Annals of Operations Research
https://doi.org/10.1007/s10479-022-04744-x

ORIGINAL RESEARCH

COVID-19 news and the US equity market interactions: An inspection through econometric and machine learning lens

Rabin K. Jana[1] · Indranil Ghosh[2] · Fredj Jawadi[3] · Gazi Salah Uddin[4] · Ricardo M. Sousa[5,6]

Accepted: 15 April 2022

Abstract
This study investigates the impact of COVID-19 on the US equity market during the first wave of Coronavirus using a wide range of econometric and machine learning approaches. To this end, we use both daily data related to the US equity market sectors and data about the COVID-19 news over January 1, 2020-March 20, 2020. Accordingly, we show that at an early stage of the outbreak, global COVID-19s fears have impacted the US equity market even differently across sectors. Further, we also find that, as the pandemic gradually intensified its footprint in the US, local fears manifested by daily infections emerged more powerfully compared to its global counterpart in impairing the short-term dynamics of US equity markets.

Keywords COVID-19 · The US equity market · Co-integration · Detrended cross-correlation analysis · Machine learning

JEL classification C22 · G1 · I1

1 Introduction

In November 2019, a new pandemic, called hereafter Coronavirus or COVID-19, was announced in Wuhan in China, inducing several cases and deaths. While this virus retained, firstly, less attention by the international community and the authorities, the World Health Organization (WHO) qualified this Coronavirus as a pandemic hereafter in March 2020 when the statistics increased rapidly, showing 118 000 cases and 4291 deaths in 114 countries (Ali et al., 2020) on 11th March 2020. Later, this pandemic was transmitted to several countries, implying an increasing number of cases and deaths. The high dimension of COVID-19 cases and deaths and its rapid contagion effects worldwide have implied an unprecedented growing fear and an intense uncertainty across populations, policymakers, and nations. To

✉ Fredj Jawadi
fredj.jawadi@univ-lille.fr

Extended author information available on the last page of the article

Published online: 08 June 2022

limit the propagation of this virus and to fight against its contagion effects, several governments announced multiple lockdown measures and imposed different social distancing rules (school closing and online courses, teleworking, travel limitation, interdiction of social activities, etc.). Accordingly, the coronavirus pandemic and these lockdown measures have implied an important economic depression for several developed and emerging countries yielding a serious economic crisis in 2020.

The financial sector has been negatively impacted by the coronavirus pandemic (Wang et al., 2020), showing important losses. How did this COVID-19 outbreak impact financial markets? One can cite at least two channels that can explain the transmission of the Coronavirus into the financial sector. On the one hand, the slowdown of real activities (less consumption, less investment, fewer transactions, etc.) has negatively impacted the firm's production, the firm's creation of value, and therefore on the firm's fundamentals affecting thus their equities' prices. On the other hand, investors, being more anxious and uncertain about this pandemic and its duration, have been less active and have reduced their investment as a matter of precaution, which has had a negative effect on their trading.

Further, unlike other financial downturns (dot bubble in 2000, subprime crisis in 2007, the global financial crisis in 2008–2009, etc.), the 2020 crash for the equity market is different as it was induced by an external shock associated with the Coronavirus pandemic. This COVID-19 crisis has typically been different across time and space. Indeed, while it had a moderate effect on China, its negative impact was much pronounced in the US, the UK, Japan, Italy, Spain, and several member states of the European Union. Accordingly, the coronavirus pandemic has appeared among the most important events for stock markets in the twenty-first century. Interestingly, the COVID-19 effects have also been time-varying. For example, with reference to the level of the US stock market on 19th February 2020, the S&P500 index lost about 34% on 23rd March 2020; however, it reached 3213 points on 9th June 2020, erasing its losses induced to be the COVID-19 outbreak. Further, the crash associated with the pandemic is quite different from those of the 2008–2009 global financial crisis and the 2000 dot bubble. Indeed, the COVID-19 has overall implied an abrupt downturn for the US stock market while the market decrease in 2008–2009 lasts about eight months against 12 months for the market decrease in 2000 (see Fig. 1 for more details).

Several economists and financial analysts (Levasseur, 2021) justify this stock market dynamics during the COVID-19 episode by two important factors. First, the COVID-19 has had a negative impact on dividends as several firms and banks either canceled or reduced the payment of their dividends. Second, given the increase of uncertainty, the risk premium has increased as investors are seeking higher returns. Overall, there was a balance between dividend cuts and risk premium increases. Further, the stimulating actions of central banks (decrease of interest rate, massive programs of Quantitative Easing and purchasing program of financial assets, etc.) have provided abundant liquidity for governments to support their economies, which has stimulated investor's optimism and has had increased the attractiveness of stocks, yielding a rapid recovery for equity markets. This rapid recovery can also be seen as a consequence of learning from the management of previous financial crises.

In the recent related literature, several on-going studies have questioned the impact of COVID-19 on the stock markets. However, these studies are inconclusive and appear less unanimous. Indeed, on the one hand, some studies pointed to the impressive financial downturn and volatility excess induced by this pandemic. For example, Jana and Das (2020) show a significant effect of the pandemic in the Chinese market. Abuzayed et al. (2021) demonstrate that the gradual surge in COVID-19 cases intensified spillovers in developed European and North American markets. An empirical study by Bentes (2021) indicates that the persistence of stock market volatility in G7 countries has not been uniform. In particular, while it was

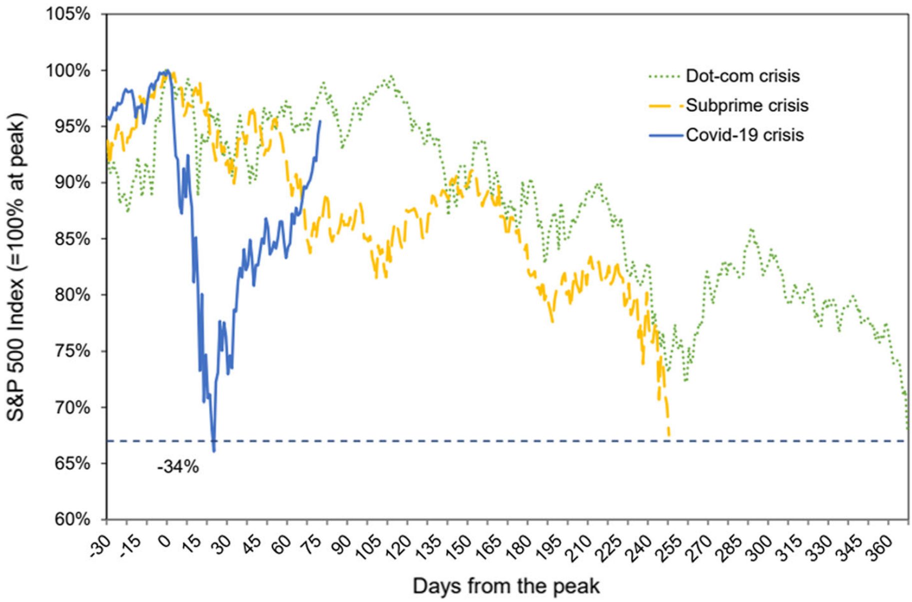

Fig. 1 The COVID-19 crash versus the Global Financial Crisis (GFC) crash. *Source* The Conversation. Available at: https://theconversation.com/decryptage-pourquoi-les-bourses-nont-presque-pas-connu-la-crise-de-la-covid-19-160936

