Inzinerine Ekonomika-Engineering Economics, 2025, 36(2), 147–160

Event Sentiment and Cross-Country Herding Spillover Effects Using Machine Learning

Wei Chen[1*], Muhammad Asim[2*], Muhammad Yasir[3], Mehboob ul Hassan[4], Adnan Shoaib[5]

[1]School of International Education*
Zhejiang Sci-Tech University, Hangzhou, China
E-mail. chenwei@zstu.edu.cn

[2]Department of Management Sciences*
COMSATS University Islamabad, Wah Campus, Pakistan
E-mail. asim.tfw@gmail.com

[3]FAST School of Management
National University of Computer and Emerging Sciences, Islamabad, Pakistan
E-mail. yasir.tfw@gmail.com

[4]Islamic Banking Center, Department of Economics
College of Business Administration, King Saud University, Riyadh, Saudi Arabia
E-mail. mehassan@ksu.edu.sa ORCID: 0000-0002-3453-695X

[5]Greater Manchester Business School
University of Greater Manchester, Gulrez 2, Islamabad, Pakistan
E-mail. adnan.shoaib@uobislamabad.com

https://doi.org/10.5755/j01.ee.36.2.36329

This research study investigates herding behaviour and its cross-country spillover effects in the UK, US, China, and Pakistan stock markets in the presence of event sentiment. We used three machine learning models for the empirical investigation: support vector regression, single-layer neural networks, and multi-layer neural networks. The daily data set of the listed stocks has been used. The results suggest a significant predictability of the Twitter sentiment of Brexit 2016 and COVID-19. Cross-country herding spillover is also evident from the UK to Pakistan and the US in the case of Brexit 2016. Similarly, there is a herding spillover effect from China to Pakistan and UK stock markets. The overall results of machine learning models are more significant than linear regression models. Furthermore, the event sentiment increases the predictability of the machine learning models. The study provides a deep insight for individual and institutional investors to take care of unpredicted events while constructing their international portfolios in these stock markets.

Keywords: *Herding Spillover; Brexit; CSAD; Market Return; Machine Learning.*

Introduction

Behavioural Finance deals with the investor's decision making by considering the psychological biases (Barberis & Thaler, 2003). In the recent literature, researchers have focused on the investigation of market anomalies especially herding behaviour. It is defined as a behaviour to imitate the peers by ignoring personal beliefs. Herding leads investors to behave in a similar pattern. This behaviour causes the asset prices to deviate from their fair values. Consequently, stock markets experience high volatility (Blasco, Corredor, & Ferreruela, 2012)

Investigating herding behaviour has constantly been a complex task. An unexpected mimic by traders momentarily causes the stock prices to diverge from their fundamental values, resulting in an enlarged spread around their mean values. Previous literature mostly focused on statistical clustering procedures and technical investigation to check for varying patterns in the prices of stock markets. Modern literature has exposed this approach and, based on the Efficient Market (EMH) hypothesis, suggests that the fluctuations in the stock market are extremely dependent on the 'news' associated with company events or any erratic economic or political [1]events. Additionally, it is established that the news pooled on social media and Internet blogging mediums formulates an overall sentiment of the public, which generates more volatility and turnover in prices of stock markets than the news available on traditional media (Khan *et al.*, 2022).

Twitter is a big Internet blogging social media platform. The sentiment analysis using Twitter data has gained a significant interest of researchers in academia for its improved forecasting abilities (Yadav, Kudale, Rao, Gupta, & Shitole, 2021a, 2021b). Modern research on stock market forecasting using sentiment analysis proposes that machine learning forecasting techniques are more robust and deliver

**Wei Chen and Muhammad Asim have contributed equally. Hence both should be considered as lead author.*

more accurate findings than historical data (Benferhat, Tabia, & Ali, 2017).

By employing investor sentiment analysis, previous studies on stock market forecasting have led the foundation of modern research to examine a stock market anomaly, i.e., herding behaviour by employing big data and machine learning models in the context of event sentiment. It is quite important to explore how the social media news sentiment causes the investor to herd in the stock markets. When investors change their decisions because of the news regarding sudden events, this behaviour leads the markets to inefficiency and consequently markets may crash. In this context, the impact of twitter-based event sentiment of Brexit 2016 and COVID-19 on herding anomaly is still underexplored. Furthermore, the social media sentiment of Brexit 2016 has never been used to investigate the cross-country herding spillover effects from the origin country, the UK, to the US, China, and Pakistan. Hence, we intend to use machine learning models to forecast herding behaviour for better estimation accuracy. In this context, the contribution of this research is multidimensional. We incorporate the Twitter-based sentiment of two major events, i.e., Brexit 2016 and COVID-19, to predict the herding behaviour and its spillover in the stock markets of the US, the UK, China, and Pakistan.

- In the first step, we apply two machine learning models, i.e. SVR (support vector regression) and SLNN (single-layer neural network), to forecast the CSAD of all four stock markets. The forecasting accuracy improved after incorporating event sentiments of Brexit and COVID-19.
- In the second step, three different machine learning models, i.e. SVR, SLNN, and multilayer neural network models (MLNN), have been used to predict the herding behaviour with and without twitter-based event sentiments.
- Using the same models, we computed the cross-country herding effects coefficients in the third step.

The use of twitter-based sentiment to predict the herding anomaly in the stock market is the first main addition to the behavioural finance literature. The second contribution of this study is the use of machine learning models that provide more robust results and provide a better insight of the stock market data sets, especially in context of herding anomaly. Thirdly, we also investigated the response of individual investors in one stock market towards the movement in another stock market in the presence of twitter-based sentiment. This aspect has been unexplored in the existing literature of herding behaviour.

Literature Review

According to traditional finance, all the available information is reflected in the stock prices and investors behave rationally while taking decisions. Hence markets are efficient (Fama, 1960). On the other hand, behavioural finance argue that market participants are exposed to certain behavioural biases and sentiments while taking investment decisions (Ferreruela & Mallor, 2021). Behavioural biases and sentiments lead the market to experience certain anomalies. These market anomalies cause the stock prices to deviate from the fair prices. Hence markets become inefficient (Rossi & Gunardi, 2018a, 2018b; Woo, Mai, McAleer, & Wong, 2020).

Herding behaviour is one of the anomalies in the stock markets, which can be defined as a behavioural tendency characterized by trading in the same direction (Nofsinger & Sias, 1999). In such a scenario, investors illogically neglect their previous opinions and start to mimic others (Devenow & Welch, 1996). The asset prices tend to deviate from their fundamental values. This deviation injects volatility to the stock market (Fei & Liu, 2021). The empirical studies concerning herding anomalies are dynamic, and the outcomes may vary as per the model used. Christie and Huang (1995) considered the Cross-Sectional Standard Deviation (CSSD) method to examine the existence of herding anomaly in the US stock market. The study used the linear regression model and found no evidence of herding existence. Likewise, Gleason, Mathur, and Peterson (2004) employed data on nine sector exchange-traded funds in the US financial market and established the non-existence of herding behaviour under risky market conditions. Contrary to this, Hwang and Salmon (2004) established a different methodology built on the cross-sectional dispersion of the factor sensitivity of asset products and suggested that the herding anomaly existed in the US and South Korean stock markets. Similarly, Klein (2013) forecasted the herding anomaly by using the Markov regime-switching model and concluded that during the market stress periods like the global financial crisis, behavioural emotions affect market prices, resulting in international herding spillover effects. Moreover, BenSaïda (2017) also came up with the same results and suggested the existence of herding behaviour in almost all the sectors of the US stock market during the market stress time period. This research study used the GJR-GARCH methodology, which delivers some deep insights into the association between herding behaviour and idiosyncratic volatility. This can be concluded that the empirical herding research is dynamic and proposes mixed results based on the model used (Balcilar, Demirer, & Hammoudeh, 2013; Chiang, Li, & Tan, 2010; Economou, Hassapis, & Philippas, 2018; Mahmud & Tiniç, 2018)..

