ARTICLE IN PRESS

Journal of Business Research xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Journal of Business Research

journal homepage: www.elsevier.com/locate/jbusres

# Herding in mutual funds: A complex network approach

Anna Maria D'Arcangelis[a], Giulia Rotundo[b,*]

[a] DEIM, Department of Economics, Engineering, Society and Business, University of Tuscia, Via del Paradiso 47, 01100 Viterbo, Italy
[b] Department of Statistical Sciences, Sapienza University of Rome, p.le A. Moro 5, 00185 Roma, Italy

ARTICLE INFO

*Keywords:*
Herding
Mutual funds
Complex networks
Regression
Asset management
Strategic asset management

ABSTRACT

The paper investigates herding in mutual funds through a complex networks approach. The detection of significant correlation coefficients constitutes the basis for the construction of the network. Some centrality measures and the assortativity are added as explanatory variables in the regression analysis of two popular indicators of herding, largely applied in finance literature. Cross-Sectional Standard Deviation and Cross-Sectional Absolute Deviation are both considered since they emphasize the bulk and the extreme values of herding. Two dummy variables designed to capture differences in investor behaviour in extreme up or down versus relatively normal markets are considered as independent variables. The results show a clear decrease of herding in stressful periods of the market. Moreover, the prevailing explanatory power of the betweenness is well evidenced, thereby highlighting the role of the network structure. In line with the literature on herding, the results also evidence a flight to safety effect.

## 1. Introduction

"Men, it has been well said, think in herds; it will be seen that they go mad in herds, while they only recover their senses slowly, and one by one" (Mobarek, Mollah, & Keasey, 2014). This quantum of wisdom in the book of Mackay (2016) pioneers the most recent developments on global risks in financial markets.

Herding is defined as an imitation behaviour resulting from individual actions leading to inefficient outcomes for the market (Bikhchandani, Hirshleifer, & Welch, 1992). The attitude of people to infer information by observing the actions of others was originally studied in psychology: a famous experiment by Asch (1952), finalized to prove the power of social pressure, revealed that within groups individuals often abandon their own private signals to rely primarily on group opinions. Several papers by Shiller (1987), Scharfstein and Stein (1990), Banerjee (1992) and Bikhchandani et al. (1992), to name the most important ones, introduced the herding behaviour into the literature on finance, so underlining the possible consequences for the informational efficiency of financial markets.

Herding has been extensively studied for speculative bubbles of single stock market indices through the detection of an oscillatory behaviour with increasingly frequent fluctuations – supposedly caused by the mass of investments by the herd - till to the burst of the bubble in a large crash. Sornette (2003) presents a review of the theory of financial crashes triggered by herding, detected through the presence of the log-periodic power law (LPPL); before, Ausloos, Ivanova, and Vandewalle (2002) gave a brief historical perspective concerning financial crashes and its modelling through LPPL, whence with predictive power (Vandewalle, Ausloos, Boveroux, & Minguet, 1998). Eventually, herding emerges as a collective phenomenon only at specific time scales, giving rise to pockets of predictability (Andersen & Sornette, 2005). Correlation breakdowns can act both as a consequence or a triggering factor in the emergence of financial crises rational bubbles (Falbo & Grassi, 2011, 2015).

Literature has indeed defined various types of herd behaviour, based on several explanations of the co-movement among returns. Herding is generally divided in two classes: either intentional herding or spurious (or unintentional) herding. The latter is typical of institutional investors, whose decisions are usually driven by fundamentals. Since financial institutions analyse the same factors every day and receive correlated private information from the same sources, they easily draw similar conclusions regarding individual stocks. Moreover, portfolio managers, to cite the agents to which we will refer in the paper, may be regarded as a homogenous group, as they share similar educational and professional backgrounds.

Co-movements may also occur due to completely different reasons. Among them, it is worth quoting the blindness to small changes, that shows as counterpart an excess of reaction (Massad & Vitting Andersen, 2018, Andersen, Nowak, Rotundo, Parrott, & Martinez, 2011, Bellenzier, Andersen, & Rotundo, 2016), or any contagion process

* Corresponding author.
*E-mail addresses:* adarcangelis@unitus.it (A.M. D'Arcangelis), giulia.rotundo@uniroma1.it (G. Rotundo).

https://doi.org/10.1016/j.jbusres.2019.11.016
Received 14 March 2019; Received in revised form 1 November 2019; Accepted 4 November 2019
0148-2963/ 

(Cinelli, Ferraro, Iovanella, & Rotundo, 2019, Foroni & Grassi, 2005). The conclusion is that, while intentional herding is considered inefficient and can intensify volatility and undermine markets causing the emergence of bubbles and crashes on financial markets, unintentional herding is generally an efficient outcome driven by fundamentals.

The financial literature has provided several methods for the detection of herding.

Christie and Huang (1995) show that, during periods of market stress, investors tend to ignore rational information and base their trading decisions on the behaviour of the market. If the investors act as a herd, then the single stock returns, ex post, are going to be quite close to the return of the market. The authors use a specific kind of standard deviation: the cross-sectional standard deviation of returns, which measures the amount of concentration around the market index. Their analysis uses the daily returns of stocks listed on NYSE and Amex from July 1962 to December 1988; their results show that herding is concentrated during periods of market stress, when single investors risk dismissing their own investment lines in favour of the consensus of the market.

Chang, Cheng, and Khorana (2000) foster the hypothesis that the number of stocks needed to provide adequate diversification is higher in presence of herding, and that herding is given by a wrong evaluation of asset prices. The authors propose a modification of the methodology of Christie and Huang (1995), able to detect herding even in mild market movements. Instead of CSSD, Chang et al. (2000) use the cross-sectional absolute deviation for the analysis of the stock markets of United States, Hong Kong, South Korea, Taiwan and Japan in the period 1963–1997. Their results did not point out to any herding in the United States and in Hong Kong, but there is a proof of partial herding in Japan and a meaningful evidence of herding in the emerging markets (South Korea and Taiwan).

Caparrelli, D'Arcangelis, and Cassuto (2004) and D'Arcangelis and Caparrelli (2002) analyse the herding in the Italian stock market since September 1988 until January 2001 by using the measure of Christie and Huang (1995). Their results do not show the presence of herding: however, in accord to the nonlinearity test of Chang et al. (2000), herding is present in extreme market events.

Previous papers on complex networks in finance suggest that some centrality measures on the holdings of portfolios of mutual funds may evidence behaviours and strategies eventually due to herding (D'Arcangelis & Rotundo, 2015, Delpini, Battiston, Caldarelli, & Riccaboni, 2018). The complex networks literature on financial time series has already shown that the overall correlation structure of the markets shrinks during recessions and enlarges during expansions (Bonanno, Caldarelli, Lillo, & Mantegna, 2003, Onnela, Chakraborti, Kaski, & Kertész, 2003, Pozzi, Aste, Rotundo, & Di Matteo, 2007). Further quantities, typical in the studies of networks, like the clustering coefficient, have been applied for studying the properties of the stock correlation networks (Huang, Zhuang, & Yao, 2009); their extensions show to be promising (Cerqueti, Ferraro, & Iovanella, 2018, Clemente, Grassi, & Hitaj, 2019).

