Measuring volatility spillovers between economic sectors for predicting financial crises

Ricardo Laborda[1], Centro Universitario de la Defensa de Zaragoza

Jose Olmo[2], Universidad de Zaragoza and University of Southampton

Abstract

This paper measures volatility spillovers between sectors of economic activity using network connectivity measures. These spillovers are an accurate proxy for the transmission of risk across sectors and are particularly informative during crisis periods. To do this, we apply the novel methodology proposed in Diebold and Yilmaz (2012) to seven economic sectors of U.S. economic activity and find that Banking&Insurance, Energy, Technology and Biotechnology are the main channels through which shocks propagate to the rest of the economy. Banking&Insurance is especially relevant during the 2007-2009 global financial crisis, while the Energy sector and Technology are especially relevant during the COVID-19 crisis. We also show that volatility spillover effects exhibit ability to predict high episodes of volatility for the S&P 500 index being useful as advance indicators of financial crises.

Keywords: Financial crises, Sectoral connectedness, Volatility spillovers, S&P 500 volatility, Random Forest.

JEL Classification: G01, G14, G17.

Declarations of interest : none

---

[1]Centro Universitario de la Defensa. Academia General Militar. Ctra. Huesca s/n. 50090. Zaragoza. Phone: +34976739844. Ricardo Laborda acknowledges financial support from CREVALOR.

[2]Corresponding author: Departamento de Análisis Económico, Universidad de Zaragoza. Gran Vía 2, 50005, Zaragoza, Spain. E-mail: joseolmo@unizar.es and Department of Economics, University of Southampton. Highfield Campus, SO17 1BJ, Southampton, UK. E-mail: J.B.Olmo@soton.ac.uk. Jose Olmo acknowledges financial support from Project PID2019-104326GB-I00 from Ministerio de Ciencia e Innovación and from Fundación Agencia Aragonesa para la Investigación y el Desarrollo (ARAID).

1. Introduction

Financial markets have recently witnessed periods of financial instability and crisis. These periods are characterised by extreme uncertainty that is reflected in significant increases in market volatility. Diebold and Yilmaz (2015) show that an increase in market volatility is a reflection of an increase in connectedness across different assets. The recent 2008 global financial crisis, the Eurozone debt crisis in 2010–2012 and the COVID-19 pandemic crisis have shown that interdependence between assets and financial markets is especially relevant during crises having the potential to trigger systemic risk episodes. The 2007-2009 global financial crisis is a clear example of such systemic events (Giglio et al, 2016). Another recent example is the reaction of global financial markets to the COVID-19 pandemic. The effects of this crisis are not fully revealed yet but have the potential to have serious negative consequences across industries and assets, see Goodell, (2020), Goodell and Huynh (2020), Conlon and McGee (2020), Conlon, Corbet, and McGee (2020) and Goodell and Goutte (2020), among many other recent contributions analysing the impact of COVID-19.

The aim of this paper is to uncover the interdependencies between sectors of economic activity in periods of heightened uncertainty. According to past empirical evidence, financial crises are preceded by an increase in market volatility. Hence, it is important to understand the mechanisms through which volatility fuels across sectors of economic activity and is reflected in asset prices. Our working hypothesis is that volatility spillovers between economic sectors have ability to predict financial crises characterised by systemic risk events. We focus on recent crises such as the 2007-2009 global financial crisis, the 2011-2012 sovereign bond crisis and the ongoing crisis originated by the COVID-19 pandemic.

To do this, our first contribution is to evaluate seven sectors of economic activity (Health Care, Pharmaceuticals, Biotechnology, Banking&Insurance, Cyclical Sector, Technology and Energy) comprising the S&P 500 index from 2003 to the present. We measure volatility spillovers between shocks to these sectors paying special attention to their dynamics and relevance during financial crisis episodes. We apply the novel methods on network connectivity and volatility spillovers in vector autoregressive settings (VAR) introduced by Diebold and Yilmaz (2009, 2012, 2014).

It is well documented that periods of increasing volatility are characterised by increasing connectedness (Diebold and Yilmaz 2009, 2012) across markets, but it is less understood whether there is information embedded in the connectedness measures that helps to predict future market volatility and, hence, financial crises. Thus, our second contribution is to assess whether increases in linkages across sectors, reflected in volatility spillovers, help to predict the likelihood of crises and systemic risk events. To do this, we apply novel techniques on machine learning, in particular, we fit random forests models for classification and prediction, see Breiman (2001) and Cutler et al (2011). This is a state-of-the-art technique in machine learning that allows us to predict extreme events without the need of imposing a parametric model relating volatility spillovers and the occurrence of crises in the S&P 500 index. The random forest methodology uses a combination of nonparametric classification and regression trees obtained from random permutations of the regressors and number of lags to predict nonparametrically the occurrence of high volatility episodes in the S&P 500 index.

In our setting, we interpret a financial crisis with a scenario in which the annualised S&P 500 daily volatility is higher than a given threshold. In the empirical application, we consider 20% but the accuracy of the predictions holds for a range of threshold values between 15% and 20%.

Our results show that volatility connectedness between the different economic sectors is time varying and ranges between 40% and 85%. For example, during the COVID-19 pandemic, we are witnessing increases in volatility spillovers about 40% compared to the levels reached at the beginning of 2020 that were around 60%, reaching a maximum of 85.30% on March 3, 2020. During this period, Energy is the main sector through which volatility spills over to other sectors of economic activity. The main reason for this is the particular characteristics of the global lockdown that drained other types of economic activity. Energy prices suffered particularly during this period and these negative shocks propagated further to the rest of sectors in the economy. Interestingly, we find that this sector is not a main driver of systemic risk events in other crises. It is well known, for example, the role of the Banking&Insurance sector during the 2007–2009 global financial crisis as the main catalyser of systemic risk events. Nevertheless, we find that the net contribution of Banking&Insurance, Energy, Technology and Biotechnology to market volatility is usually larger than for the other economic sectors particularly so during crisis events.

We use information on the magnitude and direction of volatility spillovers across sectors to predict the occurrence of extreme volatility regimes in the S&P 500 index. Despite the importance of assessing sector-specific volatilities for measuring market risk we find that it is the increase in interconnectivity between economic sectors what really drives the abnormal increases in market volatility. The choice of economic sectors rather than individual stocks is a compromise that allows us to use tractable time series econometric models such as vector autoregressive processes and work in small dimensions given by the number of sectors. Furthermore, by choosing sectors instead of stocks we assume away the effect of specific shocks on individual assets.

Shocks to economic sectors are shocks to the overall industry and hence their content is useful as a potential driver of market turmoil. The implementation of the random forest methodology allows us to assess the ability of the volatility spillovers on predicting extreme volatility for the S&P 500 index. We find strong out-of-sample predictive ability of volatility spillovers for threshold values in the range 15% to 20% characterising extreme volatility regimes. Unsurprisingly, during the ongoing COVID-19 crisis, the probability of occurrence of extreme volatility predicted out of sample by the random forest is close to one for the S&P 500 index.

The remainder of the paper is structured as follows. Section 2 discusses the importance of considering sectoral economic interdependencies for modelling financial crises and systemic risk events. Section 3 describes the dataset and reviews the methodology introduced by Diebold and Yilmaz (2012) to measure volatility spillovers between economic sectors. Section 4 presents the empirical results for the connectedness analysis. Section 5 reports the predictive exercise for the volatility of the S&P 500 index over different crisis episodes. Section 6 sets out the conclusions of our empirical study.

2. Interdependence between economic sectors and financial crises

There are different economic specifications that motivate the relevance of sector shocks and their propagation to other sectors as factors contributing to the overall aggregate volatility.

Acemoglu et al. (2012) show how idiosyncratic shocks can lead to aggregate fluctuations in the presence of intersectoral input-output linkages, depending on the importance of different sectors that act as suppliers to their immediate customers as well as their role as indirect suppliers to chains of downstream sectors. Acemoglu et al. (2015) also show how the financial network architecture can affect the likelihood of systemic failures related to contagion and counterparty risk. This financial architecture is shaped by the contribution of the different sectors to the overall economic activity. For example, in Banking&Insurance, these authors find that the larger the interdependencies between banks the higher the likelihood that a large negative idiosyncratic shock can propagate through the financial system affecting the overall risk-taking behaviour. Consequently, the interbank market acts as an amplifier of idiosyncratic shocks. Systemic risk and financial market distress are economically important as they affect the skewness of the distribution of subsequent shocks to macroeconomic variables (Giglio et al., 2016).

Although the importance of the financial sector contribution to systemic risk is well documented, other sectors such as Energy or Industrials can be also considered as risk contributors (Collet and Ielpo, 2018; Wu, 2019). The Energy sector arises as a potential shock multiplier due to the links between oil prices and the rest of economic sectors, see Hamilton (1983). The impact of shocks to this sector on other areas of economic activity varies across sectors depending on their reliance on energy consumption. For example, a positive oil supply shock that is likely to increase production costs and reduce the demand of products is likely to impact negatively on sectors that are more dependent on oil as an essential input such as the Transportation sector, Airlines, etc. Conversely, industries that obtain a significant part of their revenue from oil related products such as Oil and Gas usually exhibit a positive oil price exposure. There is also ample evidence of the time-varying relationship between oil price changes and stock returns (Filis et al., 2011). More recently, Aloui et al. (2020) have shown the impact of the COVID-19 pandemic on crude oil and natural gas S&P GS Indexes.

Another consequence of the interconnection between economic sectors is in equity valuation. Large idiosyncratic shocks from one leading sector reflecting unexpected and relevant information are likely to affect not only the volatility of the equity valuation in that sector but also other sectors' equity valuation volatilities through the spillover mechanism. The linkages between economic sectors also influence portfolio allocation due to the presence of asymmetric propagation structures, which affect the effectiveness of portfolio diversification (Zareii, 2019). Related to this is the presence of slow diffusion of information among economically linked firms. The limited information-processing capacity of investors to shock transmissions implies predictable returns (Cohen and Frazzini, 2008). This finding explains the cross-predictability among firms due to investors' late response to the occurrence of shocks. Moreover, the gradual diffusion of information across markets also explains that industry returns are able to predict market movements (Hong et al., 2007).

