CEFAGE
universidade de évora
centro de estudos e formação avançada em gestão e economia

CEFAGE-UE Working Paper
2010/06

# ON THE GLOBALIZATION OF STOCK MARKETS: AN APPLICATION OF VECM, SSA TECHNIQUE AND MUTUAL INFORMATION TO THE G7?

*Rui Menezes*[a], *Andreia Dionísio*[b] *and Hossein Hassani*[c]

[a]ISCTE *Business School*
[b]University of Évora and CEFAGE-UE
[c] Cardiff School of Mathematics, Cardiff University

# ON THE GLOBALIZATION OF STOCK MARKETS: AN APPLICATION OF VECM, SSA TECHNIQUE AND MUTUAL INFORMATION TO THE G7

**Rui Menezes[a], Andreia Dionisio[b] and Hossein Hassani[c]**

[a]ISCTE *Business School*, Av das Forças Armadas, 1649-026 Lisboa, Portugal, rui.menezes@iscte.pt
[b]University of Évora and CEFAGE-UE, Largo dos Colegiais, 2, 7000-803 Évora, Portugal, andreia@uevora.pt
[c] Cardiff School of Mathematics, Cardiff University, Cardiff, CF24 4AG, hassanih@cf.ac.uk

**Abstract**
This paper analyzes the process of stock market globalization on the basis of two different approaches: (*i*) the linear one, based on cointegration tests and vector error correction models (VECM); and (*ii*) the nonlinear approach, based on Singular Spectrum Analysis (SSA) and mutual information tests. While the cointegration tests are based on regression models and typically capture linearities in the data, mutual information and SSA are well suited for capturing global non-parametric relationships in the data without imposing any structure or restriction on the model. The data used in our empirical analysis were drawn from DataStream and comprise the natural logarithms of relative stock market indexes since 1973 for the G7 countries. The main results point to the conclusion that significant causal effects occur in this context and that mutual information and the global correlation coefficient actually provide more information on this process than VECM, but the direction of causality is difficult to distinguish in the former case. In this field, SSA shows some advantages, since it enabled us to capture the nonlinear causality in both directions. In all cases, however, there is evidence that stock markets are closely related in the long-run over the 36 years analyzed and, in this sense, one may say that they are globalized.

**Keywords**
Globalization, market integration, VECM, mutual information, SSA technique.

## 1. Introduction

Recent debates on economic globalization have triggered a substantial amount of research papers that try to determine its causes and explain the consequences of this phenomenon in terms of market performance and their ability to adjust globally to economic boosts and crises. This has been particularly relevant in the case of financial markets and even more so in the case of stock markets. Indeed, the process of globalization of international stock markets has been deeply studied and most of the times conclusions point to the evidence of globabization (Arshanapalli and Doukas, 1994; Chung and Liu, 1994; Kasa, 1992; Masih and Masih, 1997, 2002; Zhou and Sornette, 2003; Tavares, 2009, among others). However, many of these studies lack a theoretical background that supports their view of what is globalization and how it can be measured (Hamao *et al.,* 1990; Drodz *et al*., 2001).

Globalization, in its literal sense, is the process of transformation of local or regional phenomena into global ones and can be described as a process by which the world population is gradually more integrated into one sole society. That is, globalization implies uniformity in terms of tastes, behaviors, prices, goods accessibility, and much more. It is a process of interaction among the economic and social agents (people, firms, etc) driven by international trade and investment and aided by information technology that reduced significantly the geographical distance barriers and communication difficulties between people living in different parts of the world.

One important aspect of economic globalization is market integration. In the sense of Stigler (1969) and Sutton (1991) a market is "the area within which the price of an asset tends to uniformity after allowing for different transportation costs, differences in quality, marketing, etc". On the other hand, market integration refers to proportionality of price movements over time for an asset or group of assets. The economic variable *price* is, therefore, a key element in the process of market globalization and provides a suitable framework for testing market integration by looking at the price relationship of assets over time. Strictly speaking we should look at proportionality of price movements over time for a given asset sold in geographically separated markets in order to show whether these markets are integrated or not. This is what we may call *strong market integration*[1] but, in many cases, market integration only occurs in a weak or imperfect way. If this is so, one can expect nonlinearities and other types of price distortions to be present in the process of price transmission and a test of *weak market integration* can be performed on the basis of causality between prices, independently of whether they are proportional or not over time.

This definition of market integration can be mathematically expressed as a dynamic model where the long-run and the short-run effects can be clearly separated, known as the error correction mechanism. This model is quite flexible and allows for different impacts of price and returns (or log price changes) movements across markets. For example, a change in the US market, usually considered as the dominant market, may be transmitted in quite different manners to the remaining markets, in which case it is difficult to conclude that markets tend to uniformity. This is not compatible with strong market integration but fits very well in the notion of weak market integration. Indeed, the process of market globalization is complex and the nonlinear transmission of price movements must be properly accommodated within the context of stock market globalization (Bonfiglioli and Favero, 2005; Kim *et al.,* 2005; Savva, 2009; Corazza *et al*., 2010).

One advantage of the error correction model is that it allows for historical prices and returns to affect simultaneously the behavior of current stock market prices over time. Using historical prices and returns in this context is preferable to using just stock returns since the former retain both the long-run and the short-run information contained in the data, while the latter only capture the short-run information. This statement is valid under the assumption that prices are cointegrated, an issue that was extensively analyzed elsewhere (Engle and Granger, 1987; Eun and Shim, 1989; Alexander, 2008). On this basis, one can construct statistical tests to verify whether the past (and present) information contained in prices and returns of, say, market A, help to explain the behavior of prices and returns of market B. This is what we mean by Granger causality and, under this hypothesis, one can say that knowing the behavior of prices in market A allows one to explain or even predict the behavior of prices in market B.

Although there are some notable exceptions, time series models used for forecasting are often based on the restrictive assumptions of normality, stationarity and linearity. However, quite often financial and economic time series data are non-Gaussian and may be generated by processes that are nonstationary and/or nonlinear; hence, methods that do not depend on these assumptions

[1] If changes are proportional over time then the markets are said to be strongly integrated.

are likely to be useful for modelling and forecasting economic data. It is widely admitted that noçlinearity is an intrinsic and fundamental feature of foreign exchanges rates and stock returns (see for instance Hsieh (1991), Ammermann and Patterson (2003), Beine *et al.* (2008)).

Moreover, most existing models are also essentially parametric requiring the specification and estimation of models that are usually linear as the basis for analysis and forecasting. We depart from these assumptions and methods by constructing models based on Singular Spectrum Analysis (SSA), which does not embody the assumptions of normality, stationarity and linearity and is, therefore, a good candidate for modelling and forecasting these types of data. A general test based on the concept of mutual information may also, therefore, help to improve our knowledge of such co-movements. The mutual information has the advantage that it does not impose any structure or restriction to the model (Granger and Lin, 1994).

In this paper VECM, Granger causality tests, mutual information based tests and SSA causality tests will be employed in order to investigate whether the stock markets of the G7 countries are, in some way, related in the long and short-run and react in a systematic way to shocks occurring in the global market. A concise description of these methods is presented in the next section. In Section 3 we present the data set used in our empirical analysis and the main results that were obtained. Finally, Section 4 presents the main conclusions and some ideas for future research.

## 2. The econometric framework

This Section describes the methods identified in the Introduction for analyzing weak market integration: VECM, Granger causality, mutual information and the SSA technique.

### 2.1. VECM and Granger causality

As noted above, one way to analyze the extent of market integration, and thus globalization, is by using Granger causality tests (Granger, 1969) which can be defined as follows: $X_{2t}$ Granger causes $X_{1t}$ if, *ceteris paribus*, the past values of $X_{2t}$ help to improve the current forecast of $X_{1t}$, that is:

$$MSE\left(\hat{X}_{1t} \mid I_{t-1}\right) < MSE\left(\hat{X}_{1t} \mid I_{t-1} \setminus IX_{2,t-1}\right), \qquad (1)$$

where $MSE$ is the mean squared error, $I_{t-1}$ represents the set of all past and present information existing at moment $t-1$, $IX_{2,t-1}$ represents the set of all past and present information existing on $X_2$ at moment $t-1$, *i.e.*, $IX_{2,t-1} = \{X_{21}, X_{22}, \ldots, X_{2t-1}\}$, $X_{1t}$ is the value of $X_1$ at the moment $t$ ($X_{1t} \subset I_t$) and $\hat{X}_{1t}$ is a non biased predictor of $X_{1t}$. On the other hand, $X_{2t}$ instantaneously causes $X_{1t}$ in the sense of Granger if, *ceteris paribus*, the past and present values of $X_{2t}$ help to improve the prediction of the current value of $X_{1t}$, that is:

$$MSE\left(\hat{X}_{1t} \mid I_t \setminus X_{1t}\right) < MSE\left(\hat{X}_{1t} \mid I_t \setminus IX_{2,t}, X_{1t}\right). \qquad (2)$$

Given these definitions, how can we empirically implement these tests? To see this, consider the following ADL($p$, $q$) price relationship:

$$X_{1t} = \theta + \sum_{k=1}^{p} \rho_k X_{1,t-k} + \sum_{j=0}^{q} \beta_j X_{2,t-j} + v_t\,, \quad (3)$$

where $X_{it}$ ($i$ = 1, 2) denotes the relative prices (in natural logs) of asset $i$ at time $t$, $\rho_k$ captures the extent of autocorrelation in $X_{1t}$, $\beta_j$ measures the relationship between prices (in levels and lags) and $v_t$ is a white noise perturbation. One can say that $X_{2t}$ causes $X_{1t}$ if the null hypothesis that all parameters $\beta_j$ are simultaneously zero is rejected. The relationship can be bidirectional and, in this case, we say that there is a feedback relationship. If there is just one unidirectional causal relationship, then one of the markets can effectively influence the other market prices, but the reverse is not true. If the null hypothesis is not rejected in both cases, then there is no causal relationship between the underlying prices and one can say that they do not belong to the same market space. In practice, however, the Granger causality test performed in statistical software postulates as the null hypothesis that "$X_{2t}$ does not Granger cause $X_{1t}$".