imminently lower in the pre-pandemic regime, such persistence significantly increased during the pandemic. Duan et al. (2021) develop two COVID-19 sentiment indices for China based on textual data collated from official news media and the Sina Weibo blogsite to predict stock returns and turnover rates successfully. Liu et al. (2021) suggest that COVID-19 fear sentiment exacerbated stock market crash risk in China, while Mensi et al. (2021) argue that the COVID-19 pandemic induced strong spillover effects from US commodity markets to stock markets in China. O'Donnell et al. (2021) strongly emphasize that inherent growth in COVID-19 infections largely explained the changes in stock prices of six global stock indices. Takyi and Bentum-Ennin (20201) use Bayesian methods to quantify the short-term, medium to large impact of the COVID-19 pandemic on the stock market performance of 13 African countries. On the other hand, the empirical evidence shows that, unlike the global financial crisis, some stock markets did not suffer a significant impact from the COVID-19 outbreak. For example, the French stock index (CAC 40) lost about 30% in March–April 2020, but gained more than 65% thereafter.

As Fig. 1 shows, it is apparent that COVID-19 news undoubtedly had detrimental effects on financial markets across the world. Yet, despite a serious attempt to measure the influence of the pandemic on stock markets, most of the abovementioned studies focus on the impact of *country-level*, not *global*, market sentiment. Moreover, our study critically examines how fear sentiments penetrate through various *sectorial* equity indices. Thus, the present research aims to contribute to this literature gap by assessing the linkages between (US and global) COVID-19 news on different industrial sectors in the US during the first wave of the pandemic.

In order to investigate the influence of the COVID-19 outbreak, this paper aims at clarifying the impact of COVID-19 on the US equity market. Interestingly, given that the COVID-19 crisis has had an impact that varies across sectors, we propose to carry out an analysis for

20 sectorial equity indices, which might enable us to capture the effect of the COVID-19 outbreak on the wide spectrum of US business segments. These sectors have been chosen to systematically gauge the impact of the pandemic on essential consumable items, logistic, financial and real sectors, medical infrastructure, development, and internet services.[1] This is particularly relevant as the COVID-19 outbreak has had different effects that vary with the sector under consideration. For example, the information and telecommunication sector has shown a positive reaction to this outbreak in reasons of teleworking, online classes, and meetings.

Further, from a methodological point of view, we apply two different but complementary approaches to provide unbiased findings regarding the COVID-19 effects on the US equity sectors. On the one hand, a Johansen co-integration test and a detrended cross-correlation analysis (here forth, DCCA) are applied to investigate further co-movement between equity prices and COVID-19 news; Also, a nonlinear Granger causality test is carried out to check whether the COVID-19 fear-related components have had a lead-lag effect on the selected equity sectors. On other hand, we apply two ensemble Machine Learning algorithms (i.e., the gradient boosting and the random forest), which are especially designed to provide a concise characterization of causality between COVID-19 news and equity indices.

Using daily data over the period January 1, 2020–March 20, 2020, our findings reveal that, at an early stage of the outbreak, global fears had a significant effect on most sectorial equity indices and the Coronavirus stimulated the performance of a couple of sectors (telecommunication, teleworking, etc.). However, as the pandemic gradually intensified its footprint in the US, local fears became much more prominent and turned out to be the major equity drivers.

The remainder of the paper is as follows. Section 2 briefly presents the data and preliminary results. Section 3 describes the methodologies, while Sect. 4 analyses the main empirical results. Finally, Sect. 5 concludes.

2 Data and preliminary analysis

We use daily COVID-19 infections data across the world and the US from January 1, 2020, until March 20, 2020. These are denoted as Global Infections (GI) and Local Infections (LI), respectively, and they are considered as proxies for global and local COVID-19 fears and news. As for equity data, we rely on daily closing prices of 20 sectorial equity sectors of the Dow Jones Industrial Average (here forth, DJIA) and the S&P 500 indices. These sectors, which are reported in the first column of Table 1, encompass almost all major business segments in the US, ranging from daily needs to high-end innovations.

To further deepen our analysis and provide significant insights, observations are portioned into—(i) time horizon *I* (i.e., January 1, 2020 to January 31, 2020), during which local infections are substantially lower than global infections; and (ii) time horizon *II* (i.e., February 1, 2020 to March 20, 2020), during which the number of both global and local infections are significantly higher. These time horizons are denoted as TH-I and TH-II, respectively I, Table 1.

Tables 1 and 2 report the main descriptive statistics of global and local infections, and the considered equity sectors in TH-I and TH-II regimes, respectively. They reveal several interesting insights. Mean and median figures show that daily local infections were

[1] We restrict our study to the aforesaid sectors, as these are closely interlinked to livelihood and household finance. Moreover, our work predominantly focuses on the impact of the first wave of the pandemic.

Table 1 Descriptive statistics (TH-I period)

	Mean	Median	Range	Std. Dev	Skewness	Kurtosis
Global infections	463.05	18.50	2003	705.46	1.23	− 0.11376
Local infections	0.30	0.00	3	0.73	2.78	7.3795
Consumer goods	1290.80	1291.70	48.39	16.03	0.07	− 1.3699
Telecommunication	4275.40	4252.90	130.24	44.71	0.542	− 1.2751
Utility	3956	3922.80	301.33	116.18	0.28	− 1.5789
Transportation	65,351	65,319	1881.72	571.80	0.25	− 1.0283
Financial	1332.80	1330.70	38.17	11.20	0.29	− 0.93025
Customer service	1610.80	1612.20	45.89	12.72	− 0.44	− 0.52741
Bank	1092.60	1097.40	79.38	21.74	− 0.61	− 0.43451
Healthcare	1761.40	1754.50	59.07	19.46	0.47	− 1.2341
Internet commerce	1285.70	1288.90	83.87	24.41	− 0.45	− 0.81788
Internet service	310.31	312.50	22.86	6.63	− 0.72	− 0.65214
Energy	18.62	18.700	3.04	0.95	− 0.28	− 0.88391
Pharmaceutical	13,503	13,433	600.80	204.57	0.35	− 1.3666
Sustainability	1809.30	1807.30	48.59	14.80	0.18	− 1.1897
Space	362.05	363.41	24.79	6.44	− 1.23	1.0382
Auto	22,388	22,439	1308.24	353.57	− 0.67	0.0032091
IT	20,274	20,330	990.86	328.01	− 0.19	− 1.4542
Medical equipment	44,690	44,662	2222.79	642.39	0.00	− 1.0394
Agricultural	347.39	347.38	8.62	2.80	− 0.04	− 1.4752
Food and beverage	7584.20	7597.30	261.06	65.90	− 0.75	0.18175
Retail	5686.00	5689.70	198.33	53.86	− 0.45	− 0.35970

almost non-existent during the TH-I regime, despite reasonable global infections. However, daily local infections saw a major increase during the TH-II regime alongside daily global infections. By contrast, the means and medians of the 20 sectorial equity indices under consideration experienced a substantial dip in the TH-II period compared to the TH-I period.

