Dipartimento di scienze economiche, aziendali, matematiche e statistiche "Bruno de Finetti"

Research Paper Series, N. 3, 2015

A portfolio diversification strategy via tail dependence measures

Fabrizio Durante
Faculty of Economics and Management
Free University of Bozen-Bolzano, I-39100 Bolzano (Italy)

Enrico Foscolo
Faculty of Economics and Management
Free University of Bozen-Bolzano, I-39100 Bolzano (Italy)

Roberta Pappadà
Department of Economics, Business, Mathematics and Statistics "Bruno De Finetti"
University of Trieste, I-34127 Trieste (Italy)

Hao Wang
School of Economics
Jilin University, 2699 Qianjin Street, Changchun (China)

UNIVERSITÀ DEGLI STUDI DI TRIESTE

Research Paper Series

Dipartimento di Scienze Economiche, Aziendali, Matematiche e Statistiche
"Bruno de Finetti"
Piazzale Europa 1
34127, Trieste
Tel. +390405587927
Fax +390405587033
http://www.deams.units.it

EUT Edizioni Università di Trieste
Via E.Weiss, 21 - 34128 Trieste
Tel. +390405586183
Fax +390405586185
http://eut.units.it
eut@units.it

ISBN: 978-88-8303-720-7

Research Paper Series, N. 3, 2015

A portfolio diversification strategy via tail dependence measures

Fabrizio Durante
Faculty of Economics and Management
Free University of Bozen-Bolzano, I-39100 Bolzano (Italy)

Enrico Foscolo
Faculty of Economics and Management
Free University of Bozen-Bolzano, I-39100 Bolzano (Italy)

Roberta Pappadà
Department of Economics, Business, Mathematics and Statistics "Bruno De Finetti"
University of Trieste, I-34127 Trieste (Italy)

Hao Wang
School of Economics
Jilin University, 2699 Qianjin Street, Changchun (China)

ABSTRACT[1]

We provide a two-stage portfolio selection procedure in order to increase the diversification benefits in a bear market. By exploiting tail dependence-based risky measures a first-step cluster analysis is carried out for discerning between assets with the same performance during risky scenarios. Then a mean-variance efficient frontier is computed by fixing a number of assets per portfolio and by selecting only one item from each cluster. Empirical calculations on the EURO STOXX 50 prove that investing on selected index components

[1]**Corresponding author:** Fabrizio Durante, Faculty of Economics and Management, Free University of Bozen-Bolzano, piazza Universit, 1 I-39100 Bolzano (Italy), email: fabrizio.durante@unibz.it ; tel: +39 0471 013493; fax: +39 0471 013009

in trouble periods may improve the risk-averse investor portfolio performance.

KEYWORDS: Cluster Analysis, Copulas, Portfolio Selection, Tail Dependence.

1. Introduction

In recent years financial markets have been characterized by an increasing globalization and a complex set of relationships among asset returns. Moreover, it has been recognized that the linkages among different assets vary across time and that their strength tends to increase especially during crisis periods Billio et al. (2012). The presence of a stronger dependence when markets are experiencing losses is of utmost interest from a risk manager perspective. In fact, in portfolio risk analysis a usual practice for minimizing the whole risk consists of adopting some diversification techniques, an issue widely debated since the the seminal work of Grubel (1968). Namely, it has been recognized that investors can reduce the risk of their portfolios through allocating their investments in various classes of financial instruments and/or categories of assets that would move in different ways in response to the same event. In other words, diversification benefits can be achieved when the comovements among the assets is taken into account. Therefore, the portfolio diversification issue naturally poses the question of investigating the relationship between financial time series and of checking whether they can be grouped together in a way that may be helpful to portfolio selection.

To provide a suitable way to diversify a portfolio taking into account the occurrence of extreme scenarios, clustering techniques for multivariate time series have been proposed in the literature (see, e.g., Levy and Sarnat (1970); Panton et al. (1976)), mainly based on indices of global dependence like Pearson correlation coefficient (see, e.g., Mantegna (1999); Tola et al. (2008)). However, it has been recognized that the presence of dependence in the tail of the joint distribution of the markets can mitigate the diversification effects precisely when they are needed most Forbes and Rigobon (2002); Bradley and Taqqu (2004). Thus, different clustering techniques have been recently applied in order to group financial time series that are similar in extreme scenarios. Such an approach includes clustering techniques based on the tail dependence coefficient (see, e.g., De Luca et al. (2010); De Luca and Zuccolotto (2011) and Durante et al. (2015)), or conditional measure of association, like Spearman's correlation, as done in Durante et al. (2014).

