International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 5 Issue 2, January-February 2021 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

Using Multi-Dimensional Scaling Method to Compare Stock Markets

Dr. Le Nguyen Doan Khoi

Associate Professor, Department of Scientific Research Affairs, Can Tho University, Can Tho, Vietnam

ABSTRACT

This research proposed a graphical method that can defer insights possible similarities across stock markets in key America, Asia, and Europe. The multi-dimensional scaling (MDS) technique was applied to visualize dynamics of cross-dependencies across stock markets before and after introduction of the Euro. Data consists of 5541 daily returns of 13 stock markets which include two US markets, seven Europe markets and four Asia markets from 1986 to 2007.

KEYWORDS: *stock markets returns, multi-dimensional scaling, regression analysis, America, Asia, Europe, graphical method*

How to cite this paper: Dr. Le Nguyen Doan Khoi "Using Multi-Dimensional Scaling Method to Compare Stock Markets" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-5 | Issue-2, February 2021, pp.975-978, URL: www.ijtsrd.com/papers/ijtsrd35912.pdf

IJTSRD35912

1. INTRODUCTION

Recent globalization tendencies in international financial markets motivated researchers to investigate the degree and dynamics of international financial integration, since the later has important economic benefits in terms of risk-sharing and consumptions smoothing (Cochrane, 1991), potential for higher economic growth (Levine, 1997) and more efficient conduct of monetary policy (Suardi, 2001).

A number of empirical methods have been offered in the literature to measure the degree of financial integration. One approach relies on the notions of sigma- and beta-convergence borrowed from the economic growth literature (examples are Adam et al, 2002; Baele et al., 2004). The intuition behind this methodology is to measure the speed of transmission of changes in financial returns in one country on returns in other countries. The markets are perceived to be more integrated when the speed of transmission is faster.

Another methodology is based on multivariate extension of autoregressive conditional heteroskedasticity models (examples are Bollerslev, 1990; Kaminski and Peruga, 1990; Engle and Susmel, 1993). The idea is to measure time-varying correlation structure in international stock market returns and identify possible factors driving their changes over time.

MDS is a descriptive method allowing to defer insights into possible similarities across stock markets. This methodology was applied for studying correlation structure across 13 major stock markets by Groenen and Franses (2000). This study extends the dataset used in Groenen and Franses and includes the period after introduction of Euro.

This paper is organized as follows: following the introduction, the second section describes the MDS methodology. The third section provides the data. Estimation results are discussed in the fourth section. Finally, we draw conclusion in the last section.

2. Methodology

2.1. Multidimensional scaling (MDS)

MDS is a popular technique in several social sciences which aims at representing a (m x m) proximity matrix such as a correlation matrix in a graphical way. In our case, we consider a (13x13) matrix of correlations across markets. The purpose of this assignment is to measure and visualize similarity between stock markets in different countries using multidimensional scaling (MDS) method. In our application we use metric MDS, since the market proximity measures (correlations) are cardinal by nature.[1] A small distance between two points in the graph corresponds to closer association between the two objects (in our case stock markets). MDS does not impose any distributional assumptions on the data and is considered to be a descriptive method.

In most practical application, the distances are not exactly equal to one minus the relevant correlations, and hence an approximate solution needs to be found. We use minimizing stress function as an approximate solution.

[1] An alternative methodology is non-metric MDS, which assumes that the proximity measures are ordinal.

