Stock Returns and Roughness Extreme Variations: A New Model for Monitoring 2008 Market Crash and 2015 Flash Crash

Abootaleb Shirvani

1Texas Tech University, Department of Mathematics & Statistics, Lubbock TX 79409-1042, U.S.A.

Correspondence: Abootaleb Shirvani, Texas Tech University, Department of Mathematics & Statistics, Lubbock TX 79409-1042, U.S.A.

Received: April 2, 2020       Accepted: April 17, 2020       Available online: April 23, 2020
doi:10.1111/aef.v7i3.4824
URL: https://doi.org/10.1111/aef.v7i3.4824

Abstract
We use Student’s t-copula to study the extreme variations in the bivariate kinematic time series of log–return and log–roughness of the S&P 500 index during two market crashes, the financial crisis in 2008 and the flash crash on Monday August 24, 2015. The stable and small values of the tail dependence index observed for some months preceding the market crash of 2008 indicate that the joint distribution of daily return and roughness was close to a normal one. The volatility of the tail and degree of freedom indices as determined by Student’s t-copula falls down substantially after the stock market crash of 2008. The number of degrees of freedom in the empirically observed distributions falls while the tail coefficient of the copula increases, indicating the long memory effect of the market crash of 2008. A significant change in the tail and degree of freedom indices associated with the intraday price of S&P 500 index is observed before, during, and after the flash crash on August 24, 2015. The long memory effect of the stock market flash crash of August 2015 is indicated by the number of degrees of freedom in the empirically observed distributions fall while the tail coefficient of the joint distribution increases after the flash crash. The small and stable value of degrees of freedom preceding the flash crash provides evidence that the joint distribution for intraday data of return and roughness is heavy-tailed. Time-varying long-range dependence in mean and volatility as well as the Chow and Bai-Perron tests indicate non-stability of the stock market in this period.

Keywords: market crash, Copula model, tail dependence

1. Introduction
After 2008, stock market crashes became one of the most fascinating subjects in finance. The financial crisis, which lasted from September 2008 through the end of 2009, was the most severe recession since the Great Depression of the 1930s. It was a dramatic shock to the stock and housing market. The Case-Shiller home price index dropped by about 20%, and the S&P 500 index fell by about 40%. The unemployment rate rose more than 5%, and house prices fell by 16%. About four years after crisis, the S&P 500 index returned back to its 2007 levels.

Academic researchers have extensively discussed stock market crashes to identify a general framework to explain the bubble behavior of a stock price, to estimate the probable crash time, and to detect speculative stock market bubbles (Sornette, 2003; Johansen et al. 2000; Vosvrda et al. 2009; Pelea and Marinescu, 2012; and Friedman and Abraham, 2007). Why it is useful to identify, model, and analyze the stock market crashes are crucial questions for stock and risk managers. Some researchers believe that identifying a crash may help regulators and investors to distinguish warning signals according to their scale level (Abourak, 2015).

Generally, it is extremely difficult to create a statistical model for stock market crashes. One way to create a model for a stock market crash is to associate it with extreme negative returns. In this method, the researcher addressed the relationship between the extreme negative return (i.e., financial crisis) and heavy-tailed distribution.

Price movements typically follow most statistical models of the stock market. The point we would like to study is the behavior of price movement and price velocity during stock market crashes. We attempt to measure and analyze the dependence of extreme variations of price movement (return) and price velocity (roughness). Studying the behavior and the dependence of extreme variations of return and roughness (i.e., the difference of log-return) of the daily stock market is a novel way to explain and evaluate market risk before and during the potential distressed market period. We study the extreme variations in the bivariate kinematic time series of log–return and log–roughness of the S&P 500 stock index.
First, we study the extreme variations in the bivariate kinematic time series for 11 years of daily price log-return and log-roughness of the S&P 500 index (2007–2018) by using Student’s t copula. Similarly, to many financial time series, such as the daily stock returns and foreign exchange rates, the probability distributions of the daily price log-return and log-roughness for the S&P 500 index exhibit a large skewness relative to that of a normal distribution assumed in the (Black-Scholes-Merton model of option pricing; (Black and Scholes, 1973; Merton, 1973). A slow decay of correlations, as revealed by a “heavy-tail” (Rachev and Mittnik, 2000; Rachev, 2003; Shirvani et al., 2020) in distributions, is traditionally attributed to the effect of long memory in the time series. After the stock market crash of 2008, the number of degrees of freedom in the empirically observed distributions (fitted by the Student’s t-copula) fell while the tail coefficient of the copula increased, indicating the long memory effect of the crash. In general, the volatility of the tail degree of freedom values in the distributions associated with the S&P 500 index have declined substantially after 2008.

August 21, 2015, was the beginning of the bear markets and the end of one of the longest bull markets in history (Li et al., 2015). On Monday, August 24, 2015 the S&P 500 index fell 4.99% in the first minutes of trading. On Tuesday, August 25, the stock market experienced another day of sharp losses. Time-varying long-range dependence (LRD) in mean and volatility of stock returns during these shocks to market is studied by applying fractional time series models. The results show a significant change in the estimated values for LRD before and during the market shocks. After the sharp fall in price and when the market is recovering back the losses, the estimated values for LRD slightly change with low volatility. The Chow test is applied to study the presence of a structural break in the LRD in the last days of August 2015. The Chow test confirms the non-stability of trading strategy in the stock market during this period. Within one hour after observing structural break, the mean of the maximum losses for intraday traders is 1%.

