8th International Scientific Conference
"Business and Management 2014"
May 15–16, 2014, Vilnius, LITHUANIA
Section: Finance Engineering
http://www.bm.vgtu.lt

ISSN print 2029-4441 / ISSN online 2029-929X
ISBN print 978-609-457-650-8 / ISBN online 978-609-457-649-2
Article number: bm.2014.037
http://dx.doi.org/10.3846/bm.2014.037

CLUSTERING IN EUROPEAN STOCK INDICES IN CRISIS AND NON-CRISIS PERIODS

Sinem Peker[1], Bora Aktan[2]

[1]*Yaşar University, Faculty of Science and Letters,*
Deparment of Statistics
Bornova, İzmir, Turkey
Email: sinem.peker@yasar.edu.tr

[2]*University of Bahrain, College of Business Administration,*
Department of Economics & Finance
PO Box 32038 Isa Town, Kingdom of Bahrain
Email: gbora@uob.edu.bh

Abstract. Grouping the major indices of stock markets based on their homogeneities may facilitate the selection period for investors especially today's information rich financial world. This paper attempts to detect and group the homogenous stock indices in Europe both throughout the crisis and non-crisis periods. The daily index returns of leading stock exchanges over the period 03.01.2007-09.04.2013 are considered; one of the hierarchical clustering techniques so-called Ward's Method is applied and similar cases are evaluated respectively. Then, Wilcoxon signed rank test is employed for the same periods on daily index returns and meaningful differences are found.

Keywords: Cluster analysis, European stock exchanges, Index, financial crisis, non-parametric test.

JEL classification: C38, G15, C14, G11.

1. Introduction and background

Today's information-rich financial markets in which complex instruments are traded using quantitative techniques is inevitable to grasp the behaviour of prices and returns. Clustering which separates the objects or variables into homogenous groups practices in many disciplines such as medical sciences, marketing, economics, finance, and etc. (Cramer 2006). In the field of finance, unlike factor analysis is a rarely used technique. In the literature, a small number of studies employing various clustering methods to a given classification problem can be found such as Aktan (2013), Babu *et al*. (2012), Narayan *et al*. (2011), Ingram and Margetis (2010), Pattarin *et al.* (2004) among others. Pattarin *et al*. (2004), worked on equity funds in Italy by proposing and testing a classification algorithm to identify clusters of funds. Da Costa *et al.* (2005) applied hierarchical clustering and Ward methods on key stocks picked from North and South America and grouped the stocks through risk/return, earnings/price, book value/price, sales/price, sales/outstanding shares and dividend yield. Nanda *et al*. (2010) suggested an integration of cluster techniques with fuzzy ones to construct efficient portfolios via hybrid systems in Bombay stock exchange and found that k-means builds the most compact clusters. More recently, Narayan *et al.* (2011), examined share price clustering on twelve largest companies listed on Mexican stock exchange and pointed out that volume and risk impact price clustering negatively. Aktan (2013) examined if the companies listed Bahrain Bourse can be clustered based on risk and return by employing Ward's method and grouped them in three main clusters. Babu *et al.* (2012) analyzed the main clustering techniques to compare the performances and apply to 35 randomly selected stocks from a number of different sectors in India in order to be able to propose an effective method to predict the stock price movements. They indicated that the hierarchical agglomerative outperforms in terms of accuracy. Wang and Wang (2012) used hierarchical clustering to determine deceptive financial reporting. Sarlin and Peltonen (2013) proposed Self-Organizing Financial Stability Map in order to represent the condition of financial stability. They used Ward's hierarchical clustering to select a large number of Self Organizing Maps in their process. Basetto and Kalatzis (2011) examined the effect of financial constraint in the investment decision of Brazilian firms and they applied clustering technique to detect the group of firms which are used in the models. Via a hierarchical clustering, Bouvatier *et al.* (2012) were able to obtain the classification of the banking system in the investigation of the influence of banking sector structure in the explaination of the credit procyclicality. D'Urso *et al.* (2013) handled the clustering of financial time series and proposed a new approach which combines fuzziness and GARCH models. Cinca *et al.* (2005) applied Ward's hierarchical clustering to group the combination of country and size of firms with respect to financial ratios in order to detect different financial and econometric patterns. Tola *et al.* (2008) underlined the importance of clustering technique in the advancement of the reliability of the portfolio considering the ratio between predicted and realized risk. Chen and Huang (2009) applied cluster analysis to group the huge amount of equity mutual funds based on four evaluation indices in order to help investment decisions. In addition, they offered a fuzzy model which gives the optimal investment proportion of each cluster. Lange and Sauer (2005) searched the seigniorage costs of official dollarization in 15 Latin American Countries and performed cluster analysis to group countries which reflect different cases.