Additionally, the magnitude of the dispersion, range, and standard deviation has steadily grown in the TH-II regime relative to the TH-I regime. Thus, it is apparent that owing to the surge in local infections in the TH-II period, US financial markets started anticipating fears reflected in the increase in risk and the lack of investors' confidence and manifested through dispersion statistics. Moreover, the Augmented Dickey-Fuller Test (ADF) test confirms that all variables are non-stationary. Accordingly, we transform these series using a difference filter to provide stationary series.[2] Skewness and kurtosis metrics also suggest that sectorial equity returns substantially deviate from the normal distribution in both periods.

[2] For brevity, we do not provide the results of the unit root tests. However, these are available from the authors upon request.

Table 2 Descriptive statistics (TH-II period)

	Mean	Median	Range	Std. Dev	Skewness	Kurtosis
Global infections	6005.40	2837.00	33,529	8260.30	2.37	4.55
Local infections	593.44	4.50	7123	1628.80	3.08	8.14
Consumer goods	1225.50	1271.00	415.51	124.42	− 1.04	− 0.20
Telecommunication	4055.70	4168.40	1068.80	319.06	− 0.96	− 0.34
Utility	3885.80	4081.00	1382.84	390.20	− 1.26	0.33
Transportation	59,734	61,477	23,318.3	7386.50	− 0.89	− 0.49
Financial	1213.10	1273.60	537.64	175.57	− 0.90	− 0.57
Customer service	1514.40	1562.60	514.85	168.29	− 0.83	− 0.61
Bank	928.05	987.46	476.56	171.07	− 0.67	− 1.10
Healthcare	1657.30	1687.50	440.67	131.70	− 0.91	− 0.30
Internet commerce	1231.00	1273.30	447.17	139.48	− 0.92	− 0.42
Internet service	297.03	301.74	93.11	27.612	− 0.73	− 0.72
Energy	14.33	15.750	9.33	2.9103	− 0.96	− 0.98
Pharmaceutical	12,589	12,845	3546.10	1105.00	− 0.81	− 0.60
Sustainability	1723.70	1782.80	530.18	151.77	− 1.22	0.49
Space	326.46	344.77	155.84	51.152	− 0.99	− 0.40
Auto	20,143	21,099	7159.48	2029.80	− 0.86	− 0.32
IT	19,674	20,471	6241.55	1702.50	− 1.74	1.86
Medical equipment	41,595	42,482	13,178.45	4059.00	− 0.83	− 0.57
Agricultural	330.77	334.76	45.05	13.08	− 0.97	− 0.23
Food and BEVERAGE	6981.90	7146.90	1758.75	557.25	− 0.82	− 0.62
Retail	5017.40	5289.50	2288.63	731.79	− 0.99	− 0.40

3 Econometric methodology

Our proposed research framework, depicted in Fig. 2, features two major objectives. The first one investigates the association between COVID-19 fears and the movement of sectorial equity indices during TH-I and TH-II. This objective is achieved by carrying out the Johansen co-integration test and estimating the DCCA (Detrended Cross-Correlation Analysis) coefficients. The second objective is to evaluate the ability of global and local fears in explaining (and forecasting) the future figures of chosen sectorial equities. To this end, we perform a nonlinear Granger causality test, and we rely on two ensemble machine learning algorithms. We briefly discuss hereafter these modeling steps.

3.1 The Johansen co-integration test

We start by conducting the Johansen co-integration test (Johansen, 1991) to ascertain the co-movement between sectorial equity indices and global and local COVID-19 infections. The test can be thought of as a generalization of the ADF test in a multivariate environment that allows one to investigate the presence of unit roots in a linear combination of variables. Thus, it is useful in comprehending the dynamics of the equilibrium relationships between

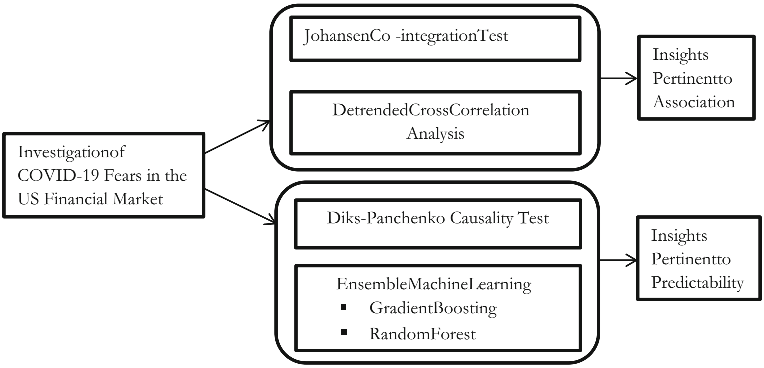

Fig. 2 The proposed research framework

equity indices and COVID-19 fears. The outcome of the co-integration test provides key insights on the co-movement pattern of underlying sectorial equity indices with global and local daily infections during TH-I and TH-II regimes, which portrays an estimate of future movements.

3.2 Detrended cross-correlation analysis (DCCA)

All time-series included in the analysis display a non-stationary dynamics. Thus, the deployment of traditional correlation measures may fail to reflect the true picture of the prevailing association between the variables of interest. Literature reports successful usage of the DCCA analysis proposed by Podobnik and Stanley (2008) in unearthing the interaction of non-stationary time series (Li et al., 2021). It can also uncover the nature of cross-correlation at different time lags.

The main rationale behind resorting to the DCCA framework in addition to the Johansen co-integration test is to find the direction of the association, which can be positive or negative at different time horizons. Thus, the DCCA methodology is appropriate for decoding the dynamic association structure of daily COVID-19 infections and sectorial equity indices across the TH-I and TH-II regimes. Being based on detrended covariances, it is a generalization of the detrended fluctuation analysis (DFA). Consequently, it can be used to investigate the power-law cross-correlation between pairs of non-stationary time series.

In this context, we adopt the DCCA cross-correlation coefficient approach by Zebende (2011), which extends the DCCA analysis. It is well-designed to measure the magnitude of the association between global and local fears with the selected financial sectors.[3]

The DCCA cross-correlation coefficient is estimated as follows:

$$\rho_{DCCA}(s) = \frac{F^2_{DCCA}(s)}{F_{DFA}\left(x_i^1\right)(s) * F_{DFA}\left(x_i^2\right)(s)} \tag{1}$$

where F_{DCCA} denotes the traditional fluctuation function derived from the DCCA analysis, while F_{DFA} represents the fluctuation function generated from DFA, and x_i^1 and x_i^2 denote the two-time series under consideration. The $\rho_{DCCA}(s)$ coefficient measures the amount

[3] Although the Johansen's co-integration test can also capture the nature of the (linear) association among the variables under consideration, it is restricted to strictly point-to-point association measurement. By contrast, the DCCA analysis can gauge the association structure at forward and backward lag directions.

of cross-correlation at the selected time-scale s, and it ranges between -1 (perfect negative association) to 1 (perfect positive association).