The aim of this contribution is to exploit such recently introduced clustering methods in order to develop a procedure to select a weighted portfolio in a group of assets. The whole methodology is intended to be used by an investor to have more insights into the relationships among different assets in crisis periods. In particular, it may serve to warn against the automatic use of standard portfolio selection procedures that may not work when the markets are expected to experiencing periods of high volatility.

The paper is organized as follows. Section 2 presents the main properties of the used cluster procedures, while Section 3 illustrates the novel methodology via an empirical application. Section 4 concludes.

2. The clustering procedure

In general, clustering procedures are based on the choice of a suitable dissimilarity matrix that expresses the relations among the time series under consideration. Traditionally, correlation–based clustering has been used to find groups in a portfolio of financial assets from the correlation matrix: see, for instance, Mantegna (1999). However, as known, Pearson's correlation is not a convenient dependence measure to be used outside the Gaussian (elliptical) distributions and, in particular, it does not give an accurate indication and understanding of the real dependence between risk exposures Embrechts et al. (2003).

In order to give an accurate estimation of the link between assets in the tail of the distribution, suitable measures of tail dependence may be used instead, as suggested, for instance, in De Luca and Zuccolotto (2011). Following these ideas, different clustering procedures have been implemented in Durante et al. (2014, 2015) in order to group time series according to their tail behaviour. These procedures are summarized below (for more details, please refer to the original papers).

Consider a matrix of d financial time series $(x_{it})_{t=1,\dots,T}$ $(i = 1, 2, \dots, d)$ representing the log–returns of different financial assets. We assume that each time series $(x_{it})_{t=1,\dots,T}$ is generated by the stochastic process $(\mathbf{X}_t, \mathcal{F}_t)$ such that, for $i = 1, \dots, d$,

$$X_{it} = \mu_i(\mathbf{Z}_{t-1}) + \sigma_i(\mathbf{Z}_{t-1})\varepsilon_{it}, \tag{1}$$

where $\mathbf{Z}_{t-1}$ depends on $\mathcal{F}_{t-1}$, the available information up to time $t-1$, and the innovations ε_{it} are distributed according to a distribution function F_i for each t. Moreover, the innovations ε_{it} are assumed to have a constant conditional distribution F_i (with mean zero and variance one, for identification) such that for every t the joint distribution function of $(\varepsilon_{1t}, \dots, \varepsilon_{dt})$ can be expressed in the form $C(F_1, \dots, F_d)$ for some copula C. Such a general model includes many multivariate time series models presented in the literature (see, for instance, Patton (2013)).

Then the following steps are implemented in order to group the time series into subgroups such that elements in each sub-group have strong tail dependence.

i. Choose a copula-based time series model (e.g. ARMA-GARCH copula model) in order to model separately the marginal behavior of each time series and the link between them.

ii. Estimate the (pairwise) tail dependence measure among the time series.

iii. Define a dissimilarity matrix by using the information contained in the tail dependence measures and apply a suitable cluster algorithm for grouping time series according to the tail behavior.

The steps are described below in detail.

a. Fitting the univariate time series

We fit an appropriate ARMA-GARCH model to each univariate time series by obtaining the estimates $\widehat{\mu}_i$ and $\widehat{\sigma}_i$ of the conditional mean and variance of these processes according to Eq. (1). Classical model selection procedures (e.g., Bayesian Information Criteria) and goodness-of-fit tests of homoscedasticity and uncorrelatedness of the residuals may confirm the adequacy of the fit. Moreover, using the estimated parametric models, we construct the standardized residuals given, for each $i = 1, \dots, d$, by

$$\widehat{\varepsilon}_{it} = \frac{x_{it} - \widehat{\mu}_i(\mathbf{Z}_{t-1})}{\widehat{\sigma}_i(\mathbf{Z}_{t-1})}.$$

The standardized residuals are converted to the pseudo–observations $z_{it} = F_i(\widehat{\varepsilon}_{it})$, where F_i may be estimated from a parametric model (Gaussian, Student–t, etc.) or by using the empirical distribution function.

