Dynamic patterns analysis meets Social Network Analysis in the modeling of financial market behavior

Caterina Liberati[1,2], and Paola Zappa[1]

[1] DEMS, University of Milano-Bicocca, Milan, ITALY

[2] Corresponding author: Caterina Liberati, e-mail: caterina.liberati@unimib.it

Abstracts

Social Network Analysis has been widely applied to detect individual drivers of counterparty selection within a number of inter-individual and inter-organizational settings. Longitudinal methods, like Stochastic Actor-Oriented Models, Separable Temporal Exponential Random Graph Models and Relational Event History Models have been proposed to identify the evolution of actors' positioning within a network of relationships. They provide a detailed description of the general tendencies shaping the network dynamics, but do not allow to isolate groups of similar behaviors. Standard cluster analysis has been used to accomplish this purpose, but with a main focus on cross-sectional data. Our proposal is to combine Social Network Analysis with a Multi-way Factor Analysis (MFA). Such an approach allows to represent the evolutions of the actors' behaviors according to the time structure of the data; it also facilitates the visual inspection of the actors' trajectories onto the compromise plan. A clustering of actors based on Multi-way factors is also applied, in order to identify similar patterns and to provide an interpretable solution. In the first phase, ego-network measures are computed to identify actors' market behaviors in their neighborhood. Then, according to Structuration des Tableaux A Trois Indices de la Statistique (STATIS) approach, a Three-way Factor Analysis is performed, in order to represent actors configurations according to their average distribution. Such solution can be reviewed as the space spanned by a linear combination of multiple factor analysis and it can be consider a "virtual" space where similar actors route paths can be highlighted. An application to the Euro Electronic Market for Interbank Deposits (e-Mid) during the recent turmoil period will be shown, in order to provide insights into the rationale and benefits of the proposed approach.

Keywords: dynamic pattern, ego-network measures, panel data, interbank liquidity market.

1. Introduction

Treatment and modeling of relational data have received increasing interest in the extant statistical literature. Applicability of a relational framework to a wide range of empirical settings (i.e., agent behaviors, market exchanges, texts, etc.) and peculiar data properties (i.e., interdependence among observations) have led to developing sophisticated methods of analysis. Particular effort has been placed in detecting the temporal evolution of relations. Within the domain of Social Network Analysis (SNA), this issue has been addressed in various ways. Stochastic Actor-Oriented Models (Snijders et al., 2010), Separable Temporal Exponential Random Graph Models (Krivitsky and Handcock, 2012) and Relational Event History Models (Butts, 2008) are recent approaches which have been proposed in order to deal with network evolution. Though different, these three classes of models follow the same logic: they assume the networks to be generated by a parsimonious set of local rules, representing actors' behavior in their neighborhood. This allows to explain the drivers of counterparties selection, while a synthetic and intuitive representation of actors' positioning across time as well as a comparison among different positions are disregarded. We propose the usage of Multi-way Factor Analysis (MFA) to explore

and describe, in a synthetic form, the dynamic patterns of a social network evolution. The main idea of such an approach is to find a common factors plan, obtained by solving the maximization problem of a common correlation matrix, which can give both information about macro similarities among waves, and evidences about micro structures relative to each year. The relevant advantage of MFA lies in performing a dimension reduction which allows an easy interpretation of the behaviors over the time span, even if the number of input variables is high. Also, the effects observed contribute all together to create the new factors: this property, typical of the factorial model in general, allows to point out inter-connections structure, their persistency over time and/or their configuration variations. This framework provides a useful tool for visualizing trajectories of individual instances that can be grouped in order to get a meaningful solution. A classification of longitudinal data is also performed, with the aim of highlighting the relevant dissimilarities among the mean paths in terms of shape and directions, and of suggesting a coherent evaluation of the phenomenon under examination.