3.3 Nonlinear Granger causality test

The conventional Granger causality test can effectively capture linear causality and identify linear lead-lag effects. Nonetheless, the test is unable to track nonlinear forms of causality. As corroborated in Tables 1 and 2, we are dealing with a relatively small set of highly volatile data observations. The presence of a linear causal structure is highly unlikely. Therefore, we rely on the nonlinear Granger causality test that Diks and Panchenko (2006) propose, which is capable of nonparametric modeling. The test is also robust to the problem of the over-rejection of the null hypothesis (Hiemstra & Jones, 1994). However, despite being able to detect bidirectional causal influences, we have deployed the test in a strictly unidirectional framework to capture the causal impact of global and local COVID-19 daily infections on 20 US sectorial equity sectors. Thus, the inspection of causality allows us to infer the predictive ability of fears owing to the first wave of the COVID-19 pandemic at local and global levels for different US sectorial equity indices.

3.4 Machine learning approach

Two ensemble machine learning algorithms—gradient boosting (Friedman, 2001) and random forest (Breiman, 2001)—are used to assess the impact of global and local COVID-19 fears on the selected equity sectors.[4] Gradient boosting comprises a series of different learning algorithms applied in a forward-stage wise manner to generate the final predictions (Schapire & Singer, 1999). As for the random forest, it is another ensemble predictive modeling algorithm that employs a series of regression trees as base learners for drawing the final prediction (Breiman, 2001).

Although nonlinear Granger causality tests allow us to uncover the explanatory capability of local and global COVID-19 fears, the machine learning approach employed in this paper carries out a rigorous forecasting exercise to duly rationalize causality. Thus, we rely on gradient boosting and random forest algorithms to yield one-day ahead estimates of closing prices of 20 US sectorial equity indices using global and local daily COVID infections as explanatory variables. We consider these forecasting exercises for both TH-I and TH-II regimes to check whether the degree of predictability varies during the two sub-sample periods. Daily global and local infections are chosen as the explanatory variables to separately predict the underlying target variables. Additionally, one-day, two-day, three-day, four-day, five-day, six-day, and seven-day lagged global and local infections are used as independent variables to account for the effect of the last week of infections.[5] Our simulations are carried out using the *'sklearn'* library of Python programming language wherein several hyper-parameters like the number of base learners, learning rate, and the number of features for branching operations are automatically set by the *'GridSearchCV'* tool of the same library.

To elicit the predictive ability of both machine learning algorithms, we compute three performance indicators, namely, the R-squared coefficient $\left(R^2\right)$, the Nash–Sutcliffe Efficiency

[4] The fundamentals of both algorithms are briefly mentioned in Appendix A.

[5] Our simulations are carried out using the *'sklearn'* library of Python programming language wherein several hyper-parameters like the number of base learners, learning rate, and the number of features for branching operations are automatically set by the *'GridSearchCV'* tool of the same library.

(NSE), and the index of agreement (IA):

$$R^2 = \frac{\sum_{t=1}^{N}\left\{\widehat{Y}_t - \overline{Y_t}\right\}^2}{\sum_{i=1}^{N}\left\{Y_t - \overline{Y_t}\right\}^2} \tag{2}$$

$$NSE = 1 - \frac{\sum_{t=1}^{N}\left\{Y_t - \widehat{Y}_t\right\}^2}{\sum_{i=1}^{N}\left\{Y_t - \overline{Y_t}\right\}^2} \tag{3}$$

$$TI = \frac{\left[\frac{1}{N}\sum_{t=1}^{N}\left(Y_t - \widehat{Y}_t\right)^2\right]^{1/2}}{\left[\frac{1}{N}\sum_{i=1}^{N}\left(\widehat{Y}_t\right)^2\right]^{1/2} + \left[\frac{1}{N}\sum_{i=1}^{N}\left(Y_t\right)^2\right]^{1/2}} \tag{4}$$

where Y_t, $\overline{Y_t}$, and $\widehat{Y}_t$ denote actual observations, their average, and their predicted values, respectively. For an effective predictive model, the values of both R^2 and NSE should be close to 1, while the TI figures should be minimum.

4 Empirical analysis

4.1 Results of co-integration test

To evaluate the co-movement pattern between the US sectorial equity indices and global and local COVID-19 infections, we start by applying the Johansen Co-integration test to both regimes (see Table 3).[6]

During the TH-I period, the global fear is co-integrated with 11 equity sectors, namely, auto, bank, consumer service, food and beverages, healthcare, IT, medical equipment, pharmaceutical, retail, transportation, and utility. Barring the auto sector, local infections in the US are also co-integrated with those ten equity sectors. Yet, consumer goods, telecommunication, financial, internet commerce, internet service, energy, sustainability, space, and agricultural sector fail to display statistically significant co-movements with the inherent scare owing to the number of global and local infections.

As the co-integration test hints at a long-run relationship, sectorial indices identified as co-moving with both global and local infection rates should exhibit the same behavioral pattern in the TH-II regime. However, the test indicates that all equity sectors significantly co-move with both global and local COVID-19 fears. Indeed, the sectors that did not exhibit co-integrated patterns in the TH-I regime appear to be interconnected with both global and local spread of infections during the TH-II period. Thus, with the increase in new infections, the underlying sectors seem to be moving in accordance with the inherent outbreak fears.

At this stage, we note that the Johansen co-integration test cannot identify the dynamic direction of the co-movement. As a result, we test whether global and local COVID-19 fears have a positive or a negative impact on sectorial equity performance using the DCCA analysis.

4.2 Results of DCCA

The DCCA enables us to identify lead-lag relationships, that is, the. the nature of the association among the variables of interest and how it evolves forward and backward. Thus, the

[6] The optimal lag length of the Johansen co-integration test is set to 2, as based on the Akaike Information Criteria (AIC).

Table 3 Johansen's co-integration test

Sector	TH-I period		TH-II period	
	Global infections	Local infections	Global infections	Local infections
Consumer goods	✗	✗	✓	✓
Telecommunication	✗	✗	✓	✓
Utility	✓	✓	✓	✓
Transportation	✓	✓	✓	✓
Financial	✗	✗	✓	✓
Consumer service	✓	✓	✓	✓
Bank	✓	✓	✓	✓
Healthcare	✓	✓	✓	✓
Internet commerce	✗	✗	✓	✓
Internet service	✗	✗	✓	✓
Energy	✗	✗	✓	✓
Pharmaceutical	✓	✓	✓	✓
Sustainability	✗	✗	✓	✓
Space	✗	✗	✓	✓
Auto	✓	✗	✓	✓
IT	✓	✓	✓	✓
Medical equipment	✓	✓	✓	✓
Agricultural	✗	✗	✓	✓
Food and beverage	✓	✓	✓	✓
Retail	✓	✓	✓	✓

✓ indicates the presence of co-integration, while ✗ denotes absence of co-integration

deployment of DCCA provides additional insights on top of the pointwise co-movement pattern detected by the Johansen's co-integration test.