Investigating the herding anomaly in the context of social media is a comparatively new phenomenon. The commencement of high volatility in stock markets combined with tranquil events results in investors accidentally trusting various information sources to support their decision-making. Modern studies propose volatility is efficiently foreseeable using the social media news sentiments (Alomari *et al.*, 2021; Jiao, Veiga, & Walther, 2020; Van Dieijen, Borah, Tellis, & Franses, 2020). Moreover, Lehrer, Xie, and Zhang (2021) have integrated deep learning tools of machine learning to calculate Twitter sentiment and determined that tools such as convolutional and recurrent neural networks expressively enhanced the volatility predicting precision of the model used. Similarly, with more evidence, social media is now categorised as an exogenous variable that can stimulate stock prices, which means that a new branch of social media network and social media sentiment analysis has appeared to get information from word-based data to expand the effectiveness of predicting models (Karlemstrand & Leckstrom, 2021). Sibande, Gupta, Demirer, and Bouri (2023) made a link between herding anomaly and the sentiment of investors in the currency market by employing Twitter data. By applying a quantile model, they determined that the herding phenomenon is specific to the regime and is predominantly prominent at times of extreme investor sentiment. The conclusions offered backing to behavioural aspects for asset pricing models and recommended that the sentiment signals in real-time can be used for forecasting probable speculative patterns in economic markets. On similar patterns, Maqsood et al. (2020) examined the effect of major local and global events by employing the dataset of Twitter on stock exchange prediction.

A recent literature review proposes that artificial intelligence techniques, such as deep learning models combined with traditional forecasting techniques, can generate new metrics in financial engineering and motivate further research ideas. Financial engineering is very important for risk management practices; therefore, the payback of using deep learning models should be practiced in finance to gain the improved accuracy of predicted models. In addition, number of recent studies have argued that the stock markets become inefficient during uncertain time periods. (See (Gaio *et al.*, 2022; Ozkan, 2021). When stock markets face uncertainty, they experience certain anomalies. Herding behaviour is one of the major such anomalies. This study, therefore, fills this gap by examining the herding anomaly by using the event-based sentiment. To accomplish this purpose, Twitter-based sentiment analysis needs to be done in the presence of unpredictable events like Brexit 2016, and COVID-19. In addition, the cross-country herding spillover effects are also unexplored, especially in the context of Brexit 2016.

Theory and Theorizing

In this section, we discuss the theoretical background of the herding anomaly investigation in detail.

Cross Sectional Absolute Deviation (CSAD) Methodology

To measure the herding behaviour in stock markets, CSAD is considered as a better measure. It is explained as under:

$$CSAD_t = \frac{1}{N}\sum_{i=1}^{N}\left|R_{i,t} - R_{m,t}\right| \quad (1)$$

It is argued that herding behaviour prevails when there is a negative significant non-linear relationship exists among CSAD and market return. (Chang et al., 2000 and Hanker et al., 2006). The following Eq.2 is used to investigate the relationship.

$$CSAD_t = \beta_o + \beta_1\left|R_{m,t}\right| + \beta_2 R_{m,t}^2 + e_t \quad (2)$$

The coefficient β_2 in above Eq.2. represents the presence or absence of herding behaviour in the stock market. A negative and significant value of this coefficient shows the presence of herding behaviour (Economou et al., 2011).

Figure 1 shows the dynamic relationship between *CSAD* and R_m. The graph of CSAD describes the basis for why the herding coefficient β_2 should be negative for herding evidence. CSAD is maximum when $R_{m,t}^* = -{\beta_1}/{2\beta_2}$.

CAPM (Capital Asset Pricing Model)

According to traditional finance, the Capital Asset Pricing Model (CAPM) presents the relationship between the expected returns $E(R_i)$ of the risky portfolios considering their systematic risk Beta. The following equation shows this relationship:

$$E(R_i) - R_f = \beta_i\left[E(R_m) - R_f\right]$$

$$E(R_i) = R_f + \beta_i\left[E(R_m) - R_f\right]$$

Here, $E(R_i)$ shows the predicted return of asset, $E(R_m)$ represents the predictable stock portfolio's return, β_i= ${COV(R_i, R_m)}/{VAR(R_m)}$ is the systematic risk. Whereas R_f shows the risk-free rate.

If the price estimation is done according to the CAPM, a gradually increasing relationship among *CSAD* and $R_{m,t}$. prevails. Furthermore, this scenario explains rationality which is unable to explain the market anomaly.

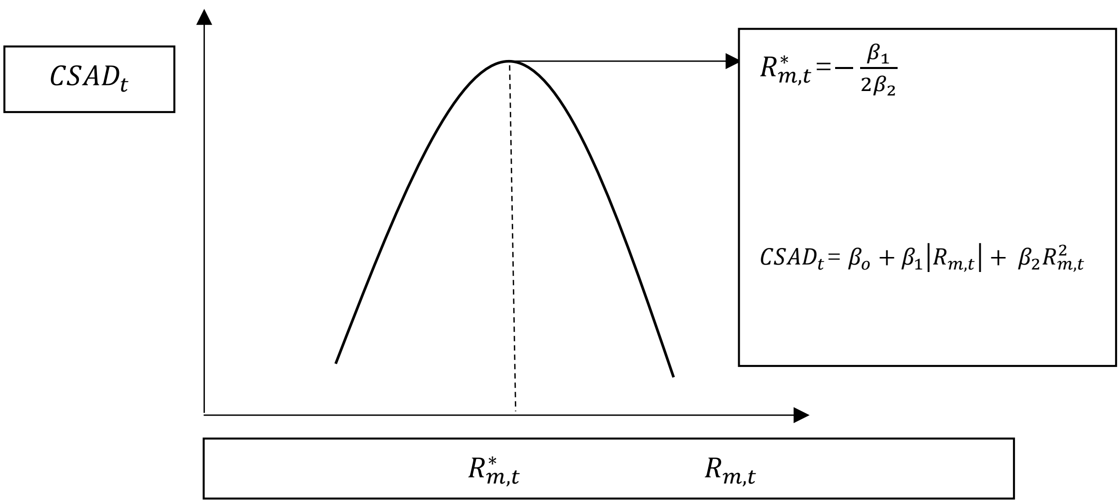

Figure 1. CSAD Quadratic Evolution Curve because of Mean Market Returns

Data and Methodology

In this part, we discuss the data description, and the methodology used for the investigation of herding behaviour in the UK, US, China, and Pakistan. We used the secondary data set comprising of individual stocks and market index. The data covers the period starting from 2006 to 2022. This period covers the two major events i.e. Brexit 2016 and COVID-19 which are intended to be investigated. We use three different machine learning models, i.e., support vector regression (SVR), single-layer neural networks (SLNN), and multi-layer neural networks (MLNN) model, for the prediction of CSAD and the presence of herding anomaly in these stock markets. We selected the stock markets of US and UK as developed markets whereas China and Pakistan stock markets have been chosen as emerging markets. In addition, the event of Brexit was mainly relevant to the UK. Whereas the origin of COVID-19 is from China. Hence, the focus of our study is to investigate the presence of herding and its spillover among these specific stock markets in presence of twitter-based event sentiments.

Data Acquisition

The following Table 1 shows the details of textual data and sources.

Table 1

Description of the Stock Data Set and Twitter Sentiment of Events

Countries	Stock Data	Duration	Event Twitter Sentiment
UK	Top 200 Companies	May 2006 to Dec 2022	Brexit 2016
China	750 stock of Shanghai stock exchange	May 2006 to Dec 2022	COVID'19
US	500 stocks listed in the S&P 500	May 2006 to Dec 2022	
Pakistan	PSX 100	May 2006 to Dec 2022	

Author's compilation

This study considered the daily data of stocks listed in the UK, China, US, and Pakistan stock markets. We used the top 200 companies of the UK, and 500 companies of the US. 750 stocks of China and the top 100 stocks of the Pakistan stock exchange. Data set ranges from Jan 2006 to Dec 2022. Daily data of individual stocks and indices has been collected from Yahoo Finance. We used Cross-Sectional Absolute Deviation (CSAD) presented by Chang et al. (2000) for the investigation of herding behaviour.

Support Vector Regression

Support vector regression is used to solve regression problems and generalize support vector machines. Furthermore, SVR has a tendency to deal; with both linear and non-linear relationships among the targeted variables (Adcock & Gradojevic, 2019; Zhong *et al.*, 2019). The SVR determines the best suitable line within a given range of data sets. Eq.3 represents the mathematical expression used to build a functional estimator using a fraction of the given dataset.

$$f(x) = w^T \emptyset(x) + b \tag{3}$$

$w \in R^n$ shows the weighted feature vector of regressors and their coefficients. Whereas b is the constant term. In our herding model, x shows the explanatory variables, i.e., market return, the square of market returns etc. SVR also sets a threshold error tolerance ε, which is like linear regression.

