Finance Research Letters 46 (2022) 102359

Contents lists available at ScienceDirect

Finance Research Letters

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

YOLO trading: Riding with the herd during the GameStop episode☆

Štefan Lyócsa[d,b,c,*], Eduard Baumöhl[a,b,e], Tomáš Výrost[a,b,e]

[a] Faculty of Commerce, University of Economics, Dolnozemská cesta 1, 852 35, Slovakia
[b] Institute of Financial Complex Systems, Masaryk University, Lipová 41a, 602 00 Brno, Czech Republic
[c] Faculty of Management, University of Prešov, Konštantínová 16, 080 01 Prešov, Slovakia
[d] Slovak Academy of Sciences, Šancova 56, 811 05 Bratislava, Slovakia
[e] Faculty of Economics, Technical University of Košice, Němcovej 32, 040 01 Košice, Slovakia

ARTICLE INFO

Keywords:
GameStop
WallStreetBets
Reddit
Short squeeze
Volatility

ABSTRACT

We explore the 2020 and early 2021 price variation of four stocks: GameStop, AMC Entertainment Holdings, Blackberry and Nokia. The four stocks were subject to a decentralized short squeeze that exploited the short positions of institutional investors. This investor movement was likely initiated by retail investors concentrated mostly around the subreddit r/WallStreetBets (WSB). We demonstrate that part of the next day's price variation can be explained by an increase in activity on the WSB subreddit relative to Google searches (terms related to the event). We discuss implications for future research.

1. Introduction

The beginning of 2020 witnessed an unprecedented decline in stock markets worldwide, induced by the coronavirus 2019 (COVID-19) pandemic and related uncertainty, fear, and extremely high volatility (Ramelli and Wagner, 2020; Baker et al., 2020; Lyócsa et al., 2020; Lyócsa and Molnár, 2020; Ashraf, 2020; Okorie and Lin, 2021). Nevertheless, major U.S. market indices recovered swiftly and ended the year 2020 in positive territory. The U.S. stock market subsequently experienced another shock related to herd behavior induced by a social media platform that seemed to have culminated in early 2021, but still persists even after the middle of the 2021. A brief and rudimentary description of this event may be found in Chohan (2021). Essentially, small-scale investors concentrated mostly around the subreddit r/WallStreetBets (WSB) initiated a short squeeze of institutional investors betting on declines of several underrated stocks, such as GME (GameStop), AMC (AMC Entertainment), BB (BlackBerry), NOK (Nokia) and a few others. The coordinated effort was accompanied by an increase in the stock price of GME from approximately 5 USD in July 2020 to approximately 10 USD in October 2020 and later from 17.25 USD in 4 January up to an intraday maximum of 483 USD on 28 January 2021. Over this period, short sellers – Wall Street giants, such as hedge funds – booked significant losses. This David-vs-Goliath narrative was primarily a driver of a large, unprecedented, crowd-sourced short squeeze.

Regardless of how appealing such a narrative has become for the general public, this type of coordinated behavior may prove to be highly disruptive for financial markets. From the perspective of financial stability, fire sales and price-mediated contagion are of particular interest due to the significant spillovers they might cause, where forced sales by one market participant tighten constraints on others and thereby lead to further forced sales (Geanakoplos, 2010; Stein, 2012; Braouezec and Wagalath, 2019; Chernenko and Sunderam, 2020), a proposition that is supported by a vast amount of research (Coval and Stafford, 2007; Jotikasthira et al., 2012; Caballero and Simsek, 2013; Hau and Lai, 2017; Barbon et al., 2019; Fricke and Fricke, 2020). However, it is still unclear whether

☆ Acknowledgment: We gratefully acknowledge the support by the Czech Science Foundation, grant no. 20-11769S.
* Corresponding author at: Institute of Financial Complex Systems, Masaryk University, Lipová 41a, 602 00 Brno, Czech Republic.
E-mail address: stefan.lyocsa@savba.sk (Š. Lyócsa).

https://doi.org/10.1016/j.frl.2021.102359
Received 20 February 2021; Received in revised form 5 July 2021; Accepted 22 July 2021
Available online 27 July 2021

such negative spillovers can be initiated from inflated prices of short-squeezed stock prices. The decentralized yet similar behavior of investors is also closely related to the strand of literature on herd behavior in financial markets (Scharfstein and Stein, 1990; Avery and Zemsky, 1998; Chang et al., 2000; Bikhchandani and Sharma, 2000; Chiang and Zheng, 2010; Cipriani and Guarino, 2014). The unique aspect of the recent events is the counterhegemonic extensive financial effort, driven by small individual investors collaborating to sabotage the short positions of large Wall Street players such as hedge funds (Chohan, 2021).

While such a decentralized people-powered initiative by retail investors might seem tempting, it undoubtedly presents a risky investing strategy. Extremely inflated prices are likely to be followed by sudden price declines, eventually leading to extreme losses for some participants. As such trading strategy seeks substantial returns irrespective of the level of risk and might be induced by crowds on a social media platforms, we refer to such a strategy as "YOLO trading",[1] which is a term that reflects the emotions and vocabulary of the WSB subreddit.

In order for the short squeeze episode initiated by retail investors to work, decentralized coordination is needed. Discussion forums, such as the WSB subreddit, provide such platforms. This study aims to verify whether activity on the WSB subreddit is related to the daily price variation of the four short squeezed stocks, GME, AMC, BB, and NOK. To disentangle activity on the WSB subreddit from the investor's attention, we control for (1) Google searches related to the given stock and short squeeze event (short squeeze-specific attention) and (2) Google searches related to the market activity in general (market-specific attention). Our analysis shows that after the discussion on the WSB subreddit intensified, the price variation of the four stocks increased above what is predicted by short-squeeze and market-specific attention.

Our study contributes to the study of Chohan (2021) and to the wide literature on herding behavior, investor attention, and market volatility by being the first to empirically verify the relationship between WSB subreddit activity and consequent stock price variation of the four stocks of concern. Our study concludes by discussing the agenda for future research, specifically related to the challenges ahead for market regulators, portfolio managers, short sellers, consequences related to trade settlement, and measurement of investors' attention and sentiment.

The remainder of the paper is structured as follows. The next section describes the data and methodology. Section 3 presents our main results, followed by some robustness checks in Section 3.3. Section 4 concludes the paper and sketches future research ideas.