2. Data

We study the Euro Electronic Market for Interbank Deposits (e-Mid). e-Mid is the EU main electronic platform for unsecured lending and records all the transactions occurred in the market among the registered banks. For each transaction, a record is produced which provides information on the identity of the lender and the borrower, the amount traded, the interest rate, the date and time of delivery.

Our analysis covers the period of the recent financial turmoil between 2006 and 2009. Hence, we focus on the changes in the patterns of exchange that have arisen from the disruption of standard market activities in the Euro area after 2007. We study the evolution in the market position, in terms of borrowing and lending, of the 91 banks active throughout the observation period and attempt to link a bank position to its lending interest rate strategy.

3. Methodological framework

The first step of our analysis consisted in representing transactions within a SNA framework. Previous findings (Liberati et al., 2012) show that banks positioning changes very gradually, although the transactions occur on a highly frequency basis (on average, 210 instances a day). Transactions were, then, aggregated over an annual basis. For each of the four years a network $N(V,E)$ was built. The network nodes V are the banks, and ties E the lending relation from one institution to another. To each tie we attached a weight $w_{ii'}$ that represents the amount of money that bank i lends to bank i' over one year. In order to describe and understand variation across banks in their market behavior, we examined the ego-network of each bank. A ego-network consists of the ties between actor i, the ego or focal actor, and its alters, all nodes to whom ego has a connection at a path length of N (here $N=1$), and of the ties among alters. This analytic approach, therefore, allows to identify the way banks are embedded in "local" social structures and to exploit the characteristics representing embeddedness as variables in our MFA.

Among the wide range of ego-network measures, we selected some corresponding to alternative definitions of market behavior and relations benefits, i.e. social capital (Putnam, 2000). Bonding social capital, which conceives market power as activity in homogeneous groups and, then, as richness of ties, was measured as node degree and node weight. As bridging social capital, which focuses on power as access to diverse resources and intermediation between structural holes, we specified effective size. For each bank i, these measure were computed on each annual network and either for "out" (all the actors to whom ties are directed from ego) or "in" (all the actors who sent ties directly to ego) neighborhood. Tab. 1 reports definitions and descriptions.

MEASURE	DESCRIPTION	DEFINITION	EFFECT
Node degree	Number of alters directly connected to ego	$\sum e_{ii'}$	Positive
Node weight	Sum of weights of the direct ties between ego and alters	$\sum w_{ii'}$	Positive
Effective size	Number of alters ego has, minus a redundancy factor, given by the average number of ties that each alter has to other alters.	$\sum_{i'} 1 - \sum_{i'} \sum_q p_{iq} m_{i'q} \qquad q \neq i, i'$ $m_{i'q} = i'$'s interaction with q divided by i''s strongest relation with anyone $p_{iq} =$ proportion of i's energy invested in relation with q	Positive

Table 1 Ego-network measures

Then, a temporal repeated monitoring of the units was realized, in order to obtain a panel data set which can be investigated in terms of trends across life span of units. When a sufficiently long term series is not available, MFA provides a suitable tool for the study of variable dynamics over various time periods (Tucker 1966; Escofier and Pagès 1984; Kiers 1989) as it summarizes the variability of complex phenomena by highlighting both the similarities/dissimilarities among the occasions. The principal feature of MFA is the comparison of different data tables (matrices) under various experimental conditions. By analogy to N-way methods, the three-way data set is denoted by X with dimensions n, p and t, corresponding to the number of rows (individuals), columns (variables) and tables (occasions), respectively. Thus, a generic element of X is x_{ijt}, where $i = 1,...,n, j = 1,..., p$ and $t = 1,...,k$ with n=91, p=6 and k=4 in our case of study.