Secondly, we report the correlation patterns observed in the tail dependence and degree of freedom indices in the bivariate kinematic time series (i.e., return and roughness) of the S&P 500 index from August 21, 2015, to August 24, 2015. On August 21, the degree of freedom index provides evidence of non-normality distribution for intraday data of return and roughness. After a sharp drop (from 14 to 7) in the first minutes of trading on August 24, the degree of freedom smoothly changes in the interval (8,9), indicating heavy-tailed distribution. The smooth changes in the degree of freedom after the flash crash coincide with the period during which the market is gaining back the losses. On August 21, a rapid rise in the tail dependence index represents the beginning of the bear markets. Smooth changes in tail dependence of return and roughness after the flash crash on August 24 indicates the long memory effect of the crash on the stock market.

This paper is organized as follows. The next section reviews the related literature in detail. Section 3 describes the data source and data validation. In Section 4, we describe our methodology for modeling the joint co-movement of the variables. We use econometrics models to check and filter the linear and the nonlinear time dependence within the individual time series. In Sections 5 and 6, we obtain four indices for bivariate time series to study the behavior of the S&P 500 index before and during the market crash. The time-varying long-range dependence in mean and volatility of stock return is studied in these sections. Section 7 summarizes our findings.

2. Literature Review

The financial crisis in the US began in 2007 and affected financial institutions in many countries. Generally, modeling and analyzing the financial crisis is crucial for stock and risk managers. The mathematical theory of financial crisis emerged after Black Monday in 1987. Johansen and Sornette (1998) studied the financial crisis in 1987 using a nonlinear log-periodic formula which is fitted to log-prices and model the stock market price fluctuations. They stated that the minimum conditions for a model of stock market return fluctuations should include time asymmetry, robustness concerning connectivity between agents, “bounded rationality,” and a probabilistic description. Feigenbaum (2001) reviewed the first differences of the logarithm of the S&P 500 (log-return) before the October 1987 crash and found the log-periodic component of this time series is not statistically significant if we exclude the last year of data before the crash. In their research, they explained the frequency distribution of drawdowns in the S&P 500 before the 1987 crash.

Prior to the financial crisis in 2008, the normality assumption for market returns was made to model the dynamics of asset returns despite the preponderance of empirical evidence that rejects normal distribution. Giacometti et al. (2007) showed that the crisis period exhibited asymmetry and heavy tails with a significant deviation from the normal distribution. Heavy tail distributions and extreme value theory vastly emerged in literature for modeling and measuring extreme risks and returns after the 2008 financial crisis. Champagnat et al. (2013) demonstrated the extreme values should be modeled by a correct heavy-tail distribution to have an accurate risk measure estimator for significant losses. Champagnat et al. (2013) studied the performance Value-at-Risk estimator using heavy tail distribution to get better understanding of extreme negative returns.
Vosvrda et al. (2009) addressed the relationship between the extreme negative return (i.e., financial crisis) and heavy-tailed distribution. They studied the differences in the behavior of the Central European stock market indices before and during the crisis using heavy-tailed distributions. The results indicated that the stock market distribution behaved more normally before the crisis period, and the heavy tails distribution better described the stock market distribution during the crisis.

The recent approach to studying the financial crisis is to measure the dependency between stock markets using copula functions. Copula functions allow us to model the dependence structure between random variables when the marginal distributions are known. The copula functions are very useful, especially when the normality assumption does not hold. Modeling the dependence structure between random variables using the copula method in finance is relatively new (Breymann et al., 2003; Silva et al., 2014; Lu et al., 2017; Siami-Namini et al., 2019; and Zhang and Jiang, 2019). Doman and Doman (2012) studied the dependence structure of stock markets during the pre- and post-crisis periods by applying Markov-switching copula model. Using tail copula dependence coefficient, they observed the lower tail dependence during and dependence in both tails after the 2008 crisis.

Although dependence structure between financial markets is widely discussed in the literature, we attempt to obtain a new look on the financial crisis by studying the extreme variation in the bivariate kinematic time series of log-return and log-roughness of the S&P 500 index. While most of the studies in the financial crisis measure the dependence structure between different indices, this novel approach allows us to analyze the dependence structure of an individual index basing on the co-movement of return and roughness. We captured the dependence structure by non-Gaussian copula functions, as the linear correlation for measuring the dependence is practically insignificant. The proposed method in this paper provides important insights for risk managers and financial planners, allowing them to determine the presence of a large dislocation in the market from its equilibrium state and the acute phases of the crisis and use them as a warning signal to distinguish a market crash.

Finally, we note there is other theoretical literature that analyzes and predicts the market crash in different ways. From the perspective of behavior finance, Harmon et al. (2015) used a collective panic measure to analyze economic market crashes. Using rational finance theory, Shirvani et al. (2019) defined a new set of derivatives together with a risk measure, which they referred to as the tail-loss ratio (TLR), to explain and evaluate market risk before and during the potential distressed market period.