2. Data and methodology

2.1. Data sources

To study the price variation of GME, AMC, BB and NOK, we use a daily sampling frequency and restrict our sample to the period from 1 October 2020 to 28 May 2021, i.e., 175 trading days. The date the activity on the WSB subreddit started cannot be established precisely, but our chosen period should be long enough to capture the initial activities on the WSB subreddit. In this study, we use three data sources. First, the prices are retrieved from Datastream. Second, WSB subreddit activity is obtained both directly from Reddit and from pushshift.io. Third, we use the Google search intensity as provided by Massicotte and Eddelbuettel (2018). The four stocks were selected because they were subjected to short squeeze events, while being among the most discussed stocks on the WSB subreddit forum. In Section 3.3. we discuss the results for other assets that also show a high level of interest on the WSB subreddit, but were not subjected to a short squeeze.

2.2. Price variation

Restricted by our use of daily data, our estimate of daily price variation, V_t for $t = 1, 2, \dots, T$, is based on an average (motivated by Patton and Sheppard (2009)[2]) of three range-based volatility estimators: PK_t of Parkinson (1980), GK_t of Garman and Klass (1980) and RS_t of Rogers and Satchell (1991) adjusted for overnight price variation J_t and annualized. Because the volatility levels (for each stock, see Fig. 2 and Table 1) changed dramatically over our sample period, we model the log of the variance.[3]

$$V_t = ln(100^2 \times (J_t + 3^{-1}(PK_t + GK_t + RS_t))) \tag{1}$$

Specifically, let O_t, H_t, L_t and C_t denote daily opening, high, low and closing prices. Next, define $h_t = ln(H_t) - ln(O_t)$, $l_t = ln(L_t) - ln(O_t)$, $c_t = ln(C_t) - ln(O_t)$, and $J_t = [ln(O_t) - ln(C_{t-1})]^2$. The Parkinson (1980) estimator is given by:

$$PK_t = \frac{(h_t - l_t)^2}{4ln2} \tag{2}$$

the Garman and Klass (1980) estimator is defined as:

$$GK_t = 0.511\left(h_t - l_t\right)^2 - 0.019\left(c_t(h_t + l_t) - 2h_t l_t\right) - 0.383c_t^2 \tag{3}$$

[1] The abbreviation YOLO is from "You Only Live Once" and as we show in our word cloud (see graphical abstract), it is one of the most frequent terms on the WSB subreddit (11th place in the frequency table, the preceding terms are "hold", "buy", "gme", "just", "like", "share", "f*ck", "stock", "sell" and "get"). Moreover, using the NRC emotion lexicon (EmoLex), we identified the two most prevalent emotions of the entire WSB subreddit — trust and anticipation. This nicely complements the spirit of our definition of "YOLO trading" as a trading strategy driven more by emotions rather than rigorous analysis.

[2] The composite range-based estimator was recently used by Lyócsa et al. (2021).

[3] The logarithmic transformation of the variance series is common in the literature (see Taylor, 2017 for a detailed discussion), as it leads to a much more symmetric distribution that is more suitable in autoregressive volatility models.

Table 1
Variables of interest — volatility, interest, Google searches.

		Mean	S.D.	Median	$\rho(1)$	min.	max.	KPSS
Panel A: Gamestop								
Volatility series	V_t	10.147	1.645	9.877	0.523	7.162	14.987	0.253
YOLO-1	$YOLO_{1,t}$	2.984	3.733	1.569	0.801	0.037	17.995	0.299
YOLO-2	$YOLO_{2,t}$	0.365	0.152	0.345	0.779	0.184	1.368	0.246
Panel B: AMC								
Volatility series	V_t	9.961	1.514	9.662	0.547	6.398	15.510	0.162
YOLO-1	$YOLO_{1,t}$	0.998	1.984	0.177	0.793	0.003	12.060	0.183
YOLO-2	$YOLO_{2,t}$	0.383	0.154	0.359	0.770	0.193	1.364	0.323
Panel C: Blackberry								
Volatility series	V_t	8.731	1.439	8.532	0.547	6.237	13.794	0.398
YOLO-1	$YOLO_{1,t}$	0.610	1.380	0.103	0.838	0.007	10.040	0.095
YOLO-2	$YOLO_{2,t}$	0.406	0.153	0.382	0.758	0.206	1.366	0.294
Panel D: Nokia								
Volatility series	V_t	7.142	1.382	6.886	0.431	4.027	13.630	0.105
YOLO-1	$YOLO_{1,t}$	0.275	1.201	0.030	0.768	0.002	11.814	0.059
YOLO-2	$YOLO_{2,t}$	0.372	0.146	0.348	0.752	0.192	1.335	0.271
Panel E: Common variables								
Volatility index (VIX)	VIX_t	22.680	4.698	21.910	0.886	15.650	40.280	0.475
General market attention	M_t	32.591	5.614	32.067	0.263	21.178	54.000	0.121

Notes: S.D. is the standard deviation, $\rho(1)$ is the first-order serial-correlation coefficient. KPSS is the test statistic of the no unit-root (in the null) hypothesis as given by Kwiatkowski et al. (1992), with long-run variance correction as in Sul et al. (2005) with a 5% critical value of 0.580.

and the Rogers and Satchell (1991) estimator is as follows:

$$RS_t = h_t(h_t - c_t) + l_t(l_t - c_t) \tag{4}$$

2.3. WSB subreddit activity

Without knowing the true intentions of retail investors, it is impossible to explicitly observe how much “YOLO trading” there is in the market. However, in this particular case, we can safely assume that these four stocks are fundamentally overvalued while at the same time short squeezed via the investment activities of retail investors who appear to coordinate their actions on the WSB subreddit. Our proxy for the extent of “YOLO trading” on the market is based on the activity on the WSB subreddit related to the given stock.[4]

To capture the intensity of discussions on the WSB subreddit, we use the total number of ticker occurrences in its comments for the day (e.g., “GME” for GameStop). Let $T_{j,t} \in \mathbb{N}$ denote the number of times a given ticker is found for the jth post on day t. The WSB subreddit activity indicator is:

$$W_t = \sum_{j=1}^{n(t)} T_{j,t} \tag{5}$$

where $n(t)$ is the number of posts on day t. The same approach is applied for the remaining three stocks with tickers “AMC”, “BB” and “NOK”. While market prices are quoted only for trading days, WSB subreddit discussions are also recorded for Saturdays and Sundays. To synchronize the data, an average over the consecutive Friday, Saturday, and Sunday is substituted for the Friday value.[5]

2.4. Google search volume intensity

Using the R package of Massicotte and Eddelbuettel (2018), we retrieve data on the Google search volume intensity (SVI_t) for a given term, where SVI_t corresponds to the relative number of daily searches. The number ranges from 0 to 100, where all values are normalized to the maximum value over the given sample period.