In multivariate cointegrated systems the Granger causality test can be performed on the basis of a VECM of the type (Sargan, 1964; Alexander, 2008):

$$\Delta \mathbf{X}_t = \boldsymbol{\alpha\beta}' \mathbf{X}_{t-1} + \sum_{k=1}^{p-1} \boldsymbol{\Gamma}_k \Delta \mathbf{X}_{t-k} + \boldsymbol{\mu} + \boldsymbol{\varepsilon}_t\,, \quad (4)$$

where $\mathbf{X}_{t-1}$ is an $i$-dimensional vector of cointegrated lagged endogenous variables representing, for instance, natural logarithms of relative asset prices (*e.g.*, stock indexes) at time $t-1$. $\Delta\mathbf{X}_t$ and $\Delta\mathbf{X}_{t-k}$ denote returns at time $t$ and $t-k$, respectively, where $\Delta$ is the operator of first difference. $\boldsymbol{\Gamma}_k$ denotes $p-1$ $i$-order matrices of short-run information parameters where each of them is associated with an $i$-dimensional vector of lagged returns up to order $p-1$. $\boldsymbol{\alpha\beta}'$ is an $i$-order matrix of long-run information parameters, where $\boldsymbol{\alpha}$ represents the adjustment speed to equilibrium and $\boldsymbol{\beta}$ contains the long-run or equilibrium coefficients. $\boldsymbol{\mu}$ is an $i$-dimensional vector of constants and $\boldsymbol{\varepsilon}_t$ denotes an $i$-dimensional vector of residuals where $\boldsymbol{\varepsilon}_t \sim \text{iid}(\mathbf{0}, \boldsymbol{\Omega})$. Note that the residuals $\boldsymbol{\varepsilon}_t$ are not serially correlated since the dynamic process linking the data is explicitly specified in the model, although they may be contemporaneously correlated.

The VECM represented in equation (4) can be interpreted as a relationship between prices and returns in a given market. What it says is that the current returns are a linear function of previous returns and historical prices. Such historical prices form a long-run equilibrium relationship, where the involved variables co-move over time independently of the existence of stochastic trends in each of them, so that their difference is stable. The long-run residuals measure the distance of the system to equilibrium at each moment $t$, which may be due to the impossibility of the economic agents to adjust instantaneously to new information or to the short-run dynamics also present in the data. There is, therefore, a whole complex adjustment process involving short-run and long-run dynamics when the variables are cointegrated.

Simple manipulation of the VECM leads to a reparameterized version where the vector $\boldsymbol{\mu}$ is multiplied by the estimated long-run residuals and the matrices $\mathbf{A}_i$ ($i = 1, \ldots, m$) contain the coefficients of the lagged returns for each variable separately. For a two cointegrated variable system and $p$ lags,[2] and noting that $\hat{u}_{t-1} = \hat{\boldsymbol{\beta}}'\mathbf{X}_{t-1}$, one has:

$$\Delta\mathbf{X}_t = \mathbf{A}_1\Delta\mathbf{X}_{1,t-j} + \mathbf{A}_2\Delta\mathbf{X}_{2,t-j} + \boldsymbol{\mu}\hat{u}_{t-1} + \boldsymbol{\varepsilon}_t, \quad (5)$$

where $\Delta\mathbf{X}_t$ represents returns or log price changes at time $t$ and $\Delta\mathbf{X}_{i,t-j}$ ($i = 1, 2$; $j = 1, \ldots, p-1$) denotes lagged returns up to $p-1$ of the $i^{\text{th}}$ variable. $\mathbf{A}_1$ and $\mathbf{A}_2$ are $[2\times(p-1)]$ matrices. $\boldsymbol{\mu}$ and $\boldsymbol{\varepsilon}_t$ are $(2\times1)$ vectors and $\hat{u}_{t-1}$ denotes the long-run residuals, where $u_t \sim \mathrm{I}(0)$. A Granger causality test can be carried out on the basis of the null hypothesis: $\delta_{i1} = \ldots = \delta_{i,p-1} = \mu_i = 0$, where the $\delta_i$ coefficients correspond to the $i^{\text{th}}$ row of $\mathbf{A}_2$. The test then compares the mean squared error under the null and under the alternative hypotheses.

### 2.2. Mutual information

The mutual information of two continuous random variables $X_1$ and $X_2$ can be defined as:

$$I(X_1, X_2) = \int_{X_1}\int_{X_2} p(x_1, x_2)\ln\frac{p(x_1, x_2)}{p(x_1)\,p(x_2)}\,dx_2 dx_1, \quad (6)$$

where $p(x_1, x_2)$ is the joint probability distribution function of $X_1$ and $X_2$ and $p(x_1)$ and $p(x_2)$ are, respectively, the marginal probability distribution functions of $X_1$ and $X_2$. In the discrete case, one just replaces the double integral by a definite double summation. Intuitively, mutual information measures the information that $X_1$ and $X_2$ share: it measures how much knowing one of these variables reduces our uncertainty about the other. Mutual information can be also expressed as:

$$\begin{aligned} I(X_1, X_2) &= H(X_1) - H(X_1 \mid X_2) \\ &= H(X_2) - H(X_2 \mid X_1) \\ &= H(X_1) + H(X_2) - H(X_1, X_2) \end{aligned}, \quad (7)$$

where $H(X_1)$ and $H(X_2)$ are the marginal entropies, $H(X_1|X_2)$ and $H(X_2|X_1)$ are the conditional entropies and $H(X_1,X_2)$ is the joint entropy of $X_1$ and $X_2$[3].

Since $H(X_1) \geq H(X_1|X_2)$, one has that $I(X_1,X_2) \geq 0$ where the mutual information is zero if and only if $X_1$ and $X_2$ are statistically independent. Therefore, the mutual information between $X_1$ and $X_2$ can be regarded as a measure of dependence between these variables, or even better, a measure of the statistical correlation between $X_1$ and $X_2$. As noted by many authors, mutual information satisfies some of the desirable properties of a good measure of dependence (Granger and Lin, 1994; Dionisio *et al.* 2004)

One of the main difficulties in estimating mutual information from empirical data lies in the fact that the relevant p.d.f. is unknown. Although one way to deal with this problem is to approximate

---

[2] Notice, however, that the number of lags can be different for each variable.

[3] For more information about Information Theory, please see Shannon, C. E. (1948). A Mathematical Theory of Communication, Bell Systems Tech., 27: 379-423, 623-656

the densities by means of histograms, an arbitrary histogram would not be the best choice because it can cause underestimation or overestimation of the empirical mutual information. In line with this, Darbellay and Wuertz (2000) used a method called marginal equiquantization and showed how to proceed to obtain a homogeneous partition. This method consists in dividing the space partition into equiprobable cells iteratively. The process of space partition stops when local independence between cells is found using a $\chi^2$ test. This is the method that will be used in this paper.

Another difficulty of using mutual information as defined in equation (6) is that it takes values between 0 and infinity, turning the comparisons between different samples a problematic task. To overcome this problem several authors (namely Granger and Lin, 1994; Granger *et al.*, 2004, Dionisio *et al.* 2004, 2006) propose and use empirically a standardized measure of mutual information, known in the literature as the global correlation coefficient, given by:

$$\lambda = \left\{1 - \exp\left[-2I\left(X_1, X_2\right)\right]\right\}^{\frac{1}{2}}. \tag{8}$$

Note that $\lambda$ captures the overall dependence, both linear and nonlinear, between $X_1$ and $X_2$. This measure varies between 0 and 1 being thus directly comparable to the linear correlation coefficient based on the relationship between the measures of information theory and the analysis of variance. According to the properties of mutual information one can construct an independence test based on the following hypotheses:

$$\begin{aligned} &H_0 : I\left(X_1, X_2\right) = 0 \\ &H_1 : I\left(X_1, X_2\right) > 0 \end{aligned}. \tag{9}$$

If $p(x_1, x_2) = p(x_1)\, p(x_2)$ then $H_0$ is not rejected and $X_1$ and $X_2$ are said to be independent. Otherwise, if $p(x_1, x_2) \neq p(x_1)\, p(x_2)$ then $H_0$ is rejected, that is, the null hypothesis of independence is rejected and the acceptance of $H_1$ implies that $X_1$ and $X_2$ share some information, which enlarges as the value of $I(X_1,X_2)$ raises. In the latter case, the higher is the value of $I(X_1,X_2)$ the more knowing one of these variables reduces the uncertainty about the other, which is quite convenient for prediction analyses.

