et al. (2011), who conclude that a hierarchical structure is strongly correlated with markets displaying erratic behaviour.

The model was then used to explore scenarios of real-world markets (Section 4.2), with findings confirming evidence in the literature, including: social media networks, which increase communication efficiency, leading to more volatile markets; echo chambers found in social media leading to higher levels of speculation; and pump-and-dump schemes being mostly successful in speculative, volatile, and chartist-dominated markets. These findings provide strong evidence of the validity and versatility of the model to explore a range of phenomena of financial markets.

**Related work:** As described in Section 2.5, our model is most closely related to the MV23 model (Meine and Vvedensky, 2023), which also extends LM99 with a hierarchical influence network. However, there are some key differences. (i) Our model's generalised network representation, which includes LM99 under some parameter configurations, allows for the gradual tuning and exploration of community effects. In contrast, MV23 presents a markedly more discrete jump in behaviour. (ii) With much simpler rules for agent behaviour transitions, our model can be more easily tailored for specific scenarios such as echo chambers and pump-and-dump schemes (see Section 4.2).

## 6. Conclusions

We have extended a well-established financial ABM (Lux and Marchesi, 1999) to introduce a new hierarchical model of social media-driven influence on trading behaviours in financial markets. Empirical evidence shows that the model can succinctly and accurately replicate several emergent phenomena of real financial markets, especially those strongly influenced by social media. To the authors' best knowledge, no other model in the literature has been shown to reproduce all these effects of social media, including the most similar model of Meine and Vvedensky (2023).

However, our model is limited by some restrictive assumptions, such as: only a single asset is traded; every trader is affected by community influences equally; and the number of traders in the market is a fixed constant. In future works, we will relax some of these assumptions. A multi-asset market will be addressed following the approach of Zhang et al. (2017), where each asset is modelled as a "topic," and each individual's interest and opinion varies depending on the topic (i.e., asset); factors that affect speculative traders will be drawn from distributions that vary between traders; and traders will switch between an *active* and *inactive* state, allowing trade volumes to vary over time. We will also focus on better understanding the causal relationships that lead to emergent phenomena, as well as exploring the simulation of other scenarios through the basic hierarchical structure presented in this paper.

## 7. Funding

This work was supported by UKRI grant [EP/Y028392/1](https://doi.org/10.1017/9781017083921) "AI for Collective Intelligence Research Hub".

## References

- Al-Yahyaee, K. H., Mensi, W., and Yoon, S.-M. (2018). Efficiency, multifractality, and the long-memory property of the bitcoin market: A comparative analysis with stock, currency, and gold markets. *Finance Research Letters*, 27:228–234. <https://doi.org/10.1016/j.frl.2018.03.017>.
- Alfarano, S., Raddant, M., and Milaković, M. (2011). A note on institutional hierarchy and volatility in financial markets. *European Journal of Finance*, 19(6):449–465. <https://doi.org/10.1080/1351847X.2011.601871>.
- Barber, B. M. and Odean, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *The Review of Financial Studies*, 21:785–818. <https://doi.org/10.1093/rfs/hhm079>.
- Bohorquez, G. (2024). A hierarchical communication network within a financial agent-based model. Master's thesis, Sch. of Computer Science, Uni. of Bristol, UK. <https://github.com/gonzalo-bo/Lux-Hierarchy>.
- Bundi, N. and Wildi, M. (2019). Bitcoin and market-(in) efficiency: a systematic time series approach. *Digital Finance*, 1:47–65. <https://doi.org/10.1007/s42521-019-00004-z>.
- Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W., and Starnini, M. (2021). The echo chamber effect on social media. *Proceedings of the National Academy of Sciences*, 118:e2023301118. <https://doi.org/10.1073/pnas.2023301118>.
- Clauset, A., Moore, C., and Newman, M. E. (2008). Hierarchical structure and the prediction of missing links in networks. *Nature*, 453:98–101. <https://doi.org/10.1038/nature06830>.
- Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1:223–236. <https://dx.doi.org/10.1088/1469-7688/1/2/304>.
- Cookson, J. A., Engelberg, J. E., and Mullins, W. (2022). Echo Chambers. *The Review of Financial Studies*, 36:450–500. <https://doi.org/10.1093/rfs/hhac058>.
- Dang, L., Chen, Z., Lee, J., Tsou, M.-H., and Ye, X. (2019). Simulating the spatial diffusion of memes on social media networks. *Int. J. Geographical Information Science*, 33(8):1545–1568. <https://doi.org/10.1080/13658816.2019.1591414>.
- Gilbert, E. and Karahalios, K. (2010). Widespread worry and the stock market. In *Proc. Int. AAAI Conference on Web and Social Media*, volume 4, pages 58–65. <https://doi.org/10.1609/icwsm.v4i1.14023>.
- Hameed, A., Ashraf, H., and Siddiqui, R. (2009). Stock market volatility and weak-form efficiency: evidence from an emerging market. *International Journal of Business and Emerging Markets*, 1:249–263. <https://doi.org/10.1504/IJBEM.2009.023445>.

- Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *Annals of Statistics*, 13:331–341. <https://doi.org/10.1214/aos/1176343247>.
- Jiao, P., Veiga, A., and Walther, A. (2020). Social media, news media and the stock market. *Journal of Economic Behavior & Organization*, 176:63–90. <https://doi.org/10.1016/j.jebo.2020.03.002>.
- Klein, T. (2022). A note on GameStop, short squeezes, and autodidactic herding: An evolution in financial literacy? *Finance Research Letters*, 46:102229. <https://doi.org/10.1016/j.frl.2021.102229>.
- Li, T., Shin, D., and Wang, B. (2021). Cryptocurrency pump-and-dump schemes. Available at SSRN 3267041. <https://doi.org/10.2139/ssrn.3267041>.
- Lund, B. (2022). The cryptocurrency “pump-and-dump”: Social media and legal and ethical ambiguity. In *Handbook of Research on Cyber Law, Data Protection, and Privacy*, pages 202–211. IGI Global. <https://doi.org/10.4018/978-1-7998-8641-9.ch012>.
- Lux, T. and Marchesi, M. (1999). Scaling and criticality in a stochastic multi-agent model of a financial market. *Nature*, 397:498–500. <https://doi.org/10.1038/17290>.
- Lux, T. and Marchesi, M. (2000). Volatility clustering in financial markets: A microsimulation of interacting agents. *Int. J. Theoretical and Applied Finance*, 3:675–702. <https://doi.org/10.1142/S0219024900000826>.
- Meine, D. C. and Vvedensky, D. D. (2023). Emergence of speculation in a hierarchical agent-based model. *Physica A: Statistical Mechanics and its Applications*, 620:128619. <https://doi.org/10.1016/j.physa.2023.128619>.
- Müller, K., Pan, Y., and Schwarz, C. (2023). Social media and stock market participation. Available at SSRN 4557783. <http://dx.doi.org/10.2139/ssrn.4557783>.
- Phillips, P. C., Shi, S., and Yu, J. (2012). Testing for multiple bubbles. *Cowles Foundation Discussion Paper No. 1843*. <http://dx.doi.org/10.2139/ssrn.1981976>.
- Phillips, P. C., Shi, S., and Yu, J. (2015). Testing for multiple bubbles: Historical episodes of exuberance and collapse in the S&P 500. *International Economic Review*, 56:1043–1078. <https://doi.org/10.1111/iere.12132>.
- Phillips, P. C., Wu, Y., and Yu, J. (2011). Explosive behavior in the 1990s Nasdaq: When did exuberance escalate asset values? *International Economic Review*, 52:201–226. <https://doi.org/10.1111/j.1468-2354.2010.00625.x>.
- Wang, X., Xiang, Z., Xu, W., and Yuan, P. (2022). The causal relationship between social media sentiment and stock return: Experimental evidence from an online message forum. *Economics Letters*, 216:110598. <https://doi.org/10.1016/j.econlet.2022.110598>.
- Watts, D. J., Dodds, P. S., and Newman, M. E. (2002). Identity and search in social networks. *Science*, 296:1302–1305. <https://doi.org/10.1126/science.1070120>.
- Zhang, Y., Lyu, T., and Zhang, Y. (2017). Hierarchical community-level information diffusion modeling in social networks. In *Proc. 40th Int. ACM SIGIR Conf. on Research and Development in Information Retrieval*, pages 753–762. <https://doi.org/10.1145/3077136.3080784>.