We retrieve two groups of search terms. The first group consists of terms that we consider specifically for the “short squeeze events”: short squeeze, short sell, call option, wall street bets, Wallstreet, Melvin Capital, to the moon, Reddit, Keith Gill, DeepF*ckingValue, Dave Portnoy, and Justin Sun.[6] Additionally, we included GME and GameStop when volatility of GameStop

[4] As was correctly noted by a reviewer, “YOLO trading” might be different from WSB subreddit interest because not all the WSB subreddit posts are related to “YOLO trading”. To account for this, our model specifications use attention (interest) proxies derived from sources other than the WSB subreddit sources, namely, from Google searches (see next Section 2.4.)

[5] On several occasions, missing data were imputed using linear interpolation.

[6] Apart from standard keywords related to the subreddit WSB, and short selling, we also added a hedge fund, Melvin Capital, one of the largest short sellers, as well as some “celebrities” connected to the event and their “nicknames”.

is of interest, AMC when the volatility of AMC Entertainment Holdings is of interest, and BB and Blackberry or NOK and Nokia, when volatility of Blackberry or that of Nokia is of interest. For each trading day, the search volume intensity is averaged across these terms, and the resulting variable is denoted as $E_t \in [0, 100]$.[7]

The second group consists of 15 terms that are related to general market conditions, while we also include tickers of stocks that are well known and are from industries similar to the four stocks subjected to the short squeeze. The terms are S&P 500, SPY, VIX, market bubble, stock market, BBY (Best Buy), AMZN (Amazon), TGT (Target), WMT (Walmart), CNK (Cinemark Holdings), IMAX, NFLX (Netflix), DIS (Disney), MSI (Motorola), and APPL (Apple). As before, we average the search volume intensity for each day, and the resulting variable is denoted as $M_t \in [0, 100]$. For both E_t and M_t, weekends are handled in the same way as with W_t.

2.5. Relative Reddit and event intensity: the YOLO variables

Regressing price variation (V_t) on the lagged activity on the WSB subreddit discussion forums (W_{t-1}) might be insufficient, as the discussions might simply be a reaction to the general attention. We instead use a ratio of:

$$YOLO_{1,t} = \frac{W_t}{E_t} \times 100^{-1} \tag{6}$$

We refer to this variable as the *relative Reddit intensity*. For days with larger values of $YOLO_{1,t}$, the activity on the WSB subreddit related to the short-squeezed stock is relatively higher compared to the attention given to the short-squeeze event found via Google searches. It follows that if the discussion on the WSB subreddit has contributed to the price variation of the given stock, then the next day's volatility should be associated with today's $YOLO_{1,t}$. Note that as both E_t and W_t differ across stocks, the $YOLO_{1,t}$ variable is also stock-specific.[8]

The $YOLO_{1,t}$ variable can be interpreted as excessive activity on the WSB subreddit, thus acting as a proxy for the unobserved "YOLO trading". This is in accordance with our definition of a "YOLO trading" that it is a trading strategy that seeks substantial returns irrespective of the level of risk and might be induced by a crowd on a social media platform.[9]

Using the same principle, we also use the following ratio, the *relative event intensity*[10]:

$$YOLO_{2,t} = \frac{E_t}{M_t} \tag{7}$$

For days with larger values of $YOLO_{2,t}$, the intensity of Google searches specifically related to the short squeeze events is relatively higher than the Google search intensity related to general market conditions. Including the variable in our specifications controls for the possible herding effect outside of the WSB subreddit community.

2.6. Model specification

To estimate the role of WSB subreddit discussions, we work within a framework of a linear regression model estimated via OLS. To better understand the role of the $YOLO_{1,t}$ and $YOLO_{2,t}$ variables, we use four specifications. Model 1 (M1) is the baseline model with:

$$V_t = \beta_0 + \beta_1 V_{t-1} + \beta_2 M_{t-1} + \beta_3 VIX_{t-1} + \epsilon_t \tag{8}$$

where β are the estimated regression parameters and ϵ_t is the error term. The persistent nature of volatility is reflected by using an autoregressive model framework adjusted with additional variables.[11] Our specification also includes the M_{t-1} term to account for the fact that it was not only the prices of GME, AMC, BB, and NOK that increased, as over the observed time frame, the U.S. market entered a bullish period that led to historical market index maximums (e.g., S&P 500 and NASDAQ). This attention might therefore partially explain the volatility levels observed for the four stocks. It might be that the attention estimated via $YOLO_{1,t}$, $YOLO_{2,t}$, and M_t can be captured via the volatility index (VIX), that is often used as a measure of general market sentiment.[12] We

[7] The stock-specific index is suppressed.

[8] Note that we could scale W_t to a range from 1 to 100, with 100 corresponding to the maximum interest on the WSB subreddit (more specifically from 0 to 100, but as W_t appears in the denominator, a correction would be applied). This way, both W_t and E_t are defined on a similar scale. If we observe lower values of W_t and higher values of E_t at the same time, this suggests that there is little interest on the WSB subreddit despite a large interest for the general public (Google Trends). As scaling will not change the interpretation or significance in an ordinary least squares (OLS) model, we use $YOLO_{1,t}$ (and $YOLO_{2,t}$) in their raw formats.

[9] If a given stock is not the subject of a short squeeze, then the same variable might have a different interpretation; it is the joint occurrence of attention and the fact that the stock was the subject of a risky trading strategy that makes the relative Reddit intensity a proxy for 'YOLO trading'.

[10] The variable $YOLO_{2,t}$ can be interpreted as excessive interest in the "GameStop episode" and short squeeze events outside of the WSB subreddit social medial platform. As this variable might also include interest on other social media platforms (or on the WSB subreddit as well), this variable might be a proxy for "YOLO trading" as well. However, it surely includes other interests as well, and therefore, it serves as a control variable. As this variable is similar in its construction to $YOLO_{1,t}$, we refer to it as $YOLO_{2,t}$.

[11] An alternative would be to use a range-based heterogeneous autoregressive (HAR) model specification that explains future volatility using average weekly and monthly range-based volatility estimates (for application at the market index level see Lyócsa et al., 2021). We elected not to use that specification for three reasons. First, given the event under study, our sample is much shorter, and therefore, the monthly volatility component is unlikely to manifest as a significant predictor. Second, our current model specification seems to sufficiently address residual serial correlation. Third, with a shorter sample, we opt for a simpler model that leads to an estimate of 5 parameters instead of the 7 parameter range-based HAR version of our model.

[12] We thank the anonymous referee for the suggestion to add VIX into our model specifications.