In order to test for independence between variables (or vectors of variables) we need to calculate the critical values of the underlying distributions. In this paper, we have simulated critical values for the null distribution or percentile approach, *i.e.*, the critical values were found through a simulation based upon a white noise for a number of sample sizes. Given that the distribution of mutual information is skewed, a percentile approach can be adopted to obtain the critical values (see the Appendix).

Despite being a good measure for a general test of independence, mutual information (and the global correlation coefficient) does not satisfy the triangle inequality and, therefore, it is not a measure of distance. Kraskov *et al.* (1996) proposed a modified mutual information-based measure which is strictly a metric. According to these authors, this modification presents some difficulties when we deal with continuous random variables. One solution for this problem consists of dividing mutual information by the sum or by the maximum of dimensions of the continuous variables in study.

### 2.3. Singular Spectrum Analysis

The SSA technique incorporates elements of classical time series analysis, multivariate statistics, multivariate geometry, dynamic systems and signal processing (Golyandina *et al.* (2001)). A thorough description of the theoretical and practical foundations of the SSA method, with many examples, can be found in Golyandina *et al.* (2001). For a comparison between SSA and other techniques for forecasting time series, see for example Hassani (2007), Hassani *et al.* (2009, 2010a). It has been shown that the results obtained by the SSA method are more accurate than those obtained by ARIMA and GARCH models (Hassani *et al*., 2009). For the SSA based causality test see (Hassani *et al.,* forthcoming). For a wide variety of applications across different types of economics and financial time series see Hassani and Thomakos (fortcomig). The SSA technique has also been used for filtering financial data and stock market data in Hassani *et al.* (2010b).

It should be noted that compared to mutual information based test, the SSA based test enables us to capture the dependence between two variables in both directions (X to Y and also Y to X) , whilst this is not the case for the mutual information based test. Moreover, compared to Granger causality test that needs to an assessment of linearity and stationarity of the series, the SSA technique does not rely on these assumptions.

To establish some notation, let the variable of interest (the 'generic' variable) be denoted $y$ and let $y_t^{(v)}$ be the $v$-th series ($v = 1, \cdots, m$) of $y$ for the period $t$. Each time series component of the multivariate system is viewed as the sum of unobservable components: the signal comprises components such as the trend, oscillations or periodic movements, and noise. The aim of MSSA is to extract the signal leaving the residual; more generally, the algorithm can also extract groups corresponding to components of the signal. The two stages to the process are decomposition and reconstruction, each of which comprises two steps. Finally, the MSSA algorithm provides forecasts via a linear recurrence formula.

Stage 1: Decomposition: Embedding and Singular Value Decomposition (SVD)

*Step 1*: Embedding
Embedding is a mapping that translates a one-dimensional time series into a multi-dimensional series through the use of subsets of the original series. The key output in this stage is the trajectory matrix, generically referred to as X. This is a matrix that is formed by taking a window of observations of length L and moving this throughout the sample. Such a procedure will be familiar from time series analysis that focuses on calculating moving averages or recursive estimation with a moving window; here moving vectors of observations are created.

To see how this works we take a window comprising the first $L_v$ observations of $Y^{(v)}$, then drop the first observation and add the ($L_v$ + 1)-th observation to create another same length window (vector). This process is continued, with the data organised into a matrix, $X^{(v)}$, of dimension $L_v \times K_v$, where $K_v = T_v - L_v + 1$. The resulting trajectory matrix $X^{(v)}$ is:

$$X^{(v)} = \begin{bmatrix} y_1^{(v)} & y_2^{(v)} & \cdots & y_{T_v-L_v+2}^{(v)} & y_{T_v-L_v+1}^{(v)} \\ y_2^{(v)} & y_3^{(v)} & \cdots & y_{T_v-L_v+1}^{(v)} & y_{T_v-L_v+2}^{(v)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ y_{L_v-1}^{(v)} & y_{L_v}^{(v)} & \cdots & y_{T_v-2}^{(v)} & y_{T_v-1}^{(v)} \\ y_{L_v}^{(v)} & y_{L_v+1}^{(v)} & \cdots & y_{T_v-1}^{(v)} & y_{T_v}^{(v)} \end{bmatrix} \tag{10}$$

In this set-up $T_v$ and $L_v$ are allowed to differ depending on v; thus, in general $X^{(v)}$ is of dimension $L_v \times K_v$ where $K_v = T_v - L_v + 1$. In practice, a common $L_v$ is chosen in the case that the trajectory matrices are stacked horizontally to obtain the system trajectory matrix. For simplicity of exposition, we also assume T = $T_1 = \ldots = T_m$ implying $K_1 = \ldots = K_m$ = K; this assumption is relaxed below.

The trajectory matrix for the system as a whole comprising $Y^{(v)}$, v = 1, …, m, is obtained by stacking the trajectory matrices horizontally to form the trajectory matrix of the multidimensional series. The resulting trajectory matrix X of dimension L×mK, is given by:

$$\text{X} = \left( X^{(1)};\ \ldots;\ X^{(v)};\ \ldots;\ X^{(m)} \right) \tag{11}$$

$$= \left( X_1^{(1)}\ \ldots\ X_K^{(1)};\ \ldots;\ X_1^{(v)}\ \ldots\ X_K^{(v)};\ \ldots;\ X_1^{(m)}\ \ldots\ X_K^{(m)} \right)$$

Notice that each of the m blocks of K columns corresponds to the trajectory matrix for a particular vintage. X is the trajectory matrix for the system of data vintages given by $Y$ = $(Y^{(1)}, \ldots, Y^{(v)}, \ldots, Y^{(m)})$ which, in this simplified case, is a vector of dimension mT×1, where $Y^{(v)} = (y_1^{(v)}, \ldots, y_T^{(v)})'$. The case where $T_v$ and so $K_v$ are not equal is easily accommodated; in that case because the individual trajectory matrices are stacked horizontally, they can be of different column dimensions. Thus, X is of dimension $L \times \sum_{v=1}^{m} K_v$ and Y is of dimension $\sum_{v=1}^{m} T_v \times 1$.

The trajectory matrix is an example of a Hankel matrix in which the diagonal elements are equal for all combinations where the sum of the row (i) and column (j) indices are equal to a constant; that is, $X_{ij} = X_{ji}$ for ij = c. Visually, the diagonals are those on a line from the South-West to the North-East of the matrix, which are referred to as the Hankel diagonals. This is a characteristic that is used in the second stage in which the original series are reconstructed using the principal components obtained in the next step.

*Step 2*: obtain the singular value decomposition, SVD, of the system trajectory matrix

The second step in stage 1 is to construct the SVD of the trajectory matrix X and represent it as a sum of d ≤ L rank-one, mutually orthogonal elementary matrices. First define the matrix C = $XX'$ and denote by $\lambda_1, \ldots, \lambda_d$ the ordered non-negative eigenvalues of C, such that $\lambda_1 \geq \ldots \geq \lambda_d \geq 0$; where d ≤ L, with d = L if all $\lambda_i \geq 0$. The corresponding eigenvectors are $\{U_i\}_{i=1}^{d}$; the factor vectors are $\{V_i\}_{i=1}^{d}$ where $V_i = X'U_i / \sqrt{\lambda_i}$ are of dimension mK× 1. The principle component vectors are $\sqrt{\lambda_i}V_i$ and the eigentriple that forms the basis of the SVD is $(\sqrt{\lambda_i}, U_i, V_i)$.

The trajectory matrix X is decomposed into the sum of d elementary matrices $X_i = \sqrt{\lambda_i}U_i V_i'$, such that:

$$X = \sum_{i=1}^{d} X_i \tag{12}$$

The matrices $X_i$ are referred to as elementary matrices, which have rank 1 and are, by construction, mutually orthogonal. The SVD given by equation (12) is optimal in the sense that among all the matrices of rank r < d, the matrix $X^{(r)} = \sum_{i=1}^{r} X_i$ provides the best approximation to the trajectory matrix X in the norm sense, such that $\| X - X^{(r)} \|$ is a minimum.

The contribution of the component $X_i$ to the expansion (11) is given by its eigenvalue $\lambda_i$ as a share in the sum of the eigenvalues, that is $\lambda_i / \sum_{j=1}^{d} \lambda_j$. The singular spectrum (hence the description singular spectrum analysis) refers to a graph of the ordered eigenvalues, $\lambda_1 \geq \ldots \geq \lambda_d \geq 0$ and is useful in deciding which principal components to include in the reconstruction step of the SSA method. If none of the eigenvalues are negative, then the singular spectrum is a graph of the L ordered eigenvalues.

