A STUDY OF EUROPEAN EQUITY MARKETS INTEGRATION USING MULTIVARIATE DATA ANALYSIS

Alexandra Horobet – Lucian Belascu – Erika Marin

Abstract

Our paper proposes a different approach regarding the study of financial integration among European equity markets, based on the first four moments of equity returns´ distributions for the 27 countries that are now EU members, calculated on an annual basis between 2000 and 2011. The first four moments are used as equity markets´ characteristics to identify their level of homogeneity using Statistical Cluster Analysis (SCA) and various clustering methods and amalgamation techniques. We observe the dynamic placements of EU equity markets in clusters, by taking into account the evolutions before 2008 and between 2008 and 2011. We find that European equity markets form a heterogeneous structure, as evidenced by returns' distributions, with a number of rather stable and homogeneous clusters – formed of mature equity markets in the EU, and with other smaller clusters, with a changing composition, which include mostly the emerging markets, which seem to be less integrated with the other EU equity markets.

Key words: equity markets, return distribution, financial integration, cluster analysis

JEL Code: F36, G15

Introduction

The European Union has been the subject of study of numerous works in financial integration, given its prominence as the best outlined framework for economic integration at the global level. It is widely believed that the introduction of a common currency is one of the underlying causes of increased financial integration, due to the removal of currency risk, with research strongly supporting this belief. For example, Yang et al. (2003) conclude that the emergence of EMU improved integration among member states, while other large EU financial markets such as the United Kingdom remained rather isolated. Kim et al. (2006) assert that the emergence of the monetary union changed investors' perceptions on the macroeconomic development and stability of member countries, which explains the dynamics of capital market integration. Kashefi (2006) studies the effect of the euro introduction on

European equity markets and finds a significant increase in correlations among stock returns between pre- and post-euro periods, which shows that diversification opportunities within EMU have decreased for the post-euro period. Nevertheless, there are authors that dispute the sufficiency of the existence of the monetary union for increased financial integration. In this context, Vo and Daly (2005) observe that European markets are far from being homogenous and would still allow investors to benefit from diversification within EMU.

The recent financial crisis raised concerns about the efficiency of the EMU and the process of financial integration in Europe. Although the current crisis has proven that an increased integration level among financial markets worldwide enhances the spillover effects of domestic shocks, the asymmetries between European markets suggest that the macroeconomic environment of these countries played a dominant role in the effects of this shock – see Lane (2012). Villaroya and de Guevara Radoselovics (2010) studied the relationship between financial integration and economic growth under the impact of the economic crisis and suggest that financial development and growth has increased since 1999, and occurred especially for the countries that were more integrated. In that sense, Eurozone had a faster growth than less integrated countries. Concerning the financial crisis, it has led to a reversal in integration since 2008, but this does not entirely explain the decline in either financial development or growth. Further on, Bartram and Wang (2011) found out that during the financial crisis European equity markets showed an increased dependence. With respect to the sovereign debt crisis that followed, the same study proves that it emphasized the heterogeneity among European countries, and Eurozone markets in particular. Therefore, the sovereign debt crisis reduced equity market integration for most industries, especially in Greece, considered as a high-risk country.

The objective of our paper is to investigate the dynamics of equity market integration in the European Union after the introduction of the euro currency, with a particular concern on the potential changes induced by the recent financial crisis. Various measures of financial integration have been used in the past, which can be divided in three broad categories, based on the approach proposed by Baele et al. (2004): price-based measures, news-based measures and quantity-based measures. While price-based measures take into account discrepancies between prices or asset returns due to assets' geographic region, the news-based measures are employed in order to distinguish the effect of new information arrival on markets from other frictions or barriers that may exist, based on the assumption that news in the form of new economic information of domestic or regional nature should have no impact on prices, while information of global nature should be more important for assets' prices and returns if markets are financially integrated. On the other hand, quantity-based measures examine the effects of frictions faced by the demand for and supply of investment opportunities, using statistics on the ease of market access across countries, for example from the perspective of cross-border lending or listings. Our research deviates from the previously proposed methods of examining financial integration, with the aim of detecting meaningful patterns in European equity markets from the perspective of their performance. We continue our previous research on capital markets' integration (Horobet and Lupu (2009), Horobet and Dumitrescu (2009)), in a new methodological setting, which brings into light various aspects of the equity markets' integration process in the European Union that have not been tackled before.