3. Data and Methodology

This section describes the data used in our empirical analysis, and provides a summary of the econometric methodology introduced in Diebold and Yilmaz (2012) to measure network dependence and, in particular, volatility spillovers.

3.1. Data

We examine annualised S&P 500 daily sectoral return volatilities according to the S&P sector classification system. We compute the logarithmic daily percentage sector returns using closing prices from Bloomberg. Data cover the period from July 20, 2003 through December 31, 2020, comprising 4546 daily observations. The choice of July 20, 2003 is motivated by data availability. This is the first period we have access to data on all subsectors of economic activity.

Table 1 reports the sectors under study and the subsectors included in each sector. We compute the volatility of each sector as the average volatility across sub-sectors. The following sectors are considered as they are the most relevant and representative of economic activity: Health Care, Pharmaceuticals, Biotechnology, Banking&Insurance, Cyclical Sector, Technology and Energy.

[Insert Table 1 about here]

We plot the seven sectors' annualised daily volatility in Figure 1 and provide summary statistics of log annualised volatilities in Table 2. The following insights emerge from these figures: first, Banking&Insurance and Energy are the most volatile, especially throughout the global financial crisis of 2007–2009 and the recent COVID-19 crisis, respectively. During the global financial crisis of 2007–2009, the volatility of the Cyclical Sector spiked remarkably. A similar observation is noted for the Technology sector during the ongoing COVID-19 pandemic. Second, the Health Care and Pharmaceuticals sectors exhibit volatilities that are, on average, lower than for the other sectors. This result is expected as these sectors are usually considered as counter-cyclical defensive sectors and, consequently, more stable under market turmoil. Third, the volatility dynamics of the Biotechnology sector do not appear to be dependent on the state of the market and are similar across bear and bull markets. Nevertheless, this volatility remains high throughout the evaluation sample, in contrast to Health Care and Pharmaceuticals, providing empirical evidence of the presence of significant sectoral idiosyncratic risks.

[Insert Table 2 and Figure 1 about here]

3.2 Measuring sectoral spillover effects

We follow the methodology introduced in Diebold and Yilmaz (2012) to examine the spillover effects between U.S. sectoral volatility market returns. These authors develop a connectedness index that is based on a decomposition of the forecast error variance from a VAR process that is invariant to the ordering of the variables in the model specification.

The procedure is as follows; first, a standard VAR model is fitted to the vector of time series. We establish an H period-ahead forecast evaluation period using data up to time t, and decompose the error variance of each forecast with respect to shocks from the same or other components at time t. Diebold and Yilmaz (2012) propose several connectedness measures based on assessing shares of forecat error variation in various locations due to shocks arising elsewhere. More formally, let

$$Y_t = [\sigma_{\text{Health Care}}, \sigma_{\text{Pharmaceuticals}}, \sigma_{\text{Biotechnology}}, \sigma_{\text{Banking\&Insurance}}, \sigma_{\text{Cyclical sectors}}, \sigma_{\text{Technology}}, \sigma_{\text{Energy}}]$$

denote a 7-dimensional time-series vector following a VAR(p) as

$$Y_t = \sum_{i=1}^{p} \Phi_i Y_{t-i} + \varepsilon_t, \tag{1}$$

where Φ_i denote the slope matrices associated to the vector of variables, and $\varepsilon \sim (0, \Sigma)$ is a vector of independent and identically distributed disturbances with variance-covariance matrix given by Σ. Assuming that the VAR process is covariance stationary, the moving-average representation exists and is given by

$$Y_t = \sum_{i=0}^{\infty} A_i \varepsilon_{t-i} \tag{2}$$

where the $N \times N$ coefficient matrices obey the folowing recursive relations:

$$A_i = \Phi_1 A_{i-1} + \Phi_2 A_{i-2} + \cdots + \Phi_p A_{i-p} \tag{3}$$

with $A_0 = I_N$ and $A_i = 0$ for $i < 0$. To make the VAR model operational, we follow Diebold and Yilmaz (2012) and replace the standard deviations σ in Y_t by estimates of the annualized volatility, computed as $\tilde{\sigma} = (0.361)^{0.5} [\ln P_{t,h} - \ln P_{t,l}]$, with $P_{t,h}$ and $P_{t,l}$ the high and low daily prices, respectively.

Diebold and Yilmaz (2012) identify financial and macroeconomic connectedness based on the H-step-ahead forecast error variance decomposition[3]. Let d_{ii} denote the share of own variance obtained from the H-step ahead forecast error variance decomposition in forecasting $Y_{i,t+H}$ due to shocks to Y_{it}, for $i = 1,2,....N$. Similarly, d_{ij} denote the share of the overall variance that is due to spillover effects obtained from the H-step-ahead forecast error variance decomposition in forecasting $Y_{i,t+H}$ due to shocks to Y_{jt}, for $i, j = 1,2,....N$ and $i \neq j$. The different contributions obtained from the H-step-ahead error variance decomposition denoted as $\theta(H)^{g}_{i,j}$ for $i, j = 1,2,....N$ can be calculated as:

$$\theta(H)^{g}_{i,j} = \frac{\sigma_{jj}^{-1} \sum_{h=0}^{H-1} (e_i' A_h \Sigma e_j)^2}{\sum_{h=0}^{H-1} (e_i' A_h \Sigma A_h' e_i)}, \quad (4)$$

where Σ is the variance-covariance matrix for the error term vector ε in the non-orthogonalized VAR model; σ_{jj} is the j^{th} diagonal element of Σ, and e_i is the selector vector with unity at its i^{th} element, and zero elsewhere. As the shocks are not necessarily orthogonal in the generalized variance decomposition, sums of the forecast error variance contributions are not necessarily unity, therefore, $\sum_{j=1}^{N} \theta(H)^{g}_{i,j} \neq 1$. Diebold and Yilmaz (2012) normalize each entry of the variance decompostion matrix by the row sum:

$$\tilde{\theta}(H)^{g}_{i,j} = \frac{\theta(H)^{g}_{i,j}}{\sum_{j=1}^{N} \theta(H)^{g}_{i,j}} = d_{ij} , \quad (5)$$

and, consequently, $\sum_{j=1}^{N} \tilde{\theta}(H)^{g}_{i,j} = 1$ and $\sum_{i,j=1}^{N} \tilde{\theta}(H)^{g}_{i,j} = N$.

[3]Diebold and Yilmaz (2012) use the generalized variance decomposition frmaework of Koop et al. (1996) and Pesaran and Shin (1998) that allows one to compute variance decompositions invariant to ordering. Therefore, the generalized variance decomposition does not require orthogonalized shocks being able to work with reduced form VAR models.

The full set of variance decompositions is a NxN matrix, $D=[d_{ij}]$ that produces a connectedness table as illustrated in Table 3, allowing us to compute the main measures of connectedness between the different variables in the VAR system.

[Insert table 3 about here]

There are N-1 pairwise directional connectedness from j to i, for each $i=1...N$, defined as

$$C^H_{i\leftarrow j} = \tilde{\theta}(H)^g_{i,j} = \mathrm{d_{ij}} \tag{6}$$

$C_{i\leftarrow j}$ denotes the fraction of the H-step-ahead forecast error variance in variable i that is due to shocks arising from variable j. Hence there are N^2-N separate pairwise directional connectedness measures. Given that, in general, $C^H_{i\leftarrow j} \neq C^H_{j\leftarrow i}$; it is sometimes convenient to define the $N(N-1)/2$ net pairwise directional connectedness measures, for all $i,j=1...N$, as

$$C^H_{ij} = C^H_{j\leftarrow i} - C^H_{i\leftarrow j}. \tag{7}$$

The off-diagonal row and column sums, labeled "from" and "to", are the 2N total directional connectedness measures, N "from others" and N "to others". The total directional connectedness from the other components to i is

$$C^H_{i\leftarrow\bullet} = \textstyle\sum_{j\neq i,j=1}^{N} \tilde{\theta}(H)^g_{i,j} \tag{8}$$

The total directional connectedness from j to others is

$$C^H_{\bullet\leftarrow j} = \textstyle\sum_{j\neq i,i=1}^{N} \tilde{\theta}(H)^g_{i,j}. \tag{9}$$

The net total directional connectedness is

$$C_i = C^H_{\bullet\leftarrow i} - C^H_{i\leftarrow\bullet} \tag{10}$$

The grand total of the off-diagonal entries in matrix *D*, see Table 3, which sums the "from" column or the "to" row, provides a measure of total connectedness. The connectedness table simply augments *D* with a rightmost column containing row sums, a bottom row containing column sums, and a bottom-right element containing the grand average, in all cases for $i \neq j$. This measure corresponds to the lower right cell of the connectedness table, as a percent of total variation

$$C^H = \frac{1}{N} \sum_{\substack{i,j=1 \\ i \neq j}}^{N} \tilde{\theta}(H)_{i,j}^{g} = \frac{1}{N} \sum_{\substack{i,j=1 \\ i \neq j}}^{N} d_{i,j} \tag{11}$$

4. Empirical results on sectoral volatility connectedness

This section reports the empirical results for the connectedness analysis discussed above for the economic sectors comprising the S&P 500 index. The variables of interest are the volatilities of each sector implying that the study of connectedness captures volatility spillovers. First, we show the static connectedness measures across economic sectors. This exercise offers a photo of the unconditional interdependencies across sectors of economic activity. The second exercise introduces the presence of dynamics in the connectedness measures via rolling window estimation.

4.1. Unconditional patterns: full-sample volatility connectedness

Table 4 shows the unconditional full-sample volatility connectedness measures, which are obtained through the generalised variance decomposition. Each of the *ij* entries in Table 4 is the estimated contribution to the forecast error variance of sector *i* coming from innovations to sector *j*. The table shows pairwise and the "from" and "to" total directional connectedness measures. The estimate of total volatility connectedness is based on a VAR model of order three and generalised variance decompositions of 12-day-ahead forecast errors (Diebold and Yilmaz, 2015). The choice of twelve days is to be consistent with

Diebold and Yilmaz (2015). Results for 6-day-ahead forecast errors are available from the authors upon request. The results are robust across specifications of the forecasting horizon.