Table 4 reports the estimated DCCA coefficients computed for a time window of 3 days. They imply that global COVID-19 fears are weakly positively associated with Financial, Internet commerce, Internet service, IT, sustainability, transportation, and utility sectors. Thus, during the first pandemic wave, these sectors remained relatively immune to fears associated with new infections reported globally. By contrast, the Pharmaceutical sector displayed a relatively strong and negative association with global fears, followed by the food and beverage sector. Therefore, the penetration of COVID-19 fears on the selected business segments remained mostly mild during the month of January 2020 (i.e., the TH-I period). These results are relevant as they point to further heterogeneity in the transmission of COVID-19 into financial sectors.

A different characterization of the empirical evidence is observed between February 1 and March 20, 2020 (i.e., the TH-II period). Indeed, for all sectors, the impact of both global and local COVID-19 fears prevailed as negative, and with substantially higher magnitudes than in the TH-I period. In particular, auto, retail, and space sectors emerged as the sectors sharing the strongest negative relationship with global fears, while the auto, IT, and retail sectors stroke as those with the largest vulnerability to local fears. Hence, the auto sector was the

Table 4 Estimates of DCCA coefficients

Sector	TH-I period		TH-II period	
	Global infections	Local infections	Global infections	Local infections
Consumer goods	− 0.0858	0.0295	− 0.5673	− 0.2866
Telecommunication	− 0.0916	0.1220	− 0.5226	− 0.2402
Utility	0.2866	0.2579	− 0.5639	− 0.2408
Transportation	0.2701	0.3850	− 0.5437	− 0.2530
Financial	0.0556	0.1180	− 0.5920	− 0.2931
Consumer service	− 0.0385	0.2608	− 0.5550	− 0.2664
Bank	− 0.1783	− 0.1826	− 0.5626	− 0.2701
Healthcare	− 0.2563	0.2039	− 0.5047	− 0.1959
Internet commerce	0.0975	0.2604	− 0.5536	− 0.2506
Internet service	0.1744	0.0973	− 0.5049	− 0.1829
Energy	− 0.0028	0.1326	− 0.5024	− 0.2458
Pharmaceutical	− 0.4923	0.3867	− 0.5557	− 0.2220
Sustainability	0.0297	0.0090	− 0.2324	− 0.1751
Space	− 0.0354	0.0537	− 0.6434	− 0.3647
Auto	− 0.2019	− 0.0262	− 0.7146	− 0.4735
IT	0.1671	− 0.2043	− 0.6832	− 0.4371
Medical equipment	− 0.1884	− 0.2833	− 0.4834	− 0.1580
Agricultural	− 0.0354	0.3741	− 0.3271	− 0.1695
Food and beverage	− 0.3247	0.2813	− 0.6078	− 0.3238
Retail	− 0.2812	0.2626	− 0.6529	− 0.3765

Cross-correlations are estimated with a window of 3 days

most negatively hit by COVID-19 fears. This can also be explained by the restrictive measures of social distancing, teleworking, and lockdown programs imposed by several governments.

Overall, the empirical evidence corroborates the view that when both global and local fears of COVID-19 were relatively low during the first pandemic wave, their impact on equity markets was somewhat muted. However, as global fears mounted, they appear to have spilled to local fears, thus, dramatically shaping the dynamics of sectorial equities in the US.

4.3 Results of nonlinear Granger causality test

To infer the (potential) nonlinear links between global and local COVID-19 fears and sectorial equity indices, we also conduct a nonlinear Granger causality test. This test is particularly relevant to assessing asymmetry, nonlinearity, and complexity in these relationships. Table 5 summarizes the outcome of the causal inspection.

Our findings show that, during the TH-I period, the consumer goods, pharmaceuticals, telecommunication, and utility sectors were affected by both fear components. Five sectors namely, healthcare, energy, sustainability, IT, and medical equipment turned out to be sensitive to global fears. Thus, it can be inferred that shocks emanating from the menacing world outbreak would immediately affect these five sectors. Hence, the degree of efficiency of the

Table 5 Nonlinear Granger causality tests

Sector	TH-I period		TH-II period	
	Global infections	Local infections	Global infections	Local infections
Consumer goods	✓	✓	✓	✓
Telecommunication	✓	✓	✓	✓
Utility	✓	✓	✓	✓
Transportation	✗	✓	✓	✓
Financial	✗	✗	✓	✓
Consumer service	✗	✗	✓	✓
Bank	✗	✗	✓	✓
Healthcare	✓	✗	✗	✓
Internet commerce	✗	✗	✓	✓
Internet service	✗	✗	✗	✓
Energy	✓	✗	✓	✓
Pharmaceutical	✓	✓	✓	✓
Sustainability	✓	✗	✓	✓
Space	✗	✗	✓	✓
Auto	✗	✗	✓	✓
IT	✓	✗	✓	✓
Medical equipment	✓	✗	✓	✓
Agricultural	✗	✗	✓	✓
Food and beverage	✗	✗	✓	✓
Retail	✗	✗	✓	✓

✓ indicates the presence of nonlinear Granger causality, while ✗ denotes absence of nonlinear Granger causality

equity sectors under investigation is not uniform, as some are more susceptible to react to the arrival of new information than others.

Additionally, the causal structure has undergone a drastic change in the TH-II period. Indeed, all sectors turned sensitive to local fears and, with the exception of healthcare and internet service, they also became responsive to global fears. Thus, with a surge in the number of COVID-19 infections, the local fear component intensified, affecting almost all equity sectors.

In order to better assess for these interactions between global, local COVID-19 news and sectorial equity indices, we apply a Machine Learning approach. This approach is particularly helpful to measure the causal penetration of COVID-19 fears in the US financial markets.

4.4 Results of machine learning tests

Gradient boosting and random forest algorithms are applied for predictive modeling. We segregate the data for both TH-I and TH-II periods into training (80%) and test (20%) sets. The predictive performance is evaluated using the determination-squared $\left(R^2\right)$, the Nash–Sutcliffe Efficiency (NSE), and the index of agreement (IA) (Ghosh et al., 2019; Jana et al., 2020, 2021).

Tables 6 and 7 report the predictive performance of gradient boosting and random forest, respectively. The performance of gradient boosting and the random forest is similar. They show the explanatory capability of global and local fears of COVID-19 during the TH-I and TH-II periods. In particular, the results suggest that the overall predictive performance of global and local infections is not impressive during the early phase of the COVID-19 pandemic (i.e., the TH-I period). Equity movements of consumer goods, energy, pharmaceutical, and utility sectors are effectively explained by global infections. However, global fears perform poorly in explaining the dynamics of auto and space equities. These two sectors are also the least sensitive depending on the local fear as well. Moreover, local infections fail to significantly capture equity sector movements except, to some extent, those of the pharmaceutical sector. Thus, during the early stage of the outbreak, most of the US equity sectors remained largely unaffected by COVID-19 fears.

In the TH-II period, though, the predictive performance has considerably improved, as witnessed by the rise in the R^2 statistic and the fall in the IA index. Indeed, retail, consumer goods, consumer service, and space are the top four sectors that are effectively predicted by global infections, whereas auto, IT, retail, and space sectors are forecasted by local infections. Consequently, the performance indicators suggest that local fears became more prominent in driving the US equity markets from February 2020 onwards.