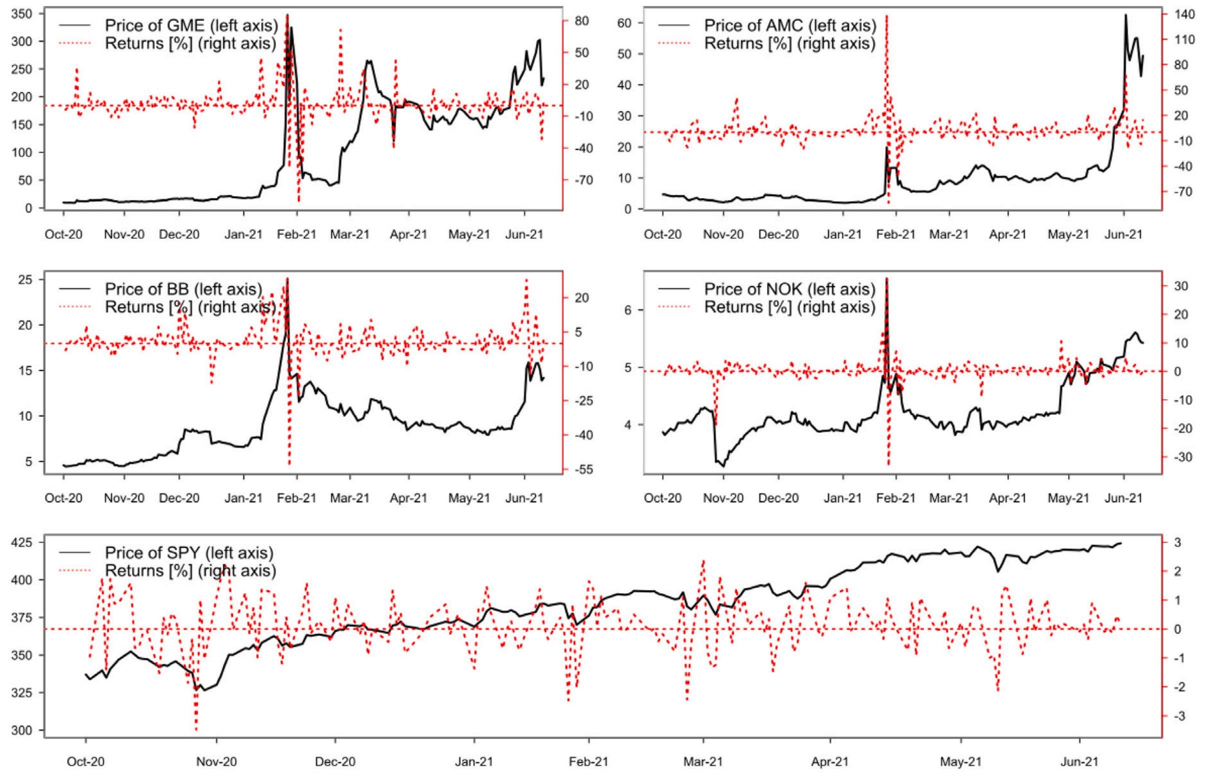

Fig. 1. Prices and returns of four stocks and the market ETF - SPY.

therefore added the VIX_{t-1} variable (which corresponds to the end-of-day VIX value) into all our model specifications. VIX_{t-1} does not correlate much with the stock's own volatility (see Table S1 in the online supplementary material), which is in line with our observation that the behavior of the four stocks is unique and differs from what is observed in the market; the correlations of VIX with GME, AMC, BB and NOK are 0.21, 0.24, 0.12 and 0.30, respectively. Next, we add $YOLO_{1,t}$, which leads to Model 2:

$$V_t = \beta_0 + \beta_1 V_{t-1} + \beta_2 M_{t-1} + \beta_3 VIX_{t-1} + \beta_4 YOLO_{1,t-1} + \epsilon_t \quad (9)$$

or we add $YOLO_{2,t}$, which leads to Model 3:

$$V_t = \beta_0 + \beta_1 V_{t-1} + \beta_2 M_{t-1} + \beta_3 VIX_{t-1} + \beta_5 YOLO_{2,t-1} + \epsilon_t \quad (10)$$

Finally, we use both variables, which leads to Model 4:

$$V_t = \beta_0 + \beta_1 V_{t-1} + \beta_2 M_{t-1} + \beta_3 VIX_{t-1} + \beta_4 YOLO_{1,t-1} + \beta_5 YOLO_{2,t-1} + \epsilon_t \quad (11)$$

3. Results

3.1. Data characteristics

We plot the price and return series for the four stocks in Fig. 1, along with the SPY ETF to facilitate comparison. There are two key observations. First, all four stocks show a similar pattern of sudden price increases in January. These increases are clearly detached from the development of the otherwise also growing market-wide index. In fact, the average return correlation between the four stocks is 0.570 for the whole sample period and only −0.047 between the four stocks and SPY ETF. However, excluding data from 2021, the average correlation of returns between the four stocks is only 0.166 and that with the market index increases to 0.256. This clearly suggests a sudden decoupling of the price development of the four stocks from the market.

Second, as opposed to the usual market-wide returns, the daily returns (dashed red line) for the four stocks are of different magnitudes. Note the right axis in Fig. 1. Daily returns outside of the ±25% range were not exceptional. These unprecedented return characteristics are also reflected in the volatility (see Table 1 and Fig. 2), which is still subject to stylized facts of right skewness and persistence, but compared to the period before January 2021, the extreme volatility is of different orders of magnitude. Similar characteristics are visible for the relative Reddit intensity (red dashed line on the right axis, $YOLO_{1,t}$ of Fig. 2), which appears to spike before or on the day the volatility spikes. This observation underlines our argument that the price variation of the four stocks was likely exacerbated by activity on the WSB discussion board.

