A statistical analysis of UK financial networks

DOI:
10.1016/j.physa.2016.12.073

A statistical analysis of UK financial networks

by

J. Chu and S. Nadarajah
School of Mathematics, University of Manchester, Manchester, UK

Abstract: In recent years, with a growing interest in big or large datasets, there has been a rise in the application of large graphs and networks to financial big data. Much of this research has focused on the construction and analysis of the network structure of stock markets, based on the relationships between stock prices. Motivated by Boginski et al. Boginski et al. (2005), who studied the characteristics of a network structure of the US stock market, we construct network graphs of the UK stock market using same method. We fit four distributions to the degree density of the vertices from these graphs, the Pareto I, Fréchet, lognormal, and generalised Pareto distributions, and assess the goodness of fit. Our results show that the degree density of the complements of the market graphs, constructed using a negative threshold value close to zero, can be fitted well with the Fréchet and lognormal distributions.

Keywords: Degree density; Fréchet distribution; Lognormal distribution; Pareto distribution

1 Introduction

Large datasets have useful applications in many different areas, for example, science, engineering, and computer science, to name just a few. Such large data sets can often be represented in terms of large graphs, comprising vertices (or nodes) connected by edges. In the most simple case, we can consider an undirected graph, say $G(V, E)$, which is defined by the set of vertices V and edges $E \subset V \times V$ connecting pairs of vertices (Boginski et al. Boginski et al. (2005)).

The applications of large graphs or networks has been studied greatly, and also spans across a number of fields. For example, in chemistry, Chou Chou (1990)) applies large graphs to enzyme kinetics and protein folding mechanisms; in ecology, Bunn et al. Bunn et al. (2000)) apply large graphs to landscape and habitat connectivity; in engineering, Dobrjanskyj and Freudenstein Dobrjanskyj and Freudenstein (1967) apply large graphs to the structural analysis of mechanisms. In recent years, there has been a rise in the application of large graphs and networks to large datasets in the area of finance. More specifically, the structure of stock markets, see Kullman et al. Kullmann et al. (2002), Jung et al. WS et al. (2006), Ping and Binghong Ping and Binghong (2006), Zhuang et al. Zhuang et al. (2007), Tabak et al. Tabak et al. (2010), Zhang et al. Zhang et al. (2010) and Vizgunov et al. Vizgunov et al. (2014).

In particular, networks constructed based on the relationships between stock prices have been shown to follow a common distributional model. Kim et al. Kullmann et al. (2002) studied cross-correlations in stock price changes among S&P 500 companies by a weighted random graph. They found that the influence-strength distribution in absolute terms follows a power law. Huang et al. Huang et al. (2009) used a threshold model to construct China's stock correlation network and studied the structural properties. After conducting a statistical analysis on the network, they also

showed that it follows a power law model. Tse et al. Tse et al. (2010) created a complex network of the US stock market and studied the correlations between closing prices of the stocks. Their results found that the network shows a scale free distribution, and variations in the stocks were heavily influenced by a small number of stocks. Boginski et al. Boginski et al. (2005) also studied the characteristics of the graph representing the structure of the US stock market, using a network representation based on cross-correlations of stock price fluctuations.

Let $P_i(t)$ denote the price of stock i, $i = 1, \ldots, N$, on day t, $t = 1, \ldots, T$, where N is the total number of stocks in the sample, and T is the total number of trading days in the time period considered. Also, let $R_i(t) = \ln\left(\frac{P_i(t)}{P_i(t-1)}\right)$ be the log one-period returns of stock i, from time $t-1$ to t. It follows that the correlation coefficient, $C_{i,j} \in [-1, 1]$ for all $i, j = 1, \ldots, N$, can be calculated as:

$$C_{i,j} = \frac{E\left(R_i R_j\right) - E\left(R_i\right) E\left(R_j\right)}{\sqrt{Var\left(R_i\right) Var\left(R_j\right)}}, \tag{1}$$

where $E\left(R_i\right)$ is the average log returns of stock i over T days, $E\left(R_i R_j\right)$ is the average of the product of the log returns of stocks i, j over T days, and $Var\left(R_i\right)$ is the variance of the log returns of stock i over T days (Boginski et al. Boginski et al. (2005)).

Define the market graph and its complement to be graphs with N vertices represented by each of the stocks. Let $\theta \in [-1, 1]$ denote a threshold value, such that an edge connects a pair of stocks (i, j) if $C_{i,j} \geq \theta$ (market), or $C_{i,j} \leq \theta$ (complement). Boginski et al. Boginski et al. (2005) used data for 6546 financial instruments traded on the US stock markets over 500 consecutive trading days in 2000-2002 and constructed graphs using various threshold values θ. They showed that under certain conditions, the degree distribution of the vertices of the market graphs follows a power law model.

This paper is motivated by Boginski et al. Boginski et al. (2005). Whilst Boginski et al. Boginski et al. (2005) consider the degree distribution of network graphs of the US stock market, we construct network graphs of the UK stock market, and fit various distributions to the degree density of the vertices in these graphs. In our analysis, we find that for the market graphs of the UK stock market, no particular model (including the Pareto I) gives a significantly better fit compared with the others. However, complements of market graphs, constructed using a threshold of θ negative and close to zero, can be fitted well with the lognormal or Fréchet distributions.

The contents of this paper are as follows. In Section 2, we give a brief overview of our dataset and the construction of our network graphs. In Section 3, we describe the models fitted and the criteria used for assessing the fit of the models. In Section 4, we present our results comparing the performances of the fitted models. In Section 5, we give a short discussion of our results and their relation to some recent related papers. In Section 6, we provide some concluding remarks.

2 Data and construction of the network graphs

The data set considered in our analysis consists of UK stock market data for stocks or financial instruments traded on three UK stock markets: FTSE100, FTSE250, and FTSEAIM, over three different time periods each of two years in length. These are the daily closing prices of each stock listed on any of the three markets, for each trading day between the dates: i) 1st January 2000 - 31st December 2001; ii) 1st January 2006 - 31st December 2007; iii) 1st January 2012 - 31st December

2013. These three periods have totals of $T = 505, 505, 506$ trading days, respectively. The stock data were obtained using the Datastream package (Datastream International Datastream International (2016)) and all analysis was performed using R (R Development Core Team R Development Core Team (2016)). Note that we consider each time period independently, and for each two year period we use only the data for stocks that were traded continuously throughout the time period - stocks for which closing prices were available for all trading days within a two year period. For each of the three periods, we have a total sample of $N = 492, 1052, 991$ stocks, respectively. Any stocks which started trading (on a market) after the first trading day or ceased trading before the last day of the two year period were omitted, and stocks with a significant amount of missing data were also omitted.

We form our network model of the UK stock market by constructing network graphs using the cross correlations of stock price movements, for each two year period. We build the graphs, using (1), according to Boginski et al. Boginski et al. (2005)'s method as described in Section 1, for a range of values of θ. To aid us in our computations, we consider the market graphs and complements in terms of their adjacency matrices - these are $N \times N$ matrices, whose elements represent the edges between stocks. The element (i, j), $i, j = 1, 2, \ldots, N$, refers to the edge (link) between stocks i and j, taking a value of zero if $C_{i,j} < \theta$ or one if $C_{i,j} \geq \theta$, for the market graph, and vice versa in the case of the complement.

Whilst Boginski et al. Boginski et al. (2005) analysed the degree distribution of the market graphs and their complements, for various values of θ, we consider the distribution of the degree density of the market graphs and complements. The degree values of the stocks could be calculated directly from the adjacency matrices, by totalling the number of edges of each stock - in other words, summing up all the elements in each row of the matrices, to obtain $N^* \leq N$ degree values for each market graph and complement. Note that in the case of the market graph, we subtract one edge from the total sum of each row to remove the 'self-correlation', correlation values of $C_{i,j} = 1$, where $i = j$, $i, j = 1, \ldots, N$, which would lead to an extra edge being accounted for, regardless of the value of θ chosen. The number of degree values, N^*, is allowed to be less than the number of stocks, as we do not consider a stock if it has a degree value (total number of edges) equal to zero.

3 Distributions fitted to the data

The Pareto I distribution has the probability density function (PDF) given by

$$f(x) = \frac{\alpha \lambda^{\alpha}}{x^{\alpha+1}}$$

for $x > \lambda > 0$ and $\alpha > 0$. Taking a very simple mathematical form, this is not a very flexible heavy tailed distribution. We consider three other heavy tailed distributions that are more flexible: the generalized Pareto distribution (GPD) given by the PDF

$$f(x) = \frac{1}{\lambda p} \left(1 + \alpha \frac{x - u}{\lambda}\right)^{-\frac{1}{\alpha}-1}$$

for $x > u$ if $\alpha \geq 0$, $u < x < u - \lambda/\alpha$ if $\alpha < 0$ and $\lambda > 0$, where u is a fixed number known as threshold and p is the probability that the data exceeds u; the lognormal distribution given by the

PDF

$$f(x) = \frac{1}{\sqrt{2\pi}\alpha x} \exp\left\{-\frac{(\ln x - \lambda)^2}{2\alpha^2}\right\}$$

for $x > 0$, $\lambda > 0$ and $\alpha > 0$; the Fréchet distribution given by the PDF

$$f(x) = \frac{\alpha\lambda^\alpha}{x^{\alpha+1}} \exp\left\{-\left(\frac{\lambda}{x}\right)^\alpha\right\}$$

for $x > 0$, $\lambda > 0$ and $\alpha > 0$. Each of these PDFs can be monotonically decreasing or unimodal. We have chosen them because of their simple mathematical forms. There are other heavy tailed distributions which are even more flexible (for example, capable of allowing for bimodality), but they come at the most of complicated mathematical forms often involving special functions. They also suffer from estimation issues like identifiability.