1 Data and research methodology

Our research uses data on daily returns' distributions of the main stock market indices of the 27 EU countries under analysis. We calculate the first four moments of daily returns' distributions and employ them as equity markets' indicators to identify their level of homogeneity – more specifically, we use in our analysis the mean to standard deviation (Mean/SD), skewness (Skew) and excess kurtosis (Kurt). For the overwhelming majority of countries data was available for each of the twelve years covered in our research, the main exception being Bulgaria (data available since 2001), Cyprus (since 2005) and Slovenia (since 2003). The data is collected from Bloomberg. Each variable has been standardized using the averages and standard deviations for each country in a specific year, as without standardization the variables with higher values would have a bigger impact on the clustering process compared to other values and would consequently bias the cluster formation.

The research methodology we employ is Statistical Cluster Analysis (SCA), whose goal resides in discovering natural clusters according to a specific internal criterion, without knowing beforehand the affiliation of entities to the identified clusters. The entities' assignment to a cluster is made by taking into account the similarity between the studied entities, according to the considered set of variables and the differentiation of entities that belong to a cluster from the ones that belong to other clusters. Our analysis is developed using Euclidian distances, and clusters are first formed using a hierarchical amalgamation algorithm and second by an integrative method – k-means algorithm. The hierarchical amalgamation is based on the Ward's method, which minimizes the sum of squares (SS) of any two clusters that can be formed at each step. We have opted for this method because it is considered to be very efficient in terms of final clustering result, although it tends to generate clusters of smaller size compared to the other amalgamation methods. The Euclidian distance D(i,K) of an observation i from cluster K, for a number of M continuous variables X_j is computed according to the following formula:

$$D(i,K) = \sqrt{\frac{1}{M}\sum_{j=1}^{M}\left(X_{ik} - \overline{X}_j^{(K)}\right)^2} \qquad (1)$$

where $\overline{X}_j^{(k)}$ is the mean for variable j and cluster K.

The k-means algorithm we employ calculates Euclidian distances from normalized quantities (i.e. values with a range between 0 and 1). For continuous variables, the distances are computed for rescaled values X_i, according to the formula below:

$$X_j^{'} = \frac{X_j - X_{\max}}{Max(X_j) - Min(X_j)} \qquad (2)$$

where $Min(X_j)$ and $Max(X_j)$ are the minimum and maximum values for variable i.

The difference between hierarchical clustering and k-means stems from the manner clusters are formed: in the hierarchical clustering algorithm clusters are formed step by step, starting with the entities that have the smallest distance and afterwards linking more and more entities together and aggregating larger and larger clusters of increasingly dissimilar entities until, in the last step, all entities are joined together; the k-means clustering algorithm, on the other hand, is based on a priori hypotheses concerning the number of clusters that may be formed based on the variables taken into account. The software used for our analysis is STATISTICA, the "Cluster Analysis" and "Generalized EM & K-means Cluster Analysis" modules. The result of this analysis will consist in identifying homogeneous groups of countries depending on the featured attributes in terms of their market performance. We are interested in observing the particular position of countries in clusters for each year, but also the transition of a country from one cluster to another over time, as an indication of capital market evolution in terms of integration. At the same time, we observe clusters' composition in two sub-periods in our time frame: 2000-2007 and 2008-2011, in order to investigate the likely impact of the financial crisis on EU capital markets' integration.