2.2. Minimizing stress function

The MDS solution relies on minimization of differences between distances in the graphical representation and dissimilarities coming from the proximity matrix. Kruskal (1964) proposed STRESS function to approximate solution for the MDS problem:

$$STRESS = L(X) = \frac{\sum_{i<j}^{13}[(1-r_{ij})-d_{ij}(X)]^2}{\sum_{i<j}^{13}(1-r_{ij})^2}$$

where r_{ij} denotes the correlation coefficient between stock markets *i* and *j*, $d_{ij}(X)$ denotes Euclidean distance in a p-dimensional space between rows *i* and *j* of the 13*xp* matrix of coordinates *X*. The sum of the individual deviations is scaled over the sum of squared distances to normalize the data. The coordinates *X* that minimize STRESS can't be found by analytical methods and should be computed by an iterative algorithm.[2]

2.3. Choice of dimensionality

The critical question in MDS analysis is to select the number of dimensions *p*. Intuitively, the larger is the number of dimensions of the distance representation, the more flexible the model is producing higher degree of fit. However, there is a risk of over fitting the model. Also, visual representation of the distances is complicated when the number of dimensions is greater than 2 and is impossible for the cases when the number of dimensions is more than 3.

One way to select the number of dimensions is "elbow criterion". For this purpose, a scatter plot is made of the STRESS values obtained in various dimensions. Then, the number of dimensions where an elbow occurs defines the dimensionality to be chosen.

2.4. Procrustes rotation

One of the properties of Euclidean distances is rotational invariance, which means that any rotation of the coordinates gives exactly the same distances. This property implies that any MDS solution can be freely rotated without affecting the STRESS. For the procedure of application of MDS to the sequential subsamples outlined above, rotational invariance implies that the points may be placed differently on the screen between two subsequent time frames, even though their distances are almost the same. To avoid this problem, a method called Procrustes rotation will be applied, which allows for the comparability of MDS outcomes for different subsample (Cliff, 1966).

The objective of Procrustes rotation is to minimize $||X_{t1}-X_{t2}*T||^2$, where *T* is the rotation matrix to be estimated, ||.|| denotes an operator of the sum of squared elements and t_1, t_2 denote the two subsamples we analyze. The rotational matrix *T* that minimizes the loss equals *QP'*, where *Q* and *P* are orthonormal matrices (i.e. *P'P=Q'Q=I*) given by the singular value decomposition $X_{t1}'X_{t2}=P\Phi Q'$, with Φ being the diagonal matrix with non-negative singular values.

3. DATA

The dataset we are employing contains an extended sample of stock market return series used in Groenen and Franses (2000). The data are obtained from Data-stream, and they measures indexes in local currencies. Our data consists of 5541 daily returns of 13 stock markets from 1986 to 2007.[3] The stock markets in our sample include two US markets, seven European markets[4] and four Asian markets (see Table 1).

Table 1: Stock markets

	Stock market	Abbreviation	Country
1.	Brussels	brus	Belgium
2.	Amsterdam	cbs	The Netherlands
3.	Frankfurt	dax	Germany
4.	New York	dj	USA
5.	London	ftse	UK
6.	Hong Kong	hs	Hong Kong
7.	Madrid	madrid	Spain
8.	Milan	milan	Italy
9.	Tokyo	nikkei	Japan
10.	Singapore	sing	Singapore
11.	Standard and Poors	sp	USA
12.	Taiwan	taiwan	Taiwan
13.	Stockholm	vec	Sweden

Source: Datastream.

Descriptive statistics of the data for the two sub-periods are displayed in Table 2. The data series exhibit excess kurtosis and skewness in most cases for both samples, which is a standard finding in the financial markets literature (see Franses and van Dijk, 2000). In the second sample, however, the magnitude of the kurtosis is somewhat smaller, which might be due to the fact that in the first sample there have been more episodes of financial turbulences. The returns are on average positive in both samples, but lower in magnitude for most of the series in the second sample, which is a relatively more tranquil period.