Each distribution was fitted by the method of maximum likelihood, and to avoid possible complications when fitting the distributions, we consider only those stocks which have at least one degree. Suppose $\{x_1, x_2, \ldots, x_n\}$ are independent observations. Then the maximum likelihood estimates for the Pareto I distribution are

$$\widehat{\lambda} = \min(x_1, x_2, \ldots, x_n)$$

and

$$\widehat{\alpha} = n\left[\sum_{i=1}^n \ln x_i - n \ln \min(x_1, x_2, \ldots, x_n)\right]^{-1}.$$

The maximum likelihood estimates for the lognormal distribution are

$$\widehat{\lambda} = \frac{1}{n}\sum_{i=1}^n \ln x_i$$

and

$$\widehat{\alpha} = \frac{1}{n}\sum_{i=1}^n \left(\ln x_i - \frac{1}{n}\sum_{j=1}^n \ln x_j\right)^2.$$

The maximum likelihood estimate of α for the Fréchet distribution is the root of

$$\begin{aligned} & \frac{n}{\alpha} + \left(\frac{1}{n}\sum_{i=1}^n x_i^{-\alpha}\right)^{-1} \sum_{i=1}^n x_i^{-\alpha} \ln x_i + \frac{1}{\alpha} \ln\left(\frac{1}{n}\sum_{i=1}^n x_i^{-\alpha}\right)\left(\frac{1}{n}\sum_{i=1}^n x_i^{-\alpha}\right)^{-1} \sum_{i=1}^n x_i^{-\alpha} \\ = \; & \frac{n}{\alpha} \ln\left(\frac{1}{n}\sum_{i=1}^n x_i^{-\alpha}\right) + \sum_{i=1}^n \ln x_i. \end{aligned}$$

The maximum likelihood estimate of λ for the Fréchet distribution is

$$\widehat{\lambda} = \left(\frac{1}{n}\sum_{i=1}^n x_i^{-\widehat{\alpha}}\right)^{-\frac{1}{\widehat{\alpha}}}.$$

The maximum likelihood estimates for the GPD are the simultaneous solutions of

$$\frac{1+\alpha}{\lambda^2}\sum_{i=1}^{n}(x_i-u)\left(1+\alpha\frac{x_i-u}{\lambda}\right)^{-1}=\frac{n}{\lambda}$$

and

$$\frac{1+\alpha}{\alpha\lambda}\sum_{i=1}^{n}(x_i-u)\left(1+\alpha\frac{x_i-u}{\lambda}\right)^{-1}=\frac{1}{\alpha^2}\sum_{i=1}^{n}\ln\left(1+\alpha\frac{x_i-u}{\lambda}\right).$$

In addition to a visual comparison of the fits of the models to the true degree density, we also assess the fitted distributions by other various criteria:

- the Akaike information criterion due to Akaike Akaike (1974) defined by

  $$\mathrm{AIC}=2k-2\ln L\left(\widehat{\boldsymbol{\Theta}}\right),$$

  where $\boldsymbol{\Theta}$ is the vector of unknown parameters, $\widehat{\boldsymbol{\Theta}}$ is the maximum likelihood estimate of $\boldsymbol{\Theta}$ and k is the number of unknown parameters;

- the Bayesian information criterion due to Schwarz Schwarz (1978) defined by

  $$\mathrm{BIC}=k\ln n-2\ln L\left(\widehat{\boldsymbol{\Theta}}\right);$$

- the Kolmogorov-Smirnov statistic (Kolmogorov Kolmogorov (1933), Smirnov Smirnov (1948)) defined by

  $$\mathrm{KS}=\sup_x\left|\frac{1}{n}\sum_{i=1}^{n}I\left\{x_i\le x\right\}-\widehat{F}(x)\right|,$$

  where $I\{\cdot\}$ denotes the indicator function and $\widehat{F}(\cdot)$ the maximum likelihood estimate of $F(x)$.

The smaller the values of these criteria the better the fit. For more discussion on these criteria, see Burnham and Anderson Burnham and Anderson (2004) and Fang Fang (2011).

We also used the chi-square goodness of fit statistic to compare of fits of the models, but its values were similar to those of the Kolmogorov-Smirnov statistic. So, we shall not report values of the chi-square goodness of fit statistic.

4 Results

In this section, we provide short graphical comparisons of the fits, for each of the three time periods considered, of the fitted distributions to the true degree density of the market graphs and their complements, generated under various threshold values of θ. We also compare numerically the AIC and BIC values, the results of the KS test for the fitted distributions, for each θ value. For each of the three time periods, we generated market graphs using threshold values

$\theta = 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4,\ 0.45, 0.5, 0.55, 0.6,$ and the complements of the market graphs using $\theta = -0.125, -0.1, -0.075, -0.05,\ -0.025.$

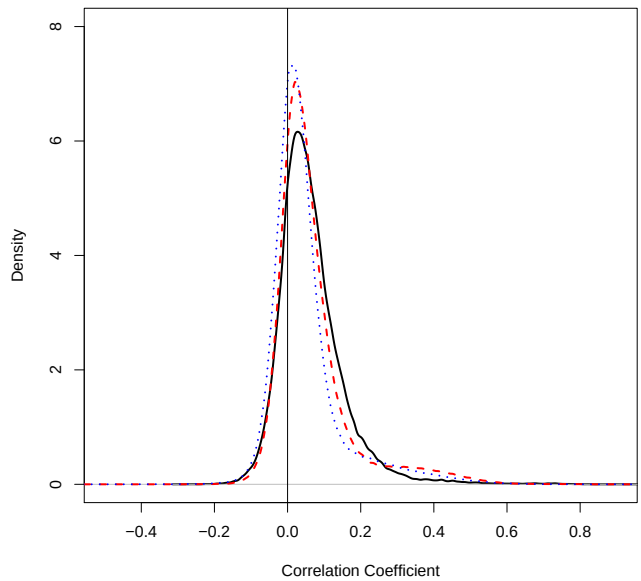

Figure 1: Comparison of the distributions of correlation coefficients, of the sample of stocks, in the three time periods: i) 1st January 2000 - 31st December 2001 (solid black line); ii) 1st January 2006 - 31st December 2007 (dashed red line); iii) 1st January 2012 - 31st December 2013 (dotted blue line).

A quick look at the distributions of the correlation coefficients of the samples of stocks (Figure 1), in each of the three time periods, reveals a similar shape to that of the equivalent distribution in Figure 1 of Boginski et al. Boginski et al. (2005). We see that all three distributions are not strictly symmetric or centred around zero, but are instead skewed slightly towards the right tail. The distributions of the coefficients for the two periods covering 2006 - 2008 and 2012 - 2014 have a higher peak density than that for the earliest period of 2000 - 2002. The density of correlations in the range of 0 to 0.2 is higher for the earliest period, whilst the density of correlations between -0.1 and 0, and in excess of 0.3, is slightly greater for the most recent two periods.

One of the key assumptions of the distributions in Section 3 is that data are independently and identically distributed and have no serial correlation. We tested for no serial correlation using Durbin and Watson Durbin and Watson (1950) Durbin and Watson (1951) Durbin and Watson (1971)'s method. This gave the p-values of 0.122, 0.056 and 0.066 for the three data periods. We tested for the independent and identical hypothesis using the difference sign, turning point and the rank tests (Brockwell and Davis Brockwell and Davis (2002), Chapter 1). The p-values of the difference sign test for the three periods were 0.122, 0.056 and 0.066. The p-values of the turning point test for the three periods were 0.064, 0.195 and 0.120. The p-values of the rank test for the three periods were 0.072, 0.174 and 0.123.

For the plots in the following subsections, the blue line denotes the Pareto I distribution; the

green line denotes the Fréchet distribution; the yellow line denotes the lognormal distribution; the purple line denotes the GPD.

4.1 Period 1: 1st January 2000 - 31st December 2001

The first period contains a total sample of $N = 492$ stocks, for which the daily closing prices on each of the $T = 505$ trading days were available. We note here that the results and plots produced are for a range of θ values smaller than that specified in Section 4, $\theta \in [0.1, 0.4]$ for the market graph and $\theta \in [-0.125, -0.025]$ for the complements. For the values of θ which were omitted, the degree density plots either had no real distinct shape or not enough data to be fitted.