Additionally, it can be inferred that during the TH-I regime, the degree of efficiency of the underlying sectors played a significant role in the forecast accuracy. Local infection rates were too weak to govern the dynamics of those equity sectors. Pharmaceutical and consumer Goods have emerged to be relatively better explained by the local fear component compared to other sectors. Thus, these US sectors appear to be more dependent on local demand. Auto, space, food and beverage, retail, and agriculture remained immune to global and local fears throughout the TH-I phase. Nevertheless, the equity performance of these sectors was highly predictable in the TH-II phase. This is largely due to the fact that no curbs were initially imposed on manufacturing and retailing operations in the initial phase of the COVID-19 infection footprint. Subsequently, lockdowns, logistic deadlocks, and supply chain disruptions have severely influenced these activities. Such high prediction accuracy during the second sub-sample period reflects the gradual downfall pattern, which our framework captures particularly well.

Finally, this characterization of the empirical evidence is confirmed by the Diebold and Mariano (2002) paired comparison test (hereforth, DM) reported in Table 8. The DM test performs a (paired) comparative predictive analysis for which it is essential to set the order of the pair constituents. Thus, the order of the variables used in the comparison of the predictive ability of global and local infections during the TH-I and TH-II regimes has been marked with an index number in parenthesis. To the extent that the test statistic is significantly positive, then, the variable indicated by the number 2 in parenthesis displays stronger predictive ability than the variable marked by the number 1. By contrast, when the test statistic is significantly negative, the variable marked with the number 1 in parenthesis exhibits larger forecasting power than the variable market with the number 2. Our results clearly show that the significance level and the sign of the test statistic indicate that the explanatory ability of global infections is superior to that of local infections during the TH-I period. However, local infections do not outperform global infections during the TH-II period. Additionally, the predictive ability of global and local infections is significantly more powerful during the TH-II period than the TH-I period.

The economic and practical implications of these findings are meaningful. As predictability statistically improved with the intensification of the pandemic, there is a marginal scope for hedging. On the one hand, business verticals that are highly interlinked with local demand

Table 6 Predictive performance of gradient boosting

Sectors	TH-I period						TH-II period					
	Global infections			Local infections			Global infections			Local infections		
	R^2	NSE	IA	R^2	NSE	IA	R^2	NSE	IA	R^2	NSE	IA
Consumer goods	0.84	0.78	0.06	0.41	0.38	0.17	0.82	0.80	0.05	0.84	0.83	0.05
Telecommunication	0.67	0.64	0.08	0.07	0.05	0.82	0.78	0.77	0.06	0.80	0.78	0.06
Utility	0.81	0.78	0.06	0.22	0.19	0.34	0.74	0.72	0.06	0.82	0.81	0.05
Transportation	0.43	0.39	0.12	0.03	0.03	0.96	0.85	0.83	0.05	0.86	0.84	0.04
Financial	0.50	0.49	0.12	0.07	0.06	0.91	0.86	0.85	0.04	0.86	0.85	0.04
Consumer service	0.43	0.39	0.13	0.20	0.19	0.33	0.84	0.83	0.04	0.86	0.84	0.04
Bank	0.56	0.54	0.09	0.14	0.13	0.39	0.84	0.83	0.04	0.82	0.82	0.05
Healthcare	0.71	0.68	0.07	0.22	0.20	0.28	0.61	0.61	0.08	0.80	0.79	0.05
Internet commerce	0.48	0.46	0.11	0.04	0.03	0.90	0.67	0.65	0.07	0.85	0.84	0.04
Internet service	0.45	0.43	0.13	0.01	0.00	0.98	0.61	0.60	0.08	0.79	0.78	0.06
Energy	0.80	0.78	0.06	0.05	0.04	0.85	0.78	0.77	0.06	0.78	0.77	0.06
Pharmaceutical	0.85	0.84	0.04	0.67	0.63	0.09	0.78	0.77	0.06	0.81	0.80	0.05
Sustainability	0.75	0.72	0.07	0.14	0.13	0.41	0.78	0.77	0.06	0.62	0.61	0.07
Space	0.04	0.02	0.88	0.00	0.00	1.00	0.89	0.86	0.03	0.90	0.89	0.03
Auto	0.02	0.01	0.94	0.00	0.00	1.00	0.66	0.65	0.07	0.88	0.86	0.04
IT	0.61	0.57	0.08	0.20	0.19	0.32	0.71	0.70	0.06	0.92	0.92	0.02
Medical equipment	0.76	0.73	0.07	0.25	0.24	0.37	0.81	0.81	0.06	0.83	0.81	0.05
Agricultural	0.33	0.30	0.15	0.07	0.07	0.89	0.79	0.78	0.06	0.82	0.81	0.06

Table 6 (continued)

Sectors	TH-I period						TH-II period					
	Global infections			Local infections			Global infections			Local infections		
	R^2	NSE	IA	R^2	NSE	IA	R^2	NSE	IA	R^2	NSE	IA
Food and beverage	0.27	0.26	0.24	0.11	0.10	0.44	0.73	0.72	0.07	0.78	0.77	0.06
Retail	0.37	0.34	0.28	0.04	0.04	0.87	0.90	0.88	0.03	0.90	0.89	0.03

Gradient boosting and random forest algorithms are applied for predictive modeling. Data for both TH-I and TH-II periods into training (80%) and test (20%) sets. The predictive performance is evaluated using the determination-squared (R^2), the Nash–Sutcliffe Efficiency (NSE), and the index of agreement (IA)

Table 7 Predictive performance of the random forest algorithm

Sectors	TH-I period						TH-II period					
	Global infections			Local infections			Global infections			Local infections		
	R^2	NSE	TI	R^2	NSE	TI	R^2	NSE	TI	R^2	NSE	TI
Consumer goods	0.83	0.78	0.06	0.40	0.38	0.17	0.81	0.80	0.06	0.84	0.83	0.05
Telecommunication	0.68	0.63	0.08	0.06	0.06	0.83	0.78	0.77	0.06	0.80	0.78	0.06
Utility	0.80	0.78	0.06	0.21	0.19	0.35	0.74	0.72	0.06	0.82	0.81	0.05
Transportation	0.42	0.40	0.12	0.03	0.03	0.97	0.84	0.83	0.05	0.85	0.84	0.04
Financial	0.49	0.48	0.12	0.07	0.06	0.92	0.86	0.84	0.04	0.85	0.85	0.04
Consumer Service	0.43	0.39	0.14	0.21	0.19	0.32	0.85	0.84	0.05	0.85	0.84	0.04
Bank	0.56	0.54	0.09	0.14	0.13	0.39	0.83	0.82	0.05	0.82	0.81	0.05
Healthcare	0.71	0.67	0.08	0.21	0.20	0.28	0.61	0.60	0.08	0.80	0.79	0.06
Internet Commerce	0.48	0.45	0.11	0.03	0.04	0.89	0.66	0.65	0.07	0.86	0.85	0.04
Internet Service	0.43	0.43	0.13	0.01	0.00	0.98	0.60	0.60	0.08	0.79	0.78	0.06
Energy	0.81	0.78	0.06	0.05	0.04	0.85	0.77	0.77	0.07	0.77	0.76	0.06
Pharmaceutical	0.84	0.83	0.04	0.66	0.62	0.09	0.77	0.76	0.07	0.81	0.80	0.05
Sustainability	0.76	0.71	0.07	0.14	0.12	0.42	0.78	0.77	0.06	0.62	0.61	0.07
Space	0.03	0.01	0.88	0.00	0.00	1.00	0.87	0.85	0.04	0.89	0.88	0.04
Auto	0.02	0.01	0.94	0.00	0.00	1.00	0.65	0.64	0.07	0.87	0.86	0.04
IT	0.62	0.56	0.08	0.20	0.19	0.32	0.70	0.69	0.07	0.92	0.91	0.03
Medical Equipment	0.76	0.73	0.07	0.25	0.24	0.38	0.80	0.79	0.06	0.82	0.81	0.05
Agricultural	0.32	0.30	0.15	0.07	0.07	0.90	0.79	0.77	0.06	0.82	0.80	0.06