Stage 2: Reconstruction, Hankelisation and Gouping

*Step 3*: diagonal averaging (block Hankelisation)
In the first stage of this step, the elementary matrices $X_i$, corresponding to the i-th principal component in the SVD, are used to (re)construct series of the same length as the original series. Note that X is in m blocks, one for each vintage of data; thus each block is Hankelised and then the m resulting $T_v \times 1$ vectors are stacked vertically into a $\sum_{v=1}^{m} T_v \times 1$ vector; if $T_v$ = T then the resulting vector is $mT \times 1$. We briefly describe the Hankelisation procedure for one of these blocks.

The Hankelisation procedure may be represented by first rearranging $X_i$ so that the sum of each Hankel diagonal is one element in a T× 1 vector, say $X_{H,i}^{'}$ and then premultiplying by a diagonal matrix H = diag$(h_i)$, with diagonal elements that are the inverse of the number of elements in the corresponding row of $X_{H,i}$. That is:

$$\tilde{X}_i = HX'_{H,i} = \begin{pmatrix} x_{1,1}^{(i)} \\ (x_{2,1}^{(i)} + x_{1,2}^{(i)})/2 \\ (x_{3,1}^{(i)} + x_{2,2}^{(i)} + x_{1,3}^{(i)})/3 \\ \vdots \\ (x_{L,K-2}^{(i)} + x_{L-1,K-1}^{(i)} + x_{L-2,K}^{(i)})/3 \\ (x_{L,K-1}^{(i)} + x_{L-1,K}^{(i)})/2 \\ x_{l,K}^{(i)} \end{pmatrix} \qquad (13)$$

where $x_{j,k}^{(i)}$ is the (i,k)-th element of $X_i$. Note that the sum of the subscripts in each row is the same; thus, the general element sums and then averages the Hankel diagonal elements for each row element. The result is a $T \times 1$ vector of the time series components corresponding to the i-th principal component of the trajectory matrix. The $T \times r_v$ matrix $\tilde{X}$ of all reconstructed components is defined with typical column vector $\tilde{X}_i$, thus $\tilde{X} = [\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_{r_v}]$.

*Step 4: grouping*

Let *r* denote the trajectory dimension resulting from grouping in the multivariate case and let $r_v$ denote the dimension of the univariate trajectory space for the v-th vintage. Stepanov and Golyandina (2005) show that $r_{\min} \leq r \leq r_{\max}$, where $r_{\min} = \max\{r_v : v = 1, \ldots, m\}$ and $r_{\max} = \sum_{v=1}^{m} r_v$. Thus, the ceiling to the multivariate trajectory dimension is simply that obtained when there are, in a sense, no common or matched components amongst the m vintages; the presence of matching components reduces the dimension of the multivariate system and indicates that the system is interrelated, which should result in gains when forecasting the series. The selection of $r < r_{\min}$ leads to a loss of precision as parts of the signals in all series will be lost. From the other side, if $r > r_{\max}$ then noise is included the reconstructed series. The selection of $r \cong r_{\min}$ (keeping $r > r_{\min}$) is a good choice for highly interrelated series sharing several common components. The selection of $r \cong r_{\max}$ is necessary when the series analysed have very little relation to each other.

The result of the grouping step is the reconstructed series, or estimated signal in this case, for all m vintages; that is:

$$\tilde{Y} = \sum_{i \in I} \tilde{X}_i \qquad (14)$$

So that $\tilde{Y}$ is the estimation counterpart of the actual series Y and of dimension $mT \times 1$ or, more generally, $\sum_{v=1}^{m} T_v \times 1$.

## 3. Data and Results

The data set used in our empirical analysis consists of five daily stock price series representing the G7 countries: US, Canada, Japan, UK, Germany, France and Italy. The data are the relative price indexes for these markets, where the base 100 was set at January, 1st 1973. The series were collected in the Datastream database and cover the period from January, 1st 1973 to January, 21st 2009, totalizing 9408 daily observations (five days per week). Figure 1 shows a graphic of the seven series in relative prices (panel a) and in the natural logarithms of relative prices (panel b).

Figure 1a. Relative price indexes for the G7 countries

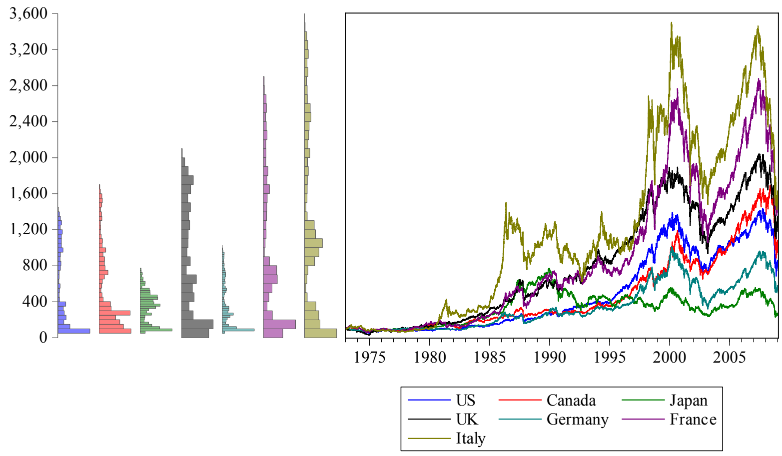
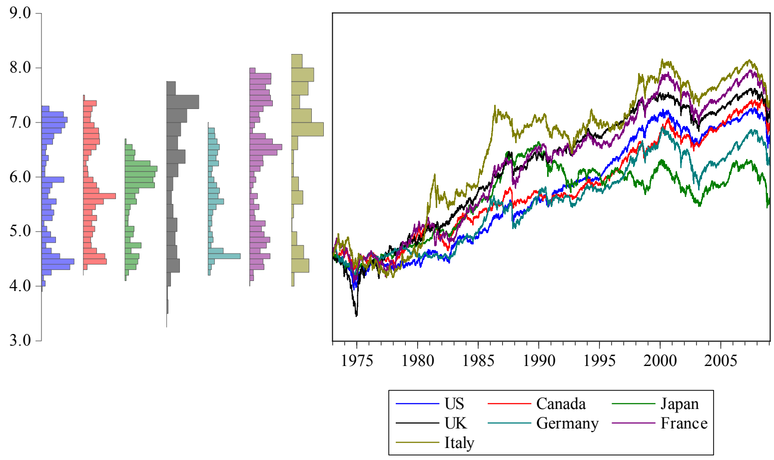



Figure 1b. Natural logarithms of relative price indexes for the G7 countries




Source: Datastream. Base 100: January, 1st 1973. 9408 data points.

A well known problem with daily data across the world is that trading days vary across markets, as they operate in different time zones. Given this and in order to correct the bias we use the procedure proposed by Beine *et al*. (2008). Since the Japanese market is the first to close amongst the stock markets under study we applied one-day lagged causality from the other six markets towards the Japanese market to include some contemporaneous causality. The US markets is the last market to close, so one-day lagged causality from the US stock market towards the other six markets includes the contemporaneous causality among those.
It is remarkable how similar the time-path pattern looks for these seven stock market indexes with market boosts and crises apparently synchronized for all the countries (panel a). Data dispersion increases substantially along time, especially after the oil crisis of the early eighties and, further on, since the end of the 20th century. Price volatility over the period was substantially higher for Italy, France and the UK than for Canada, the US, Germany and Japan. In addition, all price histograms that are shown in Figure 1a exhibit a right-hand side long tail. The series in logs (panel b) lessen volatility in the data, as expected, and the log price histograms appear flattened.

However, data dispersion does still increase over time. Some descriptive statistics of these series (in natural logarithms) are presented in Table 1:

Table 1. Descriptive statistics of the natural logarithms of relative prices

| | US | Canada | Japan | UK | Germany | France | Italy |
|---|---|---|---|---|---|---|---|
| Mean | 5.706510 | 5.779643 | 5.604483 | 6.179921 | 5.545114 | 6.226555 | 6.516560 |
| Median | 5.650874 | 5.662144 | 5.861683 | 6.434844 | 5.584004 | 6.474808 | 6.917948 |
| Maximum | 7.267135 | 7.433217 | 6.645377 | 7.621871 | 6.917379 | 7.964677 | 8.161164 |
| Minimum | 3.932218 | 4.297829 | 4.120337 | 3.446577 | 4.205439 | 4.070223 | 4.153556 |
| Std. Deviation | 1.035847 | 0.895167 | 0.668453 | 1.113543 | 0.789030 | 1.150704 | 1.238936 |
| Skewness | 0.025493 | 0.130726 | -0.634781 | -0.500333 | 0.009600 | -0.228873 | -0.592996 |
| Kurtosis | 1.508327 | 1.842078 | 2.038909 | 1.929123 | 1.653605 | 1.686334 | 1.977588 |
| | | | | | | | |
| Jarque-Bera | 873.2535 | 552.3834 | 993.9092 | 842.0589 | 710.7536 | 758.6182 | 961.1455 |
| *p*-value | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| | | | | | | | |
| Sum | 53686.84 | 54374.88 | 52726.97 | 58140.69 | 52168.43 | 58579.43 | 61307.80 |
| Sum Sq. Dev. | 10093.50 | 7538.058 | 4203.324 | 11664.46 | 5856.504 | 12455.99 | 14439.40 |
| | | | | | | | |
| *N* | 9408 | 9408 | 9408 | 9408 | 9408 | 9408 | 9408 |

Notice that all series are flatter than the Gaussian distribution and slightly skewed, therefore the J-B test statistic rejects the null hypothesis of normality for all of them. This is typical of stock market price series in the same manner as leptokurtosis and fat tails are typically observed in returns data. From this point onwards the analysis will only consider the natural logarithms data, that is, stock prices actually refer to the natural logarithms of the relative price indexes and stock returns or price changes denote the difference between log relative prices at two adjacent dates.