Such empirical evidences induced us to propose an analysis of herding behaviour of mutual funds within the framework of complex networks analysis. The theory of complex networks is increasingly used to solve problems in many fields. The interest in complex networks has arisen from specific concerns for modelling phenomena in social sciences, biology, chemistry, and was fostered by the development of quantitative methods.

The empirical question is whether the structure of the network (i.e. the network topology), summed up by the various centrality measures, is a factor that explains herding in the mutual funds market. The connection among mutual funds is obtained from the matrix of the significant Pearson correlation coefficients between returns, built following Cai and Liu (2016).[1]

Many so-called centrality measures have been developed for detecting the prominent role of the units connected through a network.

The first and the most used among them is surely the node-degree. It is a quantity pertaining to each node: the more the node is connected to the others, the higher the node-degree is. Networks in which nodes have a (mean) high node-degree are more connected than networks in which nodes have a low node-degree (in average). Centrality measures induce a ranking through the set of nodes: with respect to the node-degree, the first ranked is the most connected, the last ranked is the node with even no connection at all.

The closeness of a node is the inverse of the mean distance between the node and all the others: within this perspective, the most relevant node is the one central to the group, with the shortest distances from all the others. This measure is different from the node-degree, since it can happen that a well-connected node remains at the periphery of the entire network, thus far away from the core of the network.

The betweenness considers the role of nodes to join groups. The node with the maximal betweenness is the one which is the only connection between two different groups, which have no other connection otherwise. The betweenness has been created for emphasizing the role of "bridges" between communities.

The nodes with high eigenvector centrality are the ones for which the first neighbours have a high node-degree. Thus, they share properties of the high-degree neighbour, yet not necessarily having a high number of connections: a unique connection to a node with a very high-degree is enough to be ranked as first.

Moreover, we also consider the correlation among the node degrees, that is the assortativity. A network is said to be assortative whenever the correlation among the node-degrees is positive: highly (low) connected nodes are linked to nodes with high (low) connections as well. Obviously, the assortativity is not a centrality measure, but it considers the structure of the network as a whole.

Thereafter, we consider all the previous centrality measures and the assortativity in order to explore the relevance of the topology of the network for the detection of mutual fund herding.

Notice that other measures can characterize a network topology, like the "overlap index", introduced by Gligor and Ausloos (2008), as applied for example by Redelico and Proto (2010, 2013). For keeping this paper at a reasonable size, we postpose the pertinent numerical work for a later study.

The paper is organized as follows. Section 2 describes the sample and explains the computation of abnormal returns. Section 3 explains the network construction and the computation of centrality measures. Section 4 shows the methodology for herding estimation and evaluation. Section 5 performs a preliminary statistical analysis. Section 6 and 7 outline the results on the data sample. Conclusions follow.

[1] It is important to note that the mere use of a full correlation matrix without any further selection of the relevant elements would not allow any network study, because the network would be complete, so quantities typically measured in complex networks would be trivial. It is worth noting that in literature other approaches are proposed for selecting only some elements of the correlation matrix for giving rise to a not trivial network. For instance, Gonzalez Osorio (2016) selects the thresholds that keep the 50–80% of the data, so to maintain enough links in the network. The weakness of this approach is the relevance of the value of the threshold for any network analysis (Boginski and Pardalos (2005), Kim, Kim, and Ha (2007), Gonzalez Osorio (2016)). Alternative techniques are denoising the correlation matrix, and extracting only the minimum spanning tree, or the planar maximally filtered graphs (Bonanno et al. (2003), Heimo, Kaski, and Saramaki (2009), Onnela et al. (2003), Pozzi et al. (2007)). We do not go further in this direction, because the concept of spanning tree does not seem the best choice for the present analysis. In fact, herding is a concept closely connected to grouping, but the removal of too many correlations would affect the centrality measures.

## 2. Data and methodology for the computation of abnormal returns

A sample of 470 open-end equity mutual funds has been selected from the Bloomberg "Funds Search Engine". Daily net asset values were downloaded for the period spanning from Dec. 31st, 2006 to Dec 31st, 2017. For the same time span, we also downloaded the daily values of different stock indices for the European and Euro markets (Euro Stoxx 50, Euro Stoxx, Stoxx Europe 600). The sample has been pruned to avoid replicas in the case of the same fund with different fee structure, removing 23 funds. For each of the 447 funds, we then computed daily returns as the natural logarithm of the ratio of subsequent day's closing price. The formula for a logarithmic return is:

$$R_t = \ln\left(\frac{P_t}{P_{t-1}}\right)$$

where $R_t$ is the logarithmic return on day $t$; $P_t$ is the fund price at the end of day t and $P_{t-1}$ is the fund price at day $t-1$. The availability of returns allowed us to compute the matrix of paired Pearson correlation coefficients for all the funds in the sample. The observation of correlations higher than 99.9% induced us to remove seven funds from the selection. The final sample counts 440 funds.

To remove the potentially confounding market effect from the funds returns, following the Modern Portfolio Theory and the CAPM model, we decided to work on the residuals of the Single Index Model (Sharpe, 1963, 1964). This type of adjustment, aiming to insulate the portion of returns caused by specific factors, is widely used in the academic literature. The technique is frequently applied in the field of event studies (Campbell, Lo, & McKinlay, 1997).

For each fund in the sample, we regressed the respective daily returns against the same period returns of the market (proxied by the Euro Stoxx index). The residuals from each regression are the fund-specific returns with overall market effect removed (OLS adjusted returns).

$$\mathrm{e}_{\mathrm{it}} = \mathrm{R}_{\mathrm{it}} - [\alpha_{\mathrm{i}} + \beta_{\mathrm{i}}\mathrm{R}_{\mathrm{mt}}] \tag{1}$$

where $\mathrm{e}_{\mathrm{it}}$ is the residual, i.e. it is the return of the stock after controlling for the overall stock market and trend. $\mathrm{R}_{\mathrm{it}}$ is the daily return of the fund, $\mathrm{R}_{\mathrm{mt}}$ is the daily return of the market; $\alpha_{\mathrm{i}}$ and $\beta_{\mathrm{i}}$ are the intercept and the slope of the regression of market returns, respectively.

## 3. Network construction and the computation of centrality measures

In order to achieve the target of determining the relevance of centrality measures for herding, the first step is the network construction. We use the correlation matrix $C = (\rho_{ij})$ for building the adjacency matrix $A = (a_{ij}) \in R^{NxN}$ of the network: in this case, each mutual fund is a node of the network[2].

However, a preliminary analysis is necessary before the application of network methods. In fact, keeping the correlation matrix just as it comes out from the mere calculus would give rise to a complete network, which is not meaningful for the analysis through centrality measures. Centrality measures on networks provide a ranking of network nodes using the lack of homogeneity in the links, and the presence/absence of links. In a complete network, such differences among the nodes are not present, whence the centrality measures would not be sufficient explanatory variables.