3. Data Source, Data Validation and Preparation

The analysis of the kinematic time series associated with the S&P 500 index has been performed based on the publicly available data collected during 1,962 trading days in the period from January 3, 2003 till October 19, 2018. In our study, we used the daily closing price of the S&P 500 index for the 2008 financial crisis. The data set was collected from Yahoo Finance. Fig. 1 shows the price, daily log return, and daily log-roughness processes. The S&P 500 index is used to measure the performance of the large-cap U.S. stock market. For the year 2008, S&P 500 index fell 38.49%, its worst yearly percentage loss.
A flash crash refers to a very rapid decline in security prices occurring within an extremely short period. A flash crash in the US equities (7% drop) happened in the first minutes of August 24, 2015. To analyze the flash crash behavior on August 24, 2015, we used the intraday historical returns for the S&P 500 index from August 19, 2015 to August 28, 2015, collected from the Bloomberg Financial Markets. The midpoint of ask and bid prices was used in our analysis.

4. Methods

In our work, we used both the logarithmic return and roughness quantities for the different types of analysis as follows

\[ r(t) = \ln \frac{S(t)}{S(t-1)} , \]

and the index log–roughness,

\[ \dot{r}(t) = r(t+1) - r(t) = \ln \frac{S(t+1)}{S(t)} , \]

where \( S(t) \) is the S&P 500 index value at day \( t \).

It is important to mention that the relationship between mean logarithmic and mean simple returns is not one-to-one. The logarithmic mean returns are less than the simple mean return by an amount related to the variance of the series.

4.1 The Bivariate Student’s \( t \)-Copula

Many financial time series, such as the daily return on stocks and foreign exchange rates, exhibit fat-tails that can be attributed to the probability of large co-movement of the individuals, i.e., tail dependence. The reliable assessment of investment risk requires the study of tail dependence by using copulas. A copula is a multivariate probability distribution for which the marginal-probability distribution of each variable is uniform (Nelsen, 1999).

In line with Sklar’s Theorem (Nelsen, 1999), every cumulative bivariate distribution \( F \) with marginal distributions \( F_1 \) and \( F_2 \) can be written as

\[ F( x_1, x_2) = C(F_1( x_1), F_2( x_2)), \]

for some copula \( C \), which is uniquely determined on the interval \([0,1]^2\). Conversely, any copula \( C \) may be used to join any pair of univariate distributions \( F_1 \) and \( F_2 \) to create a joint cumulative distribution \( F \), viz., \( C(u) = F(F_1^{-1}(u_1),F_2^{-1}(u_2)) \), where \( F_1^{-1} \) and \( F_2^{-1} \) are the quantile functions of the respective marginal distributions. The copula \( C \) can be thought of as the distribution function of the component-wise probability transformed random vector \( (F_1(x_1), F_2(x_2)) \).

Copula functions allow us to model the dependence structure independently of marginal distributions (Clemente and Romano, 2004). Because both marginal distributions are a fat tail, the Student’s \( t \) distribution has been found to provide a reasonable fit to the univariate distribution of the return and roughness of stocks. The bivariate Student’s \( t \)-copula (Demart and McNeil, 2005) is defined by

\[ C_{\rho} (u_1, u_2) = \int_{-\infty}^{u_2} \int_{-\infty}^{u_1} \frac{1}{2\pi \sqrt{1-\rho^2}} \left( 1 + \frac{x^2+y^2-2\rho xy}{\nu(1-\rho^2)} \right)^{-\nu/2} dxdy \]
in which $\rho$ is the correlation coefficient; $r^{-1}$ is the quantile function of the standard univariate $t$-distribution, with $\nu$ degrees of freedom, mean of zero, and variance $v/(\nu - 2)$.

The correlation coefficient $\rho$ in (4) is estimated by using the non-parametric rank correlations between the temporal samples of the kinematic time series taken over the $t$-days window, $\hat{r}_t$ and $\hat{r}_j$, measured by the Kendall’s tau correlation coefficient (Kendall, 1938), viz.,

$$
\bar{K}_t(r_t, \hat{r}_t) = \frac{2}{n(n-1)} \sum_{i<j} sgn((r_{t,i} - r_{t,j})(\hat{r}_{t,i} - \hat{r}_{t,j})),
$$

(5)

The interpretation of Kendall’s tau correlation coefficient in terms of the probabilities of observing the agreeable and non-agreeable pairs of variables is very direct. The relationship between Kendall’s tau and the linear correlation coefficient $\rho_t$ is defined (Zeevi and Mashal, 2002) by

$$
\rho_t = \sin \left( \frac{\pi}{2} \bar{K}_t \right).
$$

(6)

4.2 Applying the ARMA – GARCH Model to Fitting the Bivariate Kinematic Time Series

The bivariate kinematic time series for the daily log–return of the S&P 500 index were tested by the Augmented Dickey–Fuller (Fuller, 1976) and KPSS tests (Kwiatkowski et al., 1992) to assess whether the series are unit root non-stationary for statistical associations based on the ranks of the data. The KPSS test to assess whether the series are unit root nonstationary Ranking data was carried out on the variables that are separately put in order and are numbered. The null hypothesis of the Augmented Dickey-Fuller test that a unit root is present in the time series sample was rejected. The p-value ($p = 0.002$) for KPSS test is about zero, leading to very strong deviance against the null hypothesis that the series is trend stationary. The assumption of whether the data under study have the skewness and kurtosis matching a normal distribution was checked by the Jarque–Bera goodness-of-fit test (Jarque and Bera, 1980) and was also rejected. In order to estimate and model dependence, it is necessary to check dependence within individual time series. The standard Ljung–Box test (Ljung and Box, 1978) was used for the lags of 5 and 10 days to test if the residuals of the model have autocorrelation. The test p-values indicate the 5%–significance of linear dependence and heteroscedasticity of variables.