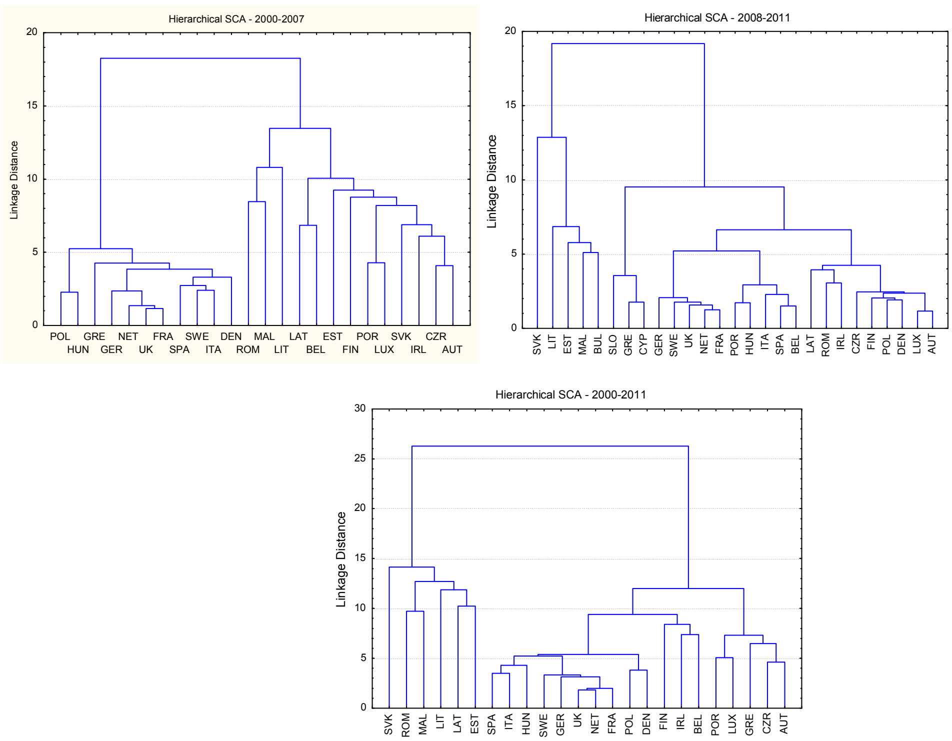
2 Results

Fig. 1 presents the results of the hierarchical clustering algorithm for the two sub-periods and the entire period[1], while Fig. 2 shows the average, minimum and maximum Euclidian

[1] All results are available from authors on request.

distances between our countries (cases), as well as the pairs of cases with the minimum and maximum distances, respectively.

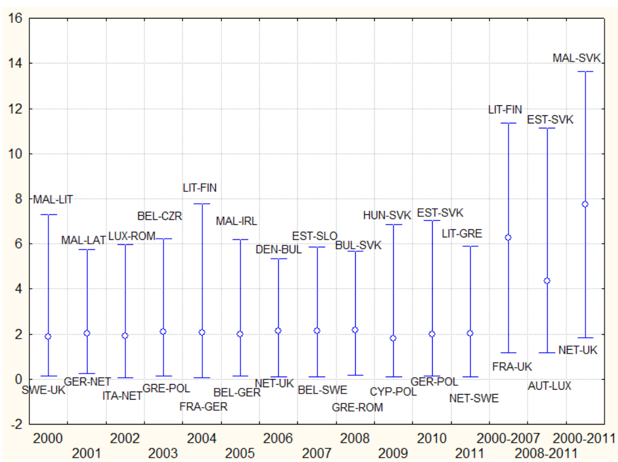
Fig. 1: Results of hierarchical clustering algorithm



Source: Authors‘ calculations

A number of interesting observations emerge from our results. First, the average Euclidian distance does not vary much between 2000 and 2011 – its lowest value is 1.8159 for 2009 and the highest is 2.1644 for 2008 -, and the same conclusion holds true for the minimum distance – the lowest value is 0.0576 for 2002 (between Italy and Netherlands) and the highest is 0.2429 for 2001 (between Germany and Netherlands). Second, there is a higher variability concerning the maximum Euclidian distances between countries: the lowest value is 5.3241, recorded in 2006 between Denmark and Bulgaria, and the highest is 7.7947, recorded in 2004 between Finland and Lithuania. Third, developed countries in the EU are generally clustered with the lowest distances in all years, although such pairs are also created between a developed and an emerging country in the EU in some years (Greece and Poland in 2003, Greece and Romania in 2008, Cyprus and Poland in 2009, and Germany and Poland in 2010). On the other hand, the highest Euclidian distances in all years belong to pairs formed of a developed and an emerging country in EU, or of two emerging countries. Moreover, when we study countries' placements in clusters over time, we generally observe countries grouped to a high extent depending on their level of economic development – developed countries tend to belong to the same clusters over the years, while emerging countries are grouped together. Nevertheless, in almost all years we observe a number of one to three countries that are somehow isolated from the others, indicating high dissimilarity in terms of market performance – for example, Lithuania in 2000, Luxembourg, Slovakia and Latvia in 2001, Romania and Malta in 2002, Lithuania and Bulgaria in 2003, Ireland and Estonia in 2005, Malta, Estonia and Bulgaria in 2008, Slovakia, Estonia and Lithuania in 2009, Slovakia in 2010, and Slovakia, Lithuania and Malta in 2011.