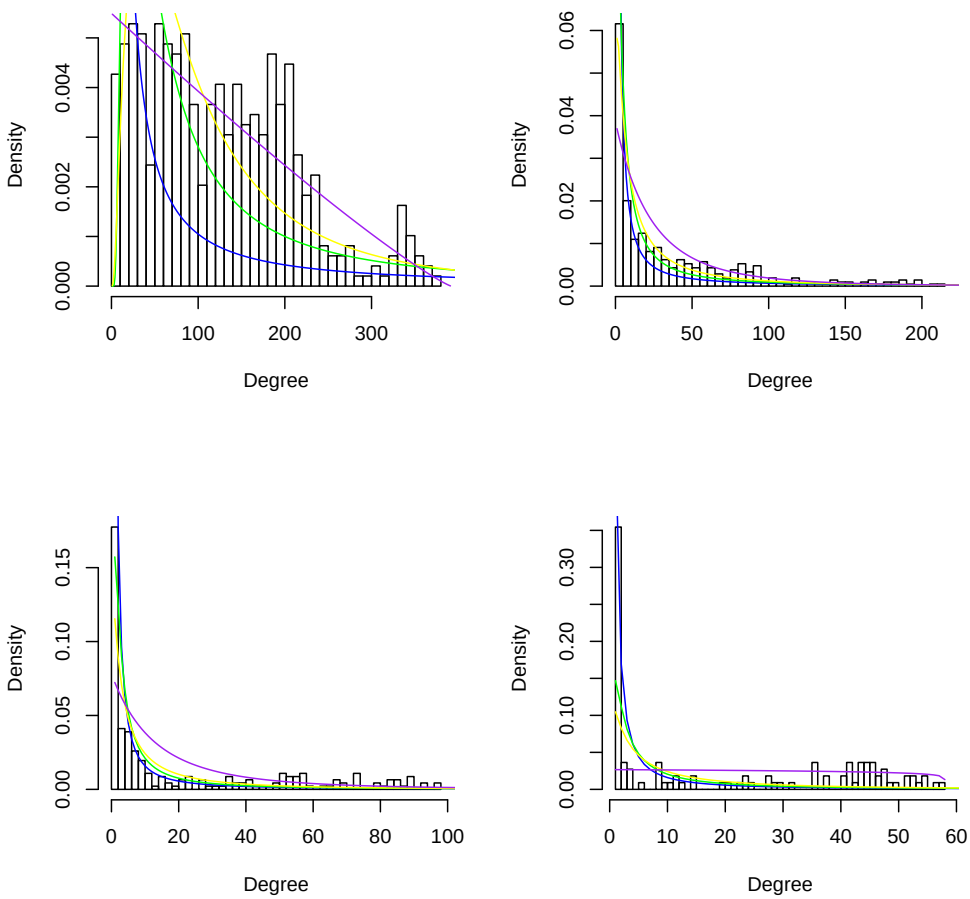

Figure 2: Comparison of the fitted distributions (using market graphs) for: $\theta = 0.1$ (top left); $\theta = 0.2$ (top right); $\theta = 0.3$ (bottom left); $\theta = 0.4$ (bottom right).

In Figure 2, for the market graph of θ close to 0, i.e. $\theta = 0.1$, it is hard to distinguish any real fit between the fitted distributions. As θ increases to 0.2, it appears that the Fréchet and lognormal distributions give the best overall fit, followed by the Pareto I distribution, and the least best fit

is given by the GPD. Increasing θ further to 0.3, the best overall fit now appears to be given by the the Pareto I distribution, and then by the Fréchet and lognormal distributions which appear to show some overestimation. However, the GPD seems to provide the best fit to the upper part of the density, where the other fitted distributions tail off and decay towards 0. With $\theta \geq 0.4$ the results are similar to that of $\theta = 0.3$, however, we note that the true density becomes more sparse with some degree ranges having a density of zero. We also found that as the chosen threshold value θ increased, the number of degrees of stocks decreased, and the total number of stocks with at least one degree also fell. The upper limit of $\theta = 0.4$ was therefore chosen, as it was found that using a threshold value above 0.4 made it difficult get a reasonable fit from any of the fitted distributions. This could be explained by the decrease in the number of nodes with at least one degree when θ exceeds 0.4, due to the low density of stocks with correlations coefficients greater than 0.4.

Table 1: AIC and BIC values for the fitted distributions (using market graphs) for $\theta \in [0.1, 0.4]$.

	AIC				BIC			
θ	Pareto I	Fréchet	Lognormal	GPD	Pareto I	Fréchet	Lognormal	GPD
0.1	6621.787	6008.05	5794.663	5653.215	6630.184	6016.447	5803.06	5661.612
0.15	5472.051	5175.919	5035.272	4820.501	5480.407	5184.275	5043.628	4828.857
0.2	3934.483	3891.716	3831.735	3531.416	3942.558	3899.792	3839.81	3539.492
0.25	2617.496	2671.312	2657.229	2358.229	2625.039	2678.855	2664.771	2365.772
0.3	1705.176	1794.992	1801.067	1519.499	1712.061	1801.877	1807.952	1526.383
0.35	1145.129	1198.558	1192.365	965.3112	1151.11	1204.539	1198.345	971.2921
0.4	859.4584	903.899	893.8395	644.3275	864.8594	909.2999	899.2404	649.7284

From Table 1, we observe that the lowest AIC and BIC values over all θ are given by the GPD, suggesting that the GPD gives the best fit out of all the fitted distributions. For $\theta \leq 0.2$, the Fréchet and lognormal distributions give the next lowest AIC and BIC values after the GPD, whereas for $\theta > 0.2$, the Pareto I distribution gives the next lowest AIC and BIC values after the GPD.

Table 2: KS test for the fitted distributions (using market graphs) for $\theta \in [0.1, 0.4]$.

	KS test			
θ	Pareto I	Fréchet	Lognormal	GPD
0.1	0.3598	0.1712	0.1180	0.05596
	(0.0000)	(5.935×10^{-13})	(2.220×10^{-06})	(0.09174)
0.15	0.2698	0.1378	0.09681	0.1113
	(0.0000)	(2.240×10^{-08})	(0.0002385)	(1.303×10^{-05})
0.2	0.1961	0.1128	0.08002	0.1861
	(2.054×10^{-14})	(4.653×10^{-05})	(0.009342)	(4.918×10^{-13})
0.25	0.1651	0.1084	0.1221	0.2315
	(5.014×10^{-08})	(0.001063)	(0.0001400)	(2.220×10^{-15})
0.3	0.2381	0.1506	0.1443	0.3020
	(8.446×10^{-12})	(5.651×10^{-05})	(0.0001335)	(0.0000)
0.35	0.2313	0.1645	0.1683	0.3979
	(2.955×10^{-07})	(0.0006981)	(0.0004820)	(0.0000)
0.4	0.2818	0.2006	0.1852	0.3609
	(5.161×10^{-08})	(0.0002852)	(0.001055)	(7.239×10^{-13})

The corresponding p-values are given in brackets below the KS statistics.

The results of the KS test (Table 2) show that for θ close to zero, the GPD and lognormal distributions produce the lowest KS statistics. For $\theta \geq 0.2$, the Fréchet and lognormal distributions give either the lowest or second lowest KS statistics. Looking closer at the p-values, we find that the majority are very small and close to zero - evidence (in these cases) at all significance levels to reject the null hypothesis that the data comes from the respective fitted model. However, when $\theta = 0.1$ the GPD gives the lowest KS statistic with a p-value of 0.09174; when $\theta = 0.2$ the lognormal distribution gives the lowest KS statistic with a p-value of 0.009342; when $\theta = 0.25$ the Fréchet distribution gives the lowest KS statistic with a p-value of 0.001063; when $\theta = 0.4$ the lognormal distribution gives the lowest KS statistic with a p-value of 0.001055, respectively. Therefore, we conclude that for $\theta = 0.1$, at the 5% significance level we fail to reject the null hypothesis that the sample degree data is drawn from a GPD.

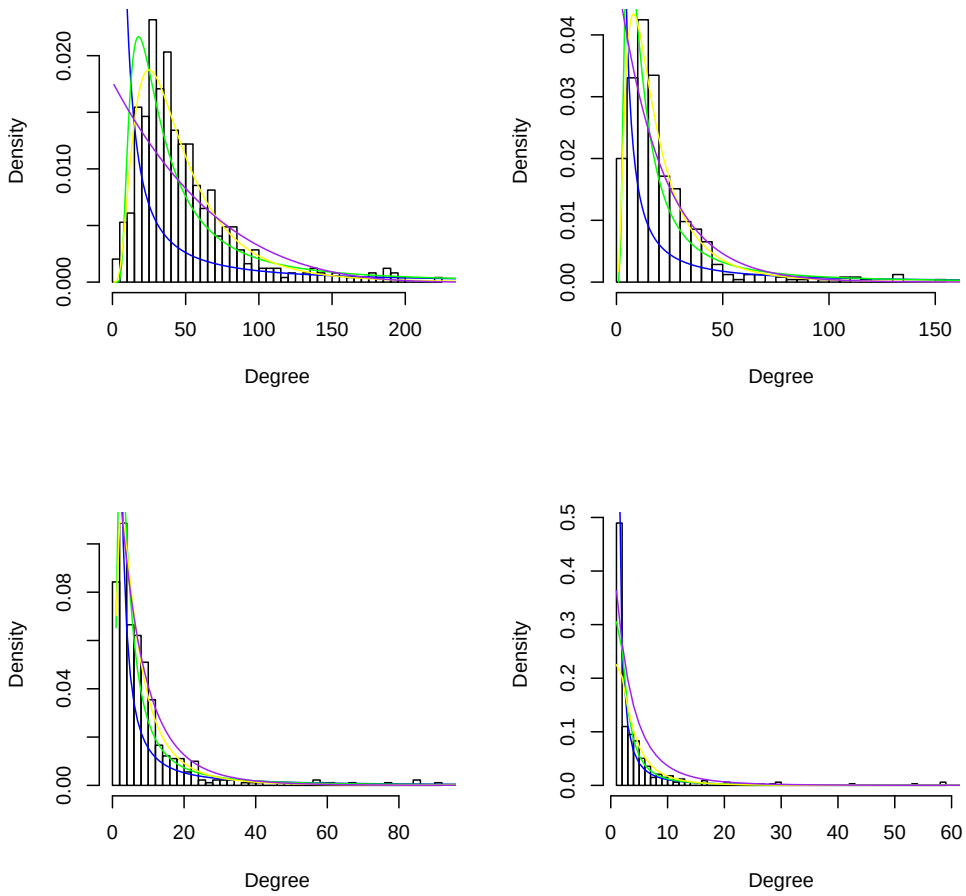

Figure 3: Comparison of the fitted distributions (using complement of market graphs) for: $\theta = -0.025$ (top left); $\theta = -0.05$ (top right); $\theta = -0.075$ (bottom left); $\theta = -0.1$ (bottom right).