Before proceeding to the analysis of market integration one should look at the (non)stationary nature of the G7 series. Unit root and stationarity tests in levels and in first differences for all the series are shown in Table 2:

Table 2. Unit root and stationarity tests in levels and in first differences

| Variable | ADF[a, c, d] | KPSS[b, c, d] |
|---|---|---|
| US[f] | -1.709328 | 0.960241 ** |
| Canada[e] | -2.806501 | 0.468607 ** |
| Japan[f] | -0.269712 | 2.549435 ** |
| UK[g] | -0.736909 | 2.320246 ** |
| Germany[e] | -1.722877 | 0.568883 ** |
| France[e] | -1.050611 | 1.038102 ** |
| Italy[g] | -0.500341 | 1.661498 ** |
| | | |
| ΔUS | -70.39091 ** | 0.244395 |
| ΔCanada | -88.91458 ** | 0.075838 |
| ΔJapan | -69.26301 ** | 0.126957 |
| ΔUK | -45.20940 ** | 0.220531 |
| ΔGermany | -92.36380 ** | 0.130575 |
| ΔFrance | -89.32861 ** | 0.186063 |
| ΔItaly | -44.66293 ** | 0.302849 |

Notes: [a] MacKinnon (1996) critical values: -3.43 (1%) and -2.86 (5%) for constant and -3.96 (1%) and -3.41 (5%) for constant and linear trend. [b] Kwiatkowski-Phillips-Schmidt-Shin (1992) critical values: 0.739 (1%) and 0.463 (5%) for constant and 0.216 (1%) and 0.146 (5%) for constant and linear trend. [c] exogenous terms in levels: constant and linear trend. [d] exogenous terms in 1st differences: constant (except for Japan in the KPSS test which is constant and linear trend). [e] 1 *lag* in levels for ADF. [f] 2 *lags* in levels for ADF. [g] 4 *lags* in levels for ADF. ** significant at 1%.

The ADF and KPSS tests are designed to capture weak stationarity with opposite null hypotheses.

In the former case the null hypothesis of nonstationarity of the variables in levels is not rejected but it is rejected at 1% for the variables in first differences. In the latter case the null hypothesis of stationarity in levels is rejected at 1% but it is not rejected in first differences. The results are, therefore, consistent in both cases and lead to the conclusion that the price series under analysis are, in fact, integrated of first order. The number of lags selected in each test was set on the basis of the BIC information criterion.

To examine the statistical existence of the structural breaks, we performed CUSUM and CUSUM-Q tests. We tested the stability of our time series by regressing it on a nonsignificant constant. The results indicate the presence of structural breaks for all the variables and some of those structural breaks seem to be related with the existence of stock market crashes and financial crisis, which reinforces the possibility of nonlinear behavior among the stock markets under study.

Given that all variables are I(1), we considered the possibility of estimating a long-run relationship between all these variables. To test for cointegration between all those series, we used the Phillips tests suggested by Gregory and Hansen (1996) because the power of the Johansen's test may be reduced substantially when the series exhibits structural breaks. The Philips test statistics are presented in Table 3.

Table 3. Phiplips cointegration tests

| | Statistics | Breakpoint |
|---|---|---|
| ADF* | | |
| C | -6.066** | (0.3911) |
| C/T | -6.282** | (0.3911) |
| C/S | -7.020** | (0.3472) |
| Z* | | |
| C | -6.155** | (0.3912) |
| C/T | -6.379** | (0.3912) |
| C/S | -7.061** | (0.3473) |
| Za* | | |
| C | -77.154** | (0.3912) |
| C/T | -82.315** | (0.3912) |
| C/S | -96.400** | (0.3091) |

Notes: [a] Critical values can be found in Gregory and Hansen (1996). ** significant at 1%.

### 3.1. Granger causality tests

The results indicate that there is at least one cointegrating vector since the null hypothesis of no cointegration is rejected at 1%. This means that the seven stock markets under analysis possibly belong to the same market space and there is a long-run equilibrium relationship linking price data along with the dynamic short-run terms denoting market returns. Altogether, these results outline the starting point for analyzing market integration on the basis of Granger causality. The Granger causality *F*-statistics are presented in Tables 4 to 6.

Table 4. Granger Causality for the log prices

| Variable | US | Canada | Japan | UK | Germany | France | Italy |
|---|---|---|---|---|---|---|---|
| US | - | 142.898 ** | 716.963 ** | 475.981 ** | 390.273 ** | 470.843 ** | 156.006 ** |
| Canada | 36.0065 ** | - | 361.477 ** | 113.420 ** | 73.5093 ** | 117.433 ** | 46.4521 ** |
| Japan | 14.3828 ** | 4.86702 ** | - | 27.3334 ** | 21.9847 ** | 24.1686 ** | 7.23094 ** |
| UK | 8.91317 ** | 3.99597 * | 233.251 ** | - | 7.33226 ** | 6.02136 ** | 3.54475 * |
| Germany | 6.03121 ** | 2.29723 | 284.715 ** | 1.13299 | - | 2.17821 | 0.32732 |
| France | 9.72877 ** | 1.51540 | 259.915 ** | 6.42979 ** | 1.54816 | - | 1.21910 |

| Italy | 1.93860 | 1.01208 | 107.911 ** | 0.61978 | 1.42051 | 7.52491 ** | - |
|---|---|---|---|---|---|---|---|

Notes: $H_0$: $X_{it}$ does not Granger cause $X_{jt}$ ($i \neq j$). 2 *lags*. 9406 observations in each series. ** significant at 1%. * significant at 5%.

Table 4 presents the Granger causality tests for the variables in levels, that is, stock prices. Recall that the test is interpreted as follows: $X_{2t}$ Granger causes $X_{1t}$ if, *ceteris paribus*, the past values of $X_{2t}$ help to improve the current forecast of $X_{1t}$, where $X_{2t}$ represents the variables in the first column and $X_{1t}$ represents the variables in the first row. One can say, therefore, that for the significant causal relationships the historical prices of the former market affect the current price of the latter, forming a dynamical long-run relationship in the global economy. As we can see, about 74% of the coefficients are statistically significant, which means that there is substantial long-run causal effects among these markets, of which many of them are feedback relationships. However, we found no causal relationship in any direction for the pairs Germany-France and Germany- Italy.

Another important result is that, in the long-run, the US causes more than is caused by other markets. To see this, note that the $F$-statistics of the former (1st row) are substantially larger than the $F$-statistics of the latter (1st column). This is consistent with the idea that the US stock market, to a greater extent, 'exports' more than 'imports' boosts and crises, being therefore the engine of the global financial world. For example, a crisis with origin in the US can spread in a broader way to other markets (as it seems in the current crisis) than a crisis with origin in Japan or even any European country. Canada shows an overall picture very similar to the US, that is, in general it causes more other markets than is caused by them, except in what refers to the US. Canada, however, appears to be caused only by the US, Japan and, to a lesser extent, the UK. Conversely, Japan is the most endogenous of the G7 markets. The European countries do not show an overall systematic pattern of causality, though the UK appears to emerge like an attractor in the EU context (but not with France) and follows the North-American markets. This is surprising insofar we would expect Germany to be the leading European stock market, given its role as the head of the European Union economy, albeit one should recognize the very important role of the London Stock Exchange in the global financial world.

Table 5 presents the Granger causality tests for the variables in first differences, that is, returns. The results show how much historical returns of one market affect the current returns of another market, making up therefore a dynamical short-run relationship in the system. Here, some 71% of the coefficients are statistically significant and we found no causal relationship in any direction only for the pair Germany-Italy (as in the long-run tests). Otherwise, the overall picture is the same as for the results in levels.

In the short-run, the North-American markets cause more other markets than are caused by them and the US leads the Canadian market. The opposite occurs for Japan as in the long-run. Again, the UK emerges as an attractor in the European Union context (except with France) but follows the North-American markets. It seems, therefore, that market causality among the G7 countries is present both in the long-run and in the short-run, affecting co-movement prices and returns.