Single Layer Neural Network

A SLNN contains an input and a single output layer. The total number of neurons or nodes in the layer used for input is equal to the independent variables provided. In our model, the input layer contains absolute market return, the square of market returns, and events sentiments of COVID-19 and Brexit 2016. Our output layer is the CSAD of the UK, US, China, and Pakistan. $x = (x_{i1}, x_{i2}, x_{i3}, \ldots \ldots . x_{in})$ represents the input layer. Each x_{in} is provided to a different node at the input layer.

The following procedure shows the process of calculating the output of the nodes from 1 to n:

$$\left(u_1 = \left(\sum i w_i x_i + b\right), u_2 = \left(\sum i w_i x_i + b\right) \ldots \ldots \ldots . u_n \left(\sum i w_i x_i + b\right)\right) \tag{4}$$

$u_1 \ldots u_n$ represent the total number of nodes in the data layers for input variables and w_i shows the respective weights. Every independent variable is multiplied to its respective weight. When we add weight inputs in every layer and added input equals zero, bias is applied so that the outcome is non-zero. Moreover, the SLNN algorithm starts learning by calibrating or adjusting the weights and continues by calculating the model error, which is the difference between the forecasted model outcome and the actual value of the CSAD of the UK, US, China, and Pakistan stock market. This process is iterated to minimize the error. In this research study, our SLNN model is trained for 20 epochs with a learning rate of 0.01 and an array size of 4. The initial weights of layers start with the Glorot uniform method, also called Xavier uniform initialization. These weights improve during the algorithm's training process by using the "Adam" weight optimizer (Glorot & Bengio, 2010; Kingma & Ba, 2014).

Multi-Layer Neural Network

The MLNN is better than classical linear regression because of its multiple-layered architectural design that involves non-linear activation functions. Inside each layer of the MLNN, different activation functions are introduced after the actual transfer function. In this research study, the architecture of MLNN includes three layers where the first and second one consists of hidden units equal to the number of independent variables "absolute market return", "square of market return", and event sentiments of Brexit 2016 and

COVID-19 followed by ReLu activation functions. Eq. (5) is the mathematical form of ReLu (Nair & Hinton, 2010).

$$ReLu(x) = \begin{cases} x & if \quad x > 0 \\ 0 & if \quad x \leq 0 \end{cases} \quad (5)$$

The result outcome of each neuron u is given in below equation (6):

$$u = \varphi\left(\sum iw_i a_i + b\right) \quad (6)$$

In the above equation (6), w_i is the weight associated with each relationship between the nodes, a_i is the input, b represents the bias term, and φ is referred to as the activation function that is ReLu on hidden layers. $(\sum iw_i a_i + b)$ represents the forecasted value of dependent variable, i.e., CSAD of the UK, US, China, and Pakistan. The evaluation metrics used in our study is as follows:

$$MSE = \frac{1}{N}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (7)$$

In the above equation (7), y_i is the actual value of a CSAD of the UK, US, China, and Pakistan stock markets, as well as $\hat{y}_i$, represents the forecasted value of CSAD.

Results and Discussion

In this section, we present the results of all the machine learning models we used to investigate herding behaviour in the UK, the US, China, and Pakistan.

Herding Behaviour in the US, UK, Pakistan and China

$$CSAD_t = \beta_o + \beta_1|R_{m,t}| + \beta_2 R_{m,t}^2 + e \quad (2)$$

$$CSAD_t = \beta_o + \beta_1|R_{m,t}| + \beta_2 R_{m,t}^2 + \beta_3 SBrexit + e \quad (8)$$

$$CSAD_t = \beta_o + \beta_1|R_{m,t}| + \beta_2 R_{m,t}^2 + \beta_3 SCOVID + e \quad (9)$$

CSAD Forecasting with and Without Sentiment

In this section, we discuss the results of CSAD forecasting with and without incorporating event sentiment of Brexit 2016 and COVID-19. Table 2 shows the results of three evaluation matrices, i.e., MAE, MSE, and RMSE. We apply two machine learning models i.e., support vector regression and single-layer neural networks to forecast the CSAD of the US, UK, Pakistan, and Chinese stock markets. We also compare our findings with the evaluation matrices provided by linear regression.

Table 2

CSAD Forecasting with and Without Sentiment of Brexit 2016 and Covid19

CSAD forecasting without sentiment			
US			
	MAE	MSE	RMSE
LR	0.3281	0.1077	0.3041
SVR	0.2785	0.1116	0.3342
SLNN	0.3040	0.1070	0.3271
UK			
LR	0.0355	0.0030	0.0550
SVR	0.0591	0.0048	0.0697
SLNN	0.0362	0.0030	0.0551
Pakistan			
LR	0.0371	0.0019	0.0447
SVR	0.0556	0.0039	0.0631
SLNN	0.0411	0.0023	0.0481
China			
LR	0.0740	0.0106	0.1032
SVR	0.0768	0.0107	0.1037
SLNN	0.0730	0.0105	0.1031
CSAD forecasting with sentiment of Brexit 2016			
US			
LR	0.3003	0.1062	0.3259
SVR	0.2903	0.1139	0.3376
SLNN	0.3001	0.1060	0.3256
UK			
LR	0.0355	0.0030	0.0550
SVR	0.0591	0.0048	0.0697
SLNN	0.0349	0.0020	0.0543
Pakistan			
LR	0.0371	0.0019	0.0446
SVR	0.0556	0.0039	0.0629
SLNN	0.0377	0.0020	0.0453
China			
LR	0.0740	0.0108	0.1030
SVR	0.0728	0.0107	0.1026
SLNN	0.0747	0.0106	0.1030

CSAD forecasting with sentiment of COVID'19			
	US		
LR	0.2959	0.1049	0.3240
SVR	0.2899	0.1127	0.3357
SLNN	0.2991	0.1054	0.3247
	UK		
LR	0.0369	0.0031	0.0513
SVR	0.0594	0.0019	0.0500
SLNN	0.0460	0.0016	0.0402
	Pakistan		
LR	0.0371	0.0020	0.0447
SVR	0.0554	0.0039	0.0627
SLNN	0.0367	0.0021	0.0460
	China		
LR	0.0739	0.0106	0.1031
SVR	0.0703	0.0100	0.1036
SLNN	0.0700	0.0100	0.1030

, ** and * means significance of coefficients at 10%, 5% and 1%.*

Table 2 shows the results of CSAD forecasting. Cross sectional absolute deviation is a divergence of individual stock returns from the market return. Herding behaviour prevails in the stock market when CSAD has a nonlinear relationship with market movement. In this context, it is quite crucial to predict the movement of CSAD overtime to foresee the prevalence of herding behaviour. The first part of Table 2 shows forecasting without considering any event sentiment. However, the second and third section presents the forecasting considering the Twitter sentiment of Brexit 2016 and COVID-19.

In the first part, it is evident from the evaluation matrices results that for the US and China, the SLNN model performs better as compared to other models. For Pakistan and the UK, LR provides better forecasted results.

When we incorporate the Twitter sentiment of the Brexit event, the results of machine learning models improve significantly, especially for the UK, US, and China. SLNN outperforms SVR and LR based on all the evaluation matrices. However, LR remains a better model for Pakistan. An interesting insight is the better performance of machine learning models for the UK, when we incorporated the sentiment of Brexit, which was not the case before. Our these findings are in line with the results reported by Maqsood et al. (2020). The only difference is that they forecasted the individual stock prices rather than the return deviations.

When we incorporate the Twitter sentiment of the COVID-19 event, the results of machine learning models improve significantly for all the stock markets. The LR performs better for the US stock market. For the US stock market, the predictability of machine learning models is more accurate when we use COVID-19 sentiment as compared to Brexit 2016. Therefore, it is argued that the US stock market remained more sensitive to COVID-19 as compared to Brexit 2016. In addition, SVR and SLNN perform better in the case of the UK and China in the presence of COVID-19 sentiment. A similar results are also reported by Yasir et al. (2023). They forecasted the time series data of crypto currency using machine learning models and event-based sentiments. They reported the better performance of machine learning models after incorporating twitter-based sentiment of multiple events. The prediction matrices show that LR remains a better model for predicting the stock market volatility of Pakistan in the presence of COVID-19 sentiment.