Table 2: Summary statistics of returns (in logs)

	Mean	Median	Minimum	Maximum	St. Dev.	Kurtosis	Skewness
Sample 1: 1986-1999 (3391 observations)							
brus	0.0005	0.0003	-0.1109	0.0813	0.0081	21.5868	-0.9613
aex	0.0004	0.0004	-0.1278	0.1118	0.0119	13.4766	-0.5965
dax	0.0004	0.0003	-0.1371	0.0729	0.0129	10.0328	-0.8869
dj	0.0005	0.0003	-0.2247	0.0842	0.0097	91.0582	-4.0352
ftse	0.0004	0.0003	-0.1303	0.0760	0.0095	19.8002	-1.3129
hs	0.0008	0.0003	-0.3820	0.1703	0.0181	76.4065	-3.5350
madrid	0.0006	0.0001	-0.0902	0.0800	0.0125	6.6578	-0.3941
milan	0.0003	0.0000	-0.5217	0.5217	0.0224	428.7501	0.5333
nikkei	0.0000	0.0000	-0.1614	0.1243	0.0138	10.4413	-0.1476
sing	0.0002	0.0000	-0.2600	0.1263	0.0135	54.8947	-2.4640
sp	0.0005	0.0004	-0.2283	0.0871	0.0101	82.9083	-3.8052
taiwan	0.0006	0.0000	-0.1029	0.1284	0.0206	2.8310	-0.0852
vec	0.0006	0.0005	-0.0831	0.1154	0.0129	6.9893	-0.0183
Sample 2: 1999-2007 (2150 observations)							
brus	0.0002	0.0003	-0.0447	0.0744	0.0100	5.7637	0.2402
aex	0.0000	0.0003	-0.0753	0.0952	0.0145	4.6560	-0.0526
dax	0.0001	0.0005	-0.0887	0.0755	0.0156	2.9529	-0.0918
dj	0.0002	0.0000	-0.0815	0.0535	0.0101	4.1517	-0.2112
ftse	0.0000	0.0000	-0.0589	0.0590	0.0112	3.1123	-0.2084
hs	0.0008	0.0000	-0.0935	0.1634	0.0143	13.4722	0.8557
madrid	0.0002	0.0003	-0.0727	0.0653	0.0132	2.7769	0.0082
milan	0.0004	0.0003	-0.0372	0.0470	0.0057	10.9947	-0.5958
nikkei	0.0001	0.0000	-0.0723	0.0722	0.0134	1.9059	-0.1329
sing	0.0004	0.0002	-0.0784	0.0544	0.0121	3.1917	-0.1731
sp	0.0001	0.0000	-0.0601	0.0557	0.0110	2.5502	0.0854
taiwan	0.0001	0.0000	-0.0994	0.0852	0.0157	3.0292	-0.0641
vec	0.0003	0.0002	-0.0869	0.0857	0.0164	2.6370	-0.0156

Source: Datastream and own estimations.

The total sample covers two sub-periods: period 1986-1999 (3391 observations), during which each European country had its own national currency, and period 1999-2007 (2150 observations), when a single currency was launched in Europe.

[2] Notice that in the above specification we use ($1-r_{ij}$) as a dissimilarity measure, rather than r_{ij} as a similarity measure.

[3] The returns are defined as the first differences in logs of stock index values.

[4] Five out of seven European countries in our sample introduced Euro in 1999. Those countries are Belgium, Germany, The Netherlands, Spain and Italy.

Using the return series, we calculated correlation matrices for two sub-periods (see Table 3). Preliminary examination of the correlation matrices for two sub-periods suggests that correlations have increased over time for most of the Eurozone member European countries (shadowed cells). This finding provides first evidence on increased interdependence across European stock markets. To plot the observed interdependence visually, in the next section we discuss the multidimensional scaling estimation results.