Actually, the definition of the correlation itself is solving the problem of moving from a complete network to a not-trivial structure when only the significant correlations are considered. The approach of Cai and Liu (2016) proposes a statistical test for detecting significant correlations. The test is based on the control of the false discovery rate and the false discovery proportion asymptotically to any predefined level $0 < \alpha < 1$. For each value in the correlation matrix, we test the null hypotheses versus the alternative one:

$$\mathrm{H}_0\colon \rho_{ij} = 0 \quad \text{versus} \quad \mathrm{H}_1\colon \rho_{ij} \neq 0$$

Setting equal to 0 all the elements verifying the null hypothesis $\mathrm{H}_0$, we obtain a not complete matrix, well suitable to be adopted as an adjacency matrix for the calculus of centrality measures.

We focus on the most classical measures: node degree, betweenness, closeness, and eigenvector centrality.

The node degree $k$ (or just degree for short) is the most classic among the centrality measures: it is a vector reporting, for each network node, the number of incoming (outgoing) links, $k = Ae$, where $e = (1,1,\ldots,1)^T \in R^N$. The degree is high for mutual funds that show a high correlation with many others. We recall that a *path* in a network is a sequence of nodes and edges such that each edge connects two successive nodes in the path.

The betweenness of each node $i$ is calculated as the number of shortest paths that need to pass through $i$ divided by all the possible shortest paths.

$$g(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

where $\sigma_{st}$ is the total number of shortest paths from node s to node t and $\sigma_{st}(v)$ is the number of those paths that pass through $v$.

In social networks, the betweenness is high for nodes that "bridge" groups. In mutual funds, it is high for the funds that have sufficiently high correlations with funds belonging to different groups, if any. In D'Arcangelis and Rotundo (2016), we remarked that the betweenness of the network of mutual funds holdings emphasize the outsiders rather than the "central" ones. In fact, nodes with high betweenness show investments in common with two nearly disjoint groups. Within that analysis, the mutual funds with high betweenness are not actually the ones prevailing on the market. In this paper, we use the correlation instead of the overlap among mutual funds holdings, so that the usefulness of the betweenness as an explanatory variable must be stated.

The closeness is a centrality measure where $C_i = 1/(\Sigma_j\, d_{ij})$ has as denominator the sum of all the distances between each node $j$ and $i$. As the inverse of the mean distance of the node from all the others, the most relevant node is the central one of the group, with short distances from all others. The measure differs from the degree, because it can happen that a well-connected node remains at the periphery of the entire network, far away from a large part of the network.

The eigenvector centrality is the eigenvector corresponding to the highest eigenvalue of the adjacency matrix.

The centrality $x_v$ of vertex v is defined as

$$x_v = \frac{1}{\lambda} \sum_{t \in M(v)} x_t = \frac{1}{\lambda} \sum_{t \in G} a_{v,t} x_t$$

where $M(v)$ is a set of the neighbours of $v$; $\lambda$ is a constant; $x_v$ is high for nodes close to hubs. In social networks, the centrality expresses the concept that there is no need to be powerful, but it is very relevant to have well connected friends. In mutual funds, nodes with high eigenvector centrality are close to funds showing many significant correlations.

It is worth remarking that the centrality measures often show a quite high level of dependence. In Valente, Coronges, Lakon, and Costenbader (2008), an extensive review on many datasets is reported. Centrality measures involve mathematical calculations over the same quantities, so it is difficult to achieve a correlation close to zero. On the opposite, measures eventually showing a too high correlation should not be considered together, because the explanatory power is already

[2] Of course, instead of using the correlation matrix, we could have used the distance $d_{ij} = [(2(1-\rho_{ij})]^{1/2}$, typically used in the Minimum Spanning Tree (MST) analysis Bonanno, Vandewalle, and Mantegna (2000). Since we do not strictly need a formalization based on the MST, it is convenient to use the correlations, directly.

well gotten by just one of them, and collinearity is a thread to linear regressions.

## 4. Herding measures and testing methods

Following Christie and Huang (1995), and Chang et al. (2000), we use two different measures for herding: the Cross-Sectional Standard Deviation (CSSD) and the Cross-Sectional Absolute Deviation (CSAD). They are both measures of dispersions; so, low values of CSSD and CSAD relate to high herding.

Christie and Huang (1995) compute $CSSD_t$ through the formula (3):

$$CSSD_t = \sqrt{\frac{\sum_{i=1}^{N} (R_{it} - R_{mt})^2}{N-1}} \tag{3}$$

where $R_{it}$ is the return of the fund $i$ at quarter $t$, $R_{mt}$ is the return of the market at quarter $t$ and N = 44 (number of quarters in the period of analysis).

Christie and Huang (1995) argue that the herding behaviour intensifies during the phases of market stress (extreme up or down movements). Therefore, they run regression (4) using two dummy variables which take the value 1 whenever the market rate of return $R_{mt}$ is on day $t$ in the upper/lower extreme tail of the distribution of market returns. Then

- ✓ $D_L = 1$ if, on day t $R_{mt}$ lies in lower tail of return distribution at 95% (2.5% in the tail) or 99% (0.5% in the tail) and 0 otherwise;
- ✓ $D_U = 1$ if, on day t $R_{mt}$ lies in upper tail of return distribution at 95% (2.5% in the tail) or 99% (0.5% in the tail) and 0 otherwise.

$$CSSD_t = \alpha + \beta_1 D_L + \beta_2 D_U + e_t \tag{4}$$

where $CCSD_t$ is the Cross-Sectional Standard Deviation at quarter $t$, $D_L$ and $D_U$ are two dummy variables.

The analysis reveals the existence of herding behaviour when $\beta_1$ and $\beta_2$ are negative and statistically significant.

Chang et al. (2000) run a test to examine the existence of non-linear relationship between dispersion and market returns. Their measure of dispersion is the result of the following formula (5)

$$CSAD_t = \frac{\sum_{i=1}^{N} |R_{it} - R_{mt}|}{N-1} \tag{5}$$

where $R_{it}$ is the return of the fund $i$ at quarter $t$, $R_{mt}$ is the return of the market at quarter $t$ and N = 44 (number of quarters in the period of analysis).

CSAD (5) is similar to CSSD (3), in the sense that, as a measure of dispersion, it signals herding when their values are low. The authors show that under CAPM assumptions, $CSAD_t$ should be a linear function of market returns.

$$CSAD_t = \alpha + \gamma_1 |R_{mt}| + \gamma_2 (R_{mt})^2 + e_t \tag{6}$$

where $CSAD_t$ is the Cross-Sectional Absolute Deviation at quarter $t$, and $R_{mt}$ is the return of the market at quarter $t$. The evidence of a non-linear relationship would be an indication of herding behaviour: it occurs if the values of the $\gamma_2$ coefficient is negative and statistically significant.