As a result, $(z_{1t}, \dots, z_{dt})_{t=1,\dots,T}$ contains the information about the link (i.e. the copula) among the time series under consideration and, as such, it can be used in order to compute dependence measures or make inference about the copula of the time series. Notice that, if the marginal model is correctly specified, $(z_{1t}, \dots, z_{dt})_{t=1,\dots,T}$ forms (asymptotically) a random sample generated by the copula C. As such, dependence measures calculated from this sample are not biased by serially dependence and/or heteroscedasticity (see, for instance, Patton (2013)).

b. Estimating a tail dependence measure

Given the pseudo–observations from the original time series, in order to quantify the degree of tail dependence of a random vector (X, Y), we adopt two different measures:

- the lower tail dependence coefficient λ_L (shortly, TDC) Joe (1997), which is a measure that depends only on the copula C linking the variables under consideration via the formula
$$\lambda_L(C) = \lim_{t \to 0^+} \frac{C(t,t)}{t}. \tag{2}$$
- the conditional Spearman's correlation ρ_α that expressed the Spearman's correlation of (X, Y) given that X and Y are both under their α–quantile (for instance, $\alpha = 0.10$).

These measures express two different ways of looking at tail dependence since they focus, respectively, on asymptotic tail dependence (λ_L) and finite tail dependence (ρ_α). As regards the estimation of these quantities we rely on two specific techniques:

- The estimator of tail dependence coefficient is based on the procedure proposed in (Frahm et al., 2005, section 3.5)). Specifically, by using the results of Genest and Segers (2009), we obtain an estimate of the extreme-value copula in the domain of attraction of C, from which the estimate of the TDC of C is derived.
- The estimator of conditional Spearman's ρ_α is based on the procedure described in Dobrić et al. (2007); Durante and Jaworski (2010) that is related to the calculation of the Spearman's correlation in a sub-sample extracted from the pseudo–observations and dependent on the threshold α.

Both these estimations are obtained via non-parametric procedures and do not require any parametric estimation of the unknown copula linking the time series of interest. For more detail about the practical calculations of these coefficients, we refer to Durante et al. (2014, 2015).

c. The clustering algorithm

Once a measure of tail dependence has been computed for all pairs extracted from the time series, we have to transform it through a monotonic function f in such a way that the obtained dissimilarity between two time series is small when their tail dependence is high, and monotonically increases when their tail dependence decreases. Thus, for $i, j = 1, \dots, d$, we define $\Delta = (\Delta_{ij})$ whose elements are given by

$$\Delta_{ij} = \sqrt{2(1 - \widehat{m}_{ij})}, \tag{3}$$

where $\widehat{m}_{ij}$ is the tail dependence measure between time series i and j, that is estimated via one of the two above procedures.

Starting from the dissimilarity matrix defined in (3) we can perform a cluster analysis of the time series by different techniques. Here, for a comparative analysis, we focus on two methods:

- The hierarchical agglomerative algorithms start from the finest possible partition (i.e. each observation forms a cluster) and, hence, each level merges a selected pair of clusters into a new cluster according to the definition of the distance between two groups. Among all the agglomerative strategies we apply the complete linkage, which defines the distance between two clusters as the maximum distance between their individual components, and give overall good performances.

- Fuzzy clustering algorithm is a partitioning method that takes into account some ambiguity in the data, which often occurs. In fact, it allows each object to belong to one or more than one cluster according to a membership coefficient, that quantifies the degree of belonging of each object to a specific group. The main advantage of fuzzy clustering is that it yields much more detailed information on the data structure compared to other partitioning techniques. In order to perform fuzzy cluster analysis we can consider FANNY algorithm Kaufman and Rousseeuw (1990), which handles either interval-scaled measurements or dissimilarities. The algorithm aims at minimizing an objective function which is a kind of total dispersion, depending on dissimilarities and membership coefficients. Once the number of clusters is chosen, the algorithm returns some general information on the type of data and the actual membership for each object in each cluster are listed. Moreover, as in our case, a crisp partition of the financial assets can be determined from the membership value of each time series.

In both methods, the optimal number of clusters is chosen by the silhouette index Kaufman and Rousseeuw (1990); Hastie et al. (2009), which reflects the within-cluster compactness and between-cluster separation of a clustering.