Moving on to the complement of the market graph, in Figure 3, we see that the differences between fits are more pronounced. It can be observed from the plots that for θ very close to zero, i.e. $\theta = -0.025, -0.05, -0.075$, the shape of the true degree density of the complement of the market graph differs from that of the original market graph. The peaks in the distributions instead occur at a 'low' number of degrees slightly greater than the minimum. For $\theta = -0.025$, the best overall fit appears to be given by the lognormal distribution, the second best fit by the Fréchet distribution, the third best fit by the GPD, with the Pareto I distribution giving the least best fit overall. These results hold true as we decrease the value (increase in absolute value) of θ to -0.05, and -0.075. The best overall fits continue to be given by the lognormal and Fréchet distributions, whilst the Pareto I distribution underestimates the density for low degree values. At $\theta = -0.1$, for low degree values the Pareto I distribution seems to provide the best fit to the true distribution. However, the best overall fit may arguably still be given by the Fréchet and lognormal distributions, whilst the GPD gives the least best fit - overestimating the density at low degree values. As was the case with the market graph, we find that the number of stocks with at least one degree decreases, but more significantly as the value of θ decreases (increases in absolute value),

which may be explained by the skew towards the right tail in the distribution of the correlation coefficients (Figure 1).

Table 3: AIC and BIC values for the fitted distributions (using complement of market graphs) for $\theta \in [-0.125, -0.025]$.

	AIC				BIC			
θ	Pareto I	Fréchet	Lognormal	GPD	Pareto I	Fréchet	Lognormal	GPD
-0.025	5538.171	4801.396	4633.709	4768.367	5546.568	4809.793	4642.106	4776.764
-0.05	4673.118	4027.107	3896.3	3935.797	4681.507	4035.496	3904.689	3944.186
-0.075	3029.539	2883.193	2827.859	2628.777	3037.762	2891.416	2836.082	2637
-0.1	1420.476	1565.515	1607.256	1256.155	1428.116	1573.155	1614.896	1263.795
-0.125	435.2945	534.822	571.4838	401.5692	441.192	540.7195	577.3813	407.4668

Table 4: KS test for the fitted distributions (using complement of market graphs) for $\theta \in [-0.125, -0.025]$.

	KS test			
θ	Pareto I	Fréchet	Lognormal	GPD
-0.025	0.4105	0.1177	0.05124	0.1689
	(0.0000)	(2.395×10^{-06})	(0.1510)	(1.289×10^{-12})
-0.05	0.3785	0.1294	0.06823	0.1179
	(0.0000)	(1.502×10^{-07})	(0.02086)	(2.412×10^{-06})
-0.075	0.2923	0.1233	0.06540	0.09016
	(0.0000)	(2.221×10^{-06})	(0.04222)	(0.001309)
-0.1	0.2849	0.1915	0.1599	0.2849
	(0.0000)	(3.698×10^{-11})	(6.518×10^{-08})	(0.0000)
-0.125	0.3972	0.2601	0.2222	0.3972
	(0.0000)	(1.030×10^{-08})	(1.799×10^{-06})	(0.0000)

The corresponding p-values are given in brackets below the KS statistics.

In comparison with Table 1 for the market graph, Table 3 shows that for θ negative and very close to 0 ($\theta \geq -0.05$), the fitted lognormal distribution produces the lowest AIC and BIC values, indicating that at these values it provides the best fit of all the distributions. However, as θ decreases (increases in absolute value) the GPD gives the lowest AIC and BIC values (and best fit), followed by the Pareto I giving the second lowest, with the lognromal and Fréchet distributions giving the highest values.

The results of the KS test (Table 4) give a clearer picture, as for all values of θ the lognormal distribution gives the lowest KS statistic. In general, the next lowest values are given by the Fréchet distribution, then the GPD, with the Pareto I distribution producing the largest KS statistics for all θ. These results are supported by the corresponding p-values from the test, as for values of $\theta = -0.025, -0.05, -0.075$ the p-values are 0.1510, 0.02086 and 0.04222, respectively, for the KS test on the fitted lognormal distribution. This indicates that at the 10% (1%) significance level for

$\theta = -0.025$ ($\theta = -0.05, -0.075$) we fail to reject the null hypothesis that the sample degree data is drawn from the lognormal distribution.

4.2 Period 2: 1st January 2006 - 31st December 2007

The second period that we examine contains a total sample of $N = 1052$ stocks, for which the closing daily prices on each of the $T = 505$ trading days are available. For this two year period, we computed and plotted the results for the whole range of θ values, as specified in Section 4.

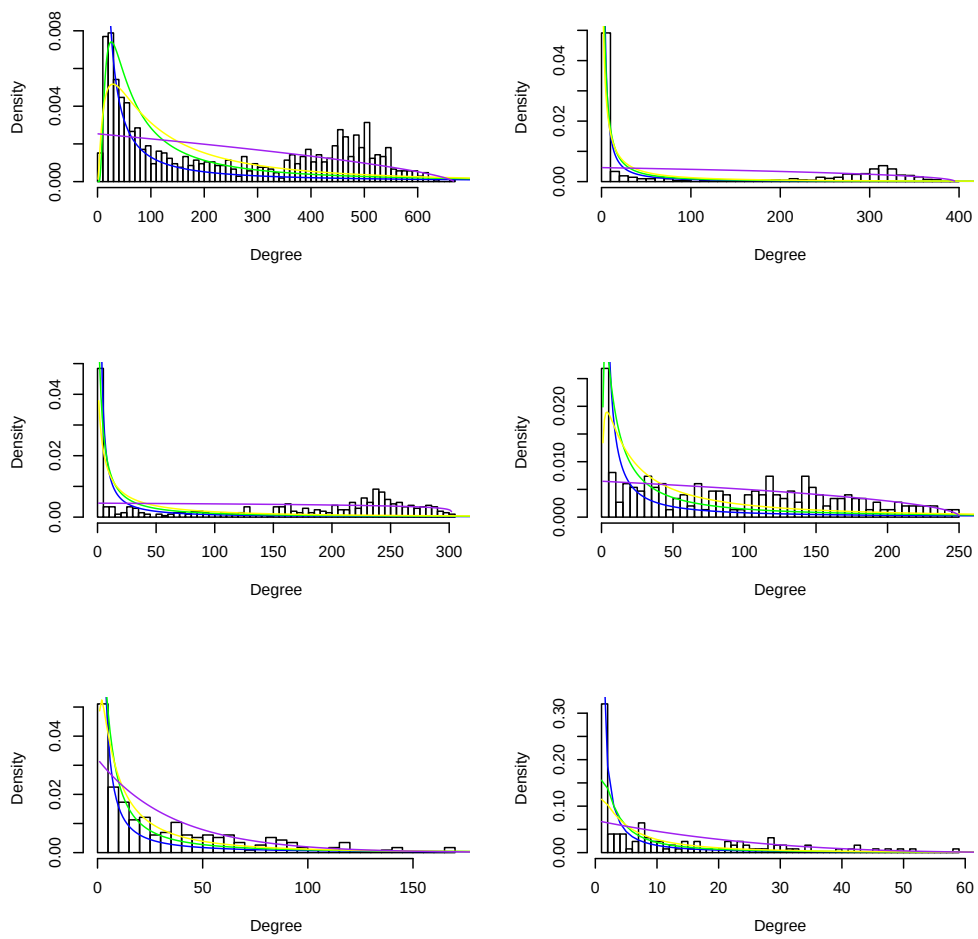

Figure 4: Comparison of the fitted distributions (using market graphs) for: $\theta = 0.1$ (top left); $\theta = 0.2$ (top right); $\theta = 0.3$ (middle left); $\theta = 0.4$ (middle right); $\theta = 0.5$ (bottom left); $\theta = 0.6$ (bottom right) .

From Figure 4, we can immediately see a difference with the distributions for the first sample period (Figure 2) - the degree density here is bimodal at low and high degree values, for θ close to zero. However, none of the fitted distributions captures this, each fitting just one of the two peaks well. For the degree density of the market graph with $\theta = 0.1$, the best fit to the lower half

of the distribution appears to be given by the Fréchet and lognormal distributions, followed by the Pareto I distribution, with the GPD giving the best fit to the upper half of the distribution. With $\theta = 0.2, 0.3$, the Pareto I, lognormal and Fréchet distributions jointly give the best fit to the lower half of the distribution, whilst the GPD continues to give the best fit over the upper part of the distribution at large degree values. At $\theta = 0.4$, the distribution is no longer bimodal, and the overall best fit seems to be given by the GPD. For $\theta = 0.5$, at low degree values the best fit appears to be given by the Fréchet, lognormal and Pareto I distributions. When θ reaches our upper limit of 0.6 we find that the GPD may give the best visual fit out of all the fitted distributions.