According to the Structuration des Tableaux A Trois Indices de la Statistique (STATIS) introduced by Escoufier (1980), first we analyzed the similarity structure of the k tables via the RV coefficient (Robert and Escoufier 1976) which measures the closeness between two variance–covariance matrices:

$$RV\left(W_t, W_{t'}\right) = \frac{\langle W_t, W_{t'} \rangle}{\langle W_t, W_t \rangle \langle W_{t'}, W_{t'} \rangle} \tag{1}$$

Geometrically, RV can also be interpreted as a scaled scalar product between two positive semi-definite matrices, therefore it is proportional to the cosine between the matrices (interstructure analysis).

Secondly a PCA was performed on the variance–covariance matrix of the composite table[1] W:

$$[L,V,L] = SVD(W) \tag{2}$$

and the coordinates (loadings) of the objects from the k-th table for f principal components on the compromise plot are obtained as shown in the following equation:

---

[1] In STATIS, variance–covariance matrix of the composite table W is obtained by summing up the cross-section individual variance-covariance matrices $W = \sum_{t=1}^{k} a_t W_t$ where a_t is a vector of optimal weights.

$$C_k = W_k LG \tag{3}$$

where G is a diagonal matrix ($f \times f$), whose non zero elements are the inverse of the square roots of the compromise eigenvalues, and matrix L ($n \times f$) contains the scores of the PCA. These partial factors were plotted as scatters of the actors relative to each occasion on the compromise plan; such configurations are composed as trajectories and clustered in terms of similarities. Following D'Urso and Vichi (1998), we obtained a compromise dissimilarity matrix as a conic combination of three dissimilarities measures: one which takes into account the instantaneous position of each pair of trajectories (trend), and two that evaluate the evolutionary behavior of the time trajectories (velocity and acceleration).

$$d_{ii'}^2 = \pi_1 \sum_{t=1}^{k} \left\| X_{i \cdot t} - X_{i' \cdot t} \right\|^2 + \pi_2 \sum_{t=1}^{k-1} \left\| V_{i \cdot t, t+1} - V_{i' \cdot t, t+1} \right\|^2 + \pi_3 \sum_{t=1}^{k-2} \left\| A_{i \cdot t, t+2} - A_{i' \cdot t, t+2} \right\|^2 \tag{4}$$

where $V_{i \cdot t, t+1} = (v_{i1t,t+1}, v_{i2t,t+1} ... v_{ipt,t+1})'$ with $v_{i1t,t+1} = \dfrac{x_{ijt+1} - x_{ijt}}{s_{t,t+1}}$ and

$A_{i \cdot t, t+2} = (a_{i1t,t+2}, a_{i2t,t+2} ... a_{ipt,t+2})'$ with $v_{i1t,t+1} = \dfrac{v_{ijt+1,t+2} - v_{ijt,t+1}}{s_{t,t+2}}$ and π_1, π_2, π_3 are weights to normalize distances.

4. Results

This section provides a discussion of the main findings of our work. The purpose of this study was to verify if the interbank money market, observed during 2006-2009, has been affected by the turmoil. In other words, we aimed to measure if and to what extent differences in lending/borrowing behaviors have become extreme and to link market behavior to pricing policies. The selected network measures capture exactly such tendencies. We computed these indexes for each year and collected them in 4 tables, composed, later, in the MFA. As a first result, AFM produced the partial analysis (relative to each occasion): the eigenvalues decomposition provides the fundamental information about the explained inertia in each year. From such decomposition we noticed a fall of variance, which got progressively less concentrated on the first factor: it dropped from the 55% of the total inertia explained in 2006 to the 46% in 2009. Hence, the contributions of the variables "Degree_i" "Weight_i" and "Effsize_i" decrease over time, and explain shares of variance gradually smaller. Moreover, the partial first PCs are fairly correlated: that means interbank market did not change over time in respect to the borrowing activity. A opposite situation was noticed for lending activity. Here the turmoil has changed the market structure. Therefore, the linear combination of the input variables highlights a shock in the banks' behaviors. Such evidences derive from a separate PCA performed on the dataset and, although they are informative from a speculative perspective, they prompted a cue about dynamics which had to be analyzed in depth in the intrastructure phase. STATIS provided, also, the compromise solution in which the first two factors explain about 70% of the total inertia. A synthetic description of the observed phenomenon was, thus, obtained. The interstructure analysis confirmed what was already highlighted in the partial PCA: the similarity among waves declines over the 4 years confirming that a shock has occurred in the market. Values collected in Tab. 2 (a) and (b), in fact, confirm our expectations: high similarity patterns between the interbank market during 2006-2007, as indicated by RV coefficient (0,859) and by the Euclidean distance value (0,530). Fall of RV coefficient relative to comparisons between structures in 2006-2008 (0,656) and in 2006-2009 (0,439) pointed out the