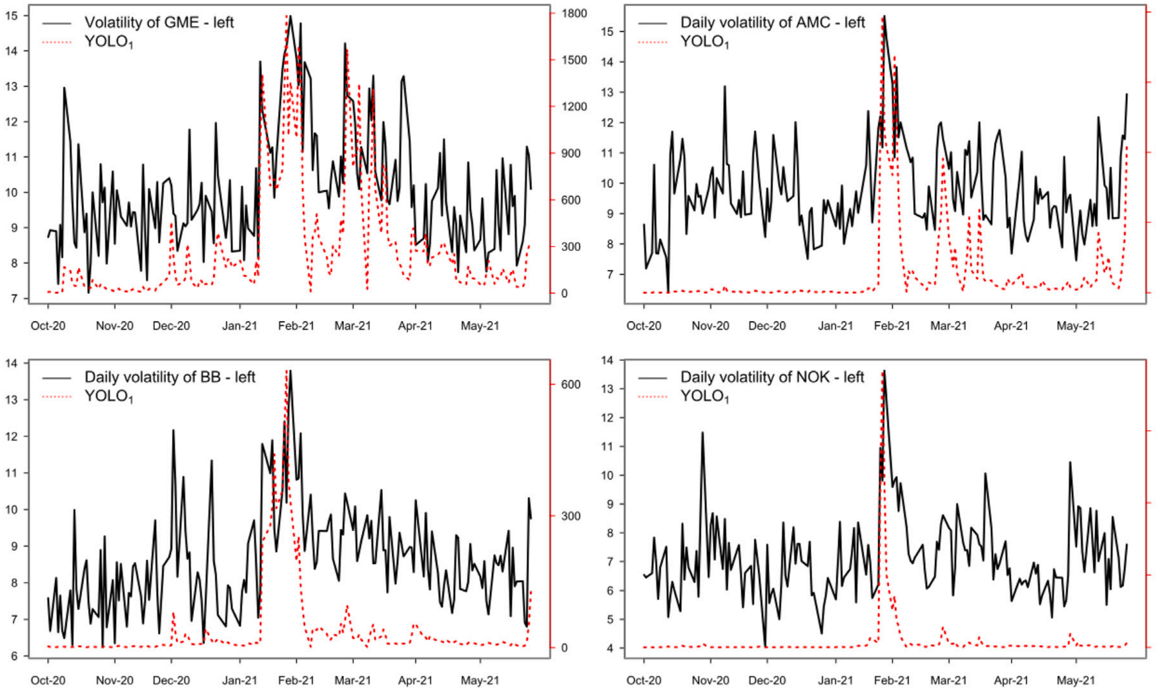

Fig. 2. Daily volatility of stocks and YOLO variable.

3.2. Model estimates

Our key results can be found in Table 2. We report estimated coefficients and significance levels with corresponding model characteristics. The latter show that despite the nonlinear nature of the event, with highly skewed volatilities, our model specification led to well-behaved residuals that show only mild and insignificant serial dependence (an exception is only the baseline model for GME), with no heteroskedasticity and no unit-root in residuals. The models, therefore, seem to adequately describe the price variation in our sample period.

The key variable of interest is the coefficient of the relative Reddit intensity ($YOLO_{1,t}$), which is positive and significant in all models where it is included (Models 2 and 4) and for all four stocks. Therefore, the greater the occurrence of ticker names on the WSB subreddit relative to the general attention as measured by Google searches (event-related search volume intensity), the higher the next day's volatility is. Our results suggest that part of the price variation is driven by excessive WSB subreddit discussions.

Interestingly, the coefficient on $YOLO_{2,t}$, which we interpret as the herding effect, is also positive and most of the time significant. Compared to $YOLO_{1,t}$, the size of the effect of $YOLO_{2,t}$ that it exerts on the next day's volatility appears to be larger.[13] This is reasonable, given that to generate large price movements, one needs more than a (albeit popular) discussion on the WSB subreddit. However, comparing R^2 between Models 2 and 3 shows that Model 2 leads to much higher R^2, which suggests that the two $YOLO_t$ variables capture different attention. This interpretation is supported by the fact that the two variables are not excessively correlated (with 0.46, 0.60, 0.44 and 0.40 correlations for GME, AMC, BB and NOK, see Table S1 in Electronic Supplementary Materials). Additionally, comparing the results of Models 2 and 3 with those of Model 4 shows that inclusion of $YOLO_{2,t}$ in the model improves the model fit only marginally.

In the baseline Model 1, the autoregressive coefficient is always significant, in line with the persistence reported in Table 1. However, the size of the coefficient declines with the inclusion of $YOLO_{1,t}$ and $YOLO_{2,t}$ variables, and when both are included (Model 4), the autoregressive coefficient is smallest; for GME and NOK, it was not even significant. It therefore appears that once our attention variables are included, the estimated persistence of volatility declines, which suggests attention-driven volatility.

Finally, our two control variables, the general market attention variable (M_t) and implied volatility (VIX_t) had mixed effects on short-squeezed volatility. General market attention was insignificant for GME, had a negative effect on AMC and a positive effect on BB, and mixed results for NOK. With the exception of Blackberry, implied volatility had a positive although not always significant effect.

[13] The comparison is made by multiplying the estimated coefficients from Table 2 with average and median values from Table 1.

Table 2
Stock price volatility model of interest and attention.

		GameStop				AMC				Blackberry				Nokia			
		M1	M2	M3	M4	M1	M2	M3	M4	M1	M2	M3	M4	M1	M2	M3	M4
Panel A: Model estimates																	
Constant		**4.41**^{d}	**7.22**^{d}	**4.49**^{d}	**6.87**^{d}	**4.87**^{d}	**7.02**^{d}	**4.75**^{d}	**6.97**^{d}	**3.94**^{d}	**5.67**^{d}	**4.03**^{d}	**5.59**^{d}	**3.54**^{d}	**5.39**^{d}	**3.09**^{c}	**4.94**^{d}
Volatility	V_{t-1}	**0.507**^{c}	**0.242**^{a}	**0.38**^{c}	0.186	**0.601**^{d}	**0.372**^{d}	**0.512**^{d}	**0.37**^{d}	**0.511**^{d}	**0.288**^{b}	**0.395**^{c}	**0.223**^{b}	**0.389**^{c}	**0.172**^{a}	**0.248**^{b}	0.107
General market attention	M_{t-1}	0.015	−0.009	0.020	−0.002	**−0.043**^{b}	**−0.052**^{d}	**−0.035**^{b}	**−0.052**^{d}	**0.039**^{c}	**0.035**^{b}	**0.045**^{c}	**0.039**^{b}	−0.013	**−0.032**^{b}	−0.006	**−0.026**^{a}
Volatility index	VIX_{t-1}	0.005	0.011	0.006	0.010	0.023	**0.029**^{a}	0.025	**0.029**^{a}	−0.041	−0.035	**−0.046**^{a}	−0.039	**0.056**^{b}	**0.063**^{b}	**0.066**^{b}	**0.068**^{b}
$YOLO_1$	$YOLO_{1,t-1}$		**0.181**^{c}		**0.155**^{c}		**0.308**^{c}		**0.302**^{b}		**0.384**^{d}		**0.351**^{d}		**0.549**^{d}		**0.495**^{c}
$YOLO_2$	$YOLO_{2,t-1}$			**2.812**^{b}	**2.092**^{a}			**1.872**^{c}	0.132			**2.017**^{d}	**1.462**^{a}			**2.657**^{b}	**1.615**^{a}
Panel B: Model characteristics																	
R^2		0.279	0.360	0.332	0.387	0.329	0.434	0.358	0.435	0.329	0.412	0.362	0.429	0.217	0.383	0.278	0.405
adj. R^2		0.267	0.345	0.316	0.369	0.317	0.421	0.343	0.418	0.317	0.398	0.347	0.412	0.203	0.369	0.261	0.387
max. VIF		1.232	2.105	1.401	2.236	1.305	1.663	1.431	2.095	1.186	1.662	1.467	1.917	1.224	1.488	1.529	1.685
		M_{t-1}	$YOLO_{1,t-1}$	V_{t-1}	VIX_{t-1}	M_{t-1}	V_{t-1}	V_{t-1}	VIX_{t-1}	M_{t-1}	V_{t-1}	V_{t-1}	V_{t-1}	M_{t-1}	V_{t-1}	V_{t-1}	V_{t-1}
First-order residual serial-correlation		−0.17	−0.08	−0.10	−0.05	−0.08	−0.01	−0.06	−0.01	−0.13	−0.08	−0.08	−0.05	−0.07	−0.06	−0.03	−0.03
Nth order residual serial-correlation test	(p-value)	0.00	0.26	0.18	0.56	0.31	0.89	0.48	0.92	0.11	0.28	0.32	0.53	0.44	0.41	0.76	0.71
Het test	(p-value)	0.36	0.68	0.48	0.52	0.61	0.29	0.20	0.26	0.36	0.16	0.48	0.38	0.31	0.47	0.59	0.56
Residual KPSS	test-stat	0.34	0.17	0.26	0.16	0.17	0.08	0.09	0.08	0.18	0.27	0.10	0.17	0.38	0.50	0.30	0.38