We adopt the herding testing method proposed in both Christie and Huang (1995) and Chang et al. (2000) and augment Eqs. (4) and (6) as follows:

$$CSSD_t = \alpha + \beta_1 D_L + \beta_2 D_U + \sum_i I(s, i)\beta_i NM_{t-1}^i + e_t \tag{7}$$

$$CSAD_t = \alpha + \gamma_1 |R_{mt}| + \gamma_2 (R_{mt})^2 + \sum_i I(s, i)\gamma_i NM_{t-1}^i + e_t \tag{8}$$

where $NM_{t-1}^i$ are the *network metrics* chosen to represent the centrality measures and the assortativity with a lag of one quarter. *I(s,i)* depends on the selected regression (s), and on the variable (*i*). $I(s,i) = 1$, if the network metrics $NM_{t-1}^i$ was used in the regression $s$; it is *0* otherwise.

Negative estimates of coefficients $\beta_1$ and $\beta_2$ are always consistent with the presence of herding behaviour in periods of stress. In presence of herding, the coefficients $\beta_i$ should be negative and statistically significant.

We also tested whether the assortativity may successfully replace the mean of a centrality measure.

## 5. Preliminary statistical analysis

The first step has been the computation of the correlation matrix on the daily returns for the 440 funds. Then, we applied the statistical test outlined in Cai and Liu (2016) to detect significant correlations. Setting α = 0.01, the 80.906% of the links are kept. This value is in good agreement with the number of relevant correlations used in Gonzalez Osorio (2016), where analyses have been performed using several thresholds, including the 80% one.[3]

In order to run the regression analyses, we constructed quarterly networks for the period January 2007-December 2017 (44 quarters). For each quarter, we used the daily returns to calculate the cross-correlations among all the funds and we applied the test of Cai and Liu (2016). Therefore, we crosschecked that none of the matrices is too full or too empty (Fig. 1(a)). We also tested the persistence of the significant correlations through quarters. Fig. 1(b) shows that the turnover rate ranges from 19% to 36%. The averages of each of the four centrality measures (degree, closeness, betweenness, and eigenvector centrality) and the assortativity have been calculated on each quarter and lagged by one period. The estimate of $CSSD_t$ (7) and $CSAD_t$ (8), used to account for herding in the funds market, completes the calculus of the variables for the set-up of the regression analyses.

### 5.1. Testing for multicollinearity

The presence of high correlations among regressors in $CSSD_t$ (7) and $CSAD_t$ (8) may cause a multicollinearity problem, which leads to erroneous empirical results: parameter estimates may be unstable. The standard error could be inflated, and consequently biased (Belsley, 1991). The removal of one or more of the explicative variables that are highly correlated with the other explicative variables can reduce poor regression results.

The literature evidenced a high level of correlations among centrality measures. As a matter of fact, they are mathematical transformations performed on the same underlying data (Valente et al., 2008). Table 1 reports the correlations for the entire dataset of the regressors used in the regression sets.

A visual inspection of Table 1 evidences concerns for multicollinearity for three measures of centrality (degree, betweenness, and closeness).

- 0.9890, for the correlation between degree and closeness
- −0.9880, for the correlation between betweenness and closeness
- −0.9584, for the correlation between degree and betweenness.

We also applied the test for Variance Inflation Factor (VIF) to measure how much the variance of the coefficients is inflated by multicollinearity[4]. The results confirm a collinearity problem among the three centrality measures stated above. Therefore, it makes sense to

[3] Of course, due to the distinct criterions, the two networks rising from the methods of Cai and Liu (2016) and Gonzalez Osorio (2016) differ each from other, even if the percentage of effective links is the same.

[4] The VIF can detect whether one regressor has a strong linear association with the remaining regressors: a rule of thumb suggests multicollinearity problems when VIF is greater than nine. In this case, the variance of the i-th regression coefficient is (all other things being equal) nine times greater than it would have been in case of independence of the other regressors. In other words, the VIF explains the amount of variance inflation due to the lack of independence.

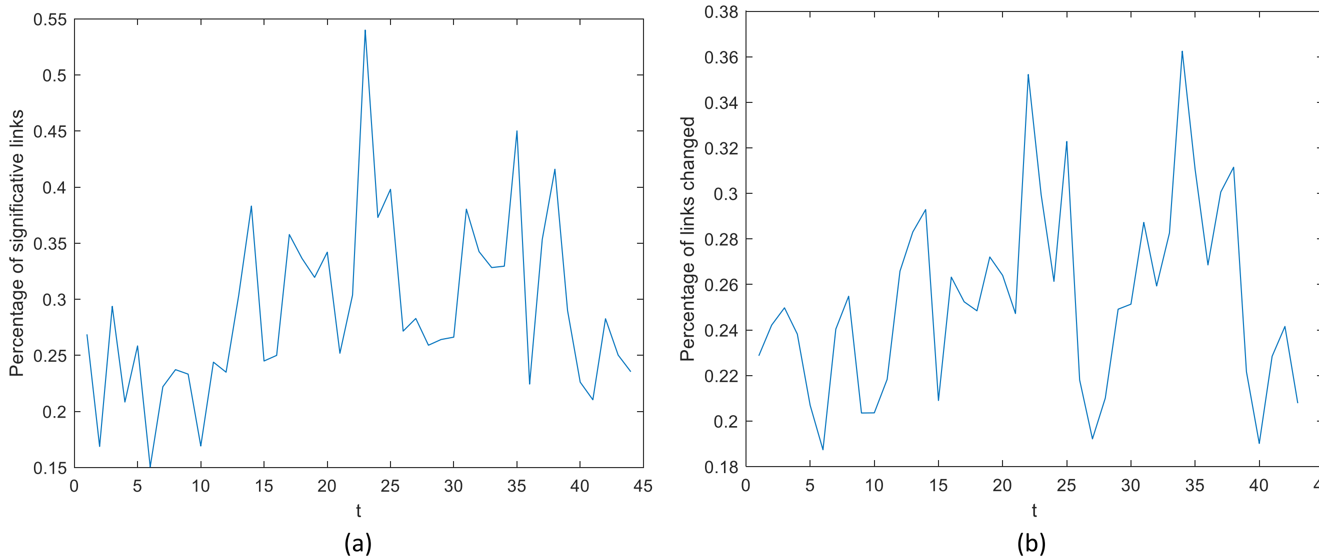


**Fig. 1.** (a) Percentages of remaining links after the application of the test of Cai and Liu (2016) for keeping the significant correlations. We remark that none of the matrices is either empty or full. (b) Percentages of links that change through quarters.