A remark is needed here. In several cases (see, for instance, Nanda et al. (2010)) the clustering algorithms are applied directly to the time series data without any previous step (pre-filtering). For our purposes, such an approach could not be applied since it can produce bias results. In fact, it has been recognized that, in order to obtain a realistic description of the dependence among assets, heteroscedastic effects should be taken into account and treated in a suitable way (see, for instance, Forbes and Rigobon (2002)).

3. The portfolio selection

Once the clustering procedure is completed, the assets have been grouped into a predefined number K of clusters. Then our possible portfolio will be selected on the basis of the following strategy:

i. Determine all possible portfolios composed by K assets such that each asset belongs to a different cluster.

ii. For these portfolios, calculate the optimal weight assigned to each of its assets with classical Markowitz portfolio selection procedure Markowitz (1952). We recall that this procedure provides a general way to maximize investor's expected utility under certain conditions, namely to produce portfolios that are able to minimize the total portfolio variance.

iii. Given all possible portfolios composed in such a way, plot the graph of their standard deviation against their expected return.

iv. Determine the portfolios that are the vertices in the convex efficient frontier of the standard deviation/expected return graph.

According his/her preference the investor could hence choose one of the portfolios that are on the convex frontier.

In other words, the procedure has the following aspects:

- It suggests to select the assets of the portfolio by taking into account the grouping structure given by the clustering algorithms. Thus, two assets from the same group (cluster) cannot be chosen in the same portfolio.

- Once the assets have been selected, their weights are determined by classical methods, like Markowitz approach.

- All the portfolios composed in the previous two steps provide a graphical representation of the possible choices of the investor (see, for instance, Figure 1). Based on his/her information, one investment strategy could be selected.

- If no preference is required by the investor, the point with the smallest risk on the convex frontier, namely the global minimum variance portfolio, can be chosen.

A remark is needed here. As clarified above, the idea of diversification by grouping assets is not new Panton et al. (1976). The further step of selecting assets taking into account group constraints and determining weights via Markowitz's approach has been used, for instance, in Hui (2005). In the latter reference, however, the groups are determined by a different methodology, that is factor analysis. Finally, De Luca and Zuccolotto (2011) suggested the idea of calculating all possible portfolios with group constraints; however, again, their clustering procedure is different since it assumes a parametric form of the dependence structure (compare with Durante et al. (2015)).

a. Illustration from EURO STOXX 50

We consider time series related to EURO STOXX 50 stock index and its components in the period from January 2, 2003 to 31 July, 2011. Moreover, as out-of-sample period, we will also show the performance of our procedure in the period from August 1, 2011 to September 9, 2011. The period has been selected due to the fact that EURO STOXX 50 was experiencing severe losses in the period (see Figure 3).

We preliminary apply a univariate Student-t ARMA(1,1)-GARCH(1,1) model to each time series of log–returns of 50 constituents of the index to remove autocorrelation and heteroscedasticity from the data and we computed the standardized residuals in order to check the adequacy of the fit.

Given the standardized residuals, we compute two kinds of tail dependence measures, namely the TDC λ_{ij} and the conditional Spearman's ρ_α (here we select $\alpha = 0.10$) for all pairs of variables. Then, we apply the hierarchical clustering algorithm (complete linkage) and the fanny algorithm to determine the cluster compositions. The results are contained in Tables 1–4, where the optimal number of clusters has been determined via silhoutte index as clarified above. As it can be seen, the compositions of the sub–groups seem to differ with respect to both the tail dependence measure and the clustering algorithms.

In order to give a more objective comparisons among the obtained clusters, a measure of agreement is needed. To this aim, we consider the Rand Index (RI) Rand (1971) and the Adjusted Rand Index (ARI) Hubert and Arabie (1985). The RI lies between 0 and 1, where the maximum value is taken when two partitions agree perfectly. Instead the ARI takes a wider range of values and, hence, increases the sensitivity of the index as a measure of agreement.

As it is noted in Table 5, if we fix the clustering method and compare the results obtained by changing the tail dependence measure, the obtained grouping compositions seem to be similar, although they do not coincide. In fact, as discussed above, the two tail dependence measures underline different aspects of tail dependence (finite and asymptotic tail behavior).

Analogously, in view of Table 6, if we fix the the tail dependence measure but allow us to use different clustering procedures, the obtained grouping compositions seem to be similar (even more than in the previous case). In other words, the effects of a changing clustering procedure seem to be less evident than those obtained by changing the tail dependence measure.