The purpose of this study is to test Ward's hierarchical clustering method in order to group the indices of major stock exchanges including UK, France, Italy, Germany, Spain, Greece, Switzerland and Russia and analyze the crisis and non-crisis periods. In addition, Europe's blue-chip index, STOXX-EUROPE 50 which covers fifty stocks from 12 Eurozone countries and STOXX-EUROPE 600 index which represents large, mid and small capitalization companies across 18 countries in Europe are included. Considering the squared distances between daily index returns, the similar and the dissimilar stock indices patterns are detected with the help of the clustering. For this reason only, this study could help not only investors but also analysts in identifying trends of major indices of Europe in turbulence and stable times. The rest of the study is organized as follows: Next section describes the data and introduces the method, third section discusses the empirical results whereas the last section gives concluding remarks.

2. Data and methodology

The daily index figures for FTSE100, CAC40, FTSEMIB, ATHEN, DAX, SMI, IBEX35, RTS, STOXX50 and 600 are obtained from Yahoo finance over the period 03.01.2007-09.04.2013. Ward's hierarchical clustering method is applied for 03.01.2007-31.08.2009, 31.08.2009-09.04.2013 and 03.01.2007-09.04.2013 periods. The differences between crisis and non-crisis periods are interpreted. The wilcoxon signed rank test is applied on daily index returns for the crisis and non-crisis periods.

In cluster analysis, the similarities among objects or variables are taken into account and the objects are distinct by satisfying the homogeneity within groups, the heterogeneity between groups (Saunders 1994; Han, Kamber 2001; Timm 2002; Larose 2005; Hair *et al*. 2007). Hierarchical clustering and k-means clustering are two well-known and widely used clustering methods in the researches. In hierarchical clustering, there are divisive and agglomerative methods. The divisive begins with one cluster that includes all objects and then the objects are separated into groups until each objects set a cluster. In agglomerative method, each object set a cluster at the beginning and then the objects are merged until all objects are put in a single group. There are several hierarchical clustering methods according to the calculations of distances between two clusters (Sharma 1996). In single linkage, complete linkage, average linkage, and centroids methods, the following distance measures are used, respectively (Han and Kamber 2001).

Minimum Distance:

$$d_{\min}(C_h, C_g) = \min_{y \in C_h, y' \in C_g} \left| y - y' \right| \quad (1)$$

Maximum Distance:

$$d_{\max}(C_h, C_g) = \max_{y \in C_h, y' \in C_g} \left| y - y' \right| \quad (2)$$

Average Distance:

$$d_{avg}(C_h, C_g) = \frac{1}{N_h N_g} \sum_{y \in C_h} \sum_{y' \in C_g} \left| y - y' \right| \quad (3)$$