Recently, financial econometrics is widely used to model the dynamic evolution of various financial instruments (Shirvani and Volchenkov, 2019; and Siami-Namini, 2019; Trindade et al., 2020). It is also used to filter out the dependencies in the time series data. Here, we use the generalized Autoregressive Moving Average (ARMA) method and the Autoregressive Conditional Heteroscedasticity (GARCH) model to filter out the linear and nonlinear temporal dependencies in the bivariate kinematic time series (Hamilton, 1994; and Engle (1982)).

The ARMA (1,1) – GARCH (1,1), model is a standard tool for modeling the volatility of returns (GARCH–model) and the conditional mean of the investigated time series (ARMA–model) by obtaining independent residual sets from the financial time series, viz.,

$$
\begin{align*}
\hat{r}_t &= \mu + \varphi (r_{t-1} - \mu) + \theta a_{t-1} + a_t \\
\epsilon_t &= \epsilon_{t-1} \alpha, \quad \epsilon_t = iid, \\
\sigma_t^2 &= \gamma + \alpha a_{t-1}^2 + \beta \sigma_{t-1}^2
\end{align*}
$$

(7)

where $r_t$ is the return value during the time period $t$; $\mu_t = E(r_t | F_{t-1})$ is the conditional mean return of the stock during the time period $t$; $\sigma_t = \text{Var}(r_t | F_{t-1})$ is the conditional variance of stock return during the time period $t$; $F_{t-1}$ denotes the information set consisting of all linear functions of the past returns available during the time period $t - 1$; $\epsilon_t$ is the standardized residual during the time period $t$, which we assume to be the iid–random variable with the standard Student’s $t$-distribution with zero mean and unit variance; $\alpha_t$ is referred to the shock of return during the time period $t$; $\alpha \geq 0, \beta \geq 0, \gamma \geq 0, \delta, \phi, \theta$ are the constant parameters of the model estimated by the rolling method.

Moving the rolling window by one observation, we have estimated the model parameters and residual sets and fitted the obtained return and roughness residual sets, $\langle \hat{r}_t, \hat{r}_j \rangle$, to the bivariate Student’s $t$-distribution. The presence of volatility clusters as well as significant autocorrelation in the vector of squared residuals would indicate conditional heteroscedasticity in the data.
4.3 Degree of Freedom (ν) and Tail Dependence (λ) Indices

The value of the degree of freedom index ν over the τ-days window was numerically estimated as the maximum of the likelihood function, viz.,

\[ \hat{\nu}_t = \arg \max \{ \sum_{i=1}^{n} \ln S(u_1, u_2, \nu, \rho_2) \} \]  \hspace{1cm} (8)

in which the density of Student’s t-copula,

\[ S(u_1, u_2, \nu, \rho_2) = \frac{\frac{\nu+2}{\nu} \left( \frac{1}{2} \right)^{\frac{\nu+2}{2}} \left( 1 + \frac{u_1^2 + u_2^2 - 2\rho u_1 u_2}{\nu-2} \right)^{-\frac{\nu+2}{2}}}{\sqrt{\pi} \left( \frac{\nu+2}{\nu} \right)^{\frac{\nu+1}{2}} \left( 1 + \frac{\rho}{\nu-2} \right) } \]  \hspace{1cm} (9)

was calculated for the pseudo-sample \((u_1, u_2)\) by

\[ (x, y) = \left( t^{-1}(u_1), t^{-1}(u_2) \right) \]  \hspace{1cm} (10)

The tail-dependence coefficient provides asymptotic measures of the dependence in the tails of the bivariate distribution of \((X_1, X_2)\). The coefficient of upper (lower) tail dependence of \(X_1\) and \(X_2\) with marginal Student-t distribution is (Demarta and McNeil, 2004)

\[ \lim_{q \to 1} P(x_1 > t^{-1}(q)|x_2 > t^{-1}(q)) = \lambda_u, \quad \lim_{q \to 0} P(x_1 < t^{-1}(q)|x_2 < t^{-1}(q)) = \lambda_l \]  \hspace{1cm} (11)

provided a limit \(\lambda_u (\lambda_l) \in [0,1]\). As Student-t is an elliptically symmetric distribution, the lower tail dependence of \(X_1\) and \(X_2\) is coincided with upper tail. Tail-dependence coefficient is limiting condition probability that both marginal distributions exceed a specific quantile given the first marginal distribution does exceed that specific quantile. This measure depends only on marginal distributions and thus on the copula \(C\) of \((X_1, X_2)\). Therefore, using (9), a simple formula for tail-dependence coefficient, obtained by Frahm et al. (2005), over the τ-days window is

\[ \lambda_\tau(v) = 2t_{\nu+1} \left( -\sqrt{\nu_\tau + 1} \frac{\sqrt{1-\rho_\tau}}{\sqrt{1+\rho_\tau}} \right) \]  \hspace{1cm} (12)