Notes: p-values are derived from a stationary bootstrap (of the raw dataset) with block lengths drawn from the geometric distribution, where the optimal block length is estimated as in Politis and White (2004) and Patton and Sheppard (2009). "a","b", "c" and "d" denote statistical significance at the 10%, 5%, 1% and 0.1% significance levels. "max. VIF" reports the maximum variance inflation factor and we also point to the variable with the max. VIF value. "Nth order residual serial-correlation test" corresponds to the p-value of the Escanciano and Lobato (2009) automatic portmanteau test for serial correlation. "Het. test" denotes the p-value of the bootstrap version of the White (1980) of Cribari-Neto (2004). The "Residual KPSS" is the test statistic of the no unit-root (in the null) hypothesis on residuals as given by Kwiatkowski et al. (1992), with long-run variance correction as in Sul et al. (2005) with a 5% critical value of 0.580.

3.3. Robustness checks

We estimate an alternative version of the model proposed above to better understand our results. First, we used a lin–lin model specification, where variance is not log-transformed. The coefficient of $YOLO_{1,t}$ had an expected, positive, and significant coefficient for all model specifications (Models 2 and 4).[14] This was not true for the coefficient of $YOLO_{2,t}$, which was now positive and significant only for GME, and negative and significant for AMC, with nonsignificant estimates for BB and NOK. As indicated by the heteroskedasticity tests, in these models, heteroskedasticity becomes a significant issue, with likely leverage points. We therefore consider our baseline model specifications with transformed volatility to be the preferred specifications.

Second, we estimated Models 1 to 4 for other assets from similar industries that were clearly not subjected to the short squeeze of retail investors but might have been subjected to extensive discussions on the WSB subreddit. Specifically, we estimated our model for Best Buy (BBY), Amazon (AMZN), IMAX, Netflix (NFLX), Motorola (MSI), Disney (DIS), and the broad U.S. market ETF (SPY). The denominator of $YOLO_{1,t}$ uses event-specific attention from Google Trends, while the numerator uses attention on the WSB subreddit forum. If the given stock/asset is not the subject of a short squeeze, which is the case of the seven assets, the resulting $YOLO_{1,t}$ variable can be interpreted as a "noisy" measure of attention on the WSB discussion forum. Given the way $YOLO_{1,t}$ is constructed, we expect the variable to not be significant for all stocks. Considering Model 4, $YOLO_{1,t}$ is significant at the usual 5% level only for 4 assets (NFLX, MSI, DIS and SPY), while $YOLO_{2,t}$ is not significant at all.[15] These results are not as consistent as what we observed in the previous section.

4. Concluding remarks and agenda for future research

We explore the 2020 and 2021 decentralized short squeeze strategies executed by retail investors on the stock prices of GameStop, AMC Entertainment Holdings, Blackberry and Nokia. If this strategy is successful, it leads to inflated prices (bubbles) and is therefore very risky. We thus refer to this strategy as "*YOLO trading*". We study whether the price variation of the four stocks can be explained by the discussion on the WSB subreddit while accounting for Google searches related to the given short squeeze events and general market conditions. Our analysis shows that this was indeed the case; as the discussion on the WSB subreddit intensified, the price variation of the four stocks increased. While our results suggest that most of the price variation is not directly related to the discussion on the WSB subreddit, it seems likely that such social network activity can activate other retail investors for the given cause, in this case, to short squeeze institutional investors. This event raises several important questions for future research.

First, the regulatory consequences will be challenging. The manipulation of stock prices was an important issue in the U.S. market up until the 1930s (Allen and Gale, 1992). Although the legal framework outlawed stock price manipulation, it is difficult to overcome the undesirable behavior of all market participants. Short squeezing belongs to this category of illegal trading practices — if we consider a short squeeze as a price manipulation. However, in the case of WSB, it might be considered "legal manipulation" protected by the 1st Amendment (i.e., freedom of speech). One of the most recent and extensive short squeeze episodes was the case of Porsche-VW (Allen et al., 2021). Porsche, in a takeover attempt, drove VW's ordinary shares to rise from approximately €200 to over €1000 in just a few days, making it briefly the most valuable company in the world (for further details, especially about the indictment and legal consequences, see, Möllers, 2015). Such events had significant ramifications in terms of distorting

[14] All results from this section are part of the electronic supplementary material.

[15] If included individually (Model 3), $YOLO_{2,t}$ is significant and positive for MSI, DIS and SPY.

price discovery and overall market efficiency and substantially increased volatility. An attempt to raise or depress stock market prices by making a false statement is illegal — how can a regulator impose high legitimacy standards in a world of easy-to-access mass trading and fast-paced growth of social media? Thus, from a legal perspective, the WSB episode constitutes a new regulatory challenge. Short selling per se might be at the center of such discussions. New studies on justifications for or restrictions on short sales should emerge, as previous research is inconclusive in highlighting all the pros and cons of short-seller participation in stock markets (e.g., Goldstein and Guembel, 2008; Beber and Pagano, 2013; Boehmer and Wu, 2013; Grullon et al., 2015; Chen et al., 2020).