presence of instability. The scatter shape has changed and/or instances have moved to different positions.

(a)	RV coefficients				(b)	Euclidean Distances[2]			
	w1	w2	w3	w4		w1	w2	w3	w4
w1	1	0,859	0,656	0,439	w1	1,000	0,530	0,829	1,059
w2	0,859	1	0,776	0,510	w2	0,530	1,000	0,669	0,990
w3	0,656	0,776	1	0,639	w3	0,829	0,669	1,000	0,850
w4	0,439	0,510	0,639	1	w4	1,059	0,990	0,850	1,000

Table 2: a)Matrix of RV coefficients among waves; b)Matrix of Euclidean distances among waves

The intrastructure analysis was performed in order to evaluate the contribution of each group of variables to the derivation of factor scores. Principal components naming and interpretation followed the classical rules of PCA, and were based on the correlation matrix among PCs and original variables (Tab. 3). The polarization of the variables in respect to the PCs highlighted a perfect distinction between the lending and borrowing activities in the market. In the light of such evidences we expect banks that take part to the exchanges on the both market sides (i.e. supplying and borrowing money) to be located in the first quadrant (I). On the specular quadrant (III) we find banks with a scant propensity to join to market activity. Banks in quadrant II combine a high tendency to lending to others with a poor propensity to borrowing money, while in quadrant IV we find financial institutions with an opposite behavior.

	Degree_i	Degree_o	Weight_i	Weight_o	Effsize_i	Effsize_o	year
f1	0,941	0,458	0,685	0,020	0,938	0,461	2006
f2	0,364	0,883	0,081	0,590	0,360	0,864	
f1	0,957	0,344	0,865	-0,018	0,963	0,345	2007
f2	0,338	0,931	0,150	0,652	0,319	0,932	
f1	0,912	0,335	0,815	-0,029	0,956	0,278	2008
f2	0,351	0,858	0,141	0,594	0,292	0,896	
f1	0,903	0,160	0,634	0,194	0,919	0,119	2009
f2	0,376	0,708	-0,036	0,522	0,282	0,699	

Table 3 Correlation matrix among Input variable and PCs of the compromise

The partial factor scores defined as in (2) can be represented onto the compromise plan in order to draw the banks' trajectories and to monitor distances covered and directions of the paths. Visual inspection of the plot points out the difficulty to perform such monitoring, especially when the number of subjects is not small (as in our case). Although longitudinal cluster algorithms identify 7 different groups for trajectories and directions, the separation among them is not always evident. For this reason and for assuring a synthetic description of the paths, we plotted for each year only the position mean of the groups. The resulting classification highlights different reactions to the crisis, coherent with the banks' positions. Clusters can be