[Insert Table 4 about here]

The first important result that emerges from Table 4 is that about 72% of the volatility forecast error variance in all the variables is due to connectedness between the variables, which highlights the importance of the volatility spillovers. The results in Table 4 show that Pharmaceuticals, Health Care, Technology and Cyclical Sectors have a “from” connectedness statistic larger than the total connectedness, due mainly to the impact of volatility shocks to the Energy sector, and to a lesser extent to shocks to the Banking&Insurance and Biotechnology sectors. The Biotechnology sector is closely connected to the Pharmaceuticals and Health Care sectors and the Banking&Insurance sector is interconnected to the Cyclical Sectors, with the latter two sectors being very dependent on the overall financing conditions. In contrast, the Energy sector has the lowest “from” connectedness reaching a value about 50%.

It is also important to emphasise that the “to” connectedness of the Energy, Banking&Insurance and Biotechnology sectors (101.35%, 99.36% and 69.51%) exceed their “from” connectedness by 51.21%, 35.52% and 6.16%, respectively. These sectors are clear net transmitters of shocks to the system. On the other hand, Pharmaceuticals, Health Care, Technology and Cyclical Sectors have negative net connectedness (−25.99%, −36.63%, −15.71% and −14.56%), indicating that these sectors are net receivers of shocks from other sectors.

4.2. Dynamics of total sectoral volatility connectedness

In this section we focus on the dynamics of the sectoral volatility connectedness over time as the above unconditional analysis may overlook important interactions across different volatility regimes. Therefore, the dynamic analysis allows us to capture cyclical and secular movements in connectedness. In our context, we are especially interested in the sectoral volatility connectedness behaviour around the global financial crisis and the COVID-19 pandemic crisis. To accomplish our objective, we carry out a dynamic analysis using 200-day rolling-sample windows (Diebold and Yilmaz, 2015) over the evaluation period, and examine graphically the connectedness plots. The choice of 200 days is for consistency with the original work of Diebold and Yilmaz (2015).

Figure 2 shows the dynamics of the total connectedness measure. This figure shows that total sectoral volatility connectedness ranges between 40% and 85% as shocks propagate across the system with different impact over time. During the calm 2004–2006 period, the index reached levels below the average conditional connectedness and then starts to increase steadily due to the tightening in monetary policy deployed by the U.S. Federal Reserve from mid-2006 onwards. Around the global financial crisis, the index reached a maximum of 78.55% on June 15, 2009, which is lower than the value subsequently attained during the Eurozone debt crisis of 2010–2012. In the latter period, sectoral volatility connectedness achieved values larger than 80%. In contrast, the period from August 2012 onwards is characterised by a decrease in the average total connectedness of the sectoral volatilities, reaching a low of 40% at the end of 2017 as equity markets entered a bullish period given by a low volatility regime.

The analysis of volatility spillovers shows a spike during the COVID-19 pandemic crisis as it increases sharply from the levels reached at the beginning of 2020 (around 60%) until reaching a maximum of 85.30% on March 3, 2020. From this day onwards, the total sectoral volatility connectedness is always above 74% even after the recovery of equity valuations, as the gloomy outlook for the real economy still remains. However, despite the sectoral volatility connectedness remains high the index dropped below 80% after Pfizer and Biotech announce publication of positive results from landmark phase three trial of their vaccine candidate, on December 10th.[4] Consequently, the sectoral volatility connectedness varies over time and is especially relevant during periods of high uncertainty.

[Insert Figure 2 about here]

4.3. Dynamics of directional sectoral volatility connectedness

More informative is the analysis of spillover effects from a directional perspective. In this case, we study the information contained in the net connectedness, measured as the difference between the "to" and "from" connectedness, using the directional connectedness plots discussed above. By doing so, we analyse the directional information that is masked under the total connectedness plot.

Figure 3 plots the net directional volatility connectedness measure for each sector as the difference between its "to" and "from" connectedness.

[Insert Figure 3 about here]

The net contribution of the directional volatility of each sector varies greatly over time. During volatile times the net directional contribution usually increases. However, there is a large variability across sectors. The Banking&Insurance sector net directional

Data from 43,448 participants, half of whom received BNT162b62 and half of whom received placebo, showed that the vaccine candidate was well tolerated and demonstrated 95% efficacy in preventing COVID19 in those without prior infection 7 days or more after the second dose.

connectedness is especially relevant during the 2007-2009 crisis as its origin was in the financial sector. During this period volatility shocks coming from the Banking&Insurance sector spread across other sectors of the S&P 500 index as financial conditions deteriorate having a subsequent adverse effect on the real economy. In contrast, during the COVID-19 crisis, this sector exhibited significantly lower levels of net directional connectedness reaching negative values after May 2020.

Figure 3 also shows that the Cyclical Sector's net directional connectedness is mainly positive during the U.S. Federal Reserve tightening of monetary policy from the mid-2006 onwards and the 2007–2009 global financial crisis. However, the net directional connectedness of this sector mainly turns into negative territory starting from the Eurozone debt crisis in 2010–2012 onwards, coinciding with a long and steady period of low interest rates. Interestingly, the COVID-19 crisis has had a negligible effect on the dynamics of this sector. The December run up of equity prices and the positive results about the efficacy of different vaccine candidates made the net directional connectedness of this sector be above zero. In contrast, the Energy sector is very sensitive to the occurrence of crisis episodes. Figure 3 shows that the net directional connectedness of this sector is especially positive throughout all the crises, but also during the 2004–2006 calm period and the period of tightening in monetary policy following from mid-2006 onwards. Therefore, the Energy sector is a net transmitter of volatility shocks to other sectors comprising the S&P 500 index. The dynamics of the net connectedness of the Energy sector during the COVID-19 crisis are particularly relevant. We observe a jump from a negative value on December 31, 2020 (−4.98%) to a maximum of 312% reached on May 6, 2020, providing an average level of net connectedness of 88% for the Energy sector during 2020. This behaviour can be partly explained by the plunge of oil prices, reaching $30 a barrel, as the coronavirus has undermined energy demand worldwide, triggering a war price among worldwide oil producers that pursued increasing shares of the market.

Figure 3 also reveals that the Biotechnology sector is a clear transmitter of shocks from 2012 onwards. However, the Biotechnology net directional connectedness is negative during the COVID-19 crisis as the Energy sector has become the unique transmitter of volatility shocks through the system. Nevertheless, there is an increasing trend of the Biotechnology net directional connectedness as the news about the availability of potential vaccines and treatments is monitored by investors. Levels of Pharmaceuticals and Health Care net directional connectedness during the COVID-19 crisis have behaved similarly to Biotechnology net directional connectedness as their businesses are closely related. The stark difference is that both sectors, Pharmaceuticals and Health Care, display mainly negative net directional connectedness during the entire evaluation sample and, consequently, these sectors are net receivers of volatility shocks. The analysis of the Technology sector also shows evidence of dynamics in the net connectedness measure across the sample, displaying negative values during both crises, the 2007–2009 global financial crisis and the COVID-19 pandemic crisis.

Overall, our analysis of volatility spillovers across U.S. economic sectors shows that Banking&Insurance, Energy, Technology and Biotechnology are the main trasnmitters of volatility spillovers over the evaluation period, with the intensity of the spillovers varying over time.

5. Predicting extreme market volatility using sectoral volatility connectedness

In this section, we report empirical evidence on the ability of sectoral volatility connectedness for predicting financial crises. We identify a crisis as a period in which market volatility exceeds 20%. Our empirical study focuses on the S&P 500 index and carries out an out-of-example exercise to predict the daily volatility of the S&P 500 index using random forests, see Breiman (2001) and Cutler et al. (2011) for a review of random forests, and Hastie et al. (2017) for an overview of machine learning methods. This novel methodology is nonparametric and allows to aggregate/ensemble the predictions of individual decision trees to reduce uncertainty in the model predictions.

5.1. Prediction using random forests

We study the ability of binary classifiers or tree-structured classification to predict out of sample the annualised S&P 500 daily volatility using S&P 500 sectoral volatility connectedness measures. As connectedness jumps whenever financial market volatility rises, we expect that the whole set of sectoral connectedness measures, which provides information about how the shocks propagate through the sectors, conveys information about future volatility. We focus on two volatility states: the high volatility state that is characterised by a volatility higher than 20% and a low volatility state below this level.

The high complexity of financial markets makes the use of tree-structured classifiers very suitable as they can handle problems characterised by high dimensionality, nonnormality and nonhomogeneity, that is, different relationships between variables in different parts of the measurement space (Breiman et al., 1984). Thus, the tree-structured classification is used not only to produce an accurate classifier, but it is also helpful if there are many variables as covariates. It can be also used to identify those variables with stronger predictive ability and their interactions. Our aim is to exploit the connectedness measures to predict out of sample extreme volatility episodes for the S&P 500 index.

A binary classifier is a rule that assigns a predicted class membership in $C = \{0,1\}$, which represents the set of classes, based on a measurement vector x in the measurement space X. Given any x in X, the rule assigns one of the classes $\{0,1\}$ to x. In our setting, we want to predict a binary variable Y (= 1 if annualised S&P 500 daily volatility is higher than 20% and 0 otherwise) from the observed vector of state variables [X_1, …, X_n] given by the set of estimated volatility spillovers among the different economic sectors. More formally, the set of state variables is comprised by the following variables: the total volatility connectedness measure, 21 total directional volatility connectedness variables and 21 total net pairwise directional volatility connectedness variables. Therefore, our state variable vector is comprised of n=43 variables. The tree-structured classifiers are constructed by repeated splits of subsets of X into two descendant subsets, beginning with X itself. The tree is made of nodes and branches: internal nodes split into two children and terminal nodes, which do not have any children. A terminal node has a class label associated with it according to the highest posterior probability such that observations that fall into the particular final node are assigned to that class. The splitting condition at each node is expressed as an "if-then-else" rule that is determined by a specific splitting criterion so that the data in each of the descendant subsets are purer.