It would be beneficial to further explore all the side effects of "YOLO trading". For example, further research might provide insights into the following:

- The WSB episode's shock propagation, in the form of return or volatility spillovers among other stocks or even different asset classes as we presented in Section 3.1, the price development of the four stocks under study exhibited a sudden decoupling from the market. Stock market contagion and comovements both have a direct impact on financial stability and direct implications from the perspective of portfolio allocation (e.g., Forbes and Rigobon, 2002; Bae et al., 2003; Pericoli and Sbracia, 2003; Bekaert et al., 2005; Diebold and Yilmaz, 2009, 2012; Wang et al., 2018; Okorie and Lin, 2021). However, with the short squeeze episode, there was no uncertainty related to whether the prices of the four stocks were close to their fundamental value — they were not. Therefore, this type of price bubble might not propagate to other assets per se.
- Short selling-related costs — with the increase in demand for protection against a short squeeze, the price of the out-of-the-money put options is likely to increase. Perhaps the usual suspects for short squeezing might be identified (e.g., Desai et al., 2006; Boehmer et al., 2008; Diether et al., 2009), which has implications for asset pricing theories and may also lead to the investigation of other financial products and procedures that mitigate investors' short squeeze exposure.
- A more efficient way of settling trades — the clearing of transactions now takes some time to process (usually $T+2$ settlement cycle applies, i.e., two business days), which might not be a problem during calm periods, but in times of market turmoil, the situation might be different. Hence, the limited ability to fulfill the collateral requirements of brokers, such as Robinhood or Interactive Brokers, forces them to employ trading restrictions. One of the top buzzwords of the 21st century emerges to provide faster and more flexible post-trade processing — blockchain (e.g., Mori, 2016; Chiu and Koeppl, 2019; Ross and Jensen, 2019).
- Sentiment and social media analysis – measuring investors' attention by internet searches (e.g., Google Trends or Wikipedia) or their sentiment by analyzing social media content (e.g., Twitter) – has attracted considerable research interest over the last decade (e.g., Preis et al., 2013; Moat et al., 2013; Hamid and Heiden, 2015; Dimpfl and Jank, 2016; Bento et al., 2020; Audrino et al., 2020). By using word-emotion lexicons (such as EmoLex), we are able to pinpoint the exact sentiment of coordinated small-scale investors' behavior over time. Is it "joy", "anticipation", or "trust" that drives hype and perhaps "fear", "anger", and "sadness" that cause downturns? We believe that this area of research is worth investigating.

Societies worldwide are witnessing polarization and discomposure, amplified by the COVID-19 pandemic and related restrictions. Moreover, since the Global Financial Crisis of 2008, seething rage against the "machine" of late-stage capitalism is growing. In recent years, we have witnessed how such negative emotions can manifest, with the Capitol Siege being just one example. In reviewing the WSB subreddit, one can easily spot considerable hate and fear, a tendency to latch onto conspiracy theories, and counterhegemonic repercussions. Apart from the direct impact on financial stability, WSB-like people-powered initiatives might dramatically increase the polarization of our societies, providing additional ammunition to both Alt-Right and Alt-Left movements. We should all keep this in mind.

CRediT authorship contribution statement

Štefan Lyócsa: Conceptualization, Methodology, Software, Writing – original draft, Writing – review & editing, Visualization, Project administration. **Eduard Baumöhl:** Conceptualization, Writing – original draft, Writing – review & editing, Funding acquisition. **Tomáš Výrost:** Conceptualization, Software, Data curation, Writing – original draft, Writing – reviewing and editing, Funding acquisition.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.frl.2021.102359. Online supplementary material contains correlations between regression model variables and all the results discussed in Section 3.3. Robustness checks.

References

Allen, F., Gale, D., 1992. Stock-price manipulation. Rev. Financ. Stud. 5 (3), 503–529.
Allen, F., Haas, M., Nowak, E., Tengulov, A., 2021. Market efficiency and limits to arbitrage: Evidence from the volkswagen short squeeze. J. Financ. Econ. (forthcoming).
Ashraf, B.N., 2020. Stock markets' reaction to COVID-19: Cases or fatalities? Res. Int. Bus. Finance 54, 101249.
Audrino, F., Sigrist, F., Ballinari, D., 2020. The impact of sentiment and attention measures on stock market volatility. Int. J. Forecast. 36 (2), 334–357.
Avery, C., Zemsky, P., 1998. Multidimensional uncertainty and herd behavior in financial markets. Amer. Econ. Rev. 88 (4), 724–748.
Bae, K.-H., Karolyi, G.A., Stulz, R.M., 2003. A new approach to measuring financial contagion. Rev. Financ. Stud. 16 (3), 717–763.