Estimation of Herding Coefficients without Event Sentiment

In this part, we estimated the herding coefficients using machine learning models, i.e., SVR, SLNN, and MLNN, and compared the findings with LR. Eq.2. has been used for the estimation of coefficients. We estimated the coefficients without any social media sentiment. The results are reported in the following Table 3.

Table 3

Herding Coefficients without Sentiment

Models	**US**		**UK**	
	$\boldsymbol{\beta_1}$	$\boldsymbol{\beta_2}$	$\boldsymbol{\beta_1}$	$\boldsymbol{\beta_2}$
LR	-0.647***	-0.480	0.543***	0.058
	(0.000)	(0.125)	(0.000)	(0.164)
SVR	-1.172***	**-0.749***	0.400***	0.105
	(0.000)	**(0.067)**	(0.000)	(0.158)
SLNN	-0.640***	**-0.494****	0.564***	0.036
	(0.000)	**(0.045)**	(0.000)	(0.487)
MLNN	0.225***	**-0.478***	0.365***	**-0.558***
	(0.000)	**(0.092)**	(0.000)	**(0.068)**
Models	**US**	**UK**	**Models**	**US**

	Pakistan		China	
LR	1.572***	0.1924	0.456***	**-0.106****
	(0.000)	(0.325)	(0.000)	**(0.042)**
SVR	1.211***	0.016	0.340**	0.020
	(0.000)	(0.235)	(0.000)	(0.689)
SLNN	1.559***	0.1290	0.434***	**-0.089****
	(0.000)	(0.125)	(0.000)	**(0.032)**
MLNN	0.225***	-0.887	0.778***	0.553
	(0.000)	(0.225)	(0.000)	(0.122)

, ** and * means significance of coefficients at 10%, 5% and 1%.*

Our results provide evidence of herding behaviour in all the stock markets except Pakistan. The herd anomaly prevails in the stock market if the coefficient of squared market return is negative and statistically significant. The herding coefficient β_2 is highly significant for the US and China when the SLNN model is used for estimation. Their values are -0.494 and -0.089 for the US and China, respectively. In addition, the MLNN model provides evidence of herding behaviour in the US and UK stock markets. The respective herding coefficients are -0.478 and -0.558, which are significant at the 10% level. Furthermore, the findings of LR in the case of China also show evidence of herding behaviour in the Chinese stock market. The respective P-values are reported in parentheses. The machine learning models provide significant evidence as compared to linear regression. A similar results are reported by Jabeen et al. (2022) in context of stock returns prediction. They used macroeconomic variables to predict the predict stock returns. They deployed multiple machine learning models and reported their significant predictability as compared to linear regression.

Estimation of Herding Coefficients with Event Sentiment of Brexit and COVID-19

In this part, we estimated the herding coefficients using machine learning models, i.e., SVR, SLNN, and MLNN, and compared the findings with LR. Eq.8 and 9 have been used for the estimation of coefficients. To estimate the herding coefficients, we incorporated the Twitter sentiment of two major events, i.e., Brexit 2016 and COVID-19. The results are reported in the following Table 4.

Table 4

Herding Coefficients of US, UK, PK, and China with Sentiment of Brexit 2016 and COVID-19

With Sentiment of Brexit 2016						
		US			**UK**	
	β_1	β_2	β_3	β_1	β_2	β_3
LR	-0.637***	**-0.470***	0.295***	0.544***	**0.057****	0.004***
	(0.000)	**(0.089)**	(0.000)	(0.000)	**(0.049)**	(0.000)
SVR	-1.170***	**-0.742***	0.226***	0.401***	**-0.105*****	0.025***
	(0.000)	**(0.058)**	(0.000)	(0.000)	**(0.000)**	(0.000)
SLNN	-0.695***	**-0.502*****	0.264	0.542***	**-0.011*****	-0.035
	(0.000)	**(0.000)**	(0.000)	(0.000)	**(0.000)**	(0.000)
MLNN	0.368***	**-0.225*****	1.258***	0.369***	**-0.129*****	2.369***
	(0.000)	**(0.000)**	(0.000)	(0.000)	**(0.000)**	(0.000)
		Pakistan			**China**	
LR	1.572***	0.192***	-0.005***	0.456***	**-0.106*****	-0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	**(0.000)**	(0.000)
SVR	1.211***	**-0.016*****	0.006**	0.340**	0.023	0.008***
	(0.000)	**(0.000)**	(0.035)	(0.000)	(0.000)	(0.000)
SLNN	1.528***	0.151	-0.024	0.412***	**-0.094*****	-0.055***
	(0.000)	(0.000)	(0.125)	(0.000)	**(0.000)**	(0.000)
MLNN	**1.336*****	**-0.456*****	2.336***	1.225***	**-0.228*****	4.669***
	(0.000)	**(0.000)**	(0.000)	(0.000)	**(0.000)**	(0.000)
With sentiment of COVID'19						
	β_1	β_2	β_3	β_1	β_2	β_3
		US			**UK**	
LR	-0.656***	**-0.570*****	0.269***	0.539***	0.057***	0.085***
	(0.000)	**(0.004)**	(0.000)	(0.000)	(0.000)	(0.000)
SVR	-1.215***	**-0.738*****	0.168***	0.412***	0.069***	0.059***
	(0.000)	**(0.000)**	(0.000)	(0.000)	(0.000)	(0.000)
SLNN	-0.705***	**-0.630*****	0.245***	0.506***	**-0.022*****	0.078***
	(0.000)	**(0.000)**	(0.000)	(0.000)	**(0.000)**	(0.000)
MLNN	1.269***	**-1.445*****	2.554***	1.396***	**-0.996*****	1.447***
	(0.000)	**(0.000)**	(0.000)	(0.000)	**(0.000)**	(0.000)
	Pakistan				**China**	
LR	1.574***	0.191	-0.009***	0.455***	**-0.106***	-0.018***
	(0.000)	(0.145)	(0.000)	(0.000)	**(0.068)**	(0.000)

	With Sentiment of Brexit 2016					
SVR	1.213 (0.000)	**-0.018***** **(0.000)**	-0.004*** (0.000)	0.338*** (0.000)	**-0.026***** **(0.000)**	-0.005*** (0.000)
SLNN	1.649 (0.000)	0.161 (0.268)	-0.014*** (0.000)	0.415*** (0.000)	**- 0.106***** **(0.000)**	-0.014*** (0.000)
MLNN	1.886*** (0.000)	**-0.225**** **(0.036)**	1.663*** (0.000)	1.776*** (0.000)	**-0.338***** **(0.000)**	0.558*** (0.000)

, ** and * means significance of coefficients at 10%, 5% and 1%.*

When we incorporate the Twitter sentiment of Brexit 2016, our findings provide evidence of herding behaviour in all the stock markets. All the models provide significant evidence of herding prevalence for the US stock market. Similarly, in the UK stock market, the herding coefficient becomes highly significant and negative for all the models when Twitter sentiment has been used. This is an interesting insight into the stock market of the UK. Similarly, the performance of machine learning models also improved for Pakistan and China after incorporating the Twitter sentiment of Brexit 2016. The overall performance of MLNN has improved as compared to other models.

The second part of the above Table 4 presents the results of herding coefficient estimations using sentiment of COVID-19. The reported results show the significant herding predictability of COVID-19, especially in the Chinese stock market. All the models strongly accept the herding hypothesis for China and the US. However, machine learning models outperform the linear regression model in the case of Pakistan and the UK. The coefficient β_3 of both Brexit 2016 and COVID-19 event sentiment is highly significant at 1%. These findings are similar to the results reported by Jabeen et al. (2022) in context of stock returns prediction. They considered economic policy uncertainty news sentiment to predict the stock returns using machine learning models and reported their significant predictability as compared to linear regression. Furthermore, their results show that the news sentiment is a significant predictor of stock return. Our results are also similar to Gaio et al. (2022). They argue that the stock markets of developed countries became inefficient during the period of crises. Similarly, Ozkan (2021) also presents the similar impact of COVID-19 on the stock markets. In this context, our results are aligning with the existing research but with a unique footprint.