In what follows, we briefly discuss the steps needed to create a tree classifier (Breiman et al., 1984). First, growing an overly large tree that gives us optimal splits for the tree (to grow a tree at a given node, we search for the best split in terms of decreasing the node impurity using the Gini diversity measure searching through all the state variables). Second, pruning the large tree using minimal cost complexity criteria to create a sequence of sub-trees that avoid the overfitting of the large tree. Third, choosing the 'right tree' using cross-validation methods or an independent tree. We choose the best tree based on the misclassification rate estimated by cross-validation using ten cross-validation subsamples. Once we have the right tree, the class labels are assigned and we compute the corresponding misclassification rate.

In order to improve the stability and predictive power of classification and regression trees, several methods that involve different ways of combining ensembles of classifiers are developed. We build on the work of Breiman (1996a, 1996b, 1996c, 2001); these methods result from growing an ensemble of trees and letting them vote for the most popular class in our classification setting. Random forest methods are a combination of tree predictors such that each tree depends on the values of a vector of state variables randomly sampled and with all the trees in the forest having the same distribution. To classify a new observation x, we use the majority vote from the random forest. In our context, at day t we consider $n^{1/2}$ variables to predict the volatility state (high or low) at time $t+1$, and consider a total of 200 trees.

5.2. Empirical prediction exercise using sectoral volatility spillovers

The data used to train the random forest decision tree cover the period from July 20, 2003 through December 19, 2017. The remaining data constitutes the out-of-sample period and cover the period January 1, 2018 through August 31, 2020, a total of 695 observations. This period reflects both a calm and a very turbulent period around the COVID-19 pandemic. During this period, daily volatility is higher than 20% a total of 27.6% of the days.

Figure 4 plots the S&P 500 annualised daily volatility together with the probability assigned by the random forest to the high volatility regime over the out-of-sample evaluation period. The blue line plots the probability predicted by the random forest using the connectedness measures as inputs of the model. The correlation between both series is 92.31%, suggesting that the volatility connectedness measures are useful for predicting the future volatility of the S&P 500 index in the high volatility regime. Throughout the

COVID-19 crisis, the predicted probability is close to one suggesting that the occurrence of a crisis is almost certain.

Figure 5 plots the receiver operating curve or ROC curve, see Fawcett (2006), that illustrates the diagnostic ability of the random forest model as its discrimination threshold is varied. The ROC curve is created by plotting the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings. The true positive rate is also known as sensitivity. The false positive rate can be calculated as (1 - specificity).[5] In our example, the curve illustrates the ability of the probability of the random forest model to predict the high volatility regime for the S&P 500 index. The area under the curve is 0.98, which shows how useful the sectoral volatility connectedness is to predict the volatility of the S&P 500 index out of sample.

[Insert Figures 4 and 5 about here]

Finally, we also compute a measure of the marginal importance of the covariates in the prediction model. To do this, we compute the variable importance (VI) statistic, see Ellies-Oury et al. (2019). These authors construct a measure of importance for each covariate X_j by estimating the response variable introducing perturbations to the covariate and computing the error due to these perturbations. The variable importance (VI) of the covariate X_j is computed as $VI_j = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \widetilde{y_i^j}\right)$, where y_i is the predicted value when the observations of the j^{th} covariate are randomly permuted in the sample and $\widetilde{y_i^j}$ is the new estimated link function. If the covariate X_j has an effect on Y, the random permutation of its observations will affect the prediction of Y by increasing the error measured in VI_j.

[5] **Sensitivity** measures the proportion of positives that are correctly identified. **Specificity** measures the proportion of negatives that are correctly identified.

The covariates with the highest VI statistic are considered the most important to predict the response variable. The VI measure is computed for every tree and the average over the entire ensemble is divided by the corresponding sample standard deviation to obtain values that are comparable across variables.

Figure 6 plots the ordered average of the variable importance measure over the out-of-sample evaluation period for each connectedness measure. This figure shows that shocks related to Technology are especially relevant for predicting the volatility of the S&P 500 index in the high volatility regime. This is so because the measures for the Technology "from", Technology "To" and Technology "net" connectedness are the first, fourth and sixth more important variables with respect to the prediction error. It is also relevant that net pairwise connectedness between Pharmaceuticals vs. Health Care and Pharmaceuticals vs. Biotechnology are the second and third more important variables in the out-of-sample analysis.

[Insert Figure 6 about here]

6. Conclusions

This paper measures volatility spillovers between sectors of economic activity using network connectivity measures. To do this, we apply the novel methodology proposed in Diebold and Yilmaz (2012) to seven economic sectors of economic activity in the U.S. Our empirical results show that Banking&Insurance, Energy, Technology and Biotechnology are the main channels through which shocks are transmitted to the rest of the economy. Banking&Insurance is especially relevant during 2007-2009 global financial crisis, while the Energy sector is especially relevant during the COVID-19 crisis. Health Care and Pharmaceuticals are, on the other hand, net receivers of shocks. Volatility spillovers are usually regime-dependent, increasing their intensity during turmoil periods. The Biotechnology sector is, however, the exception.

Volatility spillover effects are also useful for predicting crises. Using random forests algorithms and decision trees, we find that sectoral volatility spillovers help to predict the occurrence of extreme volatility regimes for the S&P 500 index. These results are particularly revealing for the recent period around the spread of the COVID-19 pandemic. During this period, using out-of-sample data, we find that the probability of observing extreme volatility in the S&P 500 index is close to one.

The results of this study can be extended in different directions. Our findings reveal the importance of measures of volatility spillovers as early indicators of financial distress, however, the predictive ability of these measures should be explored in more detail and using other prediction models as robustness exercises. We believe this is beyond the scope of this paper. A second extension is to use these measures of connectiveness to predict the dynamics of macroeconomic and financial variables. Preliminary results show the presence of cointegration between our connectivity measures and the nominal short-term interest rates. This finding suggests that volatility spillovers can be used as state variables for volatility prediction, portfolio allocation and other relevant areas in empirical finance models.

References

Acemoglu, D., Carvalho, V.M., Ozdaglar, A., Tahbaz-Salehi, A., 2012. The network origins of aggregate fluctuations. *Econometrica* 80 (5), 1977–2016.

Acemoglu, D., Ozdaglar, A., Tahbaz-Salehi, A., 2015. Systemic risk and stability in financial networks. *American Economic Review* 105 (2), 564–608.

Alizadeh, S., Brandt, M. W., & Diebold, F. X., 2002. Range-based estimation of stochastic volatility models. *Journal of Finance* 57, 1047–1092.

Aloui, D., Goutte, S., Guesmi, K., Hchaichi, R., 2020. COVID 19's Impact on Crude Oil and Natural Gas S&P GS Indexes. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3587740.

Breiman, L., 1996a. Bagging predictors. *Machine Learn*ing 26 (2), 123–140.

Breiman, L., 1996b. The heuristics of instability in model selection. *Annals of Statistics* 24, 2350–2383.

Breiman, L., 1996c. Out-of-Bag Estimation. Department of Statistics, University of California. Retrieved from ftp://ftp.stat.berkeley.edu/pub/users/breiman/.

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

Breiman, L., Friedman, J.H., Olshen, R.A. , Stone, C.G. , 1984. Classification and Regression Trees. *Wadsworth International Group*, Belmont, CA.

Cohen, L., Frazzini, A., 2008. Economic Links and Predictable Returns. *Journal of Finance* 63 (4), 1977-2011.

Collet, J.; Ielpo, F. (2018). Credit spillovers in credit markets. *Journal of Banking and Finance* 94, 267-278.

Conlon, T., Corbet, S., McGee, R.J., 2020. Are Cryptocurrencies a Safe Haven for Equity Markets? An International Perspective from the COVID-19 Pandemic. Forthcoming *Research in International Business and Finance.*

Conlon, T., McGee, R.J., 2020. Safe Haven or Risky Hazard? Bitcoin during the COVID-19 Bear Market. Forthcoming *Finance Research Letters.*

Cutler, A., Cutler, D.R., Stevens, J.R., 2011. Random Forests. *Machine Learning* 45 (1), 157-176.

Diebold, F., Yilmaz, K., 2009. Measuring financial asset return and volatility spillovers, with application to global equity markets. *The Economic Journal* 119, 158-171.

Diebold, F., Yilmaz, K., 2012. Better to give than to receive: predictive directional measurement of volatility spillovers. *International Journal of Forecasting* 28, 57-66.

Diebold, F., Yilmaz, K., 2014. On the network topology of variance decompositions: measuring the connectedness of financial firms. *Journal of Econometrics* 182, 119-134.

Diebold, F., Yilmaz, K., 2015. Financial and Macroeconomic connectedness, *Oxford University Press*.

Ellies-Oury, M.P., Chavent, M., Coranec, A., Bonnet, M., Picard, B., Saracco, J., 2019. Statistical model choice including variable selection based on variable importance: A relevant way for biomarkers selection to predict meat tenderness. *Scientific Reports* 9, Article number 10014.

Fawcett, T., 2006. An Introduction to ROC Analysis. *Pattern Recognition Letters* 27, (8), 861–874.

Filis, G., Degiannakis, S., Floros, C., 2011. Dynamic correlation between stock market and oil prices: the case of oil-importing and oil-exporting countries. *International Review Finance Analysis* 20, 152–164.

Giglio, S., Kelly, B., Pruitt, S., 2016. Systemic risk and the macroeconomy: An empirical evaluation. *Journal of Financial Economics* 119 (3), 457-471.

Goodell, J.W., 2020. COVID-19 and Finance: Agendas for Future Research. Forthcoming *Finance Research Letters.*

Goodell, J.W., Goutte,S., 2020. Co-Movement of COVID-19 and Bitcoin: Evidence from Wavelet Coherence Analysis. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3597144.

Goodell, J.W., Huynh, T.L.D., 2020. Did Congress trade ahead? Considering the reaction of US industries to COVID-19. Forthcoming *Finance Research Letters.*

Hamilton, J., 1983. Oil and the Macroeconomy since World War II. *Journal of Political Economy* 91(2), 228-48.

Hastie, T., Tibshirani, R., Friedman, J., 2017. The Elements of Statistical Learning. Data Mining, Inference, and Prediction. *Springer Series in Statistics*. Springer. Second Edition.