Table 5. Granger Causality for returns

| Variable | ΔUS | ΔCanada | ΔJapan | ΔUK | ΔGermany | ΔFrance | ΔItaly |
|---|---|---|---|---|---|---|---|
| ΔUS | - | 138.970 ** | 708.196 ** | 491.025 ** | 387.364 ** | 476.872 ** | 154.376 ** |
| ΔCanada | 30.8130 ** | - | 369.375 ** | 123.680 ** | 74.2408 ** | 128.220 ** | 47.4845 ** |
| ΔJapan | 7.60523 ** | 3.61023 * | - | 25.5639 ** | 17.2502 ** | 19.6704 ** | 5.39895 ** |
| ΔUK | 2.27861 | 3.71751 * | 238.670 ** | - | 3.34280 * | 0.07240 | 3.77099 * |
| ΔGermany | 3.37056 * | 2.36538 | 277.860 ** | 0.74040 | - | 3.50272 * | 0.25046 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| ΔFrance | 4.97310 ** | 0.80771 | 263.282 ** | 6.15885 ** | 0.91485 | - | 1.14968 |
| ΔItaly | 2.13516 | 2.54945 | 108.852 ** | 1.06984 | 2.05813 | 7.76945 ** | - |

Notes: $H_0$: $X_{it}$ does not Granger cause $X_{jt}$ ($i \neq j$). 2 *lags*. 9405 observations in each series. ** significant at 1%. * significant at 5%.

Finally, Table 6 presents the Granger causality results for the variables in first differences but where $X_{2t}$ now represents the first lag of the underlying variable. The results can be interpreted in terms of a delayed effect of returns of one market onto the current returns of another market. It should be noted the size of the $F$-statistics in this Table, where all the coefficients are significant at much less than 1%. The overall picture is, however, the same as before. Historical delayed returns worldwide have a significant impact on current returns for all the cases. In our context, historical delayed returns were only computed for one lag while one can believe that smoother but significant effects may also occur for two or more lags, though one lag computations will suffice for our purposes.

Table 6. Granger Causality for returns (lagged effects)

| Variable | ΔUS | ΔCanada | ΔJapan | ΔUK | ΔGermany | ΔFrance | ΔItaly |
|---|---|---|---|---|---|---|---|
| ΔUS($t$–1) | - | 4403.8 ** | 770.35 ** | 1237.9 ** | 1196.1 ** | 1181.1 ** | 420.91 ** |
| ΔCanada($t$–1) | 4198.8 ** | - | 525.63 ** | 961.60 ** | 761.95 ** | 851.67 ** | 314.22 ** |
| ΔJapan($t$–1) | 63.808 ** | 160.49 ** | - | 327.92 ** | 321.30 ** | 307.39 ** | 154.97 ** |
| ΔUK($t$–1) | 690.58 ** | 830.07 ** | 542.91 ** | - | 1482.9 ** | 2122.2 ** | 755.22 ** |
| ΔGermany($t$–1) | 751.19 ** | 678.05 ** | 599.03 ** | 1480.6 ** | - | 2593.0 ** | 1006.0 ** |
| ΔFrance($t$–1) | 652.79 ** | 715.79 ** | 555.44 ** | 2131.1 ** | 2589.7 ** | - | 1048.5 ** |
| ΔItaly($t$–1) | 260.01 ** | 267.33 ** | 260.49 ** | 754.62 ** | 1001.9 ** | 1054.0 ** | - |

Notes: $H_0$: $X_{it}$ does not Granger cause $X_{jt}$ ($i \neq j$). 9404 observations in each series. ** significant at 1%.

Globally, the Granger causality results point to the existence of a single global stock market leaded by the US. The UK emerges as a regional leader within the European context. Japan, however, does not emerge as a leading market within the G7 countries but this is probably due to the long-lasting economic crisis that Japan has been facing. The great surprise (or perhaps not) is the dominant position of Canada relative to many other G7 countries. Canada may benefit from its proximity to the US where, surely, intense economic relationships, some similar economic policies and firm's relationships turn up North-America as a unified financial block. The results are, overall, compatible with the definition of weak market integration introduced in this paper although do not capture nonlinearities in the data. At this point one can conclude that linear weak market integration occurs within the G7 over the period analyzed.

### 3.2. Mutual Information tests and global correlation coefficient

In order to deal with the problem of nonlinearities in the data and other complexities of stock market co-movements, we present in Table 7 the global correlation coefficient for the variables in levels, as described in Section 2. Recall that the global correlation coefficient (equation 8) is a standardized measure of mutual information taking on values between 0 and 1. The closer $\lambda$ is to 1 the more information $X_1$ and $X_2$ share, *i.e.*, knowing one of them reduces uncertainty about the other.

Table 7. Global Correlation Coefficient for log prices

| Variable | US | Canada | Japan | UK | Germany | France | Italy |
|---|---|---|---|---|---|---|---|
| US | | 0.9978 ** | 0.9907 ** | 0.9981 ** | 0.9956 ** | 0.9967 ** | 0.9948 ** |
| Canada | | | 0.9899 ** | 0.9977 ** | 0.9948 ** | 0.9972 ** | 0.9940 ** |
| Japan | | | | 0.9935 ** | 0.9870 ** | 0.9916 ** | 0.9766 ** |
| UK | | | | | 0.9971 ** | 0.9981 ** | 0.9960 ** |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Germany | | | | | | 0.9961 ** | 0.9927 ** |
| France | | | | | | | 0.9950 ** |

Notes: $H_0$: $I(X_1,X_2) = 0$. Critical values for $N(0,1)$ I data, $n \geq 2500$ (see the Appendix). ** $I(X_1,X_2)$ significant at 1%.

The major finding is that all coefficients are statistically significant, which means that substantial interaction, both linear and nonlinear, exists among the seven markets under analysis. There are very strong long-run relationships between all markets as can be seen from Table 7, where all the coefficients are higher than 0.97. However, unlike Granger causality, the global correlation tests do not provide any information about the direction of causality between the underlying variables. For the moment and until a formal test of causality based on the global correlation coefficient is available, we shall assess the direction of causality only on the basis of Granger causality tests.

Table 8. Global Correlation Coefficient for returns

| Variable | ΔUS | ΔCanada | ΔJapan | ΔUK | ΔGermany | ΔFrance | ΔItaly |
|---|---|---|---|---|---|---|---|
| ΔUS | | 0.7133 ** | 0.1448 ** | 0.4650 ** | 0.4361 ** | 0.4254 ** | 0.3464 ** |
| ΔCanada | | | 0.2135 ** | 0.4168 ** | 0.4422 ** | 0.4386 ** | 0.3531 ** |
| ΔJapan | | | | 0.2486 ** | 0.3052 ** | 0.2858 ** | 0.1966 ** |
| ΔUK | | | | | 0.5817 ** | 0.6370 ** | 0.4872 ** |
| ΔGermany | | | | | | 0.6914 ** | 0.5717 ** |
| ΔFrance | | | | | | | 0.5836 ** |

Notes: $H_0$: $I(X_1,X_2) = 0$. Critical values for $N(0,1)$ I data, $n \geq 2500$ (see the Appendix). ** $I(X_1,X_2)$ significant at 1%.

Table 8 shows the global correlation coefficient for the variables in first differences, *i.e.*, returns. It can be noted that, for returns, there are relatively strong relationships within the European and North-American blocks, where all the coefficients are higher than 0.5 except for ΔUK-ΔItaly. The global correlation between the European returns and the North-American returns is slightly lower (0.3−0.5) and the global correlation between Japan and the rest of the G7 countries is even lower (0.1−0.3 approximately) although in general higher with the European countries than with the North-American countries. Overall, when compared with the global correlation coefficient of the price variables, the figures in Table 8 are smaller but still statistically significant at 1%.

Table 9. Global Correlation Coefficient for returns (lagged effects)

| Variable | ΔUS | ΔCanada | ΔJapan | ΔUK | ΔGermany | ΔFrance | ΔItaly |
|---|---|---|---|---|---|---|---|
| ΔUS($t$−1) | - | 0.2452 ** | 0.3476 ** | 0.3469 ** | 0.3363 ** | 0.3585 ** | 0.2066 ** |
| ΔCanada($t$−1) | 0.1007 ** | | 0.2799 ** | 0.2436 ** | 0.2217 ** | 0.2378 ** | 0.1559 ** |
| ΔJapan($t$−1) | 0.1608 ** | 0.0847 ** | | 0.1017 ** | 0.2014 ** | 0.1092 ** | 0.0469 ** |
| ΔUK($t$−1) | 0.1691 ** | 0.1196 ** | 0.2429 ** | | 0.1488 ** | 0.1313 ** | 0.0977 ** |
| ΔGermany($t$−1) | 0.1916 ** | 0.2346 ** | 0.3126 ** | 0.2965 ** | | 0.2467 ** | 0.2346 ** |
| ΔFrance($t$−1) | 0.1421 ** | 0.0760 ** | 0.2535 ** | 0.2182 ** | 0.2116 ** | | 0.1834 ** |
| ΔItaly($t$−1) | 0.0881 ** | 0.0893 ** | 0.1656 ** | 0.2301 ** | 0.1775 ** | 0.2019 ** | - |

Notes: $H_0$: $I(X_1,X_2) = 0$. Critical values for $N(0,1)$ I data, $n \geq 2500$ (see the Appendix). ** $I(X_1,X_2)$ significant at 1%.