Now, if we restrict to the analysis of TDC–based cluster (the other results are quite similar), we may notice in Figure 1 the graph of standard deviation against expected returns of all 33264 portfolios composed with our procedure by using hierarchical clustering, while the same picture is obtained in Figure 2 by fuzzy clustering. In both cases, we highlight the portfolios in the convex efficient frontier.

The returns of these portfolios in the frontier are compared with the returns of naive minimum variance portfolio built from the whole set of assets and to the benchmark index EURO STOXX 50. As can be seen, the performance of the portfolios in the efficient frontier is generally better than the benchmark and, in several cases, outperforms the global minimum variance portfolio. This seems to confirm the idea that, when markets are experiencing a period of losses, a diversification strategy could be beneficial.

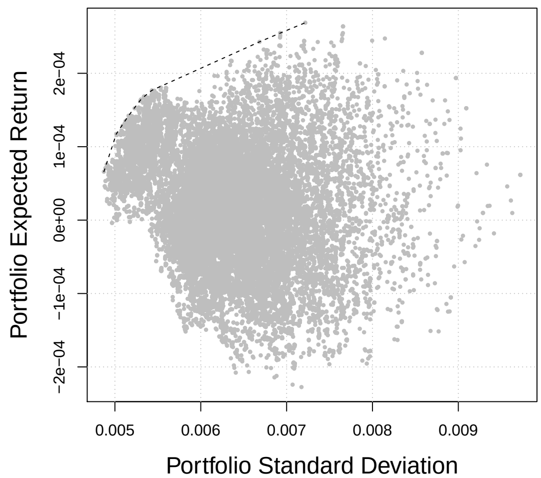

FIG. 1. Standard deviation-Expected return plot of 5-asset portfolios generated from TDC and hierarchical clustering.

The composition of the minimum variance portfolio in the convex frontier of Figures 3 and 4 are shown in Table 7. Notice that both gave large weight to one single asset, while the other are different. Anyway, as can be read from the basic statistics of the selected portfolios (Table 8), the returns of the portfolio obtained by hierarchical clustering (HC Portfolio) and the returns of portfolio obtained by fuzzy clustering (Fanny Portfolio) are quite similar. Thus, the clustering method does not have a strong influence on the overall results.

4. Conclusions

We have introduced a procedure aiming at selecting a portfolio from a group of assets in such a way that the assets are diversified in their tail behavior. Two kinds of clustering algorithms (hierarchical and fuzzy) are used in the illustration, showing that they perform in a similar way.

The procedure is expected to be used in order to have an idea about possible portfolios to be built in bearish periods.

Acknowledgments.

The first author acknowledges the support of Free University of Bozen-Bolzano via the project "'Stochastic Models for Lifetimes". The second author acknowledges the support of Free University of Bozen-Bolzano, School of Economics and Management, via the project "Handling High–Dimensional Systems in Economics". The third author acknowledges the support of the Department of Economics, Business, Mathematics and Statistics "Bruno De Finetti", University of Trieste, via the project "Metodi e modelli matematici e

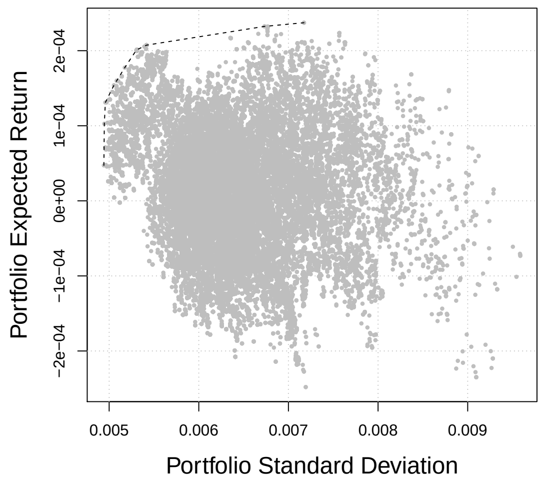

FIG. 2. Standard deviation-Expected return plot of 5-asset portfolios generated from TDC and fuzzy clustering.

statistici per la valutazione e gestione del rischio in ambito finanziario e assicurativo". The fourth author acknowledges the support of Sapienza University of Rome via the project "Portfolio optimization under extreme dependence" (prot. C26N12T5XP).