Hong, Harrison & Torous, Walter & Valkanov, Rossen, 2007. Do industries lead stock markets?. *Journal of Financial Economics* 83(2), 367-396.

Koop, G., Pesaran, M., Potter, S., 1996. Impulse response analysis in non-linear multivariate models. *Journal of Econometrics* 74, 119-147.

Pesaran, M, Shin, Y., 1998. Generalized impulse response analysis in linear multivariate models. *Economic Letters* 58, 17-29.

Wu, F. (2019). Sectoral contributions to systemic risk in the Chinese stock market. *Finance Research Letters* 31, 386-390.

Zareii, A., 2019. Network origins of portfolio risk. *Journal of Banking and Finance* 109, 1-29.

Table 1. S&P 500 Sectors.

Sector	Sub-sectors	Bloomberg RIC
Health Care	Healthcare Equipment	.SPSIHE
	Health Care Services	.SPSIHP
Pharmaceuticals	Pharmaceuticals	.SPSIPH
Biotechnology	Biotechnology	.SPBIO
Bank&Insurance	Banks	.SPSIBK
	Insurance	.SPSIINS
	Regional Banks	.SPSIRBK
	Capital Markets	.SPSICM
Cyclical sectors	Retail	.SPSIRE
	Transportation	.SPSITN
	Homebuilders	.SPHOME
Technology	Semiconductors	.SPSEMI
	Software & Services	.SPSISS
	Technology Hardware	.SPSICH
	Telecom	.SPSITE
Energy	Metals and Mining	.SPSIMM
	Oil & Gas Equipment	.SPSIOS
	Oil & Gas Exploration & Production	.SPSIOP

Table 1 reports the sectors under study and their components according to the S&P sector classification.

Table 2. Summary statistics: Log of annualized asset return volatilities by sector of economic activity.

	Health Care	Pharmaceuticals	Biotechnology	Banking & Insurance	Cyclical sectors	Technology	Energy
Mean	1.20	1.24	1.40	1.29	1.33	1.31	1.49
Median	1.17	1.23	1.39	1.23	1.28	1.28	1.45
Maximum	1.98	1.93	1.99	2.09	2.05	2.01	2.20
Minimum	0.73	0.67	1.05	0.85	0.96	0.95	1.09
Standard Deviation	0.17	0.18	0.16	0.24	0.18	0.17	0.18
Skewness	1.36	0.53	0.42	1.21	1.20	1.04	1.09
Kurtosis	2.72	0.87	0.20	1.20	1.32	1.57	1.93

Table 3. Connectedness table using Diebold and Yilmaz (2012) methodology.

	Y_1	Y_2	$\cdots$	Y_N	From Others
Y_1	$d_{11} = \tilde{\tilde{\theta}}^g_{1,1}(H)$		$\cdots$	$d_{1N} = \tilde{\tilde{\theta}}^g_{1,N}(H)$	$\sum_{j=1, j\neq 1}^{N} \tilde{\tilde{\theta}}^g_{1,j}(H)$
Y_2	$d_{21} = \tilde{\tilde{\theta}}^g_{2,1}(H)$	$d_{22} = \tilde{\tilde{\theta}}^g_{2,2}(H)$	$\cdots$	$d_{2N} = \tilde{\tilde{\theta}}^g_{2,N}(H)$	$\sum_{j=1, j\neq 2}^{N} \tilde{\tilde{\theta}}^g_{2,j}(H)$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
Y_N	$d_{N1} = \tilde{\tilde{\theta}}^g_{N,1}(H)$	$d_{N2} = \tilde{\tilde{\theta}}^g_{N,2}(H)$	$\cdots$	$d_{NN} = \tilde{\tilde{\theta}}^g_{N,N}(H)$	$\sum_{j=1, j\neq N}^{N} \tilde{\tilde{\theta}}^g_{N,j}(H)$
To Others	$\sum_{i=1, i\neq 1}^{N} \tilde{\tilde{\theta}}^g_{i,1}(H)$	$\sum_{i=1, i\neq 2}^{N} \tilde{\tilde{\theta}}^g_{i,2}(H)$		$\sum_{i=1, i\neq N}^{N} \tilde{\tilde{\theta}}^g_{i,N}(H)$	$1/N \sum_{i,j=1, i\neq j}^{N} \tilde{\tilde{\theta}}^g_{i,j}(H)$

The table reports the schematic connectedness table that proves central for understanding the various connectedness measures and their relationships. The elements of the matrix contain the variance decompositions, called the "variance decomposition matrix", $D = \left[d_{ij} = \tilde{\tilde{\theta}}^g_{i,j}(H) \right]$. The connectedness table augments the matrix D with a rightmost column containing row sums, a bottom row containing column sums, and a bottom-right element containing the grand average, in all cases for *i* different from *j*.

Table 4. Sectoral volatility connectedness.

	Pharmaceuticals	Health Care	Technology	Banking&Insurance	Cyclical	Biotechnology	Energy	From Others
Pharmaceuticals	22.59	10.13	10.18	13.72	9.06	18.95	15.38	77.41
Health Care	11.83	15.17	12.66	16.34	11.65	14.08	18.28	84.83
Technology	7.99	9.01	19.43	18.92	13.48	11.86	19.30	80.57
Banking&Insurance	6.55	6.61	10.66	36.16	13.08	8.05	18.88	63.84
Cyclical	6.83	8.19	12.99	23.87	21.79	9.50	16.83	78.21
Biotechnology	12.61	8.33	9.56	11.95	8.22	36.65	12.68	63.35
Energy	5.61	5.93	8.82	14.56	8.16	7.06	49.86	50.14
To Others	51.42	48.21	64.86	99.36	63.65	69.51	101.35	
Net	-25.99	-36.6	-15.71	**35.52**	-14.56	**6.16**	**51.21**	**71.19**

Sectoral volatility connectedness obtained through the generalized variance decomposition. The sectoral volatility connectedness is based on a vector autoregression of order 3 and generalized variance decompositions of 12-day ahead forecast errors. Each cell in the 7x7 matrix section of the table reports the relative (in percent terms) contribution of the "column" variable shocks to the variance of the forecast error for the "row" variable. Each cell in the directional "from Others" column reports the total variance of the forecast error share attribuitable to the other variables. Each cell in the directional "To Others" row reports the sum of the contributions of each variable to all other variables's variance of forecast errors. The "net" directional connectedness row reports the difference between the corresponding cells in the "To Others" and the "From Others" column. The total connectedness index is the number in the lower right corner and is equal to the average of the elements of the "From Others column" and the "To Others row".

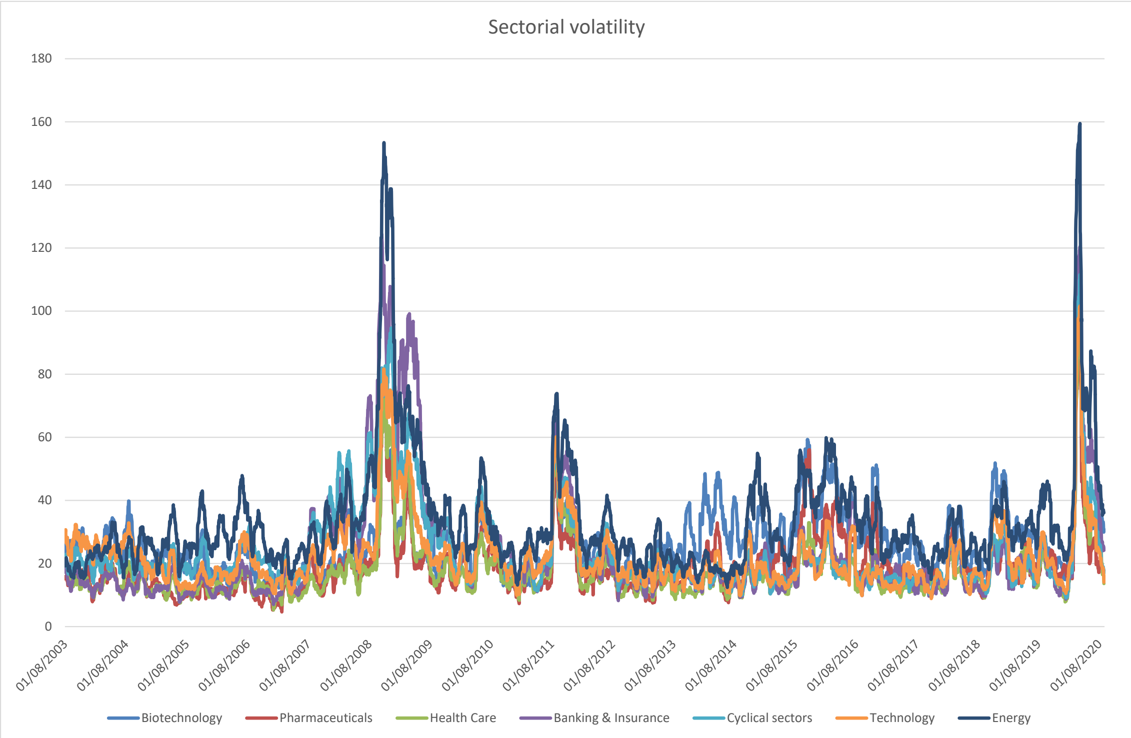


Figure 1. Annualized S&P 500 volatility for the daily returns on each sector of economic activity. We consider sectors according to S&P sector classification system: Health Care, Pharmaceutical, Biotechnology, Banking&Insurance, Cyclical sectors, Technology and Energy. The logarithmic daily percent sector returns are computed using closing prices from Bloomberg. The data consider 4486 daily observations covering the period from July 20, 2003 through December 31, 2020.