Finally, Table 9 presents the global correlation coefficient for lagged returns. While significant at 1%, all these coefficients are even smaller than those for non-lagged returns. The US historical returns present a higher global correlation with the current returns of all other markets than the historical returns of all other markets onto the current returns of the US. A similar picture occurs for Canada except in the relationship with the US and, to some extent, Germany, and conversely for Japan. The leadership within the European Union countries is now held by Germany. In fact, the global correlation of the German historical (or lagged) returns with the current returns of all other European countries is higher than the global correlation of the historical returns of all other European countries with the current German returns, and likewise regarding Canada and Japan. Germany, however, follows the US market. It seems, therefore, that the global correlation results are more in accordance with what would be expected from the general economic strength of each G7 country.

### 3.2. SSA technique

According to the results presented in Table 10 for the SSA based causality test, almost 83% of the coefficients are less than one (of which 72% are statistically significant), which means that there is a long-run causal effects among these markets. These results also indicate that there are many feedback relationships (see, for example, UK and US). Moreover, the results in the first row indicate that, in the long-run, the US causes more than is caused by other markets confirming the results obtained by Granger causality. There is also a feedback relationship between US and other countries except for Italy that the coefficient is close to one, and is not significant.

Table 10. SSA based causality test for log prices

| Variable | US | | Canada | | Japan | | UK | | Germany | | France | | Italy | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| US | | | 0.82 | ** | 0.79 | ** | 0.83 | ** | 0.89 | ** | 0.82 | ** | 0.82 | ** |
| Canada | 0.92 | | | | 0.95 | ** | 0.97 | | 0.84 | ** | 0.96 | | 1.02 | |
| Japan | 0.83 | ** | 0.85 | ** | | | 0.90 | ** | 0.91 | ** | 0.99 | | 0.91 | ** |
| UK | 0.87 | ** | 0.89 | ** | 0.84 | ** | | | 0.91 | ** | 0.78 | ** | 0.92 | ** |
| Germany | 0.85 | ** | 0.99 | | 1.02 | | 0.84 | ** | | | 0.85 | ** | 0.92 | ** |
| France | 0.95 | | 0.91 | ** | 1.05 | | 0.82 | ** | 0.84 | | | | 0.81 | ** |
| Italy | 0.99 | | 1.03 | | 0.99 | | 0.98 | | 0.97 | | 0.94 | ** | | |

Table 11 shows the SSA based causality test for the variables in first differences, that is, returns. The overall conclusion for the results in this table is that the dependence structure between markets is reduced using differencing as we saw in previous tests. Here we see some casual relationship as well, but the magnitude has been reduced.

Table 11. SSA based causality test for returns

| Variable | ΔUS | | ΔCanada | | ΔJapan | | ΔUK | | ΔGermany | | ΔFrance | | ΔItaly | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ΔUS | | | 0.87 | ** | 0.89 | ** | 0.84 | ** | 0.92 | ** | 0.89 | ** | 0.92 | ** |
| ΔCanada | 0.95 | ** | | | 0.98 | | 0.96 | | 0.92 | ** | 1.02 | | 0.98 | |
| ΔJapan | 0.94 | ** | 0.94 | ** | | | 1.10 | | 1.08 | | 1.02 | | 0.98 | |
| ΔUK | 0.95 | | 0.88 | ** | 0.97 | | | | 0.93 | ** | 0.89 | ** | 0.90 | ** |
| ΔGermany | 0.87 | ** | 1.08 | | 1.07 | | 0.89 | ** | | | 0.87 | ** | 0.91 | ** |
| ΔFrance | 0.89 | ** | 0.86 | ** | 1.11 | | 0.93 | ** | 0.91 | ** | | | 0.90 | ** |
| ΔItaly | 1.13 | | 1.01 | | 1.03 | | 1.12 | | 0.97 | | 1.02 | | | |

Finally, Table 12 presents the SSA based causality test for the variables in first differences but where $X_{2t}$ now represents the first lag of the underlying variable. The overall conclusion is the same as the results obtained by previous methods; historical delayed returns worldwide have a significant impact on current returns. However, here we only consider one lag.

Table 12. SSA based causality test for returns (lagged effects)

| Variable | ΔUS | | ΔCanada | | ΔJapan | | ΔUK | | ΔGermany | | ΔFrance | | ΔItaly | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ΔUS($t$–1) | | | 0.85 | ** | 0.84 | ** | 0.78 | ** | 0.95 | | 0.86 | | 0.91 | ** |
| ΔCanada($t$–1) | 0.94 | ** | | | 0.93 | ** | 0.92 | ** | 0.95 | | 0.98 | | 0.94 | ** |
| ΔJapan($t$–1) | 0.93 | ** | 0.92 | ** | | | 1.02 | | 1.01 | | 0.97 | | 0.96 | |
| ΔUK($t$–1) | 0.85 | ** | 0.84 | ** | 0.94 | ** | | | 0.77 | ** | 0.80 | ** | 0.85 | ** |
| ΔGermany($t$–1) | 0.83 | ** | 0.95 | ** | 0.99 | | 0.79 | ** | | | 0.86 | ** | 0.93 | ** |
| ΔFrance($t$–1) | 0.91 | ** | 0.94 | ** | 1.05 | | 0.88 | ** | 0.87 | ** | | | 0.91 | ** |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| ΔItaly($t$–1) | 1.02 | 0.98 | 0.98 | 1.10 | 0.96 | 1.03 | |

In general, the results obtained by all methodologies are consistent with the idea of weak market integration when dealing with stock exchange data and there is a single global market leaded by the US. In Europe, the UK (or Germany depending on the methodology) emerges as a regional leader. Japan however, appears to be the weakest market, in causal terms, within the G7. The mutual information and SSA based test results appear to be stronger than Granger causality where capture global relationships between stock markets without imposing any structure or constraint to the model.

Interestingly, the global correlation among stock markets is very high in levels and much smaller in first differences and in lagged first differences. Stock market prices, consequently, are more correlated than stock market returns and the latter are more correlated than lagged returns. On the basis of these results is it possible to conclude that the long-run relationship between prices incorporates more nonlinear dependencies and, probably, other complexities in the data than the short-run relationships between log price changes? It appears so. Conversely, with regard to the Granger causality results, the statistical tests for returns appear to be stronger than those for prices in levels, and even more so for lagged stock market returns. One should note that the mutual information and the SSA technique enable us to capture, in a quite global way, the total dependence and relationships between variables, being almost free from restrictions and assumptions about structure, probability distributions or errors behavior. In this context, we believe these two techniques promote more robust information than linear Granger causality tests in time series statistical relationships.

## 4. Conclusions

This paper analyzes stock market integration in the context of the global economy for the G7 countries. The theoretical background is rooted on a *new* concept of weak market integration which is defined as the causality that occurs in price transmission independently of whether this process is proportional or not over time. This allows for nonlinearities and other types of price distortions to be present in the overall process. Under proportionality of price transmission we say that strong market integration occurs. The empirical modeling of market integration based on price data is complicated by the nonstationary nature of these data sets. In order to acknowledge the nonstationarity problem, tests for unit roots and cointegration were performed prior to the empirical analysis of market integration based on Granger causality, mutual information and the SSA based tests. The unit root results are consistent with nonsationarity, and cointegration is present for the G7 stock markets over the 36-year period under analysis. It is therefore consistent to say that these markets belong to the same space, *i.e.*, they actually form a single global stock market with one long-run or equilibrium relationship linking the data.

The cointegration results obtained assure that we are not facing spurious relationships between the seven markets under analysis. Thus, market integration can be tested using the methods discussed in this paper. The results in both cases are consistent with the notion of pairwise weak market integration, since there are substantial causal effects, possibly linear and/or nonlinear, between pairs of variables. These effects occur both for prices and returns. They are also present for lagged returns relationships. Overall, the SSA based causality test results and the mutual information results appear to be more robust than the Granger causality ones. Since these technique capture global dependence relationships without imposing any assumption about the structure of the model, while the Granger causality test is deeply rooted on parametric linear regression estimation, the SSA based test and the mutual information results can be seen as a 'general case' of Granger causality, that is, the former results incorporate the latter ones. In this sense, one can say that these techniques provide more information on the process of market integration than linear Granger causality tests. It should be noted that the SSA based test enabled us to capture the causality in both directions, whilst we can not capture this using mutual information based test. Future work will look into the possibility of using a modified mutual information test that allows for separate causality effects between $X_1$ and $X_2$, and $X_2$ and $X_1$, in order to ascertaining whether causality is just one-way or, otherwise, is a feedback effect. This is important insofar endogeneity and exogeneity are central concepts for modeling market relations and the process of price/returns co-movements. In addition, we shall also look into the nature of the nonlinear relationships between stock markets, in particular with respect to the distinction between stochastic and deterministic effects and provide a robust basis to make prediction in the context of market integration.

## Acknowledgments

The authors thank the financial support provided by FCT–Fundação para a Ciência e Tecnologia, under the grants PTDC/GES/73418/2006 and FCOMP-01-0124-FEDER-007350 and program FEDER/POCI 2010.

## References

Alexander, C. (2008). *Practical Financial Econometrics – Market Risk Analysis II*, John Wiley and Sons, USA.