Table 5: AIC and BIC values for the fitted distributions (using market graphs) for $\theta \in [0.1, 0.6]$.

	AIC				BIC			
θ	Pareto I	Fréchet	Lognormal	GPD	Pareto I	Fréchet	Lognormal	GPD
0.1	14803.99	13966.86	13752.03	13438.74	14813.9	13976.78	13761.95	13448.66
0.2	8412.983	8745.172	8746.506	7907.239	8422.435	8754.624	8755.958	7916.691
0.3	5177.502	5175.583	5046.303	3962.849	5185.569	5183.649	5054.369	3970.915
0.4	3693.377	3543.815	3415.152	3021.877	3700.771	3551.209	3422.546	3029.271
0.5	2182.37	2136.479	2089.138	1901.512	2189.255	2143.364	2096.023	1908.396
0.6	869.4814	902.8605	892.2891	721.8855	875.138	908.5171	897.9457	727.5421

The comparison of the AIC and BIC values from the fitted distributions are shown in Table 5. Again, for all values of θ the GPD produces the lowest AIC and BIC values, indicating that the GPD may offer the best fit out of all the distributions fitted. In general, the next lowest AIC and BIC values are given by Fréchet and lognormal distributions, with the Pareto I distribution producing the most number of largest values across the range of θ.

Table 6: KS test for the fitted distributions (using market graphs) for $\theta \in [0.5, 0.6]$.

	KS test			
θ	Pareto I	Fréchet	Lognormal	GPD
0.5	0.2276	0.1203	0.09034	0.1724
	(8.092×10^{-11})	(0.002506)	(0.04606)	(2.174×10^{-06})
0.55	0.2033	0.1365	0.1210	0.2788
	(5.855×10^{-07})	(0.002275)	(0.009663)	(1.028×10^{-12})
0.6	0.2320	0.1600	0.1486	0.2698
	(2.865×10^{-06})	(0.003335)	(0.008037)	(2.514×10^{-08})

The corresponding p-values are given in brackets below the KS statistics.

Table 6 shows the results for the KS test for the fitted distributions using the market graphs. We note that output for $\theta < 0.5$ is omitted as the p-values corresponding to the KS statistics were found to be zero or very small and insignificant. The largest p-values, relative to all others, can be found at $\theta = 0.5, 0.55, 0.6$ for the smallest KS statistics given by the lognormal distribution, where the respective p-values are 0.04606, 0.009663 and 0.008037. From this, we conclude that at the 1% significance level, for $\theta = 0.5$, we fail to reject the null hypothesis that the sample degree data is

drawn from the lognormal distribution.

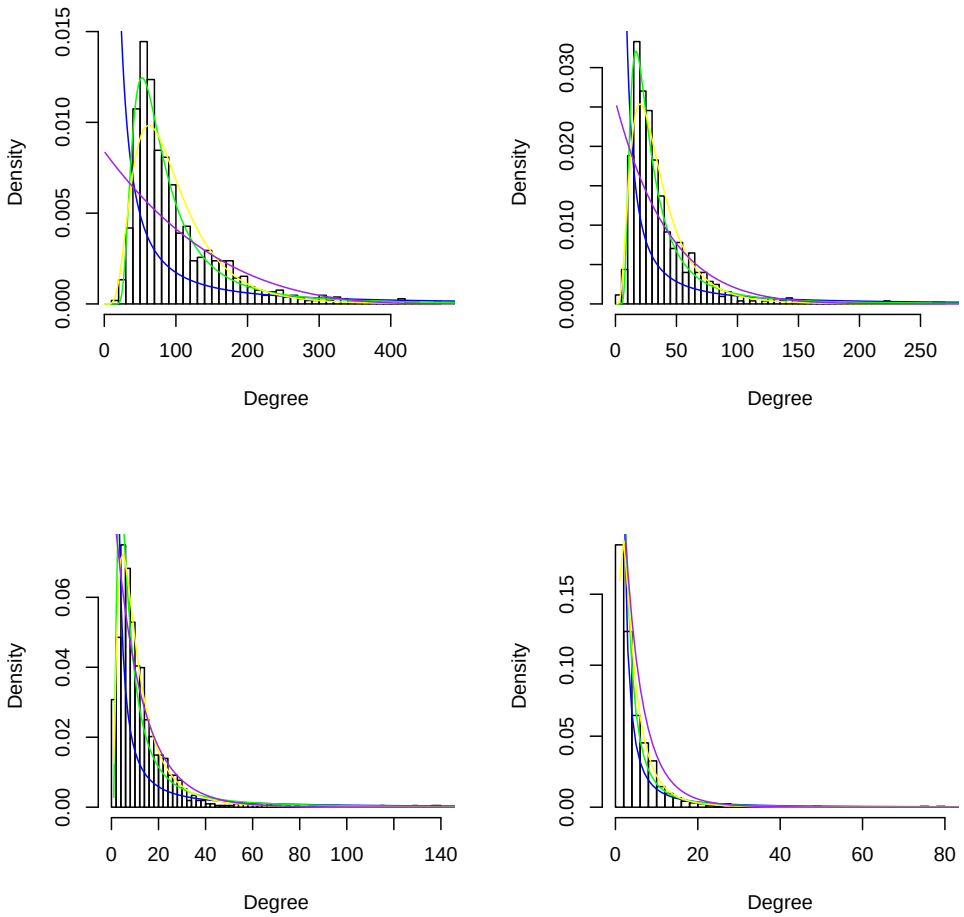

Figure 5: Comparison of the fitted distributions (using complement of market graphs) for: $\theta = -0.025$ (top left); $\theta = -0.05$ (top right); $\theta = -0.075$ (bottom left); $\theta = -0.1$ (bottom right).

In Figure 5, we notice that the results for the degree densities of the complements of the market graph, in this second period, are similar to those from the first period. Again, with θ very close to 0, the shape of the distribution differs from that when considering the market graph; the peak in the distribution of the density occurs close to but not exactly at the minimum number of degrees. With $\theta = -0.025$, it appears that the best fit overall is given by the Fréchet distribution, the second best by the lognormal distribution, the third best by the GPD, with the Pareto I distribution giving the least best fit. These results continue to hold as we decrease θ (increase in absolute value) further to -0.05 and -0.075. At $\theta = -0.1$, the fit of all the fitted distributions is much more similar, however the lognormal and Fréchet distributions still appear to provide the best overall fit. The Pareto I distribution and GPD give the least best fit overall and show signs of underestimation and overestimation, respectively.

A possible explanation for the similarity in results (relating to the complement of the market graph) between the two periods, is that the distributions of the negative correlation coefficients in

all periods, Figure 1, appear to be quite similar. Thus, due to the similarity in the distribution of the negative correlation coefficients, we may expect the complements of the market graphs (using negative θ values) in both this (and the third) period to be similar to that of the first. Therefore we may hypothesise that the respective degree densities and the shape of the fitted distributions to be very similar too.

Table 7: AIC and BIC values for the fitted distributions (using complement of market graphs) for $\theta \in [-0.125, -0.025]$.

	AIC				BIC			
θ	Pareto I	Fréchet	Lognormal	GPD	Pareto I	Fréchet	Lognormal	GPD
-0.025	12865.72	11131.17	11108.53	11688.57	12875.64	11141.09	11118.45	11698.49
-0.05	10756.18	9271.068	9201.071	9662.023	10766.1	9280.983	9210.986	9671.938
-0.075	8412.428	7449.771	7212.625	7157.561	8422.322	7459.665	7222.519	7167.455
-0.1	4114.2	4188.687	4177.259	3575.537	4123.638	4198.125	4186.697	3584.975
-0.125	1421.347	1656.218	1704.641	1249.465	1429.511	1664.382	1712.805	1257.63

From Table 7, we can also see a similar pattern to the first period in the information criteria values for the complements of the market graphs. At θ very close to 0, the lognormal distribution produces the lowest AIC and BIC values, indicating the best fit out of the fitted distributions. This is followed closely, in second, by the Fréchet distribution. As θ decreases (increases in absolute value), the GPD gives the lowest values for both (and the best fit), the Pareto I distribution follows with the second lowest values, whilst the lognormal and Fréchet distributions give the largest values for both.

Table 8: KS test for the fitted distributions (using complement of market graphs) for $\theta \in [-0.075, -0.025]$.

	KS test			
θ	Pareto I	Fréchet	Lognormal	GPD
-0.025	0.4093	0.03931	0.06791	0.2368
	(0.0000)	(0.07745)	(0.0001221)	(0.0000)
-0.05	0.3882	0.05385	0.05500	0.2100
	(0.0000)	(0.004511)	(0.003464)	(0.0000)
-0.075	0.3543	0.1245	0.06521	0.1250
	(0.0000)	(1.987×10^{-14})	(0.0002885)	(1.577×10^{-14})

The corresponding p-values are given in brackets below the KS statistics.

For the KS test in Table 8, we omit the results for $\theta < -0.075$, as the p-values corresponding to the KS statistics were found to be either zero or very small and insignificant. Those that were significant showed a slight departure from the results for the test in the first period (Table 4). For $\theta > -0.075$ we find that the Fréchet distribution produces the most number of lowest KS statistics, and the lognormal distribution giving the second most. Across the whole range of θ values the highest KS statistics are given by the Pareto I distribution. The largest p-values are produced at the θ values closest to zero (-0.025 and -0.05), however, these correspond to the lowest KS statistics

given by the Fréchet distribution, as opposed to the lognormal distribution. We note that for $\theta = -0.025, -0.05$, these p-values are 0.07745 and 0.004511, respectively, thus we conclude that for $\theta = -0.025$ at the 5% significance level, we fail to reject the null hypothesis that the sample degree data is drawn from the Fréchet distribution.