5. Numerical Analysis of 2008 Financial Crisis

5.1 Long Range Dependence in the Daily Time Series

To capture long memory (in essence hyperbolic memory) or LRD and to see where shocks in the market decay at a slower hyperbolic rate in the conditional mean of returns, we use fractionally integrated generalized autoregressive moving average model (FIARMA). The fractionally integrated autoregressive conditional heteroscedasticity (FIGARCH) model is used to capture LRD in time varying returns volatility.\(^1\) We test the fractional ARMA and fractional GARCH models, ARFIMA \((1, d, 1)\)–GARCH \((1, 1)\) and ARMA \((1, 1)\)–FIGARCH \((1, d, 1)\) (see Ling and Li, 1997; and Engle (1982)), implemented in the \(R\)–package rugarch (Ghalanos, 2018).

We filter out the linear and nonlinear temporal dependencies in times series data by applying a time series model and obtain the sample innovations. To examine the presence of long memory in the data set, instead of studying returns data, we consider their sample innovations obtained from the ARMA \((1,1)\)-GARCH \((1,1)\) with normal assumption for the innovation distribution.

\(^1\) The fractionally integrated generalized autoregressive conditional heteroscedasticity, or FIGARCH, model to capture long memory in terms of the variance is proposed by Baillie et al. (1996).
We report the results of FIARMA and fractional FIGARCH models to examine the long memory for daily data by using a rolling estimation methodology. These recursive estimated choice exercises are repeated on the following day accumulated in rolling four years periods, using data from January 4, 2000, and up to October 19, 2018. This repeating estimation methodology is used to create meaningful values and understand the stability of the estimated values of $d$. We note that $d$ is the power of the long memory fractional process, $(1 - L)^d$, and is equivalent to the Hurst Exponent $H - 0.5$.

The estimated values for $d$ in FARIMA (1, d,1)–GARCH (1,1) and ARMA (1,1)–FIGARCH (1, d,1) models for daily stock returns are shown in Figures 2 and 3, respectively. Given the estimated sequences of $d$ ranging in $(0,0.025)$ during the 2008 market crash, the FIARMA model indicates no long–memory dependence for this period. However, the FIGARCH model demonstrates long–memory dependence in the variance of the S&P 500 returns during the market crash. The estimated values for $d$ in the FIGARCH model vary in the interval $(0.7,0.9)$, confirming the long memory in the variance of the S&P 500 returns. Thus, the results show that shocks decay at an exponential rate in the mean model, and the ARFIMA (1, d,1)–GARCH (1,1) collapses to the vanilla ARMA (1,1)–GARCH (1,1). However, the FIGARCH model indicates the shocks decay more slowly than an exponential decay in modeling the volatility of the S&P 500 index.
5.2 The Tail and Degree of Freedom Indices in the Daily Bivariate Kinematic Time Series

We report on the correlation patterns observed in the daily tail dependence and daily degree of freedom indices in the bivariate kinematic time series of the S&P 500 index (2007-2018). We emphasize that $\lambda_t$ as a scalar measure of the tail dependence cannot characterize the entire dependent structure in the tail of the distribution.

The tail dependence index (12) and the degree of freedom index (8) calculated for the daily bivariate kinematic time series (return and roughness) associated with the S&P 500 index are shown in Fig. 4.a and Fig. 4.b, respectively.

![Figure 4. (a) Daily tail dependence index, and (b) Daily degree of freedom index corresponding to the bivariate kinematic time series for the S&P 500 index from January 3, 2007 to October 19, 2018.](image)

As we observe from Figures 4.a and 4.b, the daily tail dependence and daily degree of freedom indices in the bivariate kinematic time series are divided into two periods: before and after the 2008 crash. The stable and small values ($0.1 \leq \lambda \leq 0.35$) of the tail dependence index (12) observed for some months preceding the market crash of 2008 indicate that the joint distribution of daily return and roughness was close to a normal distribution. The values of degrees of freedom index (8) were between 30 and 50, providing more evidence that the daily return and roughness are jointly normally distributed in that period. In this period, the tail dependency significantly overvalued, and it was due for a correction in the market that ultimately took place in 2008.

From Figures 4.a and 4.b, we observe a large and generally rapid rise in the values of the tail dependence index (from 0.22 to 0.6) and a significant and generally rapid decline in the value of degrees of freedom index for some days preceding the market crash of 2008. It is a drop and rise of at least 10 percent in the values of tail and degree of freedom indices, respectively.

5.3 Discussion of the Daily Bivariate Kinematic Time Series

After the stock market crash of 2008, the number of degrees of freedom in the empirically observed distributions (fitted by the Student’s $t$-copula) fell while the tail coefficient of the copula increased, featuring the long memory effect of the crash. The fact is that the bivariate distribution of return and roughness is a heavy-tailed distribution, and it cannot be a normal distribution for a long time and tends to revert back to its nature, i.e., heavy tail distribution. Presently, the observed bivariate distributions of return and roughness for the S&P 500 index clearly exhibit a heavy tail.
degree of volatility in the values of both indices is lower in the past eight years than in the years preceding the financial crisis.