Mean Distance:

$$d_{mean}(C_h, C_g) = \left| \bar{y}(h) - \bar{y}(g) \right| \quad (4)$$

2.1. Ward's hierarchical method

The Ward's hierarchical method, which is also called the method of the minimum variances, is one of the most used clustering method in the applications (Sala, Bragulat 2004). In this method, the homogenity within cluster and the minimization of the loss of information from joining two groups are considered (Sharma 2007; Johnson, Wichern 2002). This method begins with N elements, each element is considered a cluster at the beginning. The similarity matrix is constructed and the most similar pair is searched where the minimum increase in the total within group error sum of squares, W_{SSE}, is satisfied (Sala, Bragulat 2004)

$$W_{SSE} = \sum_{k=1}^{C}\left(\sum_{j=1}^{P}\left(\sum_{i=1}^{N_i}(y_{ijk} - \bar{y}_{jk}(i))^2\right)\right), \quad (5)$$

where:

$$\bar{y}_{jk}(i) = \frac{1}{N_i}\sum_{i=1}^{N_i} y_{ijk} ,$$

$$\sum_{i=1}^{K} N_i = N ,$$

and C, P denote the number of groups, variables, respectively; N_i is the number of the elements in each group.

3. Application and empirical results

In this study, the Ward's algorithm is applied to cluster the leading stock indices of Europe in terms of daily index returns. Daily index returns over the examined periods are computed as logarithmic price relatives. The desciptive statitistics of daily index returns over the period of 31.08.2009-09.04.2013 are obtained in Table 1 and it is observed that the maximum standard deviation belongs to Athen index among the handled indices. The Wilcoxon signed rank test is applied on index returns over the period aforementioned and meaningful differences found are given in Table 2.

The desciptive statitistics of the daily index returns over the crisis period of 03.01.2007-31.08.2009 are evaluated and found that the maximum standard deviation belongs to RTS_Russia Index shown in Table 3. Accordingly, it is interpreted that there is a decreasing in the standard deviation of RTS_Russia Index in 31.08.2009- 09.04.2013 period when compared to 03.01.2007-31.08.2009 period.

The Wilcoxon signed rank test is applied on index returns over the crisis period meaningful differences found are given in Table 4.

Although, there are differences between "IBEX35_Spain and FTSEMIB_Italy"; "STOXX_EUROPE600 and DAX_German" index returns in the period of 03.01.2007-31.08.2009, these differences are disappeared in the period of 31.08.2009- 09.04.2013. However, the difference between "DAX_German and STOXX_EUROPE 50" is observed in both periods.

The Ward's Hierarchical results are summarized in Figure 1- 3 and are illustrated in Table 5.

Table 1. Descriptive statistics of daily share returns in the period of 31.08.2009 - 09.04.2013

	Mean	Median	Std. Deviation	Skewness	Kurtosis	Minimum	Maximum
FTSE100_ENGLAND	,00022939	,00056003	,010840172	-,179	1,873	-,047792	,050323
CAC40_FRANCE	-,00001475	,00019591	,014863925	,057	2,989	-,056346	,092208
FTSEMIB_ITALY	-,00040521	,00028365	,017656754	-,023	2,486	-,070442	,106839
RTS_RUSSIA	,00002376	,00093044	,017446928	-,440	2,332	-,090052	,068023
ATHEN_INDEX	-,00101072	-,00020350	,022554130	,322	2,380	-,073664	,134311
STOXX_EUROPE50	-,00010886	-,00025600	,015297774	,128	3,222	-,063182	,098466
DAX_GERMAN	,00034103	,00066830	,013750716	-,174	2,203	-,059947	,052104
STOXX_EUROPE600	,00021112	,00039023	,011435456	-,117	2,981	-,048853	,069066
SMI_SWISS_MARKET	,00029531	,00050050	,009668334	-,185	3,388	-,042428	,049029
IBEX35_SPAIN	-,00037567	-,00025155	,017244568	,402	4,993	-,068739	,134836