4.3 Period 3: 1st January 2012 - 31st December 2013

The final period considered contains a total sample of $N = 991$ stocks, for which the closing daily prices on each of the $T = 506$ trading days are available. For this final two year period, we also computed and plotted the results for the whole range of θ values, as specified in Section 4.

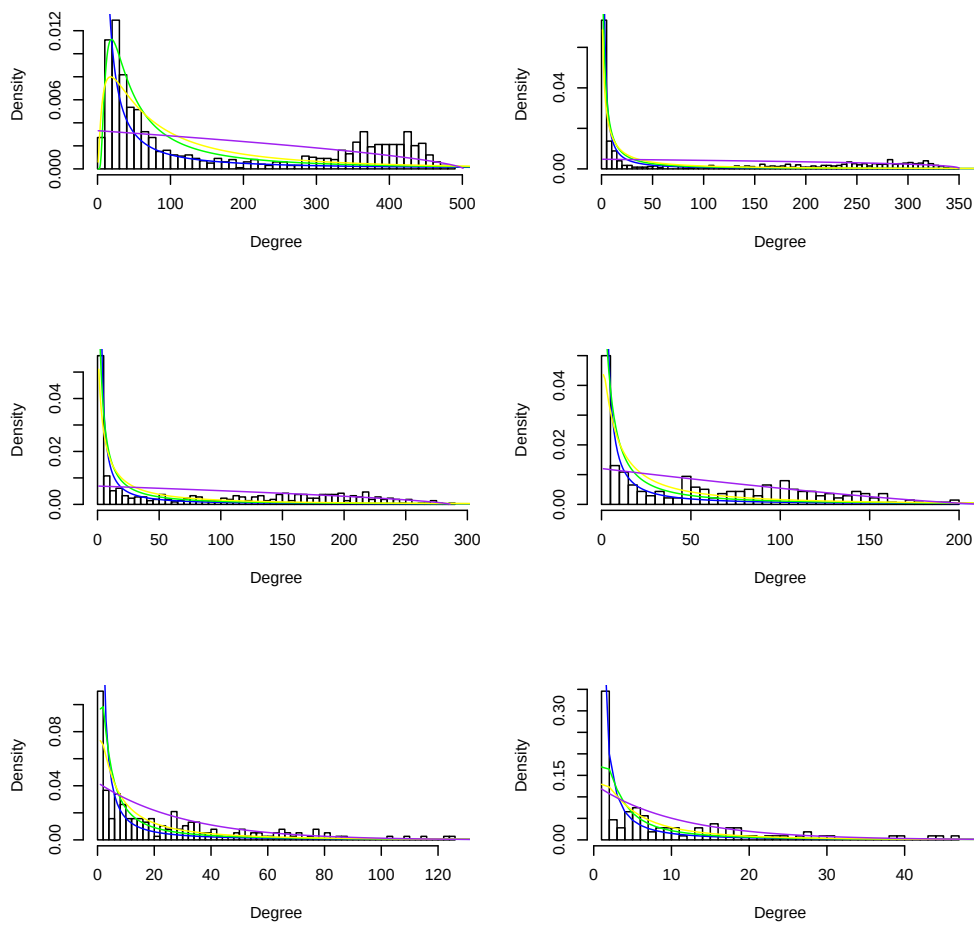

Figure 6: Comparison of the fitted distributions (using market graphs) for: $\theta = 0.1$ (top left); $\theta = 0.2$ (top right); $\theta = 0.3$ (middle left); $\theta = 0.4$ (middle right); $\theta = 0.5$ (bottom left); $\theta = 0.6$ (bottom right) .

The comparisons of the fitted distributions, in Figure 6, show very similar results to those in Figure 4 for the second period. Again, the degree density is bimodal for θ values close to 0 but loses

this property when $\theta \geq 0.4$. With $\theta = 0.1$, the best fit to the lower half of the distribution appears to be given by the Fréchet and lognormal distributions, followed by the Pareto I distribution, whilst the GPD gives the best fit to upper half of the distribution. For $\theta = 0.2, 0.3$, the Pareto I, lognormal and Fréchet distributions jointly give the best fit to the lower half of the distribution, and the GPD continues to give the best fit to the upper part of the distribution. When $\theta \geq 0.4$, the bimodal property disappears, with the overall best fit appearing to be given by the GPD. As θ approaches the upper limit of 0.6, the GPD may give the best visual fit out of all the fitted distributions.

Table 9: AIC and BIC values for the fitted distributions (using market graphs) for $\theta \in [0.1, 0.6]$.

	AIC				BIC			
θ	Pareto I	Fréchet	Lognormal	GPD	Pareto I	Fréchet	Lognormal	GPD
0.1	12940.75	12222.75	12145.49	12124.28	12950.54	12232.54	12155.29	12134.08
0.2	7353.04	7541.859	7496.998	6733.124	7362.142	7550.962	7506.101	6742.226
0.3	4721.579	4739.022	4653.492	4052.203	4729.693	4747.136	4661.606	4060.316
0.4	2922.24	2886.202	2819.461	2505.441	2929.48	2893.443	2826.702	2512.682
0.5	1626.955	1634.281	1602.635	1382.236	1633.46	1640.785	1609.14	1388.74
0.6	673.4181	696.9042	691.9642	578.6943	678.7638	702.2499	697.3098	584.04

Over all values of θ tested, the GPD again gives the lowest AIC and BIC values (Table 9), leading us to conclude that this result holds true for the market graphs in all three time periods. This indicates that according to the information criteria, the GPD would be the preferred distribution relative to the other fitted distributions. Overall, the second lowest AIC and BIC values are given by the lognormal distribution, whilst the highest values are given by the Pareto I and Fréchet distributions.

Table 10: KS test for the fitted distributions (using market graphs) for $\theta \in [0.5, 0.6]$.

	KS test			
θ	Pareto I	Fréchet	Lognormal	GPD
0.5	0.2337	0.1250	0.1145	0.2009
	(1.724×10^{-09})	(0.005121)	(0.01337)	(4.032×10^{-07})
0.55	0.2291	0.1224	0.1083	0.1903
	(4.892×10^{-07})	(0.02595)	(0.06664)	(5.512×10^{-05})
0.6	0.2056	0.1408	0.1338	0.2562
	(0.0002356)	(0.02880)	(0.04325)	(1.592×10^{-06})

The corresponding p-values are given in brackets below the KS statistics.

From the KS test, we find that in all but two of the cases, for $\theta = 0.1, 0.15$ (not shown) the lognormal distribution gives the lowest KS statistics, whilst the Fréchet distribution gives the second lowest. In general, the highest values were found to be given by the GPD for $\theta \leq 0.35$ and the Pareto I distribution for $\theta > 0.35$. However, we omit the results for $\theta < 0.5$ as the p-values of the corresponding KS statistics were found to be either zero or very small and insignificant. Where θ is large (Table 10) we find the largest p-values, as in the equivalent KS test for the second period,

Table 6. For $\theta = 0.5$ the lognormal distribution gives the lowest KS statistic with a p-value of 0.01337; for $\theta = 0.55$ the lognormal and Fréchet distributions give the lowest KS statistics with p-values of 0.06664 and 0.02595, respectively; for $\theta = 0.6$ the lognormal and Fréchet distributions give the lowest KS statistics with p-values of 0.04325 and 0.02880, respectively. Thus, for $\theta = 0.55, 0.6$ at the 1% level of significance, we fail to reject the null hypothesis that the sample degree data is drawn from the Fréchet distribution; for $\theta = 0.5, 0.6$ ($\theta = 0.55$) at the 1% (5%) level of significance, we fail to reject the null hypothesis that the sample degree data is drawn from the lognormal distribution.