Thus, we conclude that there is a significant difference between the tails and probability distributions of the returns and roughness before, during, and after the financial crisis. Thus, the crisis period has a small deviation from the normal distribution, while the pre-crisis and post-period is characterized by a significant deviation from normality in the distribution of return and roughness.

6. Numerical Analysis of 2015 Flash Crash

6.1 Long Range Dependence in the Tick Kinematic Time Series

A flash crash refers to very rapid decline in security prices occurring within an extremely short period. Friday, August 21, 2015, was the beginning of the bear markets and the end of one of the longest bull markets in history (Li et al., 2015). The S&P 500 index fell 3.9%, and the Dow Jones industrial average (DJIA) dropped 3.1%.

On Sunday night, August 23, a significant drop in equities in Asia caused a (7%) reduction in the U.S as the traders in U.S. markets wanted to sell their stocks. On Monday, August 24, 2015, sell orders overwhelmed in the first five minutes of trading, and the S&P 500 index went down from 1965.15 to 1867.01 within minutes at opening time. Although the market gained back most of the loss on August 24, the closing price was 3.66% below the opening price. These drops also happened in commodities such as oil, copper, and most of the Asian currencies (see Mitchell, 2019).

Stock market performance on Tuesday, August 25, 2015, was another day of sharp losses. The DJIA index rose 442 points in early trading but plunged again in the last hour of trading, leaving the DJIA another 204 points off from its opening level (see Egan, 2015). On Wednesday, August 26, 2015, the S&P 500 index and DJIA recovered with about 4% gains. The S&P 500 index finished up August 27, 2015, with 2% raise.

Here, we present the results of FIARMA and FIGARCH models to examine the long memory for intraday data. The data used in this Section are the intraday historical returns for the S&P 500 index from August 19, 2015, to August 28, 2015, collected from the Bloomberg Financial Markets. In our empirical analysis, we used the midpoint of ask and bid prices in our computations.

First, we adopt the ARMA (1,1)-GARCH (1,1) model to return data to obtain the sample innovations. Again, we use sample innovations to examine the presence of long memory in intraday data set by using rolling estimation methods. We note that our rolling estimation method is repeated with rolling windows of 250 observations.

The time-varying long-range dependence estimated values (d) for the FIGARCH model are plotted in Figure 5. On August 21, the estimated values of d are ranging between (0.40,0.95). The Figure 5.a shows how the long memory property carries over to the stock volatility of the underlying process on August 21. In the last hours of trading, as the price of the S&P 500 index fell, the values of d increase from 0.3 to 1. It causes a change in the shape of the distribution of the underlying process. It indicates that the feature of the return process is like a unit root and non-stationary process during the last hour of August 21.

Figure 5.b shows a significant change in the estimated values of d in the first minutes of trading on August 24. We observe the values of d sharply drop in the first minutes of trading, during the flash crash, as sell orders overwhelmed, then steeply rises to one. After the rising period for d, the market is recovering and gaining back the losses. The values of d vary slightly in the interval (0.97,1), indicating a unit root and non-stationary process that is an abnormal behavior of the return process as it should be stationary.

The low volatility of estimated values for d after significant falling near 1:00 pm on August 25 represents the market’s recovery and the presence of long memory. In general, the results reveal the long-range dependence volatility in the intraday data of the stock market before 3:00 pm on August 21 and after 1:00 pm on August 25 as d varies slightly in the interval (0.38,0.64). However, on August 24 and 25, the return time series behave as a non-stationary process with the presence of unit root.
Figure 5. The estimated values of $d$ in ARMA (1,1)–FIGARCH (1, d,1) model corresponding to intraday returns of the S&P 500 index at the midpoint of ask and bid prices. (a) the estimated values for $d$ for intraday observed return on 08/21/2015; (b) the estimated values for $d$ for intraday observed return on 08/24/2015; (c) the estimated values for $d$ for intraday observed return on 08/25/2015; (d) the estimated values for $d$ for intraday observed return on 08/26/2015; (e) the estimated values for $d$ for intraday observed return on 08/27/2015; and (f) the estimated values of $d$ for intraday observed return on 08/28/2015.

Figure 6 shows the time-varying long-range dependence in the mean of data set by applying the FIARMA model. The estimated values of $d$ vary in the interval $[0,0.52)$ indicating short-range dependence in the mean return of the S&P 500 index. The fluctuations of $d$ are divided into five sub-periods, before 3:00 pm on August 21 called period I, August 21, 3:00 pm to August 24, 10:00 am, August 24 period II, 10:00 am to August 25, 01:00 pm period III, August 25, 3:00 pm to August 27, 01:00 pm, and then after, period V.