Table 3: Correlation matrix

	brus	aex	dax	dj	ftse	hs	madrid	milan	nikkei	sing	sp	taiwan	vec
Sample 1: 1986-1999 (3391 observations)													
brus	1.0												
aex	0.4	1.0											
dax	0.5	0.7	1.0										
dj	0.2	0.4	0.3	1.0									
ftse	0.3	0.6	0.5	0.4	1.0								
hs	0.3	0.3	0.4	0.2	0.3	1.0							
madrid	0.4	0.5	0.5	0.2	0.4	0.3	1.0						
milan	0.1	0.0	0.1	0.0	0.1	0.0	0.1	1.0					
nikkei	0.3	0.3	0.3	0.1	0.3	0.3	0.3	0.0	1.0				
sing	0.4	0.3	0.3	0.2	0.3	0.4	0.3	0.1	0.3	1.0			
sp	0.2	0.4	0.3	1.0	0.4	0.2	0.2	0.0	0.1	0.1	1.0		
taiwan	0.1	0.1	0.1	0.0	0.1	0.1	0.1	0.0	0.1	0.2	0.0	1.0	
vec	0.4	0.5	0.5	0.2	0.5	0.3	0.5	0.1	0.3	0.3	0.2	0.1	1.0
Sample 2: 1999-2007 (2150 observations)													
brus	1.0												
aex	0.8	1.0											
dax	0.7	0.8	1.0										
dj	0.4	0.4	0.5	1.0									
ftse	0.7	0.8	0.7	0.4	1.0								
hs	0.3	0.3	0.3	0.1	0.3	1.0							
madrid	0.7	0.8	0.8	0.4	0.7	0.3	1.0						
milan	0.2	0.2	0.1	0.1	0.2	0.2	0.1	1.0					
nikkei	0.2	0.2	0.2	0.1	0.2	0.5	0.2	0.2	1.0				
sing	0.3	0.3	0.3	0.2	0.3	0.5	0.3	0.2	0.4	1.0			
sp	0.4	0.5	0.6	0.9	0.4	0.1	0.4	0.1	0.1	0.1	1.0		
taiwan	0.1	0.2	0.1	0.1	0.1	0.3	0.1	0.1	0.3	0.3	0.1	1.0	
vec	0.6	0.7	0.7	0.4	0.7	0.3	0.7	0.2	0.3	0.3	0.4	0.2	1.0

4. Estimation results

We begin our empirical analysis by identifying the number of dimensions of the perceptual map. For this reason we use scree plots of the STRESS measure. For the first subsample (see Figure 1) we find the "elbow" to be at the two dimensions point. The same finding holds for the second subsample (see Figure 2). The STRESS values are around 0.05, which is considered to be an acceptable level of precision by the "rule of thumb" described in Lattin, Carroll ad Green (2003). Thus, based on the "elbow" criterion we select two dimensions for our future investigation, which will allow us to represent the distances in a two-dimensional space.

Consequently, comparing the two scree plots we can observe that the magnitude of the STRESS measures for any dimension is substantially smaller for the second subsample, which implies better fit for the post-EU accession period

Figure 1: Scree plot for 1986-1999 subsample

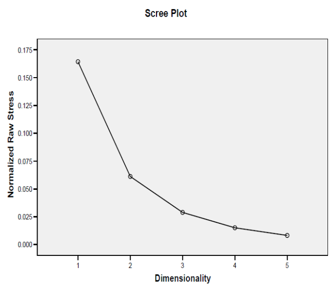
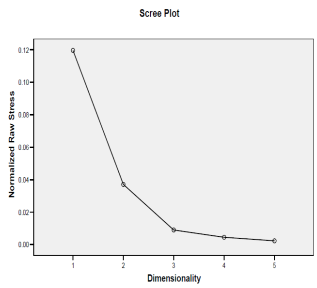


Figure 2: Scree plot for 1999-2007 subsample



Regression Analysis

To explain factors affecting co-movements in stock markets, we adopt "gravity" model approach (see Flavin, Hurley and Rousseau, 2001). These models are borrowed from the international trade literature. The key assumption of the gravity analysis is that geographical distances across international markets matter for the cross-country integration. Intuitively, the larger is the geographical distance, the lesser are the opportunities for economic integration. It is also a common practice to use variables explaining economic activity in cross-country pairs as explanatory variables affecting inter-country linkages. We use $GDP_{ij}=(GDP_i*GDP_j)^{1/2}$ as a measure of economic activity in a country-pair, where GDP_i and GDP_j denote growth rates in a gross domestic product (GDP) in countries *i* and *j*, respectively. The data on geographical distances and economic activity measures is displayed in Tables 4 and 5, respectively.