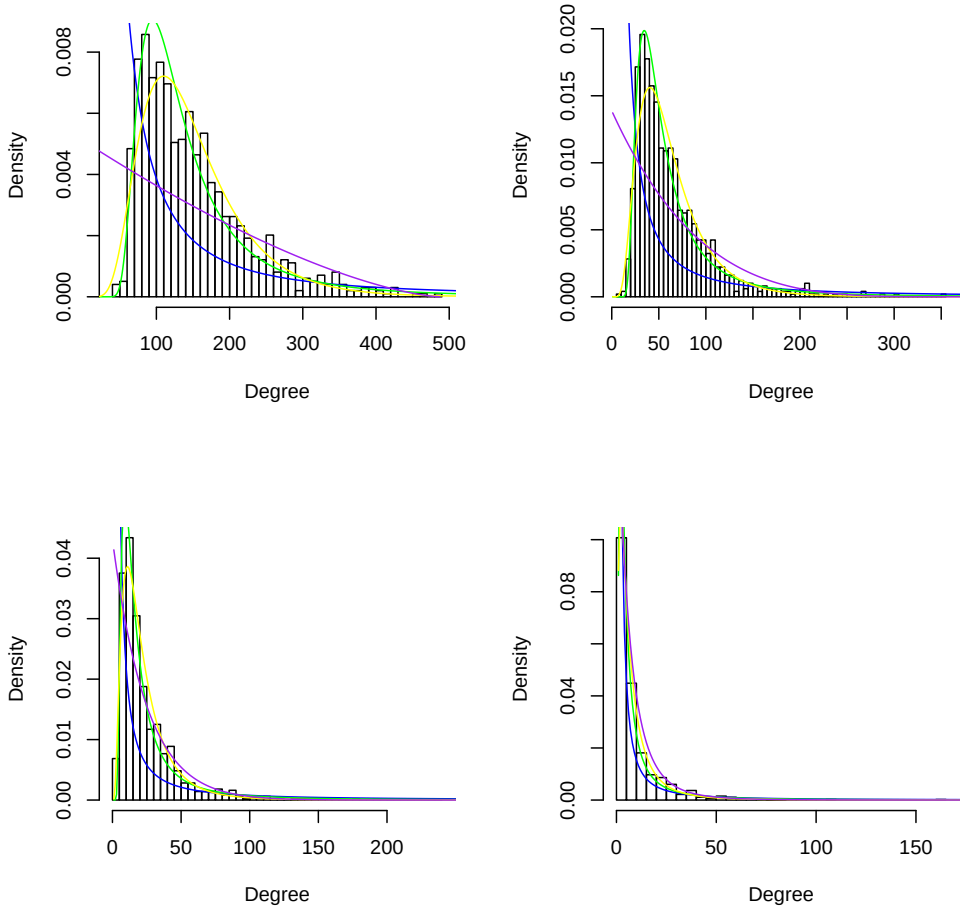

Figure 7: Comparison of the fitted distributions (using complement of market graphs) for: $\theta = -0.025$ (top left); $\theta = -0.05$ (top right); $\theta = -0.075$ (bottom left); $\theta = -0.1$ (bottom right).

As was the case for the market graph, the distributions of the complements of the market graph in the third period, Figure 7, are also very similar to those in the first and second periods over all θ. For θ close to 0, the best visual fit is given by the Fréchet distribution, second by the lognormal distribution, third by the GPD, and the least best by the Pareto I distribution. As θ decreases (increases in absolute value) these results still hold, with the Pareto I and GPD still showing some underestimation and overestimation, respectively. However, the overall fit of all the distributions

becomes much more similar. This result would seem to fit our expectations, due to the similarity in the distribution of the negative correlation coefficients in each of the three time periods, as shown in Figure 1.

Table 11: AIC and BIC values for the fitted distributions (using complement of market graphs) for $\theta \in [-0.125, -0.025]$.

	AIC				BIC			
θ	Pareto I	Fréchet	Lognormal	GPD	Pareto I	Fréchet	Lognormal	GPD
-0.025	12100.16	10986.85	10972.94	11655.52	12109.96	10996.65	10982.74	11665.32
-0.05	11073.94	9550.712	9531.659	10111.47	11083.74	9560.509	9541.456	10121.27
-0.075	9359.642	8051.588	7961.953	8238.138	9369.44	8061.385	7971.751	8247.936
-0.1	6167.447	5980.643	5930.941	5459.377	6177.111	5990.307	5940.605	5469.041
-0.125	2797.995	3100.486	3172.253	2448.657	2806.711	3109.201	3180.968	2457.373

This similarity in results continues for the information criteria, as Table 11 shows that for $\theta \geq -0.075$ the lognormal distribution gives the lowest AIC and BIC values followed by the Fréchet distribution, whilst for $\theta < -0.075$ the GPD gives the lowest AIC and BIC values. This indicates (as for the second period) that according to the information criteria, at θ negative and very close to zero, the lognormal distribution gives the best fit of all models, but away from zero the GPD gives the best fit.

Table 12: KS test for the fitted distributions (using complement of market graphs) for $\theta \in [-0.075, -0.025]$.

	KS test			
θ	Pareto I	Fréchet	Lognormal	GPD
-0.025	0.3183	0.05347	0.04761	0.2712
	(0.0000)	(0.006912)	(0.02237)	(0.0000)
-0.05	0.3917	0.04005	0.04912	0.2469
	(0.0000)	(0.08329)	(0.01677)	(0.0000)
-0.075	0.3787	0.06014	0.05607	0.1677
	(0.0000)	(0.001541)	(0.003931)	(0.0000)

The corresponding p-values are given in brackets below the KS statistics.

The results of the KS test are shown in Table 12, with the output for $\theta < -0.075$ being omitted due to the p-values of the corresponding KS statistics either being zero or very small and insignificant. We note that the results again show that the largest p-values correspond to the lowest KS statistics given by the Fréchet and lognormal distributions. In particular, for $\theta = -0.025$ the lognormal distribution gives the smallest KS statistic with a p-value of 0.02237; for $\theta = -0.05$ the Fréchet and lognormal distributions give the smallest KS statistics with p-values of 0.08329 and 0.01677, respectively. Thus, for $\theta = -0.025, -0.05$ at the 1% significance level, we fail to reject the null hypothesis that the sample degree data is drawn from a lognormal distribution; for $\theta = -0.05$ at the 5% significance level, we fail to reject the null hypothesis that the sample degree data is drawn from the Fréchet distribution.

5 Discussion

The differences between the true distributions of degree density of the market graphs using data from period 1 compared with periods 2 and 3, could be explained by the differences in the original distribution of correlation coefficients in Figure 1. The number of stocks considered in period 1 was approximately half the number in periods 2 and 3, thus the number of correlation coefficients for period 1 was approximately a quarter of that in periods 2 and 3. Further to this, the distribution of the coefficients in period 1 has a smaller range, whilst those for periods 2 and 3 have a higher peak density and a higher density of large correlation values above 0.3. This may also explain the bimodal nature of the degree densities of the market graphs in periods 2 and 3 (Figures 4 and 6).

It appears that, in the case of the market graphs, no definitive conclusion could be made about the best fitting distribution to the degree density. Visually, although we note may be misleading, the fitted distributions to the degree density of market graphs across the three periods show a similar pattern. For market graphs constructed with a low threshold θ, the Fréchet, lognormal and Pareto I distributions give a reasonable fit to the lower part of the distribution, whilst the GPD gives a reasonable fit to the upper part of the distribution. For market graphs constructed with a high threshold θ, the GPD appears to give the best fit overall.

However, for all threshold values θ in the three periods, the GPD gives the lowest AIC and BIC values indicating that according to the information criteria this would be the preferred model out of those that were fitted. In terms of the KS test, we find that the majority of tests on the fitted distributions are insignificant due to very small p-values. Although, for $\theta = 0.1$ in period 1 the GPD gives the smallest KS statistic with a p-value of almost 0.1 - agreeing with the result from the information criteria, whilst for $\theta = 0.5$ in period 2 the lognormal distribution gives the smallest KS statistic with a p-value of almost 0.05 - partially agreeing with the information criteria result (with the lognormal distribution giving the second lowest AIC and BIC values in this case).

On the other hand, we find that the true distributions of degree density of the complement of market graphs are much more similar, for each threshold value θ, across the three periods. This may be due to the similarity in the shape of the distributions of the negative correlation coefficients for each of the three periods analysed (Figure 1). Visually, the Fréchet and lognormal distributions appear to give the best fit overall to the true distributions of degree density. For θ negative and close to (further from) zero, the lognormal distribution (GPD) gives the lowest AIC and BIC values. Indeed, in periods 2 and 3 the lowest AIC and BIC values from the lognormal distribution are followed by the Fréchet distribution giving the second lowest values for both criteria, showing agreement with the visual fit.

The KS test also appears to show agreement with these results, as in period 1 the lognormal distribution gives the smallest KS statistics with statistically significant p-values, as high as 0.15, at θ values negative and close to zero. In the second period, the Fréchet distribution gives the lowest KS statistics with significant p-values, for θ negative and close to zero. In the third and final period, for θ negative and close to zero, the lognormal and Fréchet distributions give the lowest KS statistics (in comparison with the GPD and Pareto I distribution) with significant p-values.

Many recent papers that have studied financial networks and the construction of network graphs have explored the alternative statistical route of multiple decision procedures. For example, in Koldanov et al. Koldanov et al. (2013) the construction of sample market graphs for a financial network (as in our analysis) is the procedure of the identification of the true market graph, and

this identification problem is treated as a multiple decision problem of the selection of one from a set of hypotheses.

Koldanov et al. Koldanov et al. (2013) compute a simulation study in which the random vector of random variables representing the daily returns of a stock are distributed according to a multivariate normal distribution or a multivariate Student-t distribution. The main results from the study are that the method of market graph construction is i) "optimal in the class of all unbiased statistical procedures if, for the generating hypotheses testing, one uses the tests of the Neyman structures"; ii) "optimal in a restricted class of unbiased statistical procedures if, for the generating hypotheses testing, one uses the classical Pearson correlation tests" (Koldanov et al. Koldanov et al. (2013)). In the former case, edges connect two stocks (i, j) if the sample correlation $s_{i,j}$ exceeds a predetermined threshold $c_{i,j}$. However, whereas our threshold value θ was chosen arbitrarily, the threshold $c_{i,j}$ utilises the conditional distribution of $s_{i,j}$ from the Wishart distribution. In the latter case, edges connect two stocks (i, j) if the sample correlation $s_{i,j}$ exceeds a predetermined threshold c. Indeed, our method appears more similar to this case and is in fact optimal in this class for the special choice of a significance level $\alpha = 0.5$.