**Table 1**
Correlations among the regressors used in CSSD models.

| | $D_L$ | $D_U$ | DEG | BET | ASS | CLO | EIG |
|---|---|---|---|---|---|---|---|
| $D_L$ | 1 | | | | | | |
| $D_U$ | −0.0081 | 1 | | | | | |
| DEG | 0.0006 | 0.0083 | 1 | | | | |
| BET | 0.0062 | −0.0050 | −0.9584 | 1 | | | |
| ASS | 0.0088 | 0.0229 | −0.2294 | 0.2558 | 1 | | |
| CLO | −0.0038 | 0.0068 | 0.9890 | −0.9880 | −0.2628 | 1 | |
| EIG | 0.0082 | 0.0134 | 0.1174 | −0.0638 | −0.0452 | 0.0987 | 1 |

**Table 2**
AIC and BIC on the regressions (4) (4a–c).

| Model | AIC | BIC |
|---|---|---|
| $CSSD_t = \alpha + \beta_1 D_L + \beta_2 D_U + e_t$ | −17223.8 | −17206.13 |
| $CSSD_t = \alpha + \beta_1 D_L + \beta_2 D_U + \beta_3 DEG + e_t$ | −17260.3 | −17236.75 |
| $CSSD_t = \alpha + \beta_1 D_L + \beta_2 D_U + \beta_3 BET + e_t$ | −17295.8 | −17272.27 |
| $CSSD_t = \alpha + \beta_1 D_L + \beta_2 D_U + \beta_3 CLO + e_t$ | −17280.7 | −17257.1 |

keep at most one out of the three regressors, degree, betweenness and closeness: we performed various regressions and evaluated them through the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). The results in Table 2 show that the betweenness improves both the estimate of AIC and BIC, keeping as comparison the base regression (4).

In order to have a comprehensive view, we deepened the analysis to verify the robustness of our results when the single centrality measures are added to the model (9). In this case, the base model is the regression where the collinear measures are not present at all; thus, the variables are the two dummies, the assortativity and the eigenvector centrality.

$$CSSD_t = \alpha + \beta_1 D_L + \beta_2 D_U + \beta_3 ASS + \beta_4 EIG + e_t \tag{9}$$

We assume that the best explanatory variable among degree, betweenness, and closeness is the one that most reduces AIC and BIC figures when added to the regression model (9), leading to the models shown in Table 3.

Once again, the results show that the explanatory power of the betweenness is higher than that of closeness and degree. Therefore, we keep the betweenness and remove the closeness and the degree from the set of regressors.

## 6. Results of the regressions CSSD

Regressions (4) and (6) (test of herding without centralities) and subsequent (7) and (8) have been applied using the two measures of dispersion $CSSD_t$ and $CSAD_t$ with the traditional regressors of Christie and Huang (1995) set at 5% and 1%, those of Chang et al. (2000) and adding various combinations of the centrality measures (betweenness and eigenvector) and the assortativity.

### 6.1. Results on CSSD with $D_U$ and $D_L$ at 5%

Tables 4–6 sum up the results of the regressions of $CSSD_t$ (models 4 and 7) and $CSAD_t$ (models 6 and 8).

The Tables are set out as to introduce an increasing number of explanatory variables in the analysis. Table 4 starts with the regression on the $CSSD_t$ with 5% tails, without centrality measures: it is the classic test of Christie and Huang (1995).

In the same table, the next three lines report the results of the regressions with either one centrality measure or the assortativity. The lines 5–7 report the results with two out of the three network variables (betweenness, eigenvector, assortativity). The last line shows the results with all the variables in the regression.

The statistically significant and positive coefficients $\beta_1$ and $\beta_2$ in Table 4 show that herding of mutual funds is lower in stressed phases of the market. A sample fully explained by herd behaviour would be moving in accord with the market mean, the dispersion would be zero: for the sample here examined, returns dispersions increase during periods of large price changes.

This evidence goes against the theory of herding behaviour in the tail of the distribution of market returns and supports the rational asset pricing models that predict that periods of market stress induce increased levels of dispersion as individual returns differ in their sensitivity to the market return. Put differently, investors would tend to ignore information conveyed by the market dynamics in favour of the views of a subset of actors (Gębka & Wohar, 2013).

These results are in line with those of many other previous works, such as Christie and Huang (1995), Chang et al. (2000), Gleason, Mathur, and Peterson (2004), Henker, Henker, and Mitsios (2006), Ben Mabrouk and Fakhfekh (2013), among others. Obviously, the results should also be due to the subjective definition of extreme market returns at 99% or 95%, used in literature.

The observation that European stock funds seem to herd more in

**Table 3**
AIC and BIC criteria in regressions (9) and (9a–c).

| Model | AIC | BIC |
|---|---|---|
| $CSSD_t = \alpha + \beta_1 D_L + \beta_2 D_U + \beta_3 ASS + \beta_4 EIG + e_t$ | −17232.3 | −17202.82 |
| $CSSD_t = \alpha + \beta_1 D_L + \beta_2 D_U + \beta_3 ASS + \beta_4 EIG + \beta_5 DEG + e_t$ | 17266.8 | −17231.47 |
| $CSSD_t = \alpha + \beta_1 D_L + \beta_2 D_U + \beta_3 ASS + \beta_4 EIG + \beta_5 BET + e_t$ | −17299.8 | −17264.6 |
| $CSSD_t = \alpha + \beta_1 D_L + \beta_2 D_U + \beta_3 ASS + \beta_4 EIG + \beta_5 CLO + e_t$ | −17286.09 | −17250.76 |

normal stock market phases than otherwise leads to a conclusion in favour of the prevalence of similar strategic asset allocation policies. Our tests indicate that the alignment is destined to fade away when markets are in stress, a behaviour that would enhance the role of tactical asset allocation policies or a more specific conduct of fund managers.

The subsequent three lines of Table 4 show the results of the regression tests with one added centrality measure. The coefficients of the betweenness are all positive and significant, meaning that lower betweenness implies higher herding. The eigenvector coefficients are positive and significant, unlike the assortativity, which is not significant. Overall, the betweenness proves to be the best explanatory variable, as confirmed by the behaviour of the Adj. $R^2$ and by AIC and BIC figures.

In order to give a picture of the joint contribution of the two variables on herding behaviour, the subsequent tests add two explanatory regressors from the set of network variables. The results in lines 5–7 show that the betweenness is the best explanatory variable both when coupled with the eigenvector centrality and with the assortativity. Also, in this case, there is an improvement of the Adj. $R^2$ and of AIC and BIC. On the opposite, in the regression with the eigenvector centrality and the assortativity, there is no meaningful improvement of the results, and the assortativity is not significant.

The last line (the 8th) of Table 4 shows the results when all the regressors are used. The outcomes confirm the coefficients of betweenness and eigenvector centralities to be positive and significant; on the opposite, the assortativity is just below the 95% significance (Student's t = −1.94).

### 6.2. Results on CSSD with $D_U$ and $D_L$ at 1%

The $CCSD_t$ tests, with the tails cut at 1%, show results similar to the 5% case (Table 5). The returns dispersions during extreme downside moves of the market are lower than those for upside moves (the coefficient of $D_L$ is here lower than the coefficient of $D_U$). This behaviour indicates a better consensus of the market in bad times, in line with the "flight to safety" hypothesis.

The coefficients in Line 1 of Table 5 are always positive and significant, as in the same test of Table 4. The next lines show similar results: the coefficients of all the measures are positive and the assortativity is significant. When the added variables are two, the betweenness confirms its main role, while assortativity loses its significance. The reason could be that the significant correlations after the test of Cai and Liu (2016) do not have a strong relationship with the assortativity. This implies that the network topology is quite different depending on the temporal segment considered: the most correlated funds are not always the same, so they contribute to herding in different ways, depending on the time segment. Tactical asset management policies or active management may be at the basis of such instabilities.