Table 2. The Wilcoxon Signed Test Results for 31.08.2009- 09.04.2013

Pairs	Z	p
IBEX35_SPAIN - FTSE100_ENGLAND	$-2,334^{b}$	,020
DAX_GERMAN - CAC40_FRANCE	$-2,518^{c}$	,012
DAX_GERMAN - FTSEMIB_ITALY	$-2,544^{c}$	,011
ATHEN_INDEX - RTS_RUSSIA	$-2,247^{b}$	,025
STOXX_EUROPE50 - ATHEN	$-1,995^{c}$	,046
DAX_GERMAN - ATHEN_INDEX	$-2,215^{c}$	,027
STOXX_EUROPE600 - ATHEN	$-2,356^{c}$	,018
SMI_SWISS_MARKET - ATHEN	$-2,281^{c}$	,023
DAX_GERMAN - STOXX_EUROPE50	$-3,771^{c}$	,000
IBEX35_SPAIN - DAX_GERMAN	$-3,053^{b}$	,002
IBEX35_SPAIN - STOXX_EUROPE600	$-2,298^{b}$	,022

b. Based on positive ranks, c. Based on negative ranks,
p: Asymp. Sig. (2-tailed)

Table 3. Descriptive statistics of daily share returns in the period of 03.01.2007-31.08.2009

	Mean	Median	Std. Deviation	Skewness	Kurtosis	Minimum	Maximum
FTSE100_ENGLAND	-,00052792	-,00011822	,017979860	-,026	5,197	-,092646	,093842
CAC40_FRANCE	-,00081275	-,00041126	,019349445	,195	5,603	-,094715	,105946
FTSEMIB_ITALY	-,00099400	-,00012529	,019230873	,108	5,111	-,085991	,108742
RTS_RUSSIA	-,00104617	,00083799	,031580962	-,249	8,871	-,211994	,202039
ATHEN_INDEX	-,00120542	,00003592	,019056123	-,496	3,349	-,102140	,083283
STOXX_EUROPE50	-,00074361	-,00026382	,019226116	,088	5,045	-,082079	,104376
DAX_GERMAN	-,00040755	,00054060	,018585053	,291	6,386	-,074335	,107975
STOXX_EUROPE600	-,00071021	-,00018246	,017777358	,010	4,696	-,079297	,094148
SMI_SWISS_MARKET	-,00063835	-,00010287	,016485599	,204	5,489	-,081078	,107876
IBEX35_SPAIN	-,00051091	,00047063	,018824560	,007	5,115	-,095859	,101176

Table 4. The Wilcoxon Signed Test Results for 03.01.2007-31.08.2009

Pairs	Z	p
IBEX35_SPAIN - FTSEMIB_ITALY	-2,076[c]	,038
DAX_GERMAN - STOXX_EUROPE50	-2,053[c]	,040
STOXX_EUROPE600 - DAX_GERMAN	-2,041[b]	,041

b. Based on positive ranks, c. Based on negative ranks,
p: Asymp. Sig. (2-tailed)

Table 5. Ward's Hierarchical Clustering Results

	03-01-2007-09.04.2013 The number of clusters			09.04.2013-31.08.2009 The number of clusters			03.01.2007-31.08.2009 The number of clusters		
	3	4	5	3	4	5	3	4	5
FTSE100_ENGLAND	1	1	1	1	1	1	1	1	1
CAC40_FRANCE	1	2	2	1	1	2	1	1	1
FTSEMIB_ITALY	1	2	3	1	2	3	1	1	2
RTS_RUSSIA	2	3	4	2	3	4	2	2	3
GD.AT_ATHEN_INDEX	3	4	5	3	4	5	3	3	4
STOXX_EUROPE50	1	2	2	1	1	2	1	1	1
DAX_GERMAN	1	2	2	1	1	2	1	1	1
STOXX_EUROPE600	1	1	1	1	1	1	1	1	1
SMI_SWISS_MARKET	1	1	1	1	1	1	1	4	5
IBEX35_SPAIN	1	2	3	1	2	3	1	1	1