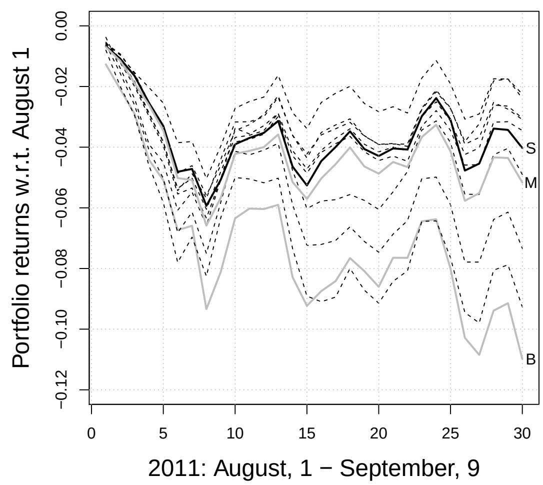

FIG. 3. 9 minimum variance portfolios on the efficient frontier of all the possible 5-asset portfolios obtained via hierarchical clustering. The returns of the minimum variance portfolio in the frontier (denoted by S) and the other portfolios in the frontier (black dotted lines) are compared with the returns of EURO STOXX 50 (denoted by B) and with the returns of global minimum variance portfolio(denoted by M) composed of all 50 assets.

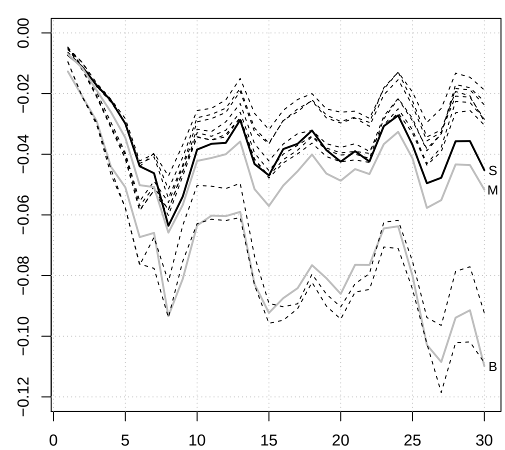

FIG. 4. 10 minimum variance portfolios on the efficient frontier of all the possible 5-asset portfolios obtained via fuzzy clustering. The returns of the minimum variance portfolio in the frontier (denoted by S) and the other portfolios in the frontier (black dotted lines) are compared with the returns of EURO STOXX 50 (denoted by B) and with the returns of global minimum variance portfolio(denoted by M) composed of all 50 assets.

Table 1. Cluster composition of EURO STOXX 50 constituents by using conditional Spearman's correlation ρ_α with $\alpha = 0.1$ and hierarchical clustering.

Cluster	Asset						
1	E.IND	F.SGE	D.BASX	D.BAYX	D.RWEX	D.SIEX	D.DTEX	D.SAPX
	H.UNIL	F.LVMH	F.CRFR	I.ISP	F.EI	E.IBE		
2	H.ASML	B.ABI	M.NOK1	D.EONX	F.FTEL	H.MT	F.BSN	F.AIR
	F.OR.F	F.SQ.F						
3	I.ENEL	F.DG.F	D.BMWX	I.ENI	F.TAL	F.UBL		
4	F.BNP	E.REP	H.ING	D.DAIX	E.SCH	F.GOB	E.BBVA	D.ALVX
	D.DBKX	D.MU2X	CRGI	I.G	I.UCG	E.TEF	H.PHIL	F.MIDI
	F.QT.F	F.GSZ						
5	D.VO3X	F.EX.F						

Table 2. Cluster composition of EURO STOXX 50 constituents by using TDC measure and hierarchical clustering.

Cluster	Asset							
1	E.IND	D.DAIX	D.BMWX					
2	H.ASML	B.ABI	D.RWEX	D.EONX	D.DTEX	D.SAPX	H.UNIL	F.CRFR
	F.EX.F	F.SQ.F	F.QT.F	F.GSZ				
3	I.ENEL	M.NOK1	F.GOB	D.BAYX	F.FTEL	CRGI		
4	F.BNP	E.REP	H.ING	E.SCH	F.SGE	F.DG.F	E.BBVA	D.ALVX
	D.BASX	D.SIEX	H.MT	F.BSN	F.LVMH	F.AIR	I.G	F.OR.F
	I.UCG	I.ISP	E.TEF	F.UBL	F.MIDI	E.IBE		
5	D.DBKX	I.ENI	D.MU2X	D.VO3X	F.TAL	H.PHIL	F.EI	

Table 3. Cluster composition of EURO STOXX 50 constituents by using conditional Speaman's correlation ρ_α with $\alpha = 0.1$ and fuzzy clustering.