## 7. Results of the regressions CSAD

Table 6 shows the output of the test of herding of Chang et al. (2000) with or without network measures as regressors.

Under the hypothesis of CAPM, $CSAD_t$ should be a linear function of $R_m$, and herding is proved when the coefficient of $R_m^2$is negative and significant: in this case, investors follow the herd. When $\gamma_2$ is not significantly different from zero, the returns behavior is in line with an equilibrium model (Bekiros, Jlassi, Lucey, Naoui, & Uddin, 2017).

The coefficients of the regressors of the base test, in line 1, are both positive and highly significant, evidencing a behaviour of fund managers known as adverse herding (Klein, 2013) or negative herding (Gębka & Wohar, 2013). The coefficient of $R_m^2$is linked to the increase of $CSAD_t$, which means less herding. Compared to the $CSSD_t$ tests, the Adj. $R^2$ is very high (around 95%).

Adding one network measure at a time, one can see that the coefficients of all the variables are always significant. The coefficients of the betweenness and the eigenvector centrality are positive, while the assortativity is now negative. This confirms the result of $CSSD_t$: the network topology is not mirroring the herding.

**Table 4**
$CSSD_t$ 5%. This table reports the estimated coefficients of the following regression model (7) $CSSD_t = \alpha + \beta_1 D_L + \beta_2 D_U + \sum_i I(s,i)\beta_i NM_{t-1}^i + e_t$, where $NM_{t-1}^i$ are the "network metrics" chosen by us to evidence the relationship between the elements in the correlation coefficients' matrix, taken with a lag of one quarter. *I(s,i)* depends on the selected regression (s), and on the variable (i). *I(s,i) = 1* if the network metric $NM_{t-1}^i$ was used in the regression *s*; it is *0* otherwise.

| | α | $D_U$ (5%) | $D_L$ (5%) | BET | ASS | EIG | Adj. $R^2$ | AIC | BIC |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.01008 | 0.02192 | 0.02321 | | | | 0.4034 | −18113.1 | −18095.4 |
| | 60.9 | *29.85* | 31.73 | | | | | | |
| 2 | 0.01196 | 0.02182 | 0.02306 | 0.00026 | | | 0.4212 | −18193.2 | −18169.6 |
| | 1.21 | *30.17* | 31.99 | *9.12* | | | | | |
| 3 | 0.0975 | 0.02191 | 0.02320 | | 0.08410 | | 0.4032 | ‘18111.2 | −18087.7 |
| | 28.98 | *29.79* | 31.63 | | *0.35* | | | | |
| 4 | 0.01007 | 0.02188 | 0.02322 | | | 0.01514 | 0.4046 | −18117.4 | −18093.9 |
| | 60.86 | *29.82* | 31.77 | | | *2.52* | | | |
| 5 | 0.00130 | 0.02190 | 0.02315 | 0.00003 | −0.49926 | | 0.4219 | −18195.3 | −18165.9 |
| | 1.32 | *30.25* | 32.08 | *9.35* | *−2.04* | | | | |
| 6 | 0.00098 | 0.02177 | 0.02307 | 2.64E-05 | | 0.01868 | 0.4232 | −18201.1 | −18171.7 |
| | 1 | *30.14* | 32.06 | *9.32* | | *3.15* | | | |
| 7 | 0.00993 | 0.02186 | 0.02320 | | 0.11226 | 0.01527 | 0.4044 | −18115.7 | −18086.2 |
| | 28.82 | *29.74* | 31.67 | | *0.47* | *2.54* | | | |
| 8 | 0.00109 | 0.02184 | 0.02316 | 0.00003 | −0.47591 | 0.01832 | 0.4238 | −18202.9 | −18167.6 |
| | 1.10 | *30.21* | 32.14 | *9.52* | *−1.94* | *3.09* | | | |

**Table 5**

CSSD 1%. This table reports the estimated coefficients of the following regression model (7) $\mathrm{CSSD_t} = \alpha + \beta_1 \mathrm{D_L} + \beta_2 \mathrm{D_U} + \sum_i \mathrm{I(s,\ i)}\beta_i \mathrm{NM}^i_{t-1} + \mathrm{e_t}$, where $\mathrm{NM}^i_{t-1}$ are the network metrics (the centrality measures and the assortativity), taken both as insulated, couples and all the three of them.

| | α | $D_U$ (1%) | $D_L$ (1%) | BET | ASS | EIG | Adj. $R^2$ | AIC | BIC |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.01172 | 0.03571 | 0.03265 | | | | 0.1675 | −17223.8 | −17206.13 |
| | 62.56 | *17.38* | 15.53 | | | | | | |
| 2 | 0.00174 | 0.03579 | 0.03254 | 0.00003 | | | 0.1899 | −17295.83 | −17272.27 |
| | 1.49 | *17.66* | 15.69 | *8.66* | | | | | |
| 3 | 0.01076 | 0.03558 | 0.0326 | | 0.75895 | | 0.1694 | −17228.97 | −17205.42 |
| | 26.55 | *17.33* | 15.52 | | *2.68* | | | | |
| 4 | 0.01171 | 0.03565 | 0.03262 | | | 0.01543 | 0.1686 | −17226.51 | −17202.96 |
| | 62.5 | *17.36* | 15.52 | | | *2.17* | | | |
| 5 | 0.0017 | 0.03577 | 0.03253 | 0.00003 | 0.14868 | | 0.1897 | −17294.09 | −17264.64 |
| | 1.46 | *17.64* | 15.68 | *8.24* | *0.51* | | | | |
| 6 | 0.00152 | 0.03572 | 0.03249 | 0.00003 | | 0.01938 | 0.1919 | −17301.46 | −17272.01 |
| | 1.3 | *17.64* | 15.68 | *8.83* | | *2.76* | | | |
| 7 | 0.0107 | 0.03551 | 0.03256 | | 0.78866 | 0.01633 | 0.1707 | −17232.26 | −17202.82 |
| | 26.4 | *17.31* | 15.51 | | *2.78* | *2.3* | | | |
| 8 | 0.00148 | 0.03569 | 0.03248 | 0.00003 | 0.17304 | 0.01951 | 0.1917 | −17299.82 | −17264.48 |
| | 1.26 | *17.62* | 15.68 | *8.39* | *0.6* | *2.78* | | | |

The regressions against the various couples of network variables and against all of them are in line with these findings. However, we can further remark that the increase in the Adj. $R^2$ and the improvement of AIC and BIC is very marginal, since they are already quite high.

## 8. Comparison with other models

We performed a cross-check of the validity of the regression results considering the approach with thresholds, used in Gonzalez Osorio (2016). As already mentioned, the application of the test of Cai and Liu (2016) to our sample keeps 80.906% of the links: in order to compare results, we performed the test also following the threshold method proposed in the literature, so to have 80.906% of the correlations above it. We note that the results are in a quite remarkable agreement with the previous ones.