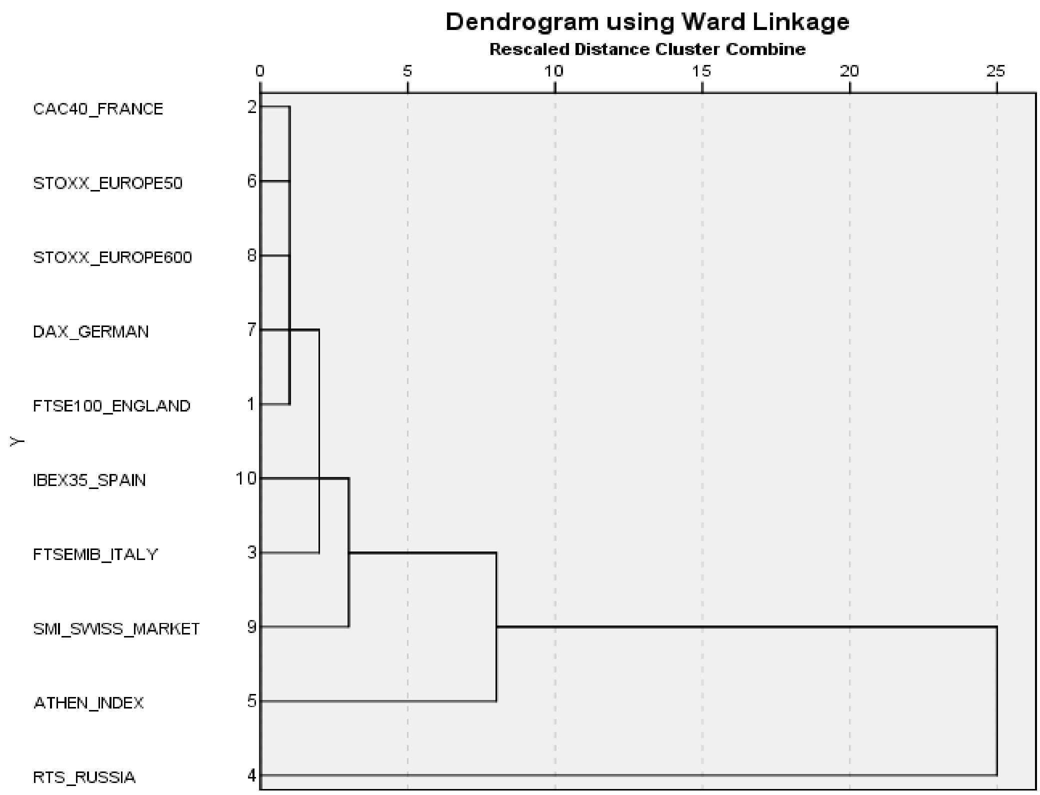

Fig. 1. Ward's Hierarchical Clustering Results in the period of 03.01.2007-31.08.2009