Fig. 2: Plot of minimum, maximum and average Euclidian distances, 2000-2011



Note: Circles indicate average Euclidian distances.
Source : Authors's calculations

The analysis over the two sub-periods confirms the results of hierarchical clustering on an annual basis. We observe (see Fig. 1) two well-defined clusters formed in 2000-2007, one that includes mostly developed markets (Germany, Netherlands, United Kingdom, France, Spain, Sweden, Italy and Denmark), and another that groups together the emerging countries of EU, with the exception of Poland and Hungary, which are closer to the first cluster, and of Austria and Luxembourg that belong to the emerging countries cluster. In the second sub-period the first cluster is maintained, while the remaining countries split in two other clusters: a group of countries joins Luxembourg and Austria, and another cluster is formed separately around Malta. The entire period shows the same developed markets' cluster (Germany, France, Netherlands, Sweden and United Kingdom) and a number of emerging markets that cluster at the last steps of the algorithm (Slovakia, Romania, Lithuania, Latvia and Estonia, joined also by Malta).

When we applied the k-Means algorithm we observe the tendency of countries to group in more than two clusters in some years, which also indicate more transitions of countries over time from one cluster to another. Tab. 1 shows clusters' members for all years, as well as for the two sub-periods and the overall period. We notice again the developed countries (France, Germany, United Kingdom, Netherlands and Spain) grouped together in all years, with a few exceptions – Sweden clusters separately in 2001, Germany in 2007, and France and Italy in 2010. At the same time, this cluster is occasionally joined by other developed or even emerging countries, but these larger clusters are not preserved over time.

Tab. 1: Clusters‘ members, 2000-2011

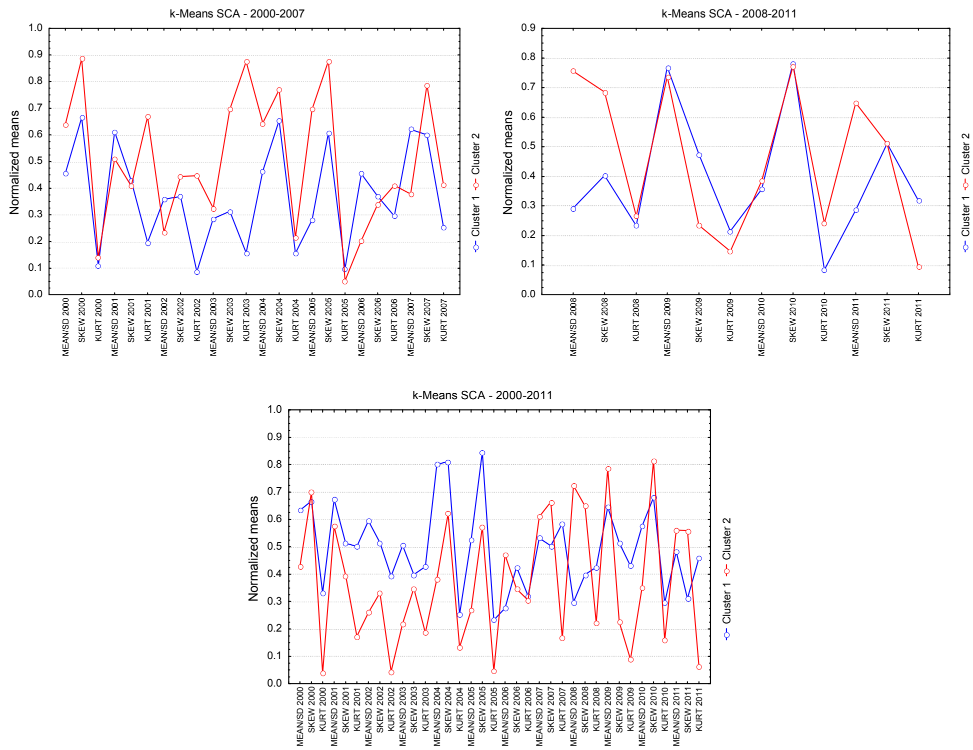
Country	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2000-2007	2008-2011	2000-2011
AUT	3	5	1	1	1	2	1	2	1	1	1	2	1	2	2
BEL	3	5	2	2	2	2	2	2	1	1	2	2	2	2	2
BUL	n/a	3	4	3	2	2	3	3	3	1	3	2	n/a	1	n/a
CYP	n/a	n/a	n/a	n/a	n/a	2	1	3	4	1	3	4	n/a	1	n/a
CZR	3	4	4	4	4	2	1	1	5	1	1	2	1	2	2
DEN	2	5	2	1	5	2	2	2	1	1	1	2	1	2	2
EST	3	5	4	1	4	1	2	1	3	2	1	2	1	1	1
FIN	3	5	2	4	3	2	2	3	2	1	1	2	1	2	2
FRA	3	5	2	1	5	2	2	2	2	1	2	3	1	2	2
GER	3	5	2	1	5	2	2	3	2	1	1	3	1	2	2
GRE	3	5	2	1	5	2	1	2	4	1	3	4	1	1	2
HUN	3	4	4	1	4	2	2	2	5	1	2	2	1	2	2
IRL	2	5	2	1	2	2	1	2	1	1	1	3	1	2	2
ITA	3	5	2	1	5	2	2	2	2	1	2	2	1	2	2
LAT	2	2	1	2	4	1	1	2	4	2	1	3	2	2	1
LIT	1	4	4	3	4	1	2	1	5	2	1	1	1	1	1
LUX	3	2	2	1	5	2	3	2	1	1	1	2	1	1	2
MAL	2	3	3	2	4	1	2	1	3	1	1	1	2	1	1
NET	3	5	2	1	5	2	2	2	2	1	1	3	1	2	2
POL	3	4	4	1	5	2	2	2	1	1	1	2	1	2	2
POR	3	5	2	4	5	2	3	1	2	1	2	2	1	2	2
ROM	2	1	3	1	1	2	2	3	4	1	2	2	1	2	1
SVK	2	2	4	1	1	2	1	1	5	2	2	3	1	2	1
SLO	n/a	n/a	n/a	n/a	n/a	2	3	3	4	1	3	1	n/a	1	n/a
SPA	3	5	2	1	5	2	2	2	2	1	2	3	1	2	2
SWE	3	4	2	1	5	2	2	2	2	1	1	3	1	2	2
UK	3	5	2	1	5	2	2	2	2	1	1	3	1	2	2
Number of clusters	**3**	**5**	**4**	**4**	**5**	2	3	3	5	2	3	4	2	2	2