Table 4: Distances between cities (in kilometers)

	Brussels	Amsterdam	Frankfurt	New York	London	Hong Kong	Madrid	Milan	Tokyo	Singapore	Taiwan	Stockholm
Brussels												
Amsterdam	173											
Frankfurt	724	653										
New York	5884	5860	6458									
London	319	358	1008	5567								
Hong Kong	9394	9278	8685	12953	9623							
Madrid	1314	1480	1918	5766	1263	10536						
Milan	697	826	858	6460	958	9349	1188					
Tokyo	9446	9286	8875	10839	9555	2886	10758	9711				
Singapore	10544	10481	9829	17761	10837	2578	11369	10250	5315			
Taiwan	9641	9512	8951	12645	9845	723	10845	9669	2219	3122		
Stockholm	1281	1125	806	6314	1431	8225	2592	1650	8167	9629	8422	

Table 5: Economic activity (product of cross-country GDP growth rates)

	Brussels	Amsterdam	Frankfurt	New York	London	Hong Kong	Madrid	Milan	Tokyo	Singapore	Taiwan	Stockholm
Sample 1												
Brussels												
Amsterdam	0.025											
Frankfurt	0.025	0.026										
New York	0.027	0.028	0.028									
London	0.023	0.024	0.024	0.026								
Hong Kong	0.038	0.039	0.039	0.043	0.036							
Madrid	0.026	0.027	0.027	0.030	0.025	0.042						
Milan	0.025	0.026	0.026	0.028	0.024	0.039	0.027					
Tokyo	0.028	0.029	0.029	0.032	0.027	0.044	0.030	0.028				
Singapore	0.043	0.044	0.044	0.049	0.041	0.068	0.047	0.044	0.049			
Taiwan	0.027	0.028	0.028	0.031	0.026	0.043	0.030	0.028	0.032	0.049		
Stockholm	0.043	0.045	0.045	0.050	0.042	0.069	0.048	0.045	0.050	0.077	0.050	
Sample 2												
Brussels												
Amsterdam	0.030											
Frankfurt	0.016	0.022										
New York	0.025	0.034	0.018									
London	0.024	0.033	0.018	0.028								
Hong Kong	0.022	0.029	0.016	0.025	0.024							
Madrid	0.030	0.040	0.022	0.034	0.033	0.029						
Milan	0.017	0.023	0.012	0.019	0.018	0.016	0.023					
Tokyo	0.017	0.023	0.013	0.019	0.019	0.017	0.023	0.013				
Singapore	0.033	0.045	0.024	0.038	0.037	0.033	0.045	0.025	0.026			
Taiwan	0.025	0.034	0.018	0.028	0.028	0.025	0.034	0.019	0.019	0.038		
Stockholm	0.028	0.038	0.021	0.032	0.031	0.028	0.038	0.021	0.022	0.042	0.032	

The regression equation takes the following form:

$$D_k = \alpha_k + \beta_{1k} * GDP_{k,ij} + \beta_{2k} * \log(DISTANCE) + \varepsilon_k$$

where index *k={1,2}* stands for the two separate regressions in two subsamples (k=1 for 1986-1999 and k=2 for 1999-2007), $GDP_{k,ij}$ stands for the economic activity variable for countries i and j in two subsamples, and DISTANCE denotes a geographical distance variable (does not vary over time).

Estimation results of regression equations are displayed in Table 6. It is important to notice that the dependent variable is measured as dissimilarity (1-correlation), so negative coefficients imply that the variable has a positive impact on the correlation.

Table 6: Regression estimation results: distances in two subsamples

	Sample 1		Sample 2	
	coefficients	p-values	coefficients	p-values
Constant	0.6567	0.1947	**-1.1197**	0.0023
GDP	0.1844	0.1114	**-0.1983**	0.0218
DISTANCE	**0.0453**	0.0637	**0.0992**	0.0000
F-test (joint significance)	4.9356	0.0097	17.4239	0.0000
R^2	0.12		0.32	
DW statistic	2.07		2.01	
# observations	78		78	

Note: Estimations are performed using OLS.