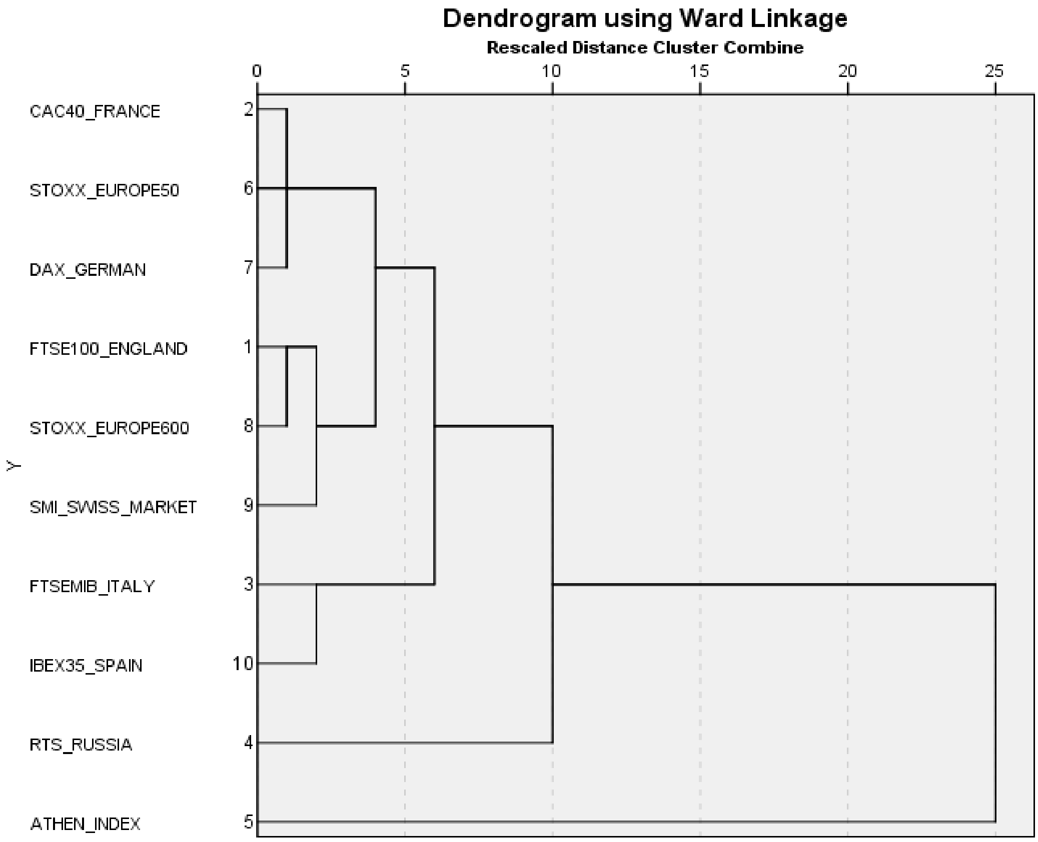

Fig. 2. Ward's Hierarchical Clustering Results in the period of 31.08.2009-09.04.2013