Source: Authors‘ calculations

In both sub-periods and the overall period countries are clearly divided in two clusters, although these clusters' composition changes: Belgium, Latvia and Malta seem to form an outlier cluster in the first sub-period, but they are linked to one of the two main clusters formed in the second sub-period. Overall, we observe two clusters: one that includes only emerging countries (Estonia, Latvia, Lithuania, Romania and Slovakia, joined by Malta) and another that groups the remaining countries.

The k-Means algorithm results provide us with a better comparison between the formed clusters, when we observe the means of the variables for each cluster (see Fig. 3). The two clusters seem to be better differentiated during the first sub-period, although no definite conclusion can be extracted in terms of the category of variables that leads to this differentiation (all three types of variables – Mean/SD, Skew and Kurt are clearly different for some years, but rather similar for other years). This might be interpreted as a general tendency of countries to display comparable market performances in the more turbulent period between 2008 and 2011, which also influences the formation of clusters for the overall period.

Fig. 3: Results of k-means clustering algorithm – graph of means



Source: Authors‘ calculations

An analysis of the distances between clusters’ centroids (see Tab. 2) shows a lower level of dissimilarity between clusters in some years (2010, 2005, 2008 and 2007 have the lowest average distances between clusters’ centroids) and a higher level of dissimilarity particularly in 2003, 2002 and 2001. At the same time, no trend in terms of the average distance evolution in time can be detected, thus indicating that countries have volatile performances from one year to the other and thus clusters tend to be closer to each other in some years, while being more distanced in some years. The distance between clusters’ centroids is higher in the first sub-period compared to the second one (1.2747 against 0.7514), while the distance for the overall period is higher than the distances for both sub-periods (1.4634), indicating that when we consider the entire period between 2000 and 2011 countries tend to split in clusters that are more dissimilar compared to 2000-2007, let alone the 2008-2011 period.