Herding Spillover from the UK to Pakistan, China, and the US without Event Sentiment

We investigated the herding spillover effect from the UK to Pakistan, China, and the US in this part. We have used Eq. 10, 11, and 12 to calculate herding coefficients without Twitter sentiment. We used LR, SVR, SLNN and MLNN; the results are reported in Table 5.

$$CSAD_{PK,t} = \beta_{o1} + \beta_{11}\left|R_{m,PK,t}\right| + \beta_{21}R^2_{m,PK,t} + \beta_{31}CSAD_{UK,t} + \beta_{41}R^2_{m,UK,t} + \varepsilon_t \quad (10)$$

$$CSAD_{C,t} = \beta_{o1} + \beta_{11}\left|R_{m,C,t}\right| + \beta_{21}R^2_{m,C,t} + \beta_{31}CSAD_{UK,t} + \beta_{41}R^2_{m,UK,t} + \varepsilon_t \quad (11)$$

$$CSAD_{US,t} = \beta_{o1} + \beta_{11}\left|R_{m,US,t}\right| + \beta_{21}R^2_{m,US,t} + \beta_{31}CSAD_{UK,t} + \beta_{41}R^2_{m,UK,t} + \varepsilon_t \quad (12)$$

Table 5

Herding Spillover from the UK to Pakistan, China, and the US without Sentiment

			Without Sentiment		
	$\boldsymbol{\beta_{o1}}$	$\boldsymbol{\beta_{11}}$	$\boldsymbol{\beta_{21}}$	$\boldsymbol{\beta_{31}}$	$\boldsymbol{\beta_{41}}$
			LR		
Pakistan	0.102*** (0.000)	0.449*** (0.000)	**-0.099*** **(0.052)**	0.125*** (0.000)	**-0.163***** **(0.000)**
China	0.160*** (0.000)	0.437*** (0.000)	**-0.103*** **(0.054)**	0.130*** (0.000)	0.054 (0.306)
US	0.513*** (0.000)	-0.356*** (0.000)	0.044 (0.792)	-0.832*** (0.000)	0.185 (0.312)
			SLNN		
Pakistan	0.098*** (0.000)	0.442*** (0.000)	-0.036 (0.503)	0.101*** (0.000)	-0.047 (0.222)
China	0.174*** (0.000)	0.432*** (0.000)	-0.065 (0.230)	0.166*** (0.000)	0.092* (0.087)
US	0.545*** (0.000)	-0.413*** (0.000)	-0.070 (0.683)	-0.840*** (0.000)	0.301 (0.103)
			SVR		
Pakistan	0.101*** (0.000)	0.691*** (0.000)	**-0.515***** **(0.000)**	0.047*** (0.012)	-0.057 (0.158)
China	0.184*** (0.000)	0.332*** (0.000)	0.018 (0.736)	0.092*** (0.000)	0.023 (0.662)
US	0.677*** (0.000)	-0.489*** (0.000)	0.011 (0.950)	-1.685*** (0.000)	0.558*** (0.000)

			MLNN		
Pakistan	0.137*** (0.000)	0.212*** (0.000)	0.060 (0.249)	-0.829*** (0.000)	0.070* (0.072)
China	0.205*** (0.000)	-0.288*** (0.000)	-0.929*** (0.000)	-0.265*** (0.000)	**-0.211*** (0.000)**
US	0.244*** (0.000)	-0.220** (0.017)	0.454*** (0.002)	-0.538*** (0.000)	0.039 (0.826)

, ** and * means significance of coefficients at 10%, 5% and 1%.*

The results reported in Table 5 show the herding spillover from the UK to Pakistan and China. The herding spillover coefficient β_{41} is highly significant for Pakistan when LR is used. Similarly, it is significant for China at 1% when the MLNN model is used. The respective values are -0.163 and -0.211, respectively. The respective P-values are reported in parenthesis below the coefficients. These findings are similar to the results reported by Yasir and Onder (2023). They also presented the evidence of herding spillover in emerging stock markets of BRICS and Turkey using structural break approach.

Herding Spillover from the UK to Pakistan, China, and the US with event sentiment of Brexit 2016

In this part, we investigated the herding spillover effect from the UK to Pakistan, China, and the US in the presence of Twitter sentiment of Brexit 2016. We have used Eq. 13, 14, and 15 to calculate herding coefficients. We used LR, SVR, SLNN and MLNN; the results are reported in Table 6.

$$CSAD_{PK,t} = \beta_{o1} + \beta_{11}\left|R_{m,PK,t}\right| + \beta_{21}R^2_{m,PK,t} + \beta_{31}CSAD_{UK,t} + \beta_{41}R^2_{m,UK,t} + \beta_{51}SBrexit + \varepsilon_t \quad (13)$$

$$CSAD_{C,t} = \beta_{o1} + \beta_{11}\left|R_{m,C,t}\right| + \beta_{21}R^2_{m,C,t} + \beta_{31}CSAD_{UK,t} + \beta_{41}R^2_{m,UK,t} + \beta_{51}SBrexit + \varepsilon_t \quad (14)$$

$$CSAD_{US,t} = \beta_{o1} + \beta_{11}\left|R_{m,US,t}\right| + \beta_{21}R^2_{m,US,t} + \beta_{31}CSAD_{UK,t} + \beta_{41}R^2_{m,UK,t} + \beta_{51}SBrexit + \varepsilon_t \quad (15)$$

Table 6

Herding Spillover from UK to Pakistan, China and US with Sentiment of Brexit 2016

			With Sentiment of Brexit 2016			
	$\boldsymbol{\beta_{o1}}$	$\boldsymbol{\beta_{11}}$	$\boldsymbol{\beta_{21}}$	$\boldsymbol{\beta_{31}}$	$\boldsymbol{\beta_{41}}$	$\boldsymbol{\beta_{51}}$
			LR			
Pakistan	0.103*** (0.000)	0.448*** (0.000)	**-0.0994** (0.0496)**	0.125*** (0.000)	**-0.163*** (0.000)**	-0.022* (0.098)
China	0.160*** (0.000)	0.431*** (0.000)	**-0.103** (0.0487)**	0.135*** (0.000)	0.054 (0.306)	0.0006 (0.963)
US	0.509*** (0.000)	-0.344*** (0.000)	**0.058 (0.735)**	-0.837*** (0.000)	0.184*** (0.000)	0.300*** (0.000)
			SLNN			
Pakistan	0.110*** (0.000)	0.414*** (0.000)	**-0.135*** (0.002)**	0.070*** (0.000)	**-0.0628* (0.098)**	-0.036*** (0.003)
China	0.166*** (0.000)	0.413*** (0.000)	**-0.087* (0.0687)**	0.145*** (0.000)	0.065 (0.212)	0.003 (0.873)
US	0.529*** (0.000)	-0.391*** (0.000)	0.009 (0.896)	-0.843*** (0.000)	0.304* (0.090)	0.292*** (0.000)
			SVR			
Pakistan	0.101*** (0.000)	0.694*** (0.000)	**-0.514*** (0.000)**	0.046** (0.020)	**-0.056* (0.078)**	0.002 (0.687)
China	0.184*** (0.000)	0.336*** (0.000)	0.011 (0.863)	0.090*** (0.000)	0.022 (0.691)	0.010 (0.503)
US	0.025*** (0.000)	0.458** (0.047)	0.029 (0.125)	0.078* (0.069)	**-0.589*** (0.006)**	0.692*** (0.000)
			MLNN			
Pakistan	0.145*** (0.000)	-0.318*** (0.000)	**-0.033 (0.513)**	0.040** (0.041)	**-0.241*** (0.000)**	-0.352*** (0.000)
China	0.196*** (0.000)	-0.127*** (0.000)	0.584*** (0.000)	-0.410*** (0.000)	0.351*** (0.000)	0.152*** (0.000)
US	0.240*** (0.000)	0.040 (0.662)	0.425** (0.010)	-0.732*** (0.000)	**-0.741*** (0.000)**	0.353*** (0.000)

, ** and * means significance of coefficients at 10%, 5% and 1%.*

The estimated results show that according to all the models, there exists strong evidence of herding spillover from the UK to Pakistan in the presence of the event sentiment of Brexit 2016. There is no evidence of herding spillover from the UK to the Chinese stock market in the presence of event sentiment of Brexit 2016. In addition, the US stock market remains sensitive towards the event sentiment of Brexit 2016. There is significant evidence of herding spillover from the UK to the US stock market in the presence of event sentiment when we use machine learning models of SVR and MLNN. The coefficient of event sentiment of Brexit 2016 β_{51} is highly significant at 1% in all the models, showing high predictability. These results are like our earlier results of CSAD forecasting with sentiment. In this context, machine learning models provide a deep insight into stock trading in this set of stock markets.