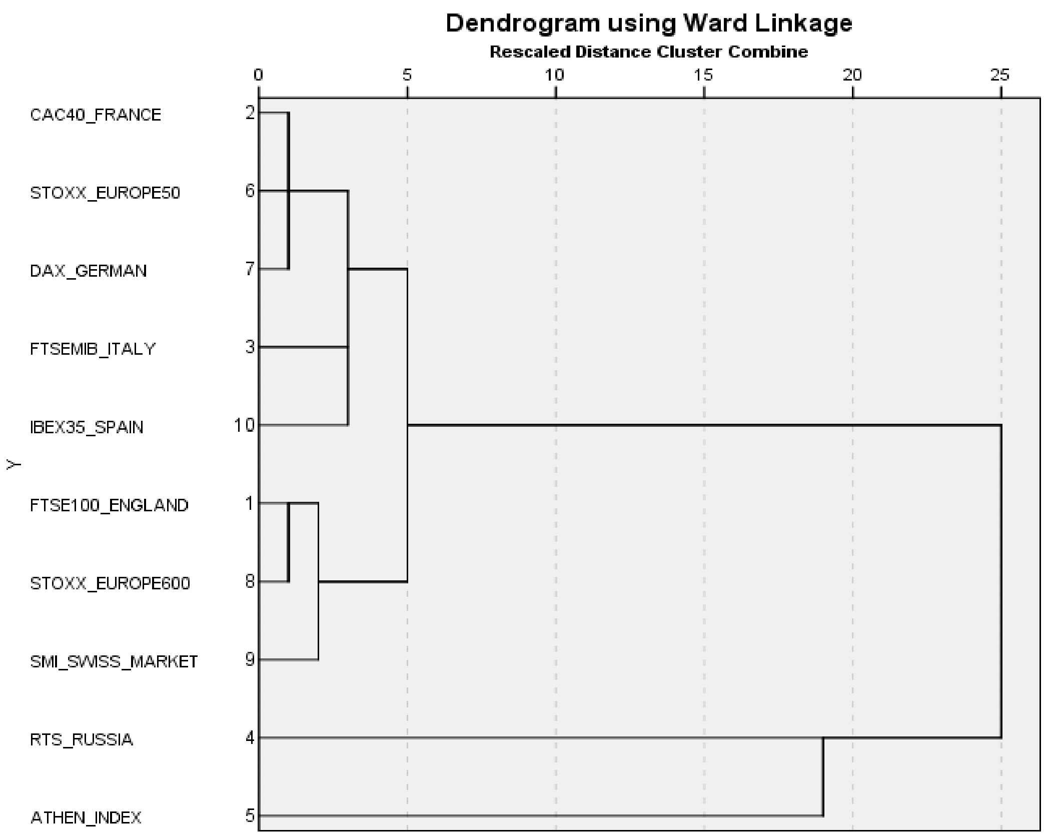

Fig. 3. Ward's Hierarchical Clustering Results in the period of 03-01-2007-09.04.2013

4. Conclusions

In this study, we discuss and examine if major stock exchange indices of Europe can be grouped in terms of daily index returns during and in the wake of crisis. Based on the analyses, RTS and ATHEN differ visibly from their peers over the both periods. In addition, results indicated that STOXX-EUROPE 600 is more similar with ENGLAND and SWISS in terms of the squared distances between daily index returns for the period of 31.08.2009- 09.04.2013 (refer to Table 5). In order to be able to see the differences between index returns, the Wilcoxon signed rank test is applied on daily index returns and meaningful differences are found in the following stock exchange pairs over the period of 31.08.2009- 09.04.2013:

- "SPAIN-ENGLAND",
- "GERMANY-FRANCE",
- "GERMANY-ITALY",
- "GREECE-RUSSIA",
- "STOXX_EUROPE50-GREECE",
- "GERMANY-GREECE",
- "STOXX_EUROPE600-GREECE",
- "SWISS-GREECE",
- "GERMANY-STOXX_EUROPE50",
- "SPAIN-GERMANY",
- "SPAIN-STOXX_EUROPE600"

Besides, the differences between "IBEX35_Spain and FTSEMIB_Italy"; "STOXX_EUROPE600 and DAX_German" are dedected in the period of 03.01.2007 - 31.08.2009. However, it is observed that these differences are disappeared in the period of 31.08.2009- 09.04.2013. Yet, the difference between "DAX_German and STOXX_EUROPE 50" returns is seen in both 03.01.2007-31.08.2009 and 31.08.2009- 09.04.2013 periods.

References

Aktan, B. 2013. An Empirical Examination of Clustering at Bahrain Bourse (BHB), *International Research Journal of Finance and Economics* 110: 149–154.

Babu, M. S.; Geethanjali, N.; Satyanarayana, B. 2012. Clustering Approach to Stock Market Prediction, *International Journal of Advanced Networking and Applications* 3(84): 1281–91.

Basetto, C. F.; Kalatzis, A. E. G. 2011. Financial distress, financial constraint and investment decision: Evidence from Brazil, *Economic Modelling* 28: 264–271. http://dx.doi.org/10.1016/j.econmod.2010.09.003

Bouvatier, V.; Villavicencio, A. L.; Mignon, V. 2012. Does the banking sector structure matter for credit procyclicality?, *Economic Modelling* 29: 1035–1044. http://dx.doi.org/10.1016/j.econmod.2012.03.012

Chen, L. H.; Huang, L. 2009. Portfolio optimization of equity mutual funds with fuzzy return rates and risks, *Expert Systems With Applications* 36: 3720–3727. http://dx.doi.org/10.1016/j.eswa.2008.02.027

Cinca, C. S.; Mar Lalinero, C.; Larraz, J. L. G. 2005. Country and size effects in financial ratios: A European perspective, *Global Finance Journal* 16: 26–47. http://dx.doi.org/10.1016/j.gfj.2005.05.003

Cramer, D. 2006. *Advanced Quantitative Data Analysis,* Open University Press, McGraw Hill Education, New York

Da Costa, N.; Cunha, J.; Da Silva, S. 2010. Stock Selection Based on Cluster Analysis, *Economics Bulletin* 13(1): 1–9.