Table 7 (continued)

Sectors	TH-I period						TH-II period					
	Global infections			Local infections			Global infections			Local infections		
	R^2	NSE	TI	R^2	NSE	TI	R^2	NSE	TI	R^2	NSE	TI
Food and Beverage	0.26	0.25	0.24	0.11	0.10	0.45	0.73	0.72	0.07	0.77	0.76	0.06
Retail	0.37	0.33	0.28	0.04	0.03	0.87	0.81	0.80	0.06	0.84	0.83	0.05

Gradient boosting and random forest algorithms are applied for predictive modeling. Data for both TH-I and TH-II periods into training (80%) and test (20%) sets. The predictive performance is evaluated using the determination-squared $\left(R^2\right)$, the Nash–Sutcliffe Efficiency (NSE), and the index of agreement (IA)

Table 8 DM Test on gradient boosting and random forest

Model	Gradient boosting				Random forest			
	GI (TH-I) (1)	LI (TH-I) (1)	GI (TH-II) (1)	LI (TH-II) (1)	GI (TH-I) (1)	LI (TH-I) (1)	GI (TH-II) (1)	LI (TH-II) (1)
GI (TH-I) (2)	–				–			
LI (TH-I) (2)	– 4.5632***	–			– 4.5627***	–		
GI (TH-II) (2)	5.8413***	6.0893***	–		5.8443***	6.0884***	–	
LI (TH-II) (2)	6.2168***/	6.4655***	0.259#	–	6.2177***	6.4651***	0.257#	–

GI stands for Global Infections; LI denotes Local Infections

***Significant at the 1% level

#Not significant at conventional levels. The order of the variables has been marked with an index number in parenthesis. To the extent that the test statistic is significantly positive (negative), then, the variable indicated by the number 2 (1) in parenthesis displays stronger predictive ability than the variable marked by the number 1(2) can benefit from effective strategic management. On the other hand, sectors that rely on external demand (thus, exports) for a major crunch of revenue may continue to suffer from the ongoing pandemic. The improvement in overall predictability in the TH-II regime indicates a deeper penetration of COVID-19 fears in the US market and a higher dependence on global and local infections. This financial markets' reaction was expected considering the timeline of the first wave unfolding and the gradual fall in the influence of other macroeconomic variables. As for the TH-I phase, it showed a marginally stronger resilience vis-à-vis daily infections, suggesting that other factors might be at play. In this context and given the novelty of the COVID-19 news' effects, the forecasting outperformance of daily and local infections in the TH-II period is unlikely to reflect a somewhat less efficient asset price discovery process compared to the TH-I period. Indeed, if anything, investors should have been able to exploit larger abnormal returns in the TH-I period than in the TH-II period in light of the unprecedented circumstances. Thus, such outperformance should capture the pure effect of daily and local infections on the dynamics of financial markets, which investors started to factor in more heavily in the TH-II period.

5 Conclusion

We assess the penetration of the first COVID-19 wave in the US equity market via econometric and machine learning approaches and using world and US daily infection counts as proxies for global and local COVID-19 fears over January 1, 2020–March 20, 2020. We show that, at an early stage of the COVID-19 outbreak in January 2020, local fears had little penetration across different sectors of US equity markets, while with the steep increase in daily US infections since February 2020, local fears became much more prominent. In fact, the application of the Johansen co-integration test and the DCCA analysis has successfully decoded the nature of the dynamic interplay. In contrast, nonlinear Granger causality tests and a machine learning modeling successfully expounded the predictive ability of global and local COVID-19 fears.

Some sectors, which were comparatively more exposed and sensitive to cross-country trade (e.g., auto, consumer goods, pharmaceuticals, and space), appeared to be the main recipients of spillovers from the global pandemic outbreak in January 2020. Manufacturing sectors like auto, space, and sectors serving daily needs (e.g., food and agriculture, beverage, and retail) were less sensitive to uncertainty owing to the initial outbreak. As new infections in the US eventually increased from February onwards, clear evidence of fear havocking the market becomes apparent. During this period, the surge in infections affected all sectors to a large extent. The rise in local infections exhibited significantly more penetration capability than global infections. Not surprisingly, we uncover a significant long-run co-movement, which is corroborated by the persistence of the prevailing situation over time. That is, while growing fears among investors were expected, our study clearly reveals that the structure of association is mostly negative. In particular, we show that equity sectors that are more reliant on domestic demand (e.g. consumer goods or pharmaceuticals) survived the initial onslaught of the COVID-19 pandemic (action plans and strategies to revamp sectors accordingly in new normal).

In this context, regulations may be framed to minimize the effects on other sectors before the local outbreak happens. Thus, it is extremely important to observe the time gap between the surge in global and local fears. It critically provides a window of opportunity for policymakers to take necessary measures to avoid financial market crashes and economic slumps. In this context, our research can be leveraged for practical policy formulation to combat such global catastrophic events. Finally, we stress that the scope of our study is restricted to the US equity market and confined to the first wave of the COVID-19 pandemic. Thus, it would be interesting to check the differences in behavioral patterns of US sectorial equity indices associated with COVID-19 news in the first and second waves. It is also possible to expand the geographical coverage to ascertain the influence of the pandemic in other developed and emerging economies. Investigating the penetration of COVID-19 fears over a longer period could also help discover the time frame and conditions associated with a significant reduction of the adverse pandemic effects.

Funding NIPE's work is financed by the National Funds of the FCT—Portuguese Foundation for Science and Technology within the project "UIDB/ECO/03182/2020".

Appendix A

Ensemble machine learning

The algorithms associated with the two ensemble machine learning algorithms, i.e. the gradient boosting (Friedman, 2001) and the random forest (Breiman, 2001) are described below.

Gradient boosting

The gradient boosting is a variant of the classical boosting algorithm that mimics the same principle of identifying training samples via the determination of gradient-driven error rates. Classical regression trees are used as base learners for carrying out the learning operation in each stage sequentially in a forward direction.

The steps of the algorithm are as follows:

Step 1. For T training samples $\{(x_1, y_1), (x_2, y_2), \ldots, (x_T, y_T)\}$ in set S, initialize *1/d* as the weight of each sample, where d is the cardinality of S.