The statistical procedure of stock selection by using the Sharpe ratio is considered in Koldanov et al. (2015). The problem of stock selection is presented as a multiple decision problem of choosing (by observations) one hypothesis from a set of hypotheses. In the study, two well known multiple testing statistical procedures are compared, these are the Holm step down procedure and the Hochberg step up procedure. Given a set of observations of a random variable describing a characteristic of stocks in a financial market, stocks are selected if the condition $Sh_i > Sh_0$ is satisfied, that is, the Sharpe ratio of a stock (Sh_i) exceeds a threshold value (Sh_0). The conditional risk of the two methods as a function of the threshold is explored. For a real financial market which shows a concentration of Sharpe ratios, it is shown to be more appropriate to use the Hochberg procedure for stock selection; in all other cases, the Holm procedure may be better suited. Arguably, this could be viewed as an alternative method for stock selection to what could be achieved through the method used in our paper. By constructing the complements of the market graphs for threshold values θ negative and less than zero, one could select stocks and build a portfolio of stocks which are interconnected in these graphs. For example, this may enable one to choose stocks whose returns move in opposite directions, which would offset each other and help to diversify risk in an investment portfolio.

Kalyagin et al. Kalyagin et al. (2016) propose a general method for multivariate network construction based on measures of association. In relation to Kalyagin et al. Kalyagin et al. (2016), our analysis can be categorised as the construction of threshold graphs for UK financial networks, with the aim of recovering the reference threshold graph from observations. This provides the simplest statistical procedure for the identification of the reference threshold graph (a subgraph of the reference network). Due to the measure of association chosen in our method, the network graphs constructed can be classed as Pearson correlation networks. Kalyagin et al. Kalyagin et al. (2016) show that the problem of identifying the reference threshold graph, from observations of attributes of nodes in a multivariate network, can be solved in terms of a multiple decision problem — again, as the selection of one from many hypotheses. In their study, they conclude that for a random vector of nodes in a network with a fixed multivariate distribution, it is possible to construct various reference networks which relate to a particular measure of association, e.g. Pearson correlation network, Kendall correlation network, Spearman correlation network, etc.

6 Conclusions

In our analysis, we have fitted four distributions: i) Pareto I distribution; ii) Fréchet distribution; iii) lognormal distribution; iv) generalised Pareto distribution, to the degree density of financial network graphs constructed from correlation coefficients of pairs of stocks in the stock markets, according to various threshold values of $\theta \in [-1, 1]$, over three different time periods. Our data set consisted of the daily closing prices of stocks listed on one of three UK stock markets: FTSE100; FTSE250; FTSEAIM, over three time periods each of two years in length covering: 2000 - 2002; 2006 - 2008; 2012 - 2014.

We followed Boginski et al. Boginski et al. (2005)'s method to construct the network graphs, using positive threshold values of θ for market graphs and negative threshold values θ for the complement of the market graphs. Although our data set consists of UK stock data, as opposed to Boginski et al. Boginski et al. (2005) who considered US stock data, our analysis covers the same time period as Boginski et al. Boginski et al. (2005) in addition to two more recent periods.

Our results show that for the degree density of market graphs, constructed using positive values of the threshold θ, none of the fitted models performs significantly better than others. Observing the plots of the distributions, it appears that (in general) for low values of θ the Fréchet, lognormal and Pareto I distributions give a reasonable fit to the lower part of the true distribution, whilst the GPD gives a reasonable fit to the upper part. For higher values of θ the GPD may give the best fit overall. However, according to the information criteria, the GPD would be the preferred model out of the fitted distributions, for all θ across the three time periods.

On the contrary, for the degree density of the complement of the market graphs, constructed using negative values of the threshold θ, the Fréchet and lognormal distributions appear to give the best fit with θ close to zero. This is supported by the fact that they give the lowest and second lowest AIC and BIC values when θ is close to zero. Furthermore, the lowest (and significant) KS statistics are also either given by the lognormal and Fréchet distributions for θ close to zero.

This work could be extended to analyse stock market data for other countries, in particular for the US. For example, we note that for our data set the distributions of the degree density of market graphs in the second and third periods differ from those in the first. In particular, when using a low threshold θ the true distributions in periods 2 and 3 appear to be bimodal. It would therefore be interesting to see whether this pattern can be found, over the same time periods, when using data from other international stock markets, such as the US, or whether this may be due to the dependence and correlation between stocks in the UK markets only. In addition, it would be of interest to test whether the Fréchet and lognormal distributions fit well to the degree density of the complement of market graphs, for other stock data too.

Also, one measure of the goodness of fit used was the information criteria, the Akaike and Bayesian information criteria. Both of these measures are dependent on the number of estimated parameters in the model, k, however, in our analysis our four fitted models each have $k = 2$ estimated parameters. A natural extension to this could be to extend the number of distributions fitted (including those with a different number of parameters) and utilise a greater number of goodness of fit tests.

Finally, we acknowledge that our results only scratch the surface in terms of network analysis of stock market graphs. With regards to the UK stock market network, the next step would be

to evaluate and analyse some of the many network measures, such as network disparity; graph centrality; closeness centrality; domination power; clustering coefficients; minimum spanning tree. This would allow for a deeper understanding of the complexities of the UK stock market network. Another step is to use more flexible distributions to account for bimodality of some of the data sets.

Acknowledgments

The authors would like to thank the referee and the Editor for careful reading and comments which improved the paper.

References

H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19:716–723, 1974.

V. Boginski, S. Butenko, and P. Pardalos. Statistical analysis of financial networks. *Computational Statistics and Data Analysis*, 48:431–443, 2005.

P. Brockwell and R. Davis. *Introduction to Time Series and Forecasting*. Springer Verlag, New York, 2002.

A. Bunn, D. Urban, and T. Keitt. Landscape connectivity: A conservation application of graph theory. *Journal of Environmental Management*, 59:265–278, 2000.

K. Burnham and D. Anderson. Multimodel inference: Understanding aic and bic in model selection. *Sociological Methods and Research*, 33:261–304, 2004.

K. Chou. Applications of graph theory to enzyme kinetics and protein folding kinetics. steady and non-steady-state systems. *Biophysical Chemistry*, 35:1–24, 1990.

Datastream International. *Datastream International*. 2016.

L. Dobrjanskyj and F. Freudenstein. Some applications of graph theory to the structural analysis of mechanisms. *Journal of Engineering for Industry*, 89:153–158, 1967.

J. Durbin and G. Watson. Testing for serial correlation in least squares regression i. *Biometrika*, 37:409–428, 1950.

J. Durbin and G. Watson. Testing for serial correlation in least squares regression ii. *Biometrika*, 38:159–178, 1951.

J. Durbin and G. Watson. Testing for serial correlation in least squares regression iii. *Biometrika*, 58:1–19, 1971.

Y. Fang. Asymptotic equivalence between cross-validations and akaike information criteria in mixed-effects models. *Journal of Data Science*, 9:15–21, 2011.

W. Huang, X. Zhuang, and S. Yao. A network analysis of the chinese stock market. *Physica A: Statistical Mechanics and its Applications*, 388:2956–2964, 2009.

V. Kalyagin, A. Koldanov, and P. Pardalos. On multivariate network analysis of statistical data sets with different measures of association. *Annals of Mathematics and Artificial Intelligence*, 76:83–92, 2016.

A. Koldanov, P. Koldanov, V. Kalyagin, and P. Pardalos. Statistical procedures for the market graph construction. *Computational Statistics and Data Analysis*, 68:17–29, 2013.

A. Koldanov, V. Kalyagin, and P. Pardalos. Step down and step up statistical procedures for stock selection with sharp ratio. *Lecture Notes in Computer Science*, 9432:26–36, 2015.

A. Kolmogorov. Sulla determinazione empirica di una legge di distribuzione. *Giornale dell'Istituto Italiano degli Attuari*, 4:83–91, 1933.

L. Kullmann, J. Kertesz, and K. Kaski. Time-dependent cross-correlations between different stock returns: A directed network of influence. *Physical Review E*, 66, 2002.

L. Ping and W. Binghong. An approach to hang seng index in hong kong stock market based on network topological statistics. *Chinese Science Bulletin*, 51:624–629, 2006.

R Development Core Team. *R: A Language and Environment for Statistical Computing*, 2016.

G. Schwarz. Estimating the dimension of a model. *Annals of Statistics*, 6:461–464, 1978.

N. Smirnov. Table for estimating the goodness of fit of empirical distributions. *Annals of Mathematical Statistics*, 19:279–281, 1948.

B. Tabak, T. Serra, and D. Cajueiro. Topological properties of stock market networks: The case of brazil. *Physica A: Statistical Mechanics and Its Applications*, 389:3240–3249, 2010.

C. Tse, J. Liu, and F. Lau. A network perspective of the stock market. *Journal of Empirical Finance*, 17:659–667, 2010.

A. Vizgunov, B. Goldengorin, V. Kalyagin, A. Koldanov, P. Koldanov, and P. Pardalos. Network approach for the russian stock market. *Computational Management Science*, 11:45–55, 2014.

J. WS, S. Chae, J. Yang, and M. HT. Characteristics of the korean stock market correlations. *Physica A: Statistical Mechanics and its Applications*, 361:263–271, 2006.

J. Zhang, Y. Chen, and D. Zhai. Network analysis of shanghai sector in chinese stock market based on partial correlation. In *Proceedings of the 2nd IEEE International Conference on Information Management and Engineering (ICIME)*, pages 321–324, 2010.

X. Zhuang, Z. Min, and S. Chen. Characteristic analysis of complex network for shanghai stock market. *Journal of Northeastern University (Natural Science)*, 2007.