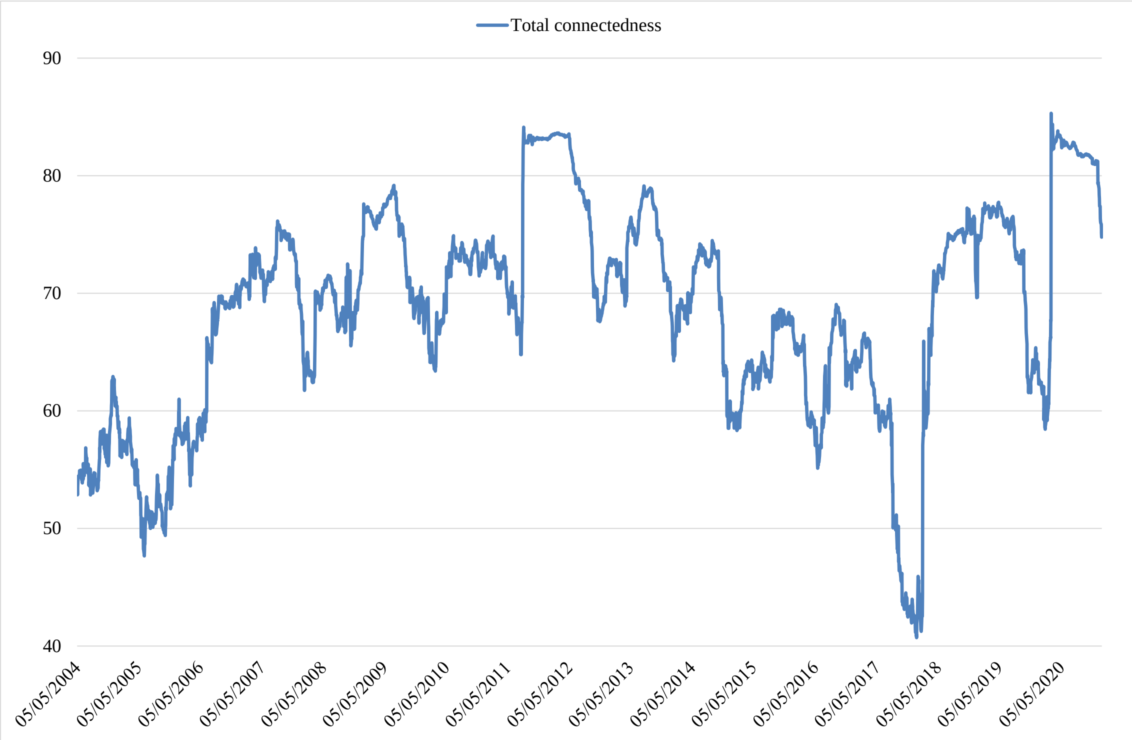
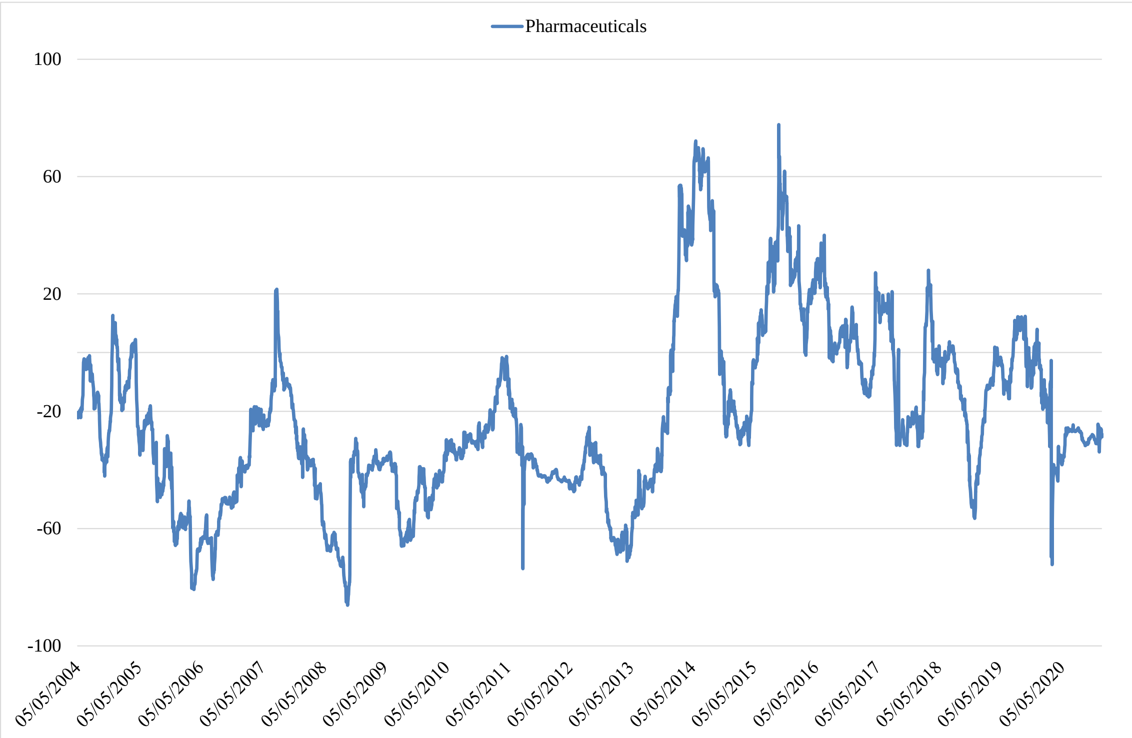
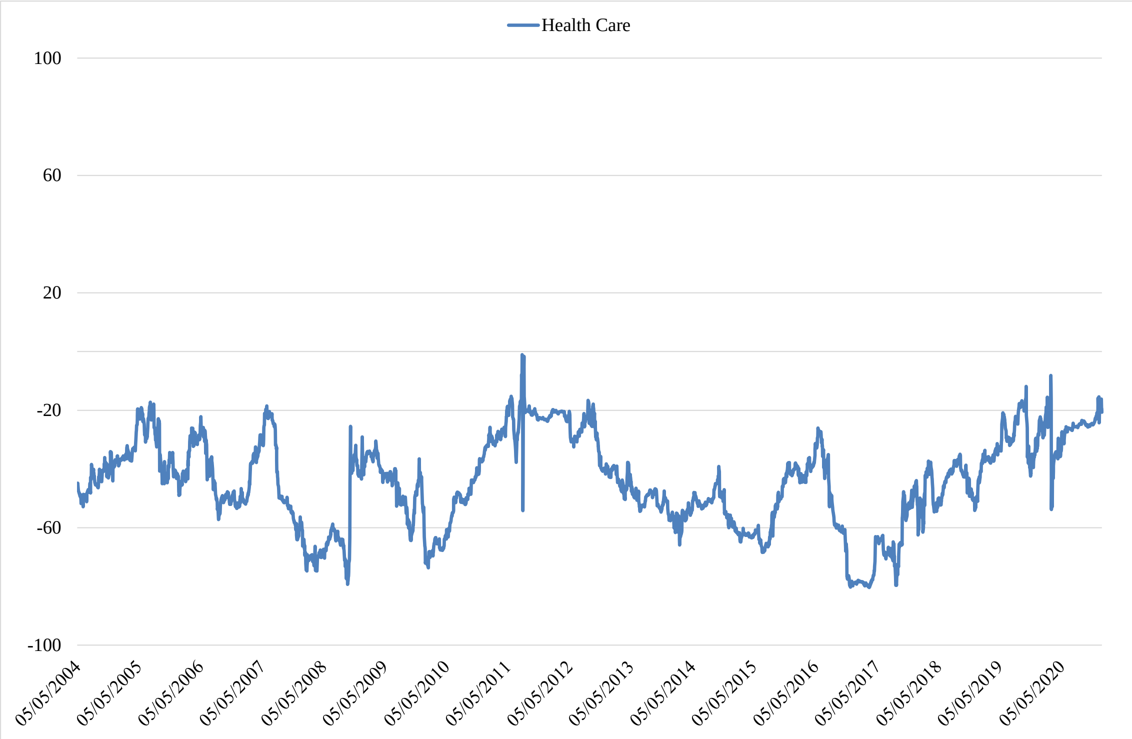
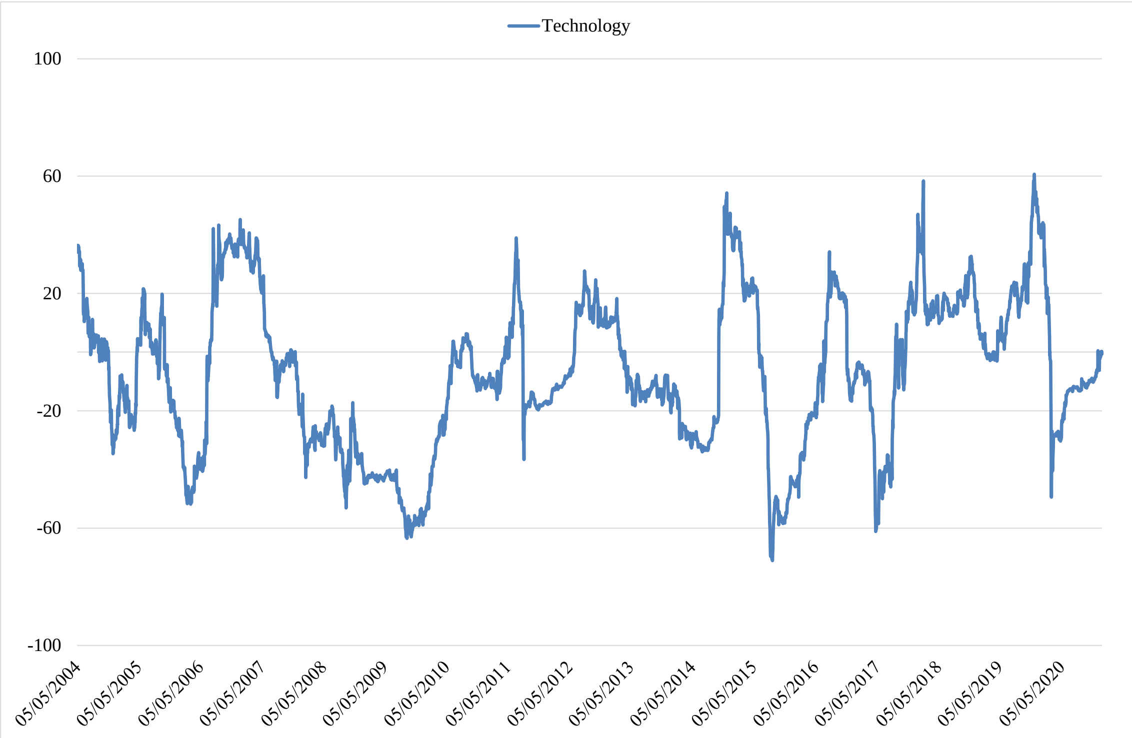
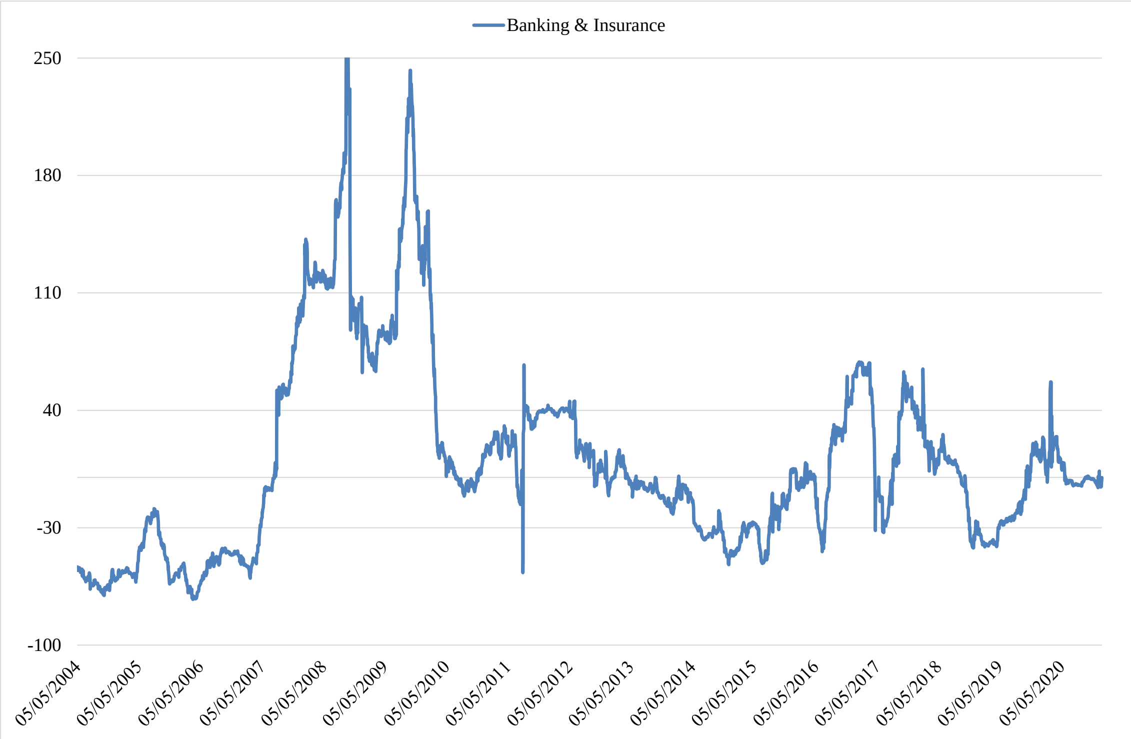