---

[2] The Euclidean distance among individual covariance matrices relative to a different waves (W_k) can be computed as follows:

$$d(W_k, W_{k'}) = \sqrt{2 \cdot (1 - RV(W_k, W_{k'}))}$$

distinguished in respect to the length of the paths covered: those with long trajectories consist of banks that change their pricing behavior during the observation period (c1, c6, c7). By contrast, clusters with short routes (often almost full circles) contain banks with more stable and less aggressive behaviors (c2, c3, c4, c5).
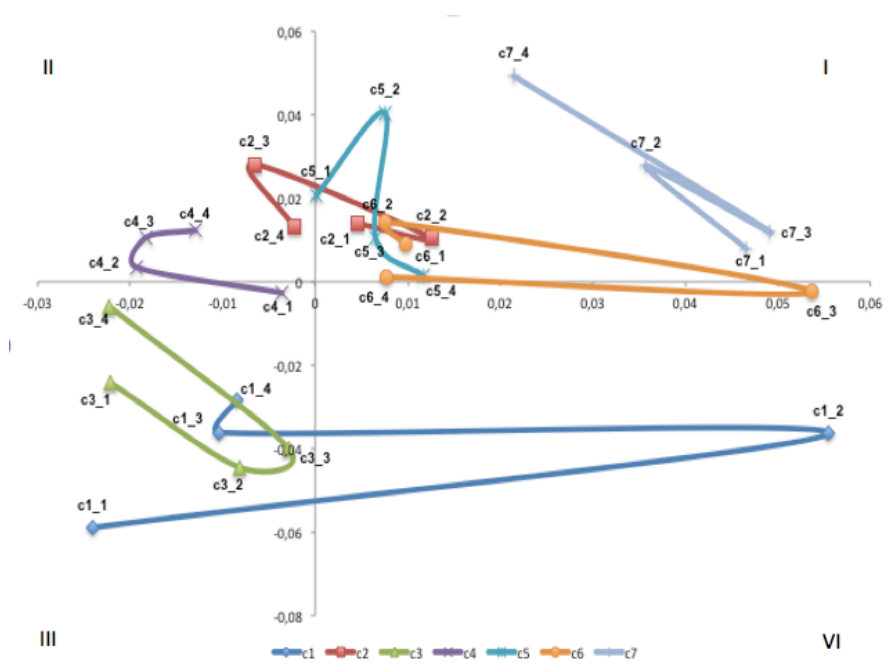

Figure 1- Clusters Trajectories onto the compromise plan

References

- Butts, C.T. (2008) "A relational event framework for social action". *Sociological Methodology* 38(1): 155-200.
- D'Urso, P. and Vichi, M. (1998) "Dissimilarities between trajectories of a three-way longitudinal data set". In: Rizzi A Vichi M (eds.) *Advances in data science and classification.* Springer, Heidelberg.
- Escoufier, Y. (1980) "L'analyse conjointe de plusieurs matrices", in Jolivet et al. (eds.), *Biométrie et Temps*, Société Francaise de Biométrie.
- Escofier, B. and Pagès, J. (1984) "L'analyse factorielle multiple", in *Cahiers du Bureau Univ. Recherche Operat.,* Série Recherche 423.
- Kiers, H.A.L. (1989) "INDSCAL for the analysis of categorical data". In: R. Coppi and S. Bolasco (eds.), *Analysis of multiway data matrices.* Amsterdam: Elsevier Science Publishers. pp. 155-167.
- Krivitsky, P.N. and Handcock, M.S. (2010) "A separable model for dynamic networks". arXiv preprint arXiv:1011.1937.
- Liberati, C., Marzo, M., Zagaglia, P. and Zappa, P. (2012). "Structural distortions in the Euro interbank market: The role of 'key players' during the recent market turmoil". *Quaderni DSE Working Paper No. 841.*
- Putnam, R. (2000) *Bowling Alone: The Collapse and Revival of American Community*. Simon & Schuster, New York.
- Robert, P. and Escoufier, Y. (1976) "A Unifying Tool for Linear Multivariate Statistical Methods: The RV-Coefficient". *Applied Statistics* 25 (3): 57–265.
- Snijders, T.A.B., Van de Bunt, G.G. and Steglich, C.E.G. (2010) "Introduction to stochastic actor-based models for network dynamics". *Social Networks* 32(1): 44-60.
- Tucker, L.R. (1966) "Some mathematical notes on three-mode factor analysis". *Psychometrika*, 31: 279-311.