The regression outcome suggests that distance is significant explanatory variable driving dissimilarity between stock indexes: the larger is the distance, the less is the correlation. The impact of distance almost doubled for the second subsample, rising from 0.4% to 0.09%, suggesting more pronounced geographical clasterization over time.

The economic activity measure is not significant in the first sample, but becomes significant in the second sample. The coefficient is negative, which implies that country pairs with higher level of economic activity also exhibit higher stock market correlation. This finding is in-line with the intuition behind "gravity" models.

5. CONCLUSION

This paper proposed MDS methodology on the extended sample of Groenen and Franses (2000) to study similarity between 13 stock markets. The extension of the sample allowed us to separate the effect of introduction of Euro on interdependencies between stock markets. Another difference from the Groenen and Franses (2000) is that we have applied “gravity” equation methodology to identify factors driving distances between stock markets.

The estimation results suggest that two dimensions are reasonably sufficient for explaining substantial part of the distances between markets in two subsamples (before and after introduction of Euro). Regression analysis predicts that the size of the similarity is significantly affected by geographical distance between stock markets in both subsamples – the closer the markets; the higher is the correlation between stock returns. The impact of distance is growing over time, with the elasticity coefficient being two times higher in the second subsample. Economic activity was found to have a significant impact only in the second sample – the larger is the growth rate for a give country pair, the more correlated the stock markets are.

REFERENCES

[1] Adam, K., T. Japelli, A. Menichini, M. Padula, and M. Pagano (2002), “Analyze, Compare, and Apply Alternative Indicators and Monitoring Methodologies to Measure the Evolution of Capital Market Integration in the European Union”, *Report to the European Commission*

[2] Baele, L., A. Fernando, P. Hordahl, E. Krylova, and C. Monnet (2004), “Measuring Financial Integration in the Euro Area”, *ECB Occasional Paper Series*, No. 14, Frankfurt

[3] Bollerslev, T. (1990), “Modeling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized Arch Model”, *The Review of Economics and Statistics* 72: 498–506

[4] Cliff, N. (1966), “Orthogonal Rotation to Congruence”, *Psychometrika* 31: 33-42

[5] Cochrane, J. (1991), “A Simple Test of Consumption Insurance”, *Journal of Political Economy* 99: 957–976

[6] Flavin, Hurley and Rousseau (2001), “Explaining Stock Market Correlation: A Gravity Model Approach”, paper presented at Irish Economic Association Conference, Portumna 2001

[7] Engle and Susmel (1993), “Common Volatility in International Equity Markets”, *Journal of Business and Economic Statistics* 11: 167-176

[8] Franses, P., and D. van Dijk (2000), Non-Linear Time Series Models in Empirical Finance, *Cambridge University Press, UK*

[9] Groenen, P. and P. Franses (2000), “Visualizing Time-varying Correlations across Stock markets”, *Journal of Empirical Finance* 7: 155-172

[10] Kaminski, G. and R. Peruga (1990), “Can a Time-Varying Risk Premium Explain Excess Returns in the Forward Market for Foreign Exchange?”, *Journal of International Economics* 28: 47-70

[11] Kruskal, J. (1964), “Multidimensional Scaling by Optimizing Goodness of Fit to a Nonmetric Hypothesis”, *Psychometrica* 29: 1-27

[12] Lattin, Carroll and Green (2003), Analyzing Multivariate Data, Duxburry Press

[13] Levine, R. (1997), “Financial Development and Economic Growth: Views and Agenda”, *Journal of Economic Literature* 35: 688–726

[14] Suardi, M. (2001), “EMU and Asymmetries in the Monetary Policy Transmission”, *Economic paper 157, European Commission, Brussels*