Baker, S.R., Bloom, N., Davis, S.J., Kost, K.J., Sammon, M.C., Viratyosin, T., 2020. The Unprecedented Stock Market Impact of COVID-19. Technical Report, National Bureau of Economic Research.
Barbon, A., Di Maggio, M., Franzoni, F., Landier, A., 2019. Brokers and order flow leakage: Evidence from fire sales. J. Finance 74 (6), 2707–2749.
Beber, A., Pagano, M., 2013. Short-selling bans around the world: Evidence from the 2007–09 crisis. J. Finance 68 (1), 343–381.
Bekaert, G., Harvey, C.R., Ng, A., 2005. Market integration and contagion. J. Bus. 78 (1), 39–69.
Bento, A.I., Nguyen, T., Wing, C., Lozano-Rojas, F., Ahn, Y.-Y., Simon, K., 2020. Evidence from internet search data shows information-seeking responses to news of local COVID-19 cases. Proc. Natl. Acad. Sci. 117 (21), 11220–11222.
Bikhchandani, S., Sharma, S., 2000. Herd behavior in financial markets. IMF Staff Pap. 47 (3), 279–310.
Boehmer, E., Jones, C.M., Zhang, X., 2008. Which shorts are informed? J. Finance 63 (2), 491–527.
Boehmer, E., Wu, J., 2013. Short selling and the price discovery process. Rev. Financ. Stud. 26 (2), 287–322.
Braouezec, Y., Wagalath, L., 2019. Strategic fire-sales and price-mediated contagion in the banking system. European J. Oper. Res. 274 (3), 1180–1197.
Caballero, R.J., Simsek, A., 2013. Fire sales in a model of complexity. J. Finance 68 (6), 2549–2587.
Chang, E.C., Cheng, J.W., Khorana, A., 2000. An examination of herd behavior in equity markets: An international perspective. J. Bank. Financ. 24 (10), 1651–1679.
Chen, Y., Kelly, B., Wu, W., 2020. Sophisticated investors and market efficiency: Evidence from a natural experiment. J. Financ. Econ. 138 (2), 316–341.
Chernenko, S., Sunderam, A., 2020. Do fire sales create externalities? J. Financ. Econ. 135 (3), 602–628.
Chiang, T.C., Zheng, D., 2010. An empirical analysis of herd behavior in global stock markets. J. Bank. Financ. 34 (8), 1911–1921.
Chiu, J., Koeppl, T.V., 2019. Blockchain-based settlement for asset trading. Rev. Financ. Stud. 32 (5), 1716–1753.
Chohan, U.W., 2021. Counter-hegemonic finance: The gamestop short squeeze. Available At SSRN.
Cipriani, M., Guarino, A., 2014. Estimating a structural model of herd behavior in financial markets. Amer. Econ. Rev. 104 (1), 224–251.
Coval, J., Stafford, E., 2007. Asset fire sales (and purchases) in equity markets. J. Financ. Econ. 86 (2), 479–512.
Cribari-Neto, F., 2004. Asymptotic inference under heteroskedasticity of unknown form. Comput. Statist. Data Anal. 45 (2), 215–233.
Desai, H., Krishnamurthy, S., Venkataraman, K., 2006. Do short sellers target firms with poor earnings quality? Evidence from earnings restatements. Rev. Account. Stud. 11 (1), 71–90.
Diebold, F.X., Yilmaz, K., 2009. Measuring financial asset return and volatility spillovers, with application to global equity markets. Econ. J. 119 (534), 158–171.
Diebold, F.X., Yilmaz, K., 2012. Better to give than to receive: Predictive directional measurement of volatility spillovers. Int. J. Forecast. 28 (1), 57–66.
Diether, K.B., Lee, K.-H., Werner, I.M., 2009. Short-sale strategies and return predictability. Rev. Financ. Stud. 22 (2), 575–607.
Dimpfl, T., Jank, S., 2016. Can internet search queries help to predict stock market volatility? Eur. Financial Manag. 22 (2), 171–192.
Escanciano, J.C., Lobato, I.N., 2009. An automatic portmanteau test for serial correlation. J. Econometrics 151 (2), 140–149.
Forbes, K.J., Rigobon, R., 2002. No contagion, only interdependence: measuring stock market comovements. J. Finance 57 (5), 2223–2261.
Fricke, C., Fricke, D., 2020. Vulnerable asset management? The case of mutual funds. Journal of Financial Stability 100800.
Garman, M.B., Klass, M.J., 1980. On the estimation of security price volatilities from historical data. J. Bus. 53 (1), 67–78.
Geanakoplos, J., 2010. The leverage cycle. NBER Macroeconomics Annual 24 (1), 1–66.
Goldstein, I., Guembel, A., 2008. Manipulation and the allocational role of prices. Rev. Econom. Stud. 75 (1), 133–164.
Grullon, G., Michenaud, S., Weston, J.P., 2015. The real effects of short-selling constraints. Rev. Financ. Stud. 28 (6), 1737–1767.
Hamid, A., Heiden, M., 2015. Forecasting volatility with empirical similarity and Google Trends. J. Econ. Behav. Organ. 117, 62–81.
Hau, H., Lai, S., 2017. The role of equity funds in the financial crisis propagation. Rev. Finance 21 (1), 77–108.
Jotikasthira, C., Lundblad, C., Ramadorai, T., 2012. Asset fire sales and purchases and the international transmission of funding shocks. J. Finance 67 (6), 2015–2050.
Kwiatkowski, D., Phillips, P.C., Schmidt, P., Shin, Y., 1992. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? J. Econometrics 54 (1–3), 159–178.
Lyócsa, v., Baumöhl, E., Výrost, T., Molnár, P., 2020. Fear of the coronavirus and the stock markets. Finance Res. Lett. 36, 101735.
Lyócsa, v., Molnár, P., 2020. Stock market oscillations during the corona crash: The role of fear and uncertainty. Finance Res. Lett. 36, 101707.
Lyócsa, v., Molnár, P., Vỳrost, T., 2021. Stock market volatility forecasting: Do we need high-frequency data? Int. J. Forecast.
Massicotte, P., Eddelbuettel, D., 2018. Gtrendsr: Perform and display google trends queries. R Package Version 1 (2).
Moat, H.S., Curme, C., Avakian, A., Kenett, D.Y., Stanley, H.E., Preis, T., 2013. Quantifying wikipedia usage patterns before stock market moves. Sci. Rep. 3 (1), 1–5.
Möllers, T.M., 2015. The takeover battle Volkswagen/Porsche: the Piëch-Porsche clan–family clan acquires a majority holding in Volkswagen. Capital Markets Law Journal 10 (3), 410–430.
Mori, T., 2016. Financial technology: Blockchain and securities settlement. J. Secur. Oper. Custody 8 (3), 208–227.
Okorie, D.I., Lin, B., 2021. Stock markets and the COVID-19 fractal contagion effects. Finance Res. Lett. 38, 101640.
Parkinson, M., 1980. The extreme value method for estimating the variance of the rate of return. J. Bus. 53 (1), 61–65.
Patton, A.J., Sheppard, K., 2009. Optimal combinations of realised volatility estimators. Int. J. Forecast. 25 (2), 218–238.
Pericoli, M., Sbracia, M., 2003. A primer on financial contagion. J. Econ. Surv. 17 (4), 571–608.
Politis, D.N., White, H., 2004. Automatic block-length selection for the dependent bootstrap. Econometric Rev. 23 (1), 53–70.
Preis, T., Moat, H.S., Stanley, H.E., 2013. Quantifying trading behavior in financial markets using Google Trends. Sci. Rep. 3 (1), 1–6.
Ramelli, S., Wagner, A.F., 2020. Feverish stock price reactions to COVID-19. Rev. Corp. Finance Stud. 9 (3), 622–655.
Rogers, L.C.G., Satchell, S.E., 1991. Estimating variance from high, low and closing prices. Ann. Appl. Probab. 1 (4), 504–512.
Ross, O., Jensen, J., 2019. Assets under Tokenization: Can Blockchain Technology Improve Post-Trade Processing? In: ICIS 2019 Proceedings.
Scharfstein, D.S., Stein, J.C., 1990. Herd behavior and investment. Am. Econ. Rev. 465–479.
Stein, J.C., 2012. Monetary policy as financial stability regulation. Q. J. Econ. 127 (1), 57–95.
Sul, D., Phillips, P.C., Choi, C.-Y., 2005. Prewhitening bias in HAC estimation. Oxford Bull. Econ. Stat. 67 (4), 517–546.
Taylor, N., 2017. Realised variance forecasting under Box-Cox transformations. Int. J. Forecast. 33 (4), 770–785.
Wang, G.-J., Xie, C., Stanley, H.E., 2018. Correlation structure and evolution of world stock markets: Evidence from pearson and partial correlation-based networks. Comput. Econ. 51 (3), 607–635.
White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 48 (4), 817–838.