Cluster	Asset							
1	E.IND	E.REP	M.NOK1	H.MT	F.UBL	F.EI		
2	H.ASML	F.BNP	H.ING	E.SCH	E.BBVA	D.BAYX	D.BMWX	D.RWEX
	I.ENI	D.MU2X	D.DTEX					
3	I.ENEL	D.EONX	CRGI	F.BSN	F.CRFR	E.TEF	E.IBE	F.GSZ
4	B.ABI	F.SGE	F.DG.F	D.BASX	F.FTEL	D.VO3X	D.SAPX	H.UNIL
	F.EX.F	F.AIR	F.OR.F	F.SQ.F				
5	D.DAIX	F.GOB	D.ALVX	D.DBKX	D.SIEX	F.TAL	F.LVMH	I.G
	I.UCG	I.ISP	H.PHIL	F.MIDI	F.QT.F			

TABLE 4. Cluster composition of EURO STOXX 50 constituents by using TDC measure and fuzzy clustering.

Cluster	Asset							
1	E.IND	F.DG.F	D.BMWX	H.MT	F.LVMH	I.ISP	E.TEF	E.IBE
	F.GSZ							
2	H.ASML	B.ABI	F.SGE	D.EONX	D.DTEX	H.UNIL	F.BSN	F.EX.F
	F.AIR	F.UBL	F.QT.F					
3	I.ENEL	M.NOK1	D.BAYX	I.ENI	F.FTEL	D.SAPX	CRGI	F.TAL
	F.CRFR	H.PHIL	F.EI	F.SQ.F				
4	F.BNP	E.REP	H.ING	E.SCH	F.GOB	E.BBVA	D.ALVX	D.BASX
	D.DBKX	D.SIEX	I.G	F.OR.F	I.UCG	F.MIDI		
5	D.DAIX	D.RWEX	D.MU2X	D.VO3X				

TABLE 5. Rand Index and Adjusted Rand Index between cluster compositions obtained from a hierarchical clustering (respectively, fuzzy clusteing) applied on different tail dependence measures.

	Hierarchical Clustering	Fuzzy Clustering
RI	0.62	0.68
ARI	0.03	0.02

TABLE 6. Rand Index and Adjusted Rand Index between cluster compositions obtained, respectively, by hierarchical clustering and by fuzzy clustering, according to different tail dependence measures.

	$\rho_{0.10}$	TDC
RI	0.68	0.73
ARI	0.08	0.26

TABLE 7. Composition of the minimum variance portfolio selected by hierarchical cluster procedure and fuzzy cluster procedure using TDC measure.

Hierarchical Clustering					
Assets	F.EI	F.FTEL	F.BSN	D.SAPX	E.IND
Weights	0.41	0.25	0.15	0.14	0.05
Fuzzy Clustering					
Assets	F.EI	H.UNIL	E.TEF	D.RWEX	I.G
Weights	0.43	0.20	0.15	0.13	0.09

Table 8. Basic statistics related to the log–returns of selected minimum variance portfolios in the convex frontier by hierarchical clustering and fuzzy clustering. Period: August 1, 2011 – September 9, 2011.

	Mean	S.D.	Skewness	5% VaR	5% E.S.	Sharpe Ratio
HC Portfolio	0.0001	0.0049	-0.0426	-0.0077	-0.0116	0.0115
Fanny Portfolio	0.0001	0.0051	0.0296	-0.0077	-0.0118	0.0108
Naive MVP	0.0001	0.0048	-0.1527	-0.0076	-0.0114	0.0149
EUROSTOXX 50	-0.0001	0.0071	0.0636	-0.0111	-0.0170	-0.0125

REFERENCES

Billio, M., Getmansky, M., Lo, A., Pelizzon, L.: Econometric measures of connectedness and systemic risk in the finance and insurance sectors. J. Fin. Econ. **104** (3), 535–559 (2012)