Herding Spillover from China to the US, Pakistan, and the UK without Event Sentiment

We investigated the herding spillover effect from China to the US, Pakistan, and the UK in this part. We have used Eq. 16, 17, and 18 to calculate herding coefficients without any Twitter sentiment. We used LR, SVR, SLNN, and MLNN; the results are reported in Table 7

$$CSAD_{us,t} = \beta_{o1} + \beta_{11}\left|R_{m,us,t}\right| + \beta_{21}R^2_{m,us,t} + \beta_{31}CSAD_{c,t} + \beta_{41}R^2_{m,c,t} + \varepsilon_t \quad (16)$$

$$CSAD_{pk,t} = \beta_{o1} + \beta_{11}\left|R_{m,pk,t}\right| + \beta_{21}R^2_{m,pk,t} + \beta_{31}CSAD_{c,t} + \beta_{41}R^2_{m,c,t} + \varepsilon_t \quad (17)$$

$$CSAD_{UK,t} = \beta_{o1} + \beta_{11}\left|R_{m,UK,t}\right| + \beta_{21}R^2_{m,UK,t} + \beta_{31}CSAD_{c,t} + \beta_{41}R^2_{m,c,t} + \varepsilon_t \quad (18)$$

Table 7

Herding Spillover from China to the US, Pakistan, and UK without Sentiment

	Without Sentiment				
	$\boldsymbol{\beta_{o1}}$	$\boldsymbol{\beta_{11}}$	$\boldsymbol{\beta_{21}}$	$\boldsymbol{\beta_{31}}$	$\boldsymbol{\beta_{41}}$
			LR		
US	0.609***	-0.468***	**-0.311*** **	-0.809***	0.024
	(0.000)	(0.000)	**(0.000)**	(0.000)	(0.762)
Pakistan	0.108***	0.451***	-0.076	0.036***	**-0.037* **
	(0.000)	(0.000)	(0.138)	(0.006)	**(0.060)**
UK	0.067***	0.525***	0.060	0.047***	0.030*
	(0.000)	(0.000)	(0.274)	(0.000)	(0.077)
			SLNN		
US	0.608***	-0.499*** (0.000)	**-0.361** **	-0.760***	-0.112
	(0.000)		**(0.023)**	(0.000)	(0.168)
Pakistan	0.102***	0.454***	-0.058	0.020	**-0.059*** **
	(0.000)	(0.000)	(0.253)	(0.134)	**(0.002)**
UK	0.061***	0.518***	-0.003	0.043***	0.049***
	(0.000)	(0.000)	(0.957)	(0.000)	(0.004)
			SVR		
US	0.765***	-0.631	**-0.362** **	-1.295***	-0.082
	(0.000)	(0.000)	**(0.026)**	(0.000)	(0.331)
Pakistan	0.108***	0.671	**-0.448*** **	0.001	-0.016
	(0.000)	(0.000)	**(0.000)**	(0.964)	(0.419)
UK	0.127***	0.402***	0.105	0.001	0.009
	(0.000)	(0.000)	(0.105)	(0.936)	(0.654
			MLNN		
US	0.237***	-0.254***	0.664***	-0.185***	-0.040
	(0.000)	(0.001)	(0.000)	(0.000)	(0.681)
Pakistan	0.136***	-0.042	0.315***	-0.036***	0.134***
	(0.000)	(0.107)	(0.000)	(0.009)	(0.000)
UK	0.064***	-0.124***	**-0.978*** **	0.223***	0.107***
	(0.000)	(0.000)	**(0.000)**	(0.000)	(0.000)

, ** and * means significance of coefficients at 10%, 5% and 1%.*

The results reported in Table 7 show the herding spillover from China to the US, Pakistan, and the UK. The herding spillover coefficient β_{41} is significant for Pakistan when LR and SLNN are used. β_{41} is -0.037 and -0.059 for linear regression and SLNN model. In addition, they are significant at 10% and 1%, respectively. None of the models provide evidence of herding spillover from China to the US and UK stock markets.

Herding Spillover from China to the US, Pakistan, and the UK with Event Sentiment of COVID-19

In Table 8, we present the results of herding spillover effect from China to the US, Pakistan, and the UK. We incorporated the twitter-based sentiment of COVID-19.

$$CSAD_{us,t} = \beta_{o1} + \beta_{11}\left|R_{m,us,t}\right| + \beta_{21}R^2_{m,us,t} + \beta_{31}CSAD_{c,t} + \beta_{41}R^2_{m,c,t} + \beta_{51}SCOVID + \varepsilon_t \quad (19)$$

$$CSAD_{pk,t} = \beta_{o1} + \beta_{11}\left|R_{m,pk,t}\right| + \beta_{21}R^2_{m,pk,t} + \beta_{31}CSAD_{c,t} + \beta_{41}R^2_{m,c,t} + \beta_{51}SCOVID + \varepsilon_t \quad (20)$$

$$CSAD_{UK,t} = \beta_{o1} + \beta_{11}\left|R_{m,UK,t}\right| + \beta_{21}R^2_{m,UK,t} + \beta_{31}CSAD_{c,t} + \beta_{41}R^2_{m,c,t} + \beta_{51}SCOVID + \varepsilon_t \quad (21)$$

Table 8

Herding Spillover from China to Pakistan, the US, and the UK with Sentiment of COVID-19

			With the Sentiment of COVID-19		
	$\boldsymbol{\beta_{11}}$	$\boldsymbol{\beta_{21}}$	$\boldsymbol{\beta_{31}}$	$\boldsymbol{\beta_{41}}$	$\boldsymbol{\beta_{51}}$
			LR		
Pakistan	0.449***	-0.072	0.036***	**-0.037****	-0.011
	(0.000)	(0.159)	(0.007)	**(0.050)**	(0.186)
US	-0.483***	**-0.396****	-0.792***	0.0399	0.237***
	(0.000)	**(0.012)**	(0.000)	(0.625)	(0.000)
UK	0.518***	0.059	0.053***	0.036**	0.088***
	(0.000)	(0.271)	(0.000)	(0.032)	(0.000)
`					
			SLNN		
Pakistan	0.398***	**-0.109****	0.033***	-0.021	-0.023
	(0.000)	**(0.042)**	(0.009)	(0.302)	(0.009)
US	-0.524***	**-0.432*****	-0.733***	-0.061	0.222***
	(0.000)	**(0.000)**	(0.000)	(0.415)	(0.000)
UK	0.529***	**0.024**	0.035***	**-0.072*****	0.092***
	(0.000)	**(0.667)**	(0.004)	**(0.000)**	(0.000)
			SVR		
Pakistan	0.674***	**-0.456*****	0.003	-0.015	0.005
	(0.000)	**(0.000)**	(0.808)	(0.441)	(0.545)
US	-0.649***	**-0.386*****	-1.285***	-0.062	0.181***
	(0.000)	**(0.017)**	(0.000)	(0.460)	(0.000)
UK	0.416***	0.080	0.013	0.013	0.063***
	(0.000)	(0.206)	(0.313)	(0.496)	(0.000)
			MLNN		
Pakistan	0.237***	**-0.103****	0.165***	**-0.005*****	-0.087***
	(0.000)	**(0.041)**	(0.000)	**(0.005)**	(0.000)
US	0.134*	**0.326****	-0.178***	0.475***	0.677***
	(0.080)	**(0.021)**	(0.000)	(0.000)	(0.000)
UK	0.0020	0.006	-0.425***	**-0.270*****	0.201***
	(0.918)	(0.902)	(0.000)	**(0.000)**	(0.000)

, ** and * means significance of coefficients at 10%, 5% and 1%.*

The estimated results show that there exists strong evidence of the herding spillover from China to Pakistan in the presence of the event sentiment of COVID-19. Herding spillover coefficient β_{41} is negative and significant for LR and MLNN models. In addition, there is a herding spillover from China to the UK stock market in the presence of COVID-19 sentiment, according to the estimated results of SLNN and MLNN. The findings suggest that machine learning models provide better results than LR. In this context, machine learning models provide a deep insight view of the stock trading in these sets of stock markets.