## 9. Conclusions

In this paper, we inquire about the role of the structure of the network based on correlations among mutual funds returns as proxy for the presence of herding. The network has been built keeping the significant correlations, only. We consider the node degree, the closeness, the betweenness and the eigenvector centrality as network variables to explain herding; we also calculated the assortativity, although it is not a centrality measure.

A first test of collinearity led us to discard the node degree and the closeness from the set of regressors, as they are highly correlated to the betweenness. Regressions (4) and (6) (test of herding without centrality) and subsequent (7) and (8) have been applied using the two measures of dispersion ($CSSD_t$ and $CSAD_t$) with the traditional regressors of Christie and Huang (1995) and of Chang et al. (2000) and various combinations of betweenness, eigenvector centrality and assortativity.

None of the standard measures of dispersion -$CSSD_t$ and $CSAD_t$ - supports the classical hypothesis of herding behaviour. Our data, on a sample of 440 stock mutual funds focused on European stock market, show an opposite result, namely that managers move more in synchronization with the market mean during normal stock market phases. Such a result can be interpreted as a proof of the prevalence of similar strategic asset allocation policies in normal times, whereas the asset allocation strategies fall out of alignment during periods of market stress. It is likely that, in these periods, tactical asset allocation or more active or specific conducts of mutual fund managers take a prominent role.

In the $CSSD_t$ test with $D_L$ and $D_U$ at 1% (very extreme market movements), the significance of the $D_L$ is lower than the significance of the $D_U$: for the mutual funds under observation, returns dispersions during extreme downside moves are much lower than those for upside moves. This behaviour indicates a better consensus of the market in bad times, in line with the “flight to safety” effect evidenced in the literature on herding (Demirer & Kutan, 2006).

**Table 6**

$CSAD_t$. This table reports the estimated coefficients of the regression model (8): $CSAD_t = \alpha + \gamma_1\ |R_{mt}| + \gamma_2 (R_{mt})^2 + \sum_i I(s,\ i)\gamma_i NM^i_{t-1} + e_t$ where $CSAD_t$ is the cross-sectional absolute deviation and $R_{mt}$ is the return of the market at quarter t. $NM^i_{t-1}$ are the “network metrics” taken with a lag of one quarter. *I(s,i)* depends on the selected regression(s), and on the variable (i). *I(s,i) = 1* if the network metric $NM^i_{t-1}$ was used in the regression *s*; it is *0* otherwise.

| | α | $\vert R_m\vert$ | $(R_m)^2$ | BET | ASS | EIG | Adj. $R^2$ | AIC | BIC |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.00164 | 0.92681 | 1.33447 | | | | 0.9473 | −24746.2 | −24728.5 |
| | *21.26* | *100.18* | *7.47* | | | | | | |
| 2 | 0.00017 | 0.9242 | 1.33289 | 4.36E-06 | | | 0.9478 | −24771.6 | −24748.0 |
| | *0.57* | *100.25* | *7.5* | *5.24* | | | | | |
| 3 | 0.00191 | 0.92861 | 1.31831 | | −0.22223 | | 0.9475 | −24754.5 | −24730.9 |
| | *16.79* | *100.37* | *7.39* | | *−3.20* | | | | |
| 4 | 0.00164 | 0.92702 | 1.32120 | | | 0.00654 | 0.9476 | −24758.4 | −24734.9 |
| | *21.24* | *100.45* | *7.41* | | | *3.77* | | | |
| 5 | 0.00023 | 0.92630 | 1.30824 | 0.00001 | −0.33371 | | 0.9483 | −24791.6 | −24762.1 |
| | *0.79* | *100.76* | *7.39* | *6.27* | *−4.69* | | | | |
| 6 | 0.00008 | 0.92430 | 1.31823 | 0.00001 | | 0.00718 | 0.9482 | −24786.9 | −24757.4 |
| | *0.28* | *100.57* | *7.44* | *5.53* | | *4.16* | | | |
| 7 | 0.00189 | 0.92871 | 1.30638 | | −0.21052 | 0.00629 | 0.9477 | −24765.6 | −24736.2 |
| | *16.64* | *100.61* | *7.34* | | *−3.04* | *3.63* | | | |
| 8 | 0.00015 | 0.92634 | 1.29473 | 0.00001 | −0.32490 | 0.00694 | 0.9485 | −24805.8 | −24770.4 |
| | *0.51* | *101.05* | *7.33* | *6.51* | *−4.58* | *4.03* | | | |

Next, we tested the relevance of various network measures as explanatory variables for the herding behaviour. The coefficients of the betweenness always prove the main role of the centrality measure for herding, a result that holds for both $CSSD_t$ and $CSAD_t$. The eigenvector centrality is less significant, with a similar behaviour. The role of the assortativity is less stable in the $CSSD_t$ analysis and, in general, it has the weakest coefficient. Therefore, the correlation structure of the node degrees does not seem to play a key role in the herding behaviour.

## References


Andersen, J. V., Nowak, A., Rotundo, G., Parrott, L., & Martinez, S. (2011). "Price-Quakes" shaking the world's stock exchanges. *PloS ONE, 6*(11), e26472.

Andersen, J. V., & Sornette, D. (2005). A mechanism for pockets of predictability in complex adaptive systems. *EPL Europhysics Letters, 70*(5), 697.

Asch, S. (1952). *Social psychology.* Englewood Cliffs, NJ: Prentice Hall.

Ausloos, M., Ivanova, K., & Vandewalle, N. (2002). Crashes: Symptoms, diagnoses and remedies. *Empirical science of financial fluctuations* (pp. 62–76). Japan: Springer.

Banerjee, A. (1992). A simple model of herd behaviour. *The Quarterly Journal of Economics, 57*(3), 797–817.

Bellenzier, L., Andersen, J. V., & Rotundo, G. (2016). Contagion in the world's stock exchanges seen as a set of coupled oscillators. *Economic Modelling, 59*, 224–236.

Bekiros, S., Jlassi, M., Lucey, B., Naoui, K., & Uddin, G. S. (2017). Herding behavior, market sentiment and volatility: Will the bubble resume? *North American Journal of Economics & Finance, 42*(2017), 107–131.

Belsley, D. (1991). *Conditioning diagnostics: Collinearity and weak data in regression.* New York: Wiley.

Ben Mabrouk, H., & Fakhfekh, M. (2013). Herding during market upturns and downturns: International evidence. *The IUP Journal of Applied Finance, 19*(2), 5–26.

Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A theory of fads, fashion, custom and cultural change as informational cascades. *Journal of Political Economy, 100*, 992–1026.

Boginski, B., & Pardalos, P. (2005). Statistical analysis of financial networks. *Computational Statistics & Data Analysis, 48*(2), 431–443.

Bonanno, G., Vandewalle, N., & Mantegna, R. N. (2000). Taxonomy of stock market indices. *Physical Review E, 62*, R7615(R).