The estimated $d$ parameters are varying in $[0,1,0.3)$ indicating the presence of LRD in the mean of return process on August 21. A sharp decrease, near 03:00 pm from 0.3 to zero, reveals an abnormal behavior in the stock return process as it was the beginning of the bear market. After a few minutes, a significant and general rapid rise in the value of $d$, starting from 0 on August 21 and ending up to 0.5 at 10:00 am, is observed. This steep rise referred to a signification decline in the price of S&P 500 index in the last hours of trading on August 21 and the first minutes of trading on August 24. The estimated values of $d$ varying in $[0.45,0.51)$ features the long memory effect of the flash crash in the period II. The volatility of the estimated values of $d$ indicates that the market is recovering the losses.
As the market experienced another shock on Tuesday, August 25, we observe a significant falling near 1:00 pm from 0.5 to zero in the value of $d$, demonstrating a substantial change in the market behavior. The values of $d$ varying slightly between 0 near 1:10 pm August 25, to 0.1 near 01:00 pm on August 27. It represents a convalescent period of the market. On August 27, the market tends to revert back to its nature, i.e., the fact that the return process for high-frequency data in the short term has LRD in the mean.

To examine the stability of estimated $d$ in the FARIMA and FIGARCH models from August 21 to August 28, we apply the Bai-Perron\(^2\) and Chow tests.\(^3\) To check the presence of breaks in the estimated value for $d$ in the FARIMA and FIGARCH models, we performed the Bai-Perron test on data set using the R-package “\textit{strucchange}” (Zeileis et al., 2003). Figures 7.a and 7.b exhibit the breakpoints in regression models with structural changes result. The dotted vertical lines indicated the breakpoints. As we observed, the test shows strong evidence for breaks in estimated values of $d$ in both models. In particular, plots 8.a and 8.b shows the residual sum of squares (RSS) and Bayesian information criteria (BIC) as the number of breakpoints is varied between zero and five. The plots show the decrease in Bayesian information criteria (BIC) and the residual sum of square as we increase the number of breakpoints.

\(^2\) Bai et al. (2003) proposed a method to construct confidence intervals for the break dates, test statistics and model selection procedures using a dynamic programming algorithm based on the Bellman principle.

\(^3\) The Chow test, proposed by Chow (1960), evaluate the structural break by testing the null hypothesis that the true coefficients of two regression model are the same. This test assesses the presence of an instantaneous change over time in the parameters of regression models.
We use the rolling method to assess the structural break behavior in the estimated values of $d$ using the Chow test. We select the first hour as in-sample data and move the window by one observation and repeat the Chow test. For each window, we split the sample data into two sub-periods, the first and second thirty minutes observation. After estimating the parameters for each of the sub-period, we test the null hypothesis that the correct coefficients of the two regression models are the same with the help of the F-test (see Hansen, 2001).
The time-varying p-values of the Chow test for FIARMA are plotted in Figure 9. At a 95% confidence level, the percentage of the structural break existence is 70%. As we observe from Figure 9 when the market is recovering the losses, after 10:00 am on August 24, and 01:00 pm on August 25 (after the market experiences shocks), structural breaks are not observed.

Figure 10 significantly exhibits structural breaks in the long-range dependence in FIGARCH models as in 92% of time the p-values are less than 0.05. Thus, it confirms that the Chow test rejects the stability of estimated values for $d$. These results indicate that the financial crash affected the trading partners during the sample period and caused a significant change in trading strategy in the market, particularly on August 24 and 25.

To evaluate the effect of structural breaks on the financial markets, we calculate the maximum drawdown occurs one hour after each break. Maximum drawdown (MDD) is the maximum observed loss from a peak to a trough of a portfolio before a new peak is attained. It is an indicator of downside risk over a specified time (see Hayes, 2019; and Pedersen and Rudholm-Alfvin, 2003). The formula for MDD during a specified period of time [0; T]

$$MDD = \sup_{t \in [0, T]} \left[ \sup_{s \in [0, t]} \left( \frac{S_s - S_t}{S_s} \right) \right].$$

(13)

Figure 10. Box plot of maximum drawdown for S&P 500 index returns after observing structural breaks during the sample period from August 21, 2015 to August 28, 2015.
Figure 11 displays the box plot of MDD (most significant movement from a high point to a low point) after each structural break in long-range dependence in both models (FIARMA and FIGARCH) during the sample period of the flash crash. The maximum value, mean, and standard deviation of MDD in FIARMA are 0.0354, 0.0101, and 0.0071, respectively. These values indicate that the markets continued to fluctuate and experience losses during this period. For the FIGARCH model, the obtained values of the maximum, mean, and standard deviation of MDD are 0.0354, 0.0109, and 0.0076, respectively.

6.2 The Tail and Degree of Freedom Indices in the Intraday Bivariate Kinematic Time Series

In this section, we report on the correlation patterns observed in the intraday tail dependence and degree of freedom indices in the bivariate kinematic time series (return and roughness) of the S&P 500 index from August 21, 2015, to August 24, 2015.

The degree of freedom index (8) of return and roughness associated with the intraday data of S&P index for August 21 and 24 are shown in Figure 12.a and 12.b, respectively. The degree of freedom index on August 21 varies between 12 and 24 with a “decreasing trend”, providing evidence of non-normality distribution for intraday data of return and roughness on August 21.