In the existing literature, number of studies have been done to predict the stock returns and herding. (Economou, Kostakis, & Philippas, 2011) used linear regression to investigate the herding behaviour in the stock markets. Public perception, mood of investors and stock returns are closely connected (Nisar & Yeung, 2018). Investors' sentiment leads the stock prices to deviate from the intrinsic values. When the sentiment of investors wanes, the mispricing of stock gets settled (Bae, Karolyi, & Stulz, 2003). After this price adjustment, investor sentiment and the future stock returns exhibit a negative relationship. Hence, investors' sentiment acts as a strong predictor of stock returns. Investors sentiments also change because of any news on social media platforms because they are either users or connected with the active users. In this context, our results of herding behaviour and its cross-country spillover effects are in line with the above argument. In the existing literature, most of the studies have focused on the predictability of events towards stock returns movements. Whereas this research study contributes to the existing literature by predicting the presence of herding behaviour and its spillover across the other markets.

Conclusion

This research study investigates the presence of herding behaviour in the stock markets of the US, UK, China, and Pakistan. For this purpose, we use multiple machine learning models and investigate the herding behaviour in these sets of markets along with the herding spillover effects among the markets. We incorporate the Twitter-based sentiment of two major events, i.e., Brexit 2016 and COVID-19, to predict the herding behaviour in these stock markets. For this purpose, we used daily stock price data from May 2006 to December 2022 for the UK, China, US, and Pakistan stock markets.

In the first step, we apply two machine learning models, i.e., SVR and SLNN, to forecast the CSAD of all four stock markets. The forecasting accuracy improved after incorporating event sentiments of Brexit and COVID-19. In the second step, three different machine learning models, i.e., support vector regression, single-layer neural network, and multilayer neural network models, have been used to predict the herding behaviour with and without event sentiments mentioned above. Our results provide evidence of herding behaviour in all the stock markets except Pakistan. In addition, when we incorporated the event sentiment of Brexit 2016 in our models, the herding coefficients became highly significant in all the markets, especially the UK. Similarly, herding significance increased in both the Chinese and UK stock markets when COVID-19 sentiment was incorporated. Using the same models, we computed the cross-country herding coefficients in the third step. Our machine learning models provide significant evidence of herding spillover from the UK to Pakistan and US stock markets in the presence of event sentiment of Brexit 2016. Similarly, a cross-country herding effect exists from the Chinese stock market to Pakistan and the UK when the estimation is done using machine learning models. In this context, machine learning models provide a deep insight view of the stock trading in these sets of stock markets, especially the trading patterns in the presence of mega events like Brexit 2016 and COVID-19. This research study contributes to the existing body of knowledge in multiple ways. The use of twitter-based sentiment to predict the herding anomaly in the stock market is the first main addition to the behavioural finance literature. The second contribution of this study is the use of machine learning model that provide more robust results and provide a better insight of the stock market data sets especially in context of herding anomaly. Thirdly, we also investigated the response of individual investors in one stock market towards the movement in another stock market in presence of twitter-based sentiment. This aspect has been unexplored in the existing literature of herding behaviour.

This research study provides an insight for individual and institutional investors to take care of unpredicted events while constructing their international portfolios in these sets of stock markets. Furthermore, policymakers should also take care of unnecessary panic caused by unpredicted events on social media platforms to control anomalies and crashes.

The study has some limitations in the form of data availability of individual stocks and market indices. Furthermore, a limited number of events have been examined to investigate their impact on herding behaviour. In future research, multiple events of different natures can be examined. In addition, macroeconomic variables can be added that may improve the predictability of machine learning models and provide a further insight of stock markets.

Acknowledgment: The authors appreciate the support of the Researchers Supporting Project number (RSPD2025R997), King Saud University, Riyadh, Saudi Arabia.

Appendix

Descriptive Statistics

	Market Return				CSAD			
	China	**US**	**UK**	**Pakistan**	**China**	**US**	**UK**	**Pakistan**
Mean	0.0002	0.00015	0.00015	0.00016	0.0179	0.4994	0.0131	0.01580
Std. Dev.	0.0007	0.0006	0.00053	0.0004	0.0089	0.4160	0.0067	0.0084
Jarque-Bera	464838.9	9391993	7390701	820198.1	4708.745	422.060	109533.3	56455.99
Probability	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Observations	4109	4109	4109	4109	4109	4109	4109	4109

****, * show significance at 1% and 10%*

References:

Adcock, R., & Gradojevic, N. (2019). Non-fundamental, non-parametric Bitcoin forecasting. *Physica A: Statistical Mechanics and its Applications, 531*, 121727. https://doi.org/10.1016/j.physa.2019.121727

Alomari, M., Al Rababa'a, A. R., El-Nader, G., Alkhataybeh, A., & Rehman, M. U. (2021). Examining the effects of news and media sentiments on volatility and correlation: Evidence from the UK. *The quarterly review of economics and finance, 82*, 280-297. https://doi.org/10.1016/j.qref.2021.09.013

Bae, K.-H., Karolyi, G. A., & Stulz, R. M. (2003). A new approach to measuring financial contagion. *The Review of Financial Studies, 16*(3), 717-763. https://doi.org/10.1093/rfs/hhg012

Balcilar, M., Demirer, R., & Hammoudeh, S. (2013). Investor herds and regime-switching: Evidence from Gulf Arab stock markets. *Journal of International Financial Markets, Institutions and Money, 23*, 295-321. https://doi.org/10.1016/j.intfin.2012.09.007

Barberis, N., & Thaler, R. (2003). A survey of behavioral finance. *Handbook of the Economics of Finance, 1*, 1053-1128. https://doi.org/10.1016/S1574-0102(03)01027-6

Benferhat, S., Tabia, K., & Ali, M. (2017). *Advances in Artificial Intelligence: From Theory to Practice: 30th International Conference on Industrial Engineering and Other Applications of Applied Intelligent Systems, IEA/AIE 2017, Arras, France, June 27-30, 2017, Proceedings, Part II* (Vol. 10351): Springer. https://doi.org/10.1007/978-3-319-60045-1

BenSaïda, A. (2017). Herding effect on idiosyncratic volatility in US industries. *Finance Research Letters, 23*, 121-132. https://doi.org/10.1016/j.frl.2017.03.001

Blasco, N., Corredor, P., & Ferreruela, S. (2012). Does herding affect volatility? Implications for the Spanish stock market. *Quantitative Finance, 12*(2), 311-327. https://doi.org/10.1080/14697688.2010.516766

Chiang, T. C., Li, J., & Tan, L. (2010). Empirical investigation of herding behavior in Chinese stock markets: Evidence from quantile regression analysis. *Global Finance Journal, 21*(1), 111-124. https://doi.org/10.1016/j.gfj.2010.03.005

Christie, W. G., & Huang, R. D. (1995). Following the pied piper: do individual returns herd around the market? *Financial Analysts Journal, 51*(4), 31-37. https://doi.org/10.2469/faj.v51.n4.1918

Devenow, A., & Welch, I. (1996). Rational herding in financial economics. *European economic review, 40*(3-5), 603-615. https://doi.org/10.1016/0014-2921(95)00073-9

Economou, F., Hassapis, C., & Philippas, N. (2018). Investors' fear and herding in the stock market. *Applied Economics, 50*(34-35), 3654-3663. https://doi.org/10.1080/00036846.2018.1436145

Economou, F., Kostakis, A., & Philippas, N. (2011). Cross-country effects in herding behaviour: Evidence from four south European markets. *Journal of International Financial Markets, Institutions and Money, 21*(3), 443-460. https://doi.org/10.1016/j.intfin.2011.01.005

Fama, E. F. (1960). Efficient market hypothesis. *Diss. PhD Thesis, Ph. D. dissertation*.