Ammermann, P. and Patterson, D. (2003). The cross-sectional and cross-temporal universality of nonlinear serial dependencies: Evidence from world stock indices and the Taiwan stock exchange. *Pacific-Basin Finance Journal, 11* (2), 175-195.

Arshanapalli, B. and Doukas, J. (1994). Common Stochastic Trends in a System of Eurocurrency Rates, *Journal of Banking and Finance, 18*, 1047 -1061.

Beine, M.; Capelle-Blancard, G. and Raymond H. (2008). International nonlinear causality between stock markets, *The European Journal of Finance, 14* (8), 663-686.

Bonfiglioli, A. and Favero, C. (2005). Explaining co-movements between stock-markets: The case of US and Germany, *Journal of International Money and Finance, 24*, 1299-1316.

Chung, P.J. and D.J. Liu (1994), Common Stochastic Trends in Pacific Rim Stock Markets, *The Quarterly Review of Economics and Finance , 34* (3): 241–259.

Corazza, M; Malliaris, A.G. and Scalco, E. (2010). Nonlinear Bivariate Comovements of Asset Prices: Methodology, Tests and Applications. *Computational Economics, 35*, 1-23.

Darbellay, G. e D. Wuertz, 2000b. The Entropy as a Tool for Analysing Statistical Dependence's in Financial Time Series, Physica A, 287, 429-439.

Dionísio, A., R. Menezes e D.A. Mendes, 2004b. Mutual Information: a Measure of Dependency for Nonlinear Time Series, Physica A, 344, 326-329.

Dionísio, A., Menezes, R., Mendes, D. (2006) "Entropy-Based Independence Test", *Nonlinear Dynamics,* vol 44, 351-357.

Drożdż, S.; Grümmer, F.; Ruf. F. and Speth J. (2001). Towards identifying the world stock market cross-correlations: DAX versus Dow Jones Physica A, 294, 226-234

Engle, R.F. and Granger, C. (1987). Cointegration and error-correction: Representation, estimation and testing, *Econometrica, 55*, 251-276.

Eun, C.S. and Shim, S. (1989). International transmission of stock market movements, *Journal of Financial and. Quantitative. Analysis,. 24* , 241–256

Golyandina, N., Nekrutkin, V., and Zhigljavsky, A. (2001). Analysis of Time Series Structure: SSA and related techniques, Chapman & Hall/CRC, New York – London.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica, 37, (3*): 424–438

Granger, C. e J. Lin, 1994. Using the Mutual Information Coefficient to Identify Lags in Nonlinear Models, Journal of Time Series Analysis, 15, 4, 371-384.

Granger, C., E. Maasoumi e J. Racine, 2004. A Dependence Metric for Possibly Nonlinear Processes, Journal of Time Series Analysis, 25, 5, 649-669.

Gregory, A. and Hansen, B. (1996). Residual-based tests for cointegration in models with regime shifts, *Journal of Econometrics, 70,* 99-126.

Hamao et al., 1990. Y. Hamao, R.W. Masulis and V. Ng , Correlations in price changes and volatility across international stock markets. *Review of Financial Studies* **3** (1990), pp. 281–307.

Hassani, H. (2007). Singular Spectrum Analysis: Methodology and Comparison. *Journal of Data Science. 5,* 239-257.

Hassani, H; Heravi, H; and Zhigljavsky, A. (2009). Forecasting European Industrial Production with Singular Spectrum Analysis, *InternationalJjournal of Forecasting*, 25, 103-118.

Hassani, H; Thomakos, D. (2010). A Review on Singular Spectrum Analysis for Economic and Financial Time Series, *Statistics and Its Interface*, Forthcoming

Hassani, H; Soofi, A; and Zhigljavsky, A. (2010a). Predicting Daily Exchange Rate with Singular Spectrum Analysis., *Nonlinear Analysis: Real World Applications, 11*, 2023-2034.

Hassani, H.; Dionisio, A., Ghodsi, M. (2010b), "The effect of noise reduction in measuring the linear and nonlinear dependency of financial markets", *Nonlinear Analysis: Real World Applications*, vol. 11, 492-502.

Hsieh, D. (1991). Chaos and nonlinear dynamics: Application to financial markets. *Journal of Finance*, 46, 1839-77.

Kasa*, K*. (1992). Common Stochastic Trends in International Stock Markets, *Journal* of Monetary *Economics*, *29*, *95*-124

Kim, S.J.; Moshirian, F. and Wu, E. (2005). Dynamic stock market integration driven by the European Monetary Union: An empirical Analysis, *Journal of Banking and Finance, 29*, 2475-2502.

Kraskov, A., H. Stogbauer e P. Grassberger, 2003. Estimating Mutual Information, preprint http://www.arxiv:cond-mat/0305641.

Kwiatkowski, D., Phillips, P.C.B.; Schmidt, P.and Shin, Y. (1992) Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root*?, Journal of Econometrics, 54*, 159-178.

MacKinnon, J.G. (1996). Numerical Distribution Functions for Unit Root and Cointegration Tests, *Journal of Applied Econometrics, 11*: 601-618.

Masih, A. M. M. and Masih, R. (1997): “Dynamic Linkages and the Propagation Mechanism Driving Major International Stock Markets: An Analysis of the Pre-and Postcrash Eras”, *Quarterly Review of Economics and Finance* 37: 859-885

Masih, A.M.M and Masih, R. (2002), “Propagative causal price transmission among international stock markets: evidence from the pre- and post globalization period”, Global Finance Journal, Vol. 13, pp. 63-91.

Savva, C. (2009). International stock markets interactions and conditional correlations, *Journal of International Financial Markets, Institutions and Money, 19*, 645-666.

Sargan, J.D. (1964) in *Econometric Analysis for National Economic Planning*, edited by Hart, P.E., G. Mills and J.K. Whitaker (Colston Papers, V. 16).

Stepanov, D., Golyandina, N. (2005) Variants of the “Caterpillar”-SSA method for forecasting multidimensional time series. In: *Proceedings of the IV International conference “System indentification and control problems*” SICPRO’05, Moscow, p. 1831-1848 (in Russian).

Stigler, G.J. (1969). *The Theory of Price*, Macmillan, London.

Sutton, J. (1991). *Sunk Costs and Market Structure*, MIT Press, Cambridge.

Tavares, José, 2009. "Economic integration and the comovement of stock returns," Economics Letters, , 103(2), 65-67.

Zhou, W. and Sornette, D. (2003). Evidence of a world-wide stock market log-periodic anti-bubble since mid-2000, Physica *A 330*, *543*-83

## Appendix

This Appendix lists the 90th, 95th and 99th percentiles of the empirical distribution of the mutual information for the process $Y_t = \varepsilon_t$ with $\varepsilon_t \sim N(0,1)$, on the basis of 5000 simulations for each critical value. This methodology was used as proposed by Granger, Maasoumi and Racine (2004). According to these authors, the critical values can be used as the base to test for time series serial independence. DF denotes the degrees of freedom for the mutual information, which corresponds to the dimension ($d$) of the analyzed vectors.

$N = 100$

| DF | Percentiles | | |
|---|---|---|---|
| | 90 | 95 | 99 |
| 2 | 0.0185 | 0.0323 | 0.0679 |
| 3 | 0.1029 | 0.1232 | 0.1933 |
| 4 | 0.1059 | 0.1260 | 0.1722 |
| 5 | 0.2290 | 0.2580 | 0.3261 |
| 6 | 0.6639 | 0.7528 | 0.9663 |
| 7 | 0.8996 | 0.9731 | 1.1586 |
| 8 | 1.3384 | 1.3839 | 1.5024 |
| 9 | 1.9030 | 1.9352 | 2.0142 |
| 10 | 2.5266 | 2.5571 | 2.6181 |

$N = 1000$

| DF | Percentiles | | |
|---|---|---|---|
| | 90 | 95 | 99 |
| 2 | 0.0019 | 0.0041 | 0.0071 |
| 3 | 0.0133 | 0.0191 | 0.0311 |
| 4 | 0.0340 | 0.0399 | 0.0568 |
| 5 | 0.0708 | 0.0865 | 0.1128 |
| 6 | 0.2119 | 0.2430 | 0.3046 |

| | | | |
|---|---|---|---|
| 7 | 0.3635 | 0.3954 | 0.4688 |
| 8 | 0.4041 | 0.4414 | 0.5252 |
| 9 | 0.3865 | 04114 | 0.4640 |
| 10 | 0.6418 | 0.6585 | 0.6942 |

$N \geq 2500$

| DF | Percentiles | | |
|---|---|---|---|
| | 90 | 95 | 99 |
| 2 | 0.0008 | 0.0015 | 0.0030 |
| 3 | 0.0054 | 0.0078 | 0.0129 |
| 4 | 0.0134 | 0.0171 | 0.0251 |
| 5 | 0.0556 | 0.0648 | 0.0797 |
| 6 | 0.1203 | 0.1376 | 0.1738 |
| 7 | 0.2181 | 0.2418 | 0.2884 |
| 8 | 0.3938 | 0.4217 | 0.4719 |
| 9 | 0.3175 | 0.3409 | 0.4024 |
| 10 | 0.2931 | 0.3124 | 0.3477 |