Bradley, B., Taqqu, M.: Framework for analyzing spatial contagion between financial markets. Financ. Lett. **2**(6), 8–16 (2004)

Panton, D.B., Lessig, V.P., Joy, O.M.: Comovement of international equity markets: A taxonomic approach. J. Financ. Quant. Anal. **11**(3), 415–432 (1976)

De Luca, G., Rivieccio, G., Zuccolotto, P.: Combining random forest and copula functions: A heuristic approach for selecting assets from a financial crisis perspective. Intell. Sys. Acc. Fin. Mgmt. **17**(2), 91–109 (2010)

De Luca, G., Zuccolotto, P.: A tail dependence-based dissimilarity measure for financial time series clustering. Adv. Data Anal. Classif. **5**(4), 323–340 (2011)

Dobrić, J., Frahm, G., Schmid, F.: Dependence of stock returns in bull and bear markets. Discussion Papers in Statistics and Econometrics 9/07, University of Cologne, Department for Economic and Social Statistics (2007). URL `http://ideas.repec.org/p/zbw/ucdpse/907.html`

Durante, F., Jaworski, P.: Spatial contagion between financial markets: a copula-based approach. Appl. Stoch. Models Bus. Ind. **26**(5), 551–564 (2010)

Durante, F., Pappadà, R., Torelli, N.: Clustering of financial time series in risky scenarios. Adv. Data Anal. Classif. **8**(4), 359–376 (2014)

Durante, F., Pappadà, R., Torelli, N.: Clustering of time series via non-parametric tail dependence estimation Stat. Papers **56**, 701–721 (2015)

Embrechts, P., Lindskog, F., McNeil, A.J.: Modelling dependence with copulas and applications to risk management. In: S. Rachev (ed.) Handbook of heavy tailed distributions in finance, pp. 329–384. Elsevier, Amsterdam (2003)

Forbes, K.J., Rigobon, R.: No contagion, only interdependence: measuring stock market comovements. J. Financ. **57**(5), 2223–2261 (2002)

Frahm, G., Junker, M., Schmidt, R.: Estimating the tail-dependence coefficient: Properties and pitfalls. Insurance Math. Econom. **37**(1), 80–100 (2005)

Genest, C., Segers, J.: Rank-based inference for bivariate Extreme Value copulas. Ann. Statist. **37**(5B), 2990–3022 (2009)

Grubel, H.: Internationally diversified portfolios: Welfare gains and capital flows. Am. Econ. Rev. **58**(5), 1299–1314 (1968)

Hastie, T., Tibshirani, R., Friedman, J.: The elements of statistical learning, second edn. Springer Series in Statistics. Springer, New York (2009). Data mining, inference, and prediction

Hubert, L., Arabie, P.: Comparing partitions. J. Classif. **2**, 193–218 (1985)

Hui, T.K.: Portfolio diversification: a factor analysis approach. Appl. Finan. Econ. **15**(12), 821–834 (2005)

Joe, H.: Multivariate models and dependence concepts, *Monographs on Statistics and Applied Probability*, vol. 73. Chapman & Hall, London (1997)

Kaufman, L., Rousseeuw, P.: Finding groups in data. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons Inc., New York (1990)

Levy, H., Sarnat, M.: International diversification of investment portfolios. Am. Econ. Rev. **60**(4), 668–675 (1970)

Mantegna, R.: Hierarchical structure in financial markets. Euro. Phys. J. B **11**(1), 193–197 (1999)

Markowitz, H.: Portfolio selection. J. Finance **7**(1), 77–91 (1952)

Nanda, S., Mahanty, B., Tiwari, M.: Clustering Indian stock market data for portfolio management. Expert Systems with Applications **37**(12), 8793–8798 (2010)

Patton, A.J.: Copula methods for forecasting multivariate time series. In: G. Elliott, A. Timmermann (eds.) Handbook of Economic Forecasting, vol. 2, pp. 899–960. Elsevier, Oxford (2013)

Rand, W.M.: Objective criteria for the evaluation of clustering methods. J. Amer. Statist. Assoc. **66**(336), 846–850 (1971)

Tola, V., Lillo, F., Gallegati, M., Mantegna, R.: Cluster analysis for portfolio optimization. J. Econom. Dynam. Control **32**(1), 235–258 (2008)