The volatility of the degree of freedom in the first hours of trading on August 21 is high compared to the last hours of trading when the price of the S&P 500 index fell. Our results for August 24 indicate a swift fall in the degree of freedom index. It is observed that the degree of freedom sharply drops from 14 to 7 in the first minutes of trading on August 24, during the flash crash. It is because of an overwhelming of sell orders and rapid changes in the price process from 9:30 to 10:00 am. After a sharp decrease, near 10:00 am, the degree of freedom slightly varies in the interval (8,9) providing evidence that the joint distribution of daily return and roughness is heavy-tailed. During this period, the market is gaining back the losses. Thus, the joint distribution of return and roughness is a heavy tail Student-t distribution with the degree of freedom less than 10.

To see the tail behavior and tail dependence of return and roughness, we calculate the tail index by using (12). Figure 13.a shows the variation of the tail dependence index on August 21. A significant and generally rapid rise in the value of tail index near 1:15 pm on August 21 represents a change in the behavior of the price process of the market. On August 21, the S&P 500 index fell 3.9%, and it was the end of one of the longest bull markets in history. The volatility of tail dependence index falls within the last hours of trading compared to the market’s opening hours.

Figure 11. Intraday time-varying estimated values of degree of freedom index corresponding to bivariate kinematic time series for the S&P500 index on (a) August 21, 2015 and (b) August 24, 2015

To see the tail behavior and tail dependence of return and roughness, we calculate the tail index by using (12). Figure 13.a shows the variation of the tail dependence index on August 21. A significant and generally rapid rise in the value of tail index near 1:15 pm on August 21 represents a change in the behavior of the price process of the market. On August 21, the S&P 500 index fell 3.9%, and it was the end of one of the longest bull markets in history. The volatility of tail dependence index falls within the last hours of trading compared to the market’s opening hours.
A rapid rise in the value of the tail dependence index (from 0.28 to 0.52) is observed in the first minutes of trading on August 24, featuring the flash crash. There is a significant difference between the joint probability distributions of the returns of roughness before and after the financial flash crash. After 9:00 am on August 24, the joint distribution is characterized by a heavy tail distribution and significant deviation from normality in the joint distribution of return and roughness with a stable degree of freedom ranging in (8.0, 9.5), while the apparent volatility of the degree of freedom is higher in the day before the flash crash.

6.3 Discussion of the Intraday Bivariate Kinematic Time Series

Thus, we observe that there is a significant difference between the degree of freedom, tail dependence, and the joint distribution of the return and roughness before, during, and after the August 24 flash crash. Before the crisis period, the joint distribution behaves as a semi-heavy tail distribution. The joint distribution has Student-t with 14 degrees of freedom at the first minutes of trading, and after 30 minutes, it changes to a very heavy tail Student-t distribution with 7 degrees of freedom. After the crisis period, the joint distribution is characterized by a significant deviation from normality with significantly overvalued tail dependency. The FIARMA and FIGARCH models indicate the long memory for intraday data during the sample period of the flash crash.

7. Conclusion

We have studied the bivariate kinematic time series for the 11 years of daily price return and roughness for the S&P 500 index and found that the probability distributions exhibit a significant negative skewness relative to that of a normal distribution that might be attributed to the effect of long memory in the time series. The long memory effect of the stock market crash of 2008 is featured by that the number of degrees of freedom in the empirically observed distributions fell while the tail coefficient of the joint distribution increased after the crash, indicating the presence of a heavy-tail in distributions traditionally attributed to the effect of long memory in the time series. These results significantly feature the long memory effect of the crash.

We reported on the correlation patterns observed in the intraday tail dependence and degree of freedom indices in the bivariate kinematic time series (return and roughness) of the S&P 500 index from August 21, 2015, to August 24, 2015. The volatility of the degree of freedom index in the Student’s t-distribution fitting the intraday bivariate kinematic time series (return and roughness) associated with the intraday price of the S&P 500 index fell down substantially during the stock market flash crash of August 2015 with low fluctuation after it. However, the tail coefficient of the joint distribution increased after the flash crash. The results indicated a significant difference between the joint probability distributions of the returns and roughness before (semi-heavy tail distribution), after (heavy tail distribution) the financial flash crash. Time-varying long-range dependence in mean and volatility significantly change before and during the flash crash. The stable value of long-range dependence after each drop in the price of the S&P 500 index represents the recovering period of the losses. We believe our finding has the important implications for risk managers.
and investors, because of it can help them to determine the presence of large dislocation in the market from its equilibrium stat.

The study examines the dependence structure of return and roughness using stock market data and asymmetric copula method. Research purposes were expanded to meet the primary objective of this study, and results were derived through a proper procedure to reach the conclusions. The author has identified a few limitations, and he has given three suggestions for future researchers to overcome those limitations. The study was conducted for S&P 500 stock price data, so the results can be generalized for other indices, especially for Real Estate indices. Also, we suggest future researchers conduct studies on major world indices too. Secondly, as the crisis period exhibited asymmetry, we recommend reviewing the dependence structure by applying asymmetric copula functions. Also, the researchers can study the three-phase space of the stock market: position, velocity, and acceleration by implementing Mahalanobis distance and structural break.

References


Copyrights
Copyright for this article is retained by the author(s), with first publication rights granted to the journal.
This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.