Tab. 2: Distances between clusters' centroids, 2000-2011

Clusters	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2000-2007	2008-2011	2000-2011
1 and 2	1.2238	0.9384	0.3495	0.8094	0.2936	0.5679	0.3573	0.6375	0.3548	0.6605	0.4313	0.7268	1.2747	0.7514	1.4634
1 and 3	1.1869	0.8593	1.1462	0.7702	1.2507		0.7861	0.7744	0.7298		0.5050	0.9048			
1 and 4		0.6655	0.3871	0.3175	0.3447				0.2838			0.8508			
1 and 5		0.7641			0.5454				0.5512						
2 and 3	0.3766	0.8718	1.1274	0.9146	1.0502		0.8691	0.3840	1.0227		0.4715	0.2989			
2 and 4		0.6862	0.4359	0.9408	0.4322				0.4559			0.5168			
2 and 5		0.5004			0.3342				0.4841						
3 and 4		0.8093	1.0111	0.8338	1.3955				0.6317			0.7221			
3 and 5		0.8670			1.1972				0.8832						
4 and 5		0.2253			0.4662				0.5780						
Average distance	**0.9291**	**0.7017**	**0.7429**	**0.7644**	**0.7310**	**0.5679**	**0.6708**	**0.5986**	**0.5975**	**0.6605**	**0.4692**	**0.6700**	**1.2747**	**0.7514**	**1.4634**

Source: Authors‘ calculations

Conclusion

The results we obtained are homogeneous and consistent, leading us towards concluding that the European equity market is a heterogeneous structure, at least in terms of performance, with a number of rather stable clusters that move in time towards more homogeneity – formed of mature capital markets in the EU (France, Germany, United Kingdom, Sweden, Netherlands and Spain), and with other clusters, which seem to be less integrated with the other EU capital markets. An interesting result refers to the fact that emerging countries are typically clustered together over the years, but after 2008 some of them leave the traditional emerging countries' cluster in order to join clusters formed mostly of developed markets. These results suggest that financial integration in EU is a process led by the mature markets, while emerging countries are following, but with different paces. At the same time, countries tend to display higher similarity in more turbulent times, but this result should not be necessarily interpreted as indicating a higher level of integration – only increased dependence and analogous behaviour of equity markets in periods of crisis.

References

Baele, L., Ferrando, A., Hördahl, P., Krylova , E., & Monnet, C. European Central Bank, (2004). Measuring financial integration in the Euro area (Occasional Paper Series No. 14). Retrieved from website: http://www.ecb.int/pub/pdf/scpops/ecbocp14.pdf

Bartram, S., & Wang, Y. H. (2011). European financial market integration in the wake of the sovereign debt crisis: An industry analysis. European Finance eJournal, 6, doi: 10.2139/ssrn.1863797

Horobeţ, A., & Dumitrescu, S. (2009). On the causal relationships between monetary, financial and real macroeconomic variables: evidence from Central and Eastern Europe. Economic Computation and Economic Cybernetics Studies and Research, 43(3), 77-93.

Horobeţ, A., & Lupu, R. (2009). Are Capital Markets Integrated? A Test of Information Transmission within the European Union. Romanian Journal of Economic Forecasting, 10(2), 64-80.

Kashefi, J. (2006). The effect of the euro on European equity markets and international diversification. Journal of International Business Research, 5(2), 1-21.

Kim, S.-J., Moshirian, F., & Wu, E. (2006). Evolution of International Stock and Bond Market Integration: Influence of the European Monetary Union. Journal of Banking and Finance, 30(5), 1507-1534.

Lane, P. R. (2012). The European Sovereign Debt Crisis. Journal of Economic Perspectives, 26(3), 49–68.

Villarroya, J. M., & de Guevara Radoselovics, J. F. (2010). Financial crisis, financial integration and economic growth: The European case. Fundacion BBVA, Bilbao. Retrieved from http://www.fbbva.es/TLFU/dat/DT-02-2010.pdf

Vo, X. V., & Daly, K. J. (2005). European equity markets integration - implications for US investors. Research in International Business and Finance, 19(1) 155–170.

Yang, J., Min, I., & Li, Q. (2003). European stock market integration: Does EMU matter?. Journal of Business Finance & Accounting, 30(9&10), 1256-1276.

Contact

Alexandra Horobet
Bucharest University of Economic Studies
6 Piata Romana, Bucharest-1, Romania
alexandra.horobet@gmail.com

Lucian Belascu
Lucian Blaga University of Sibiu
17 Calea Dumbravii, Sibiu, Romania
lucian_belascu@yahoo.com

Erika Marin
Bucharest University of Economic Studies
6 Piata Romana, Bucharest-1, Romania
erika_tusa@yahoo.com