Figure 2. Sectoral volatility connectedness (200-day rolling-sample windows) over the period July 20, 2003 through December 31, 2020.

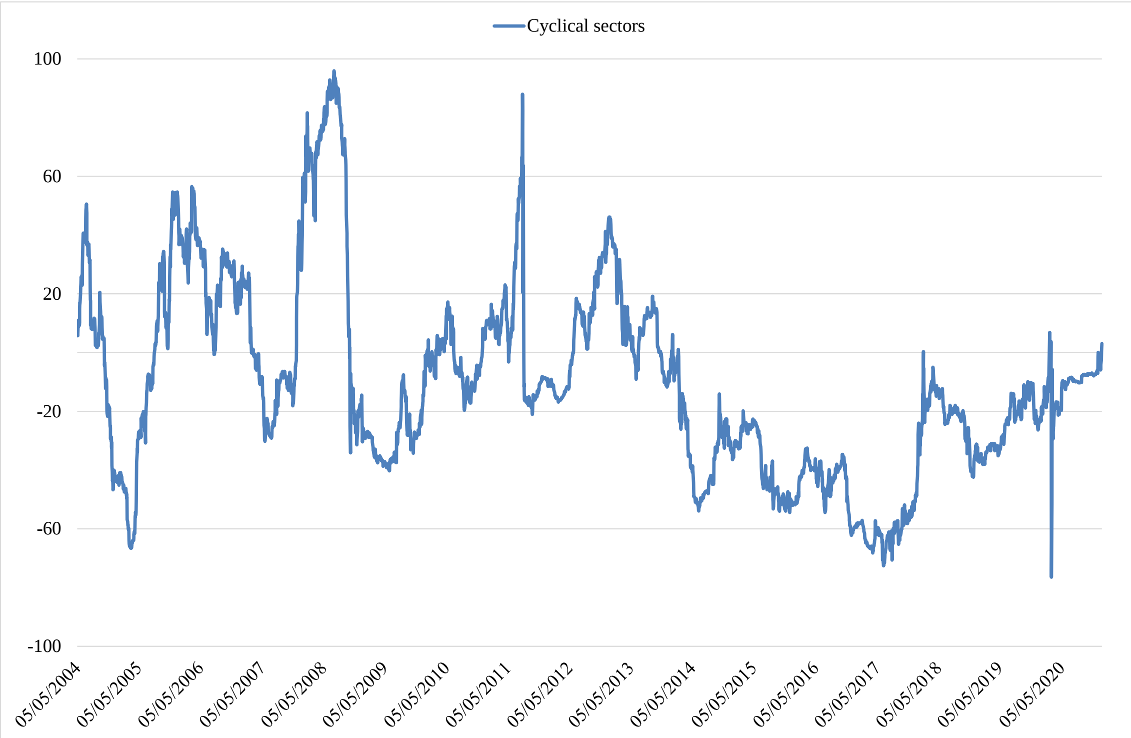



Health Care
100
60
20
-20
-60
-100
05/05/2004
05/05/2005
05/05/2006
05/05/2007
05/05/2008
05/05/2009
05/05/2010
05/05/2011
05/05/2012
05/05/2013
05/05/2014
05/05/2015
05/05/2016
05/05/2017
05/05/2018
05/05/2019
05/05/2020

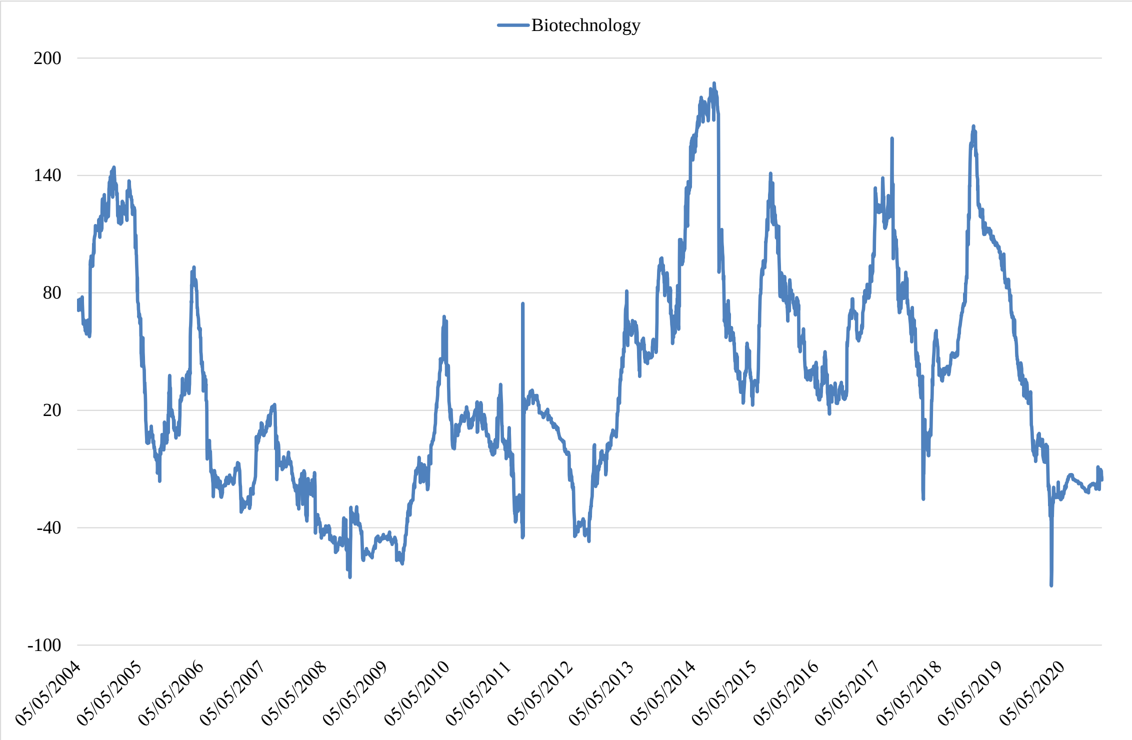

Technology
100
60
20
-20
-60
-100
05/05/2004
05/05/2005
05/05/2006
05/05/2007
05/05/2008
05/05/2009
05/05/2010
05/05/2011
05/05/2012
05/05/2013
05/05/2014
05/05/2015
05/05/2016
05/05/2017
05/05/2018
05/05/2019
05/05/2020



Banking & Insurance
250
180
110
40
-30
-100
05/05/2004
05/05/2005
05/05/2006
05/05/2007
05/05/2008
05/05/2009
05/05/2010
05/05/2011
05/05/2012
05/05/2013
05/05/2014
05/05/2015
05/05/2016
05/05/2017
05/05/2018
05/05/2019
05/05/2020



Cyclical sectors
100
60
20
-20
-60
-100
05/05/2004
05/05/2005
05/05/2006
05/05/2007
05/05/2008
05/05/2009
05/05/2010
05/05/2011
05/05/2012
05/05/2013
05/05/2014
05/05/2015
05/05/2016
05/05/2017
05/05/2018
05/05/2019
05/05/2020