Bonanno, G., Caldarelli, G., Lillo, F., & Mantegna, R. N. (2003). Topology of correlation-based minimal spanning trees in real and model markets. *Physical Review E, 68*(4), 046130.

Cai, T. T., & Liu, W. (2016). Large-scale multiple testing of correlations. *Journal of the American Statistical Association, 111*(513), 229–240. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4894362/https://doi.org/10.1080/01621459.2014.999157.

Campbell, J., Lo, A. H., & McKinlay, C. (1997). *The econometrics of financial markets.* Princeton, NJ: Princeton University Press.

Caparrelli, F., D'Arcangelis, A. M., & Cassuto, A. (2004). Herding in the Italian stock market: A case of behavioral finance. *The Journal of Behavioral Finance, 5*(4), 222–230.

Cerqueti, C., Ferraro, G., & Iovanella, A. (2018). A new measure for community structures through indirect social connections. *Expert Systems with Applications, 114*(30), 196–209.

Chang, E. C., Cheng, J. W., & Khorana, A. (2000). An examination of herd behavior in equity markets: An international perspective. *Journal of Banking & Finance, 24*(10), 1651–1679.

Christie, W. G., & Huang, R. D. (1995). Following the pied piper: Do individual returns herd around the market? *Financial Analysts Journal, 51*(4), 31–37.

Cinelli, M., Ferraro, G., Iovanella, A., & Rotundo, G. (2019). Assessing the impact of incomplete information on the resilience of financial networks. *Annals of Operations Research*. https://doi.org/10.1007/s10479-019-03306-y.

Clemente, G. P., Grassi, R., & Hitaj, A. (2019). Asset allocation: New evidence through network approaches. *Annals of Operations Research*. https://doi.org/10.1007/s10479-019-03136-y.

D'Arcangelis, A. M., & Caparrelli, F. (2002). La competizione tra i fondi comuni, la performance ed il comportamento del gestore bancario. *Bancaria, 6*, 104.

D'Arcangelis, A. M., & Rotundo, G. (2015). Mutual funds relationships and performance analysis. *Quality & Quantity, 49*(4), 1573–1584.

D'Arcangelis, A. M., & Rotundo, G. (2016). Complex networks in finance. *Complex networks and dynamics* (pp. 209–235). Springer International Publishing.

Delpini, D., Battiston, S., Caldarelli, G., & Riccaboni, M. (2018). The network of U.S. mutual fund investments: Diversification, similarity and fragility throughout the global financial crisis. *Working Paper arXiv*.

Demirer, R., & Kutan, A. M. (2006). Does herding behavior exist in Chinese stock markets? *Journal of International Financial Markets, Institutions and Money, 16*(2), 123–142.

Falbo, P., & Grassi, R. (2011). Market dynamics when agents anticipate correlation breakdown. *Discrete Dynamics in Nature and Society*. https://doi.org/10.1155/2011/959847 Article ID 959847.

Falbo, P., & Grassi, R. (2015). Does expectation of correlation breakdown in financial market fulfill itself? *Discrete Dynamics in Nature and Society,* 8. https://doi.org/10.1155/2015/263908 Article ID 263908.

Foroni, I., & Grassi, R. (2005). The contagion process in a financial model: A synergetic approach. *PU.M.A. 16*(4), 377–398.

Gębka, B., & Wohar, M. E. (2013). International herding: Does it differ across sectors? *Journal of International Financial Markets, Institutions and Money, 23*, 55–84.

Gleason E.R., K. C., Mathur, I., & Peterson, M. A. (2004). Analysis of intraday herding behavior among the sector ETFs. *Journal of Empirical Finance, 11*(5), 681–694.

Gligor, M., & Ausloos, M. (2008). Clusters in weighted macroeconomic networks: The EU case. Introducing the overlapping index of GDP/capita fluctuation correlations. *The European Physical Journal B, 63*(4), 533–539.

Gonzalez Osorio, D. H. (2016). Essays on mutual fund performance. Ph.D. Dissertation Thesis. < http://hdl.handle.net/10138/162991 > .

Heimo, T., Kaski, T., & Saramaki, J. (2009). Maximal spanning trees, asset graphs and random matrix denoising in the analysis of dynamics of financial networks. *Physica A, 388*, 145–156.

Henker, J., Henker, T., & Mitsios, A. (2006). Do investors herd intraday in Australian equities? *International Journal of Managerial Finance, 2*(3), 196–219.

Huang, W. Q., Zhuang, X.-T., & Yao, S. (2009). A network analysis of the Chinese stock market. *Physica A, 388*(14), 2956–2964.

Kim, K., Kim, S. Y., & Ha, D. H. (2007). Characteristics of networks in financial markets. *Computer Physics Communications, 177*, 184–185.

Klein, A. (2013). Time-variations in herding behavior: Evidence from a Markov switching SUR model. *Journal of International Financial Markets, Institutions & Money, 26*, 291–304.

Mackay, C. (2016). *Extraordinary popular delusions and the madness of the crowds.* London: Createspace Independent Pub.

Massad, N., & Vitting Andersen, J. (2018). Three different ways synchronization can cause contagion in financial markets. *Risks, 6*, 104. https://doi.org/10.3390/risks6040104.

Mobarek, A., Mollah, S., & Keasey, K. (2014). A cross-country analysis of herd behavior in Europe. *Journal of International Financial Markets, Institutions and Money, 32*, 107–127.

Onnela, J. P., Chakraborti, A., Kaski, K., & Kertesz, J. (2003). Dynamic asset trees and Black Monday. *Physica A, 324*(1), 247–252.

Pozzi, F., Aste, T., Rotundo, G., & Di Matteo, T. (2007). Dynamical correlations in financial systems. In Microelectronics, MEMS, and nanotechnology, international society for optics and photonics, 68021E–68021E.

Redelico, F. O., & Proto, A. N. (2010). Nonlinear time series into complex networks scheme. *International Journal of Bifurcation and Chaos, 20*(02), 413–417.

Redelico, F. O., & Proto, A. N. (2013). Complex networks topology: The statistical self-similarity characteristics of the average overlapping index. *Advanced dynamic modeling of economic and social systems* (pp. 163–174). Berlin, Heidelberg: Springer.

Scharfstein, D., & Stein, J. (1990). Herd behavior and investment. *The American Economic Review, 80*(3), 465–479.

Sharpe, W. F. (1963). A simplified model for portfolio analysis. *Management Science, 9*(2), 277–293.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance, 19*(3), 425–442.

Shiller, R. J. (1987). Investor behavior in the October 1987 stock market crash: Survey evidence. NBER Working Paper 2446.

Sornette, D. (2003). Critical market crashes. *Physics Reports, 378*(1), 1–98.

Valente, T. W., Coronges, K., Lakon, C., & Costenbader, E. (2008). How correlated are network centrality measures? *Connect (Tor), 28*(1), 16–26.

Vandewalle, N., Ausloos, M., Boveroux, P., & Minguet, A. (1998). How the financial crash of October 1997 could have been predicted. *The European Physical Journal B-Condensed Matter and Complex Systems, 4*(2), 139–141.