Step 2. For the base learner i, execute the following:

Step 2.1. Bootstrapped samples are picked to form the training set, S_i.

Step 2.2. Training the model M_t using S_i.

Step 2.3. Compute the error of M_t, $\mathrm{Err(M_t)} = \sum_{\mathrm{i}=1}^{\mathrm{T}} \mathrm{p_i^t sign(||h_t(x_i) - y_i||-\Delta)}$, where $\mathrm{h_t(x_i)}$ is the predicted value of i-th sample by t-th modeler, $\mathrm{y_i}$ is the actual response of the i-th sample, Δ is a real number, and $\mathrm{p_i^t} = \frac{\mathrm{w_i^t}}{\sum_{\mathrm{i}=1}^{\mathrm{T}} \mathrm{w_i^t}}$, $\mathrm{w_i^t}$ is the weight of i-th sample for t-th modeler.

Step 2.4. Repeat steps 2.1-2.3, if the error is greater than 0.5; else, perform step 2.5.

Step 2.5. Update weights of each sample in S_i.

Step 2.6. Normalize the weights.

Step 3. Compute each base learner's accuracy.

Step 4. Obtain the final prediction from the weighted outcomes of the base learners.*

Random forest

In the random forest, decision trees are constructed based on a randomly selected subset of the training dataset. At each node of the chosen tree, the best feature for splitting the operation is identified based on a randomly selected subset of features. The steps of the algorithm are mentioned below:

Step 1. Draw a bootstrapped sample from training data.
Step 2. For each sample, build a regression tree. At each node of the tree:
Step 2.1. Randomly select a subset of features from all available features.
Step 2.2. Identify the best feature.
Step 2.3. Continue until the tree is fully grown.
Step 3. Generate the final prediction.

References

Abuzayed, B., Bouri, E., Al-Fayoumi, N., & Jalkh, N. (2021). Systemic risk spillover across global and country stock markets during the COVID-19 pandemic. *Economic Analysis and Policy, 71*(1), 180–197.

Ali, S. A., Baloch, M., Ahmed, N., Ali, A. A., & Iqbal, A. (2020). The outbreak of Coronavirus Disease 2019 (COVID-19): An emerging global health threat. *Journal of Infection and Public Health, 13*(4), 644–646.

Bentes, S. R. (2021). How COVID-19 has affected stock market persistence? Evidence from the G7's. *Physica a: Statistical Mechanics and Its Applications, 581*(1), 126210.

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32.

Diebold, F. X., & Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business & Economic Statistics, 20*(1), 134–144.

Diks, C., & Panchenko, V. (2006). A new statistic and practical guidelines for nonparametric Granger causality testing. *Journal of Economic Dynamics and Control, 30*(9–10), 1647–1669.

Duan, Y., Liu, L., & Wang, Z. (2021). COVID-19 sentiment and the Chinese stock market: Evidence from the Official News Media and Sina Weibo. *Research in International Business and Finance, 58*(1), 101432.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics, 29*(5), 1189–1232.

Ghosh, I., Jana, R. K., & Sanyal, R. K. (2019). Analysis of temporal pattern, causal interaction and predictive modeling of financial markets using nonlinear dynamics, econometric models and machine learning algorithms. *Applied Soft Computing, 82*, 105553.

Hiemstra, C., & Jones, J. D. (1994). Testing for linear and nonlinear granger causality in the stock price-volume relation. *The Journal of Finance*, *49*(5), 1639–1664.

Jana, R. K., & Das, D. (2020). Did Bitcoin act as an antidote to the Chinese equity market and booster to altcoins during the Novel Coronavirus outbreak? Retrieved from https://ssrn.com/abstract=3544794

Jana, R. K., Ghosh, I., Das, D., & Dutta, A. (2021). Determinants of electronic waste generation in Bitcoin network: Evidence from the machine learning approach. *Technological Forecasting and Social Change, 173*, 121101.

Jana, R. K., Ghosh, I., & Sanyal, M. K. (2020). A granular deep learning approach for predicting energy consumption. *Applied Soft Computing, 89*, 106091.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica, 59*(6), 1551–1580.

Levasseur, M. (2021). Décryptage: pourquoi les bourses n'ont (presque) pas connu la crise de la Covid-19. *The Conversation*, 19 May.

Li, j., Wu, X., Zhang, L., & Feng, Q. (2021). Research on the portfolio model based on Mean-MF-DCCA under multifractal feature constraint. *Journal of Computational and Applied Mathematics*, *386*, 113264.

Liu, Z., Huynh, T. L. D., & Dai, P. F. (2021). The impact of COVID-19 on the stock market crash risk in China. *Research in International Business and Finance, 57*(1), 101419.

Mensi, W., Reboredo, J. C., & Ugolini, A. (2021). Price-switching spillovers between gold, oil, and stock markets: Evidence from the USA and China during the COVID-19 pandemic. *Resources Policy, 73*(1), 102217.

O'Donnell, N., Shannon, D., & Sheehan, B. (2021). Immune or at-risk? Stock markets and the significance of the COVID-19 pandemic. *Journal of Behavioral and Experimental Finance, 30*(1), 100477.

Podobnik, B., & Stanley, H. E. (2008). Detrended cross-correlation analysis: A new method for analyzing two non-stationary time series. *Physical Review Letters, 100*, 084102.

Schapire, R. E., & Singer, Y. (1999). Improved boosting algorithms using confidence-rated predictions. *Machine Learning, 37*(3), 297–336.

Wang, L., Wang, Y., Ye, D., & Liu, Q. (2020). Review of the 2019 Novel Coronavirus (COVID-19) based on current evidence. *International Journal of Antimicrobial Agents, 55*(6), 105948.

Zebende, G. F. (2011). DCCA cross-correlation coefficient: Quantifying level of cross-correlation. *Physica a: Statistical Mechanics and Its Applications, 390*(4), 614–618.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Rabin K. Jana[1] · Indranil Ghosh[2] · Fredj Jawadi[3] · Gazi Salah Uddin[4] · Ricardo M. Sousa[5,6]

Rabin K. Jana
rkjana@iimraipur.ac.in

Indranil Ghosh
indranil@imthyderabad.edu.in

Gazi Salah Uddin
gazi.salah.uddin@liu.se

Ricardo M. Sousa
rjsousa@eeg.uminho.pt; rjsousa@alumni.lse.ac.uk

1 Operations & Quantitative Methods Area, Indian Institute of Management Raipur, Atal Nagar, CG 493661, India

2 IT & Analytics Area, Institute of Management Technology Hyderabad, Shamshabad, Hyderabad, Telangana 501218, India

3 Univ. Lille, ULR 4999 - LUMEN, 59000 Lille, France

4 Department of Management and Engineering, Linköping University, 581 83 Linköping, Sweden

5 Department of Economics and Centre for Research in Economics and Management (NIPE), University of Minho, Campus of Gualtar, 4710-057 Braga, Portugal

6 London School of Economics and Political Science, LSE Alumni Association, Houghton Street, London WC2A 2AE, UK