Biotechnology
200
140
80
20
-40
-100
05/05/2004
05/05/2005
05/05/2006
05/05/2007
05/05/2008
05/05/2009
05/05/2010
05/05/2011
05/05/2012
05/05/2013
05/05/2014
05/05/2015
05/05/2016
05/05/2017
05/05/2018
05/05/2019
05/05/2020

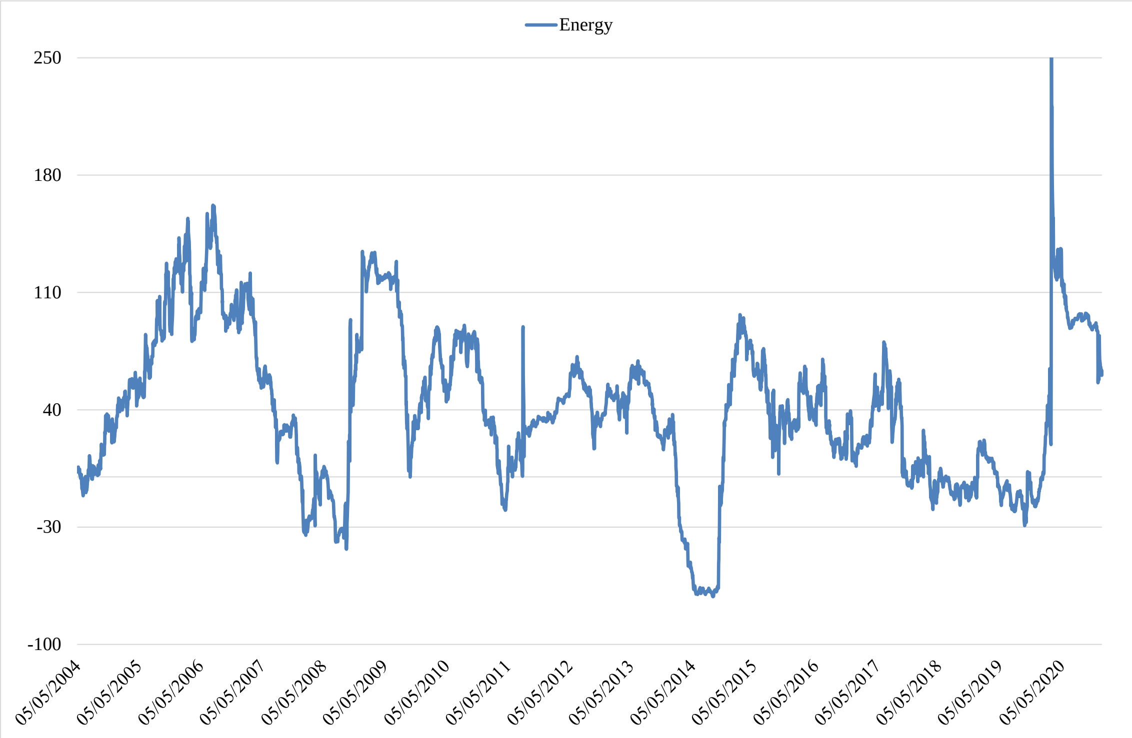

Energy
250
180
110
40
-30
-100
05/05/2004
05/05/2005
05/05/2006
05/05/2007
05/05/2008
05/05/2009
05/05/2010
05/05/2011
05/05/2012
05/05/2013
05/05/2014
05/05/2015
05/05/2016
05/05/2017
05/05/2018
05/05/2019
05/05/2020

Figure 3. Net connectedness Index (200-day rolling-sample windows) for the seven sectors of U.S. economic activity comprising the S&P 500 index. Net directional volatility connectedness is the difference between its “to” and “from” connectedness.

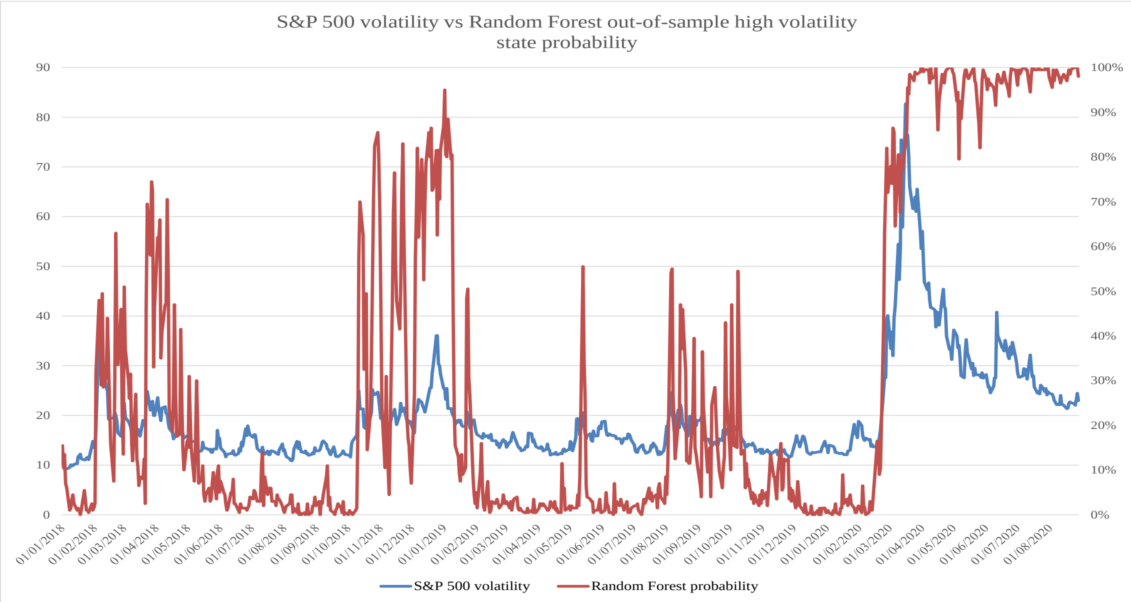


Figure 4. S&P 500 volatility is represented in red. The blue line plots the probability predicted by the random forest using the connectedness measures as inputs of the model. The model is trained using data from July 20, 2003 through December 19, 2017. The remaining data constitutes the out-of-sample period and spans from January 1, 2018 to December t 31, 2020.

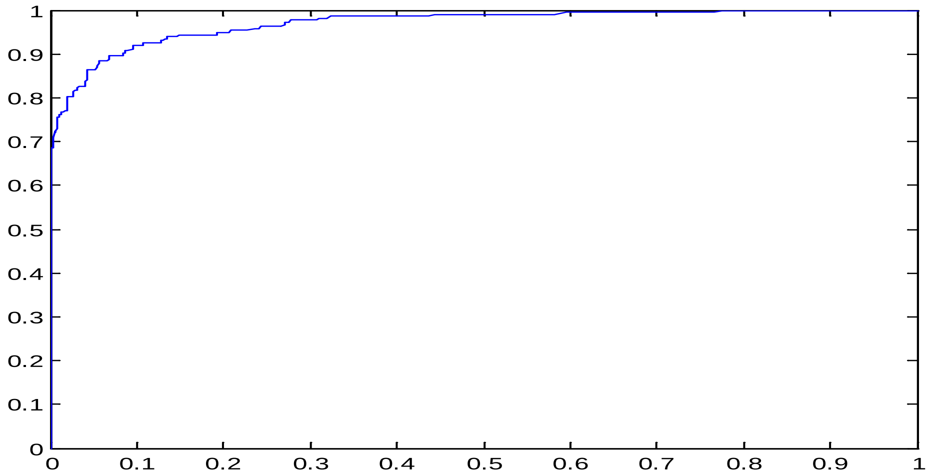


Figure 5. The ROC curve illustrates the ability of the probability of the random forest model to predict the high volatility regime for the S&P 500 index. The model is trained using data from July 20, 2003 through December 19, 2017. The remaining data constitutes the out-of-sample period and cover the period January 1, 2018 through December 31, 2020.

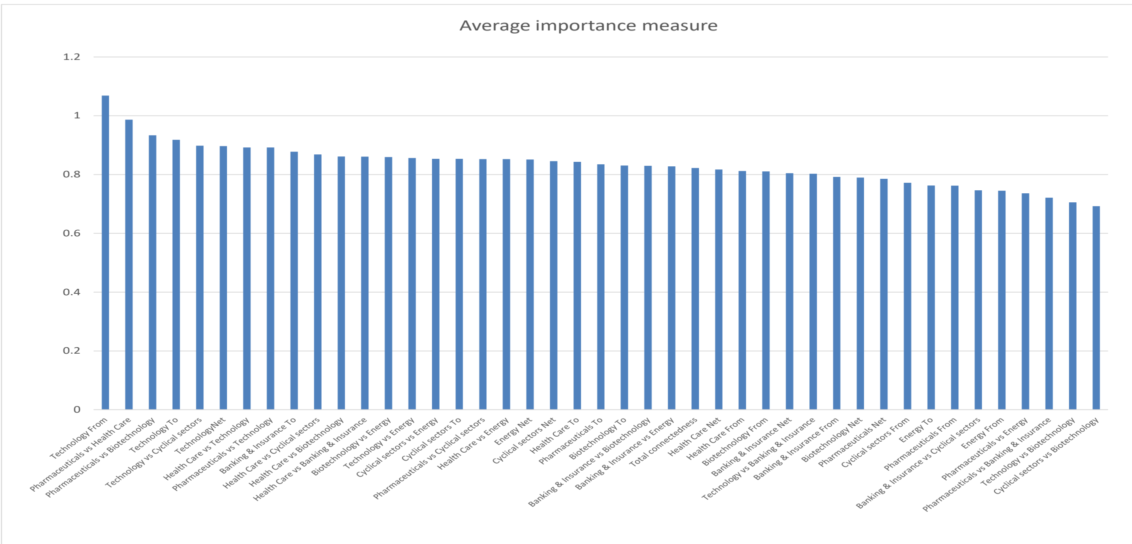


Figure 6. Average of the variable importance (IV) obtained from the prediction errors of the random forest model fitted to the data. The model is trained using data from July 20, 2003 through December 19, 2017. The remaining data constitutes the out-of-sample period and cover the period January 1, 2018 through December 31, 2020. The VI measure is computed for every tree and the average over the entire ensemble is divided by the corresponding sample standard deviation. The random forest is comprised by 200 decision trees.