Fei, T., & Liu, X. (2021). Herding and market volatility. *International Review of Financial Analysis, 78*, 101880. https://doi.org/10.1016/j.irfa.2021.101880

Ferreruela, S., & Mallor, T. (2021). Herding in the bad times: The 2008 and COVID-19 crises. *The North American Journal of Economics and Finance, 58*, 101531. https://doi.org/10.1016/j.najef.2021.101531

Gaio, L. E., Stefanelli, N. O., Júnior, T. P., Bonacim, C. A. G., & Gatsios, R. C. (2022). The impact of the Russia-Ukraine conflict on market efficiency: Evidence for the developed stock market. *Finance Research Letters, 50*, 103302. https://doi.org/10.1016/j.frl.2022.103302

Gleason, K. C., Mathur, I., & Peterson, M. A. (2004). Analysis of intraday herding behavior among the sector ETFs. *Journal of empirical Finance, 11*(5), 681-694. https://doi.org/10.1016/j.jempfin.2003.06.003

Glorot, X., & Bengio, Y. (2010). *Understanding the difficulty of training deep feedforward neural networks.* Paper presented at the Proceedings of the thirteenth international conference on artificial intelligence and statistics.

Hwang, S., & Salmon, M. (2004). Market stress and herding. *Journal of empirical Finance, 11*(4), 585-616. https://doi.org/10.1016/j.jempfin.2004.04.003

Jabeen, A., Yasir, M., Ansari, Y., Yasmin, S., Moon, J., & Rho, S. (2022). An Empirical Study of Macroeconomic Factors and Stock Returns in the Context of Economic Uncertainty News Sentiment Using Machine Learning. *Complexity, 2022*. https://doi.org/10.1155/2022/4646733

Jiao, P., Veiga, A., & 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

Karlemstrand, R., & Leckström, E. (2021). Using twitter attribute information to predict stock prices. *arXiv preprint arXiv:2105.01402*. https://doi.org/10.1007/s12652-020-01839-w

Khan, W., Ghazanfar, M. A., Azam, M. A., Karami, A., Alyoubi, K. H., & Alfakeeh, A. S. (2022). Stock market prediction using machine learning classifiers and social media, news. *Journal of Ambient Intelligence and Humanized Computing*, 1-24.

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Klein, A. C. (2013). Time-variations in herding behavior: Evidence from a Markov switching SUR model. *Journal of International Financial Markets, Institutions and Money, 26*, 291-304. https://doi.org/10.1016/j.intfin.2013.06.006

Lehrer, S., Xie, T., & Zhang, X. (2021). Social media sentiment, model uncertainty, and volatility forecasting. *Economic Modelling, 102*, 105556. https://doi.org/10.1016/j.econmod.2021.105556

Mahmud, S. F., & Tiniç, M. (2018). Herding in Chinese stock markets: a nonparametric approach. *Empirical Economics, 55*, 679-711. https://doi.org/10.1007/s00181-017-1281-y

Maqsood, H., Mehmood, I., Maqsood, M., Yasir, M., Afzal, S., Aadil, F., . . . Muhammad, K. (2020). A local and global event sentiment based efficient stock exchange forecasting using deep learning. *International Journal of Information Management, 50*, 432-451. https://doi.org/10.1016/j.ijinfomgt.2019.07.011

Nair, V., & Hinton, G. E. (2010). *Rectified linear units improve restricted boltzmann machines.* Paper presented at the Icml.

Nisar, T. M., & Yeung, M. (2018). Twitter as a tool for forecasting stock market movements: A short-window event study. *The journal of finance and data science, 4*(2), 101-119. https://doi.org/10.1016/j.jfds.2017.11.002

Nofsinger, J. R., & Sias, R. W. (1999). Herding and feedback trading by institutional and individual investors. *The Journal of finance, 54*(6), 2263-2295. https://doi.org/10.1111/0022-1082.00188

Ozkan, O. (2021). Impact of COVID-19 on stock market efficiency: Evidence from developed countries. *Research in international business and finance, 58*, 101445. https://doi.org/10.1016/j.ribaf.2021.101445

Rossi, M., & Gunardi, A. (2018a). Efficient market hypothesis and stock market anomalies: Empirical evidence in four European countries. *Journal of Applied Business Research, 34*(1), 183. https://doi.org/10.19030/jabr.v34i1.10111

Rossi, M., & Gunardi, A. (2018b). Efficient market hypothesis and stock market anomalies: Empirical evidence in four European countries. *Journal of Applied Business Research (JABR), 34*(1), 183-192. https://doi.org/10.19030/jabr.v34i1.10111

Sibande, X., Gupta, R., Demirer, R., & Bouri, E. (2023). Investor sentiment and (anti) herding in the currency market: evidence from Twitter feed data. *Journal of Behavioral Finance, 24*(1), 56-72. https://doi.org/10.1080/15427560.2021.1917579

Van Dieijen, M., Borah, A., Tellis, G. J., & Franses, P. H. (2020). Big data analysis of volatility spillovers of brands across social media and stock markets. *Industrial Marketing Management, 88*, 465-484. ttps://doi.org/10.1016/j.indmarman.2018.12.006

Woo, K.-Y., Mai, C., McAleer, M., & Wong, W.-K. (2020). Review on efficiency and anomalies in stock markets. *Economies, 8*(1), 20. https://doi.org/10.3390/economies8010020

Yadav, N., Kudale, O., Rao, A., Gupta, S., & Shitole, A. (2021a). Twitter sentiment analysis using supervised machine learning *Intelligent Data Communication Technologies and Internet of Things* (pp. 631-642): Springer. https://doi.org/10.1007/978-981-15-9509-7_51

Yadav, N., Kudale, O., Rao, A., Gupta, S., & Shitole, A. (2021b). *Twitter sentiment analysis using supervised machine learning.* Paper presented at the Intelligent data communication technologies and internet of things: Proceedings of ICICI 2020. https://doi.org/10.1007/978-981-15-9509-7_51

Yasir, M., Attique, M., Latif, K., Chaudhary, G. M., Afzal, S., Ahmed, K., & Shahzad, F. (2023). Deep-learning-assisted business intelligence model for cryptocurrency forecasting using social media sentiment. *Journal of Enterprise Information Management, 36*(3), 718-733. https://doi.org/10.1108/JEIM-02-2020-0077

Yasir, M., & Önder, A. Ö. (2023). Time-varying herding spillover for emerging countries: evidence from BRIC countries and Turkey. *Review of Behavioral Finance, 15*(5), 709-728. https://doi.org/10.1108/RBF-10-2021-0218

Zhong, H., Wang, J., Jia, H., Mu, Y., & Lv, S. (2019). Vector field-based support vector regression for building energy consumption prediction. *Applied Energy, 242*, 403-414. https://doi.org/10.1016/j.apenergy.2019.03.078

Authors Biography

Wei Chen, Dr., is an academician currently working as a Lecturer at the School of International Education, Zhejiang Sci-Tech University, Hangzhou, China. Dr. Wei Chen has published various articles in the top-tier journals related to Economics, Finance, and Supply Chain Management.

Muhammad Asim is a Ph.D. Scholar at the Department of Management Sciences, COMSATS University Islamabad, Wah Campus, Pakistan. Mr. Asim holds a vast experience of 15 years working at the National Bank of Pakistan. Currently, he is working as Assistant Vice President at the National Bank of Pakistan. He has published his research in the Review of Behavioural Finance.

Muhammad Yasir (Corresponding Author) holds a Ph.D. in Economics from EGE University Izmir, Turkey. He is currently employed as an Assistant Professor (Economics) at FAST School of Management, National University of Computer and Emerging Sciences, Islamabad, Pakistan since Feb 2021. His key research interests are Behavioural Finance and time series prediction using advanced econometric techniques and machine learning models.

Mehboob ul Hassan, Dr., is a full Professor in the Department of Economics, College of Business Administration, King Saud University, Riyadh, Kingdom of Saudi Arabia. He has vast teaching and research experience of over 20 years in different International Universities. He served as Associate Dean in the School of Business Administration at Al-Dar University College, UAE, from May 2013 to August 2015. Before that, he also served as Dean of Business Administration at Sindh Madressatul Islam University, Karachi, Pakistan 2012-2013. He has published several articles in top-tier journals.

Adnan Shoaib, Dr., is currently serving as an Associate Professor at Greater Manchester Business School, University of Bolton Islamabad Campus. He completed his Doctorate Degree in Finance from Bahria University Islamabad in 2017. His area of special interest is “Financial Modelling Dynamics, Investment management, International Financial Management, Data mining and Financial Econometrics.

The article has been reviewed.

Received in February 2024; accepted in December 2024.