D'Urso, P.; Cappelli, C.; Di Lallo, D.; Massari, R. 2013. Clustering of financial time series, *Phsica A* 392: 2114–2129. http://dx.doi.org/10.1016/j.physa.2013.01.027

Hair, J. F.; Black, B.; Babin, B.; Anderson R.E., Tatham, R. L. 2007. *Multivariate Data Analysis*, Prentice Hall.

Han, J.; Kamber, M. 2001. *Data Mining: Concepts and Techniques*, Morgan Kaufmann.

Ingram, M.; Margetis, S. 2010. A Practical Method to Estimate the Cost of Equity Capital for a Firm Using Cluster Analysis, *Managerial Finance* 36(2): 160–67. http://dx.doi.org/10.1108/03074351011014569

Johnson, R. A.; Wichern, D. W. 2002. *Applied Multivariate Statistical Analysis*, Prentice-Hall.

Lange, C.; Sauer, C. 2005. Dollarization in Latin America: seigniorage costs and policy implications, *The Quarterly Review of Economics and Finance* 45: 662–679. http://dx.doi.org/10.1016/j.qref.2004.01.002

Larose, D. T. 2005. *Discovering Knowledge in Data: An introduction to Data Mining*, Wiley

Nanda, S. R.; Mahanty, B.; Tiwari, M. K. 2010. Clustering Indian Stock Market Data for Portfolio Management, *Expert Systems with Applications,* 37: 8793–98. http://dx.doi.org/10.1016/j.eswa.2010.06.026

Narayan, P. K.; Narayan, S.; Popp, S.; D'Rosario, M. 2011. Share Price Clustering in Mexico, *International Review of Financial Analysis* 20: 113–19. http://dx.doi.org/10.1016/j.irfa.2011.02.003

Patterin, F.; Paterlini, S.; Tommaso, M. 2004. Clustering Financial Time Series: An Application to Mutual Funds Style Analysis, *Computational Statistics and Data Analysis* 47: 353–72. http://dx.doi.org/10.1016/j.csda.2003.11.009

Sala, C. H.; Bragulat, E. J. 2004. 'A program to perform Ward's clustering method on several regionalized variables', *Computers & Geosciences* 30: 881–886. http://dx.doi.org/10.1016/j.cageo.2004.07.003

Sarlin, P.; Peltonen, T. A. 2013. Mapping the state of Financial Stability, *Journal of International Markets, Institutions&Money* 26: 46–76.

Saunders, J. 1994. Cluster Analysis, *Journal of Marketing Management* 10: 13–28. http://dx.doi.org/10.1080/0267257X.1994.9964257

Sharma, S. 1996. *Applied Multivariate Techniques'*, Wiley.

Timm, N. H. 2002. *Applied Multivariate Analysis*, Springer.

Tola, V.; Lillo, F.; Gallegati, M.; Montegna, R. N. 2008. Cluster analysis for portfolio optimization, *Journal of Economic Dynamics & Control,* 32: 235–258. http://dx.doi.org/10.1016/j.jedc.2007.01.034

Wang, B.; Wang, X. 2012. Deceptive Financial Reporting Detection: A Hierarchical Clustering Approach Based on Linguistic Features, *Procedia, Engineering* 29: 3392–3396. http://dx.doi.org/10.1016/j.proeng.2012.01.500