ELSEVIER

Physica A 301 (2001) 397–406

PHYSICA A

www.elsevier.com/locate/physa

Measures of globalization based on cross-correlations of world financial indices

Sergei Maslov*

Department of Physics, Brookhaven National Laboratory, Upton, NY 11973, USA

Received 9 July 2001

Abstract

The cross-correlation matrix of daily returns of stock market indices in a diverse set of 37 countries worldwide was analyzed. Comparison of the spectrum of this matrix with predictions of random matrix theory provides an empirical evidence of strong interactions between individual economies, as manifested by three largest eigenvalues and the corresponding set of stable, non-random eigenvectors. The observed correlation structure is robust with respect to changes in the time horizon of returns ranging from 1 to 10 trading days, and to replacing individual returns with just their signs. This last observation confirms that it is correlations between signs and not absolute values of fluctuations, which are mostly responsible for the observed effect. Negative changes in the index are somewhat more correlated than the positive ones. Also, in our data set the reaction of Asian stock indices to changes in European and American ones persists for about 3 days. © 2001 Elsevier Science B.V. All rights reserved.

PACS: 89.65.Gh; 89.75.Fb; 05.40.Ca; 89.70.+c

In spite of the tremendous importance that current public opinion places on issues of globalization of the world's economy, its sources and consequences remain poorly understood. Large downturns and collapses of the economic and financial situation in one country are routinely blamed on recent events in other countries. This point of view is reinforced by sensational newspaper headlines like "Latin American markets catch the asian flu". The level of globalization of a diverse set of 50 developed countries and key emerging markets worldwide was recently measured and reported in the

* Tel.: +1-631-344-3742; fax: +1-631-344-2918.
E-mail address: maslov@bnl.gov (S. Maslov).

0378-4371/01/$ - see front matter © 2001 Elsevier Science B.V. All rights reserved.
PII: S0378-4371(01)00370-3

A.T. Kearney/Foreign Policy Magazine Globalization Index™ [1]. The factors selected to contribute to this index are extremely diverse and include, for example, volumes of inward- and outward-directed foreign investments, the amount of international travel and phone calls, number of servers of the World Wide Web, etc. Among other things, globalization is expected to manifest itself in the dynamics of financial indices of stock markets in different countries. Indeed, it is reasonable to expect that a significant coupling of the economy of a given country to the rest of the world (e.g. through foreign investments), would make its stock index more susceptible to changes in the world economic climate.

In this work, we suggest a simple measure of the level of financial globalization of a given country based on the analysis of cross-correlations between stock market indices in different countries and regions of the world. The main object of our study is the $N \times N$ empirical correlation matrix C_{ij} of index price fluctuations in a large number of individual countries ($N = 37$ in our study). The matrix is constructed by applying the formula $C_{ij} = (1/T) \sum_{t=1,\mathrm{T}} \delta x_i(t) \delta x_j(t)$ to the set of normalized local currency returns of individual indices, recorded over a period of T trading days. A return $\delta X_i(t)$ of the stock index $S_i(t)$ with the time horizon Δt is usually defined as $\delta X_i(t) = \ln S_i(t+\Delta t) - \ln S_i(t) \simeq (S_i(t+\Delta t) - S_i(t))/S_i(t)$. Different markets are characterized by different volatilities of their stock market indices. In order to be able to detect similarity in the pattern of returns in different countries one needs to exclude volatility effects by using normalized returns $\delta x_i(t)$. $\delta x_i(t)$ is constructed by offsetting each $\delta X_i(t)$ by its empirical average value $\langle \delta X_i \rangle$, and normalizing it by its empirical variance (volatility): $\delta x_i(t) = (\delta X_i(t) - \langle \delta X_i \rangle)/\sqrt{\langle \delta X_i^2 \rangle - \langle \delta X_i \rangle^2}$. The matrix C_{ij} defined this way has the property that in the absence of correlations and in the limit $T \gg N$ it approaches the unity matrix. However, in real life, the number of trading days T in one's dataset is always finite. As a result an empirically measured matrix C_{ij} is always dressed by a substantial amount of noise. It is exactly the task of separating any real correlations present in the signal from this spurious noise, that makes the analysis of real world data a non-trivial task. Mainstream economics literature was mostly devoted to a detailed analysis of these correlations for just a pair of stock indices, e.g. those of New York and Tokyo stock exchanges see e.g. Refs. [2,3], but even if a large correlation matrix was considered, little effort was made to reliably separate the signal from the noise. As is common in statistics, the job of looking for correlations gets much simpler when the number of correlated signals with similar properties is large. In our case, this corresponds to a large number of country indices N. The spectral analysis of the correlation matrix followed by a comparison of its spectrum with predictions of the random matrix theory (RMT) is a useful tool, which allows one to detect even weak correlations between multiple signals. This method was recently successfully applied towards quantifying reproducible correlations between price fluctuations of hundreds of stocks traded on the stock exchange of a single country [4,5], and two countries [6,7]. In this work, we go one step further and apply the techniques pioneered in Refs. [4,5] to a large number of stock market indices in a geographically diverse set of countries. An alternative method of analyzing the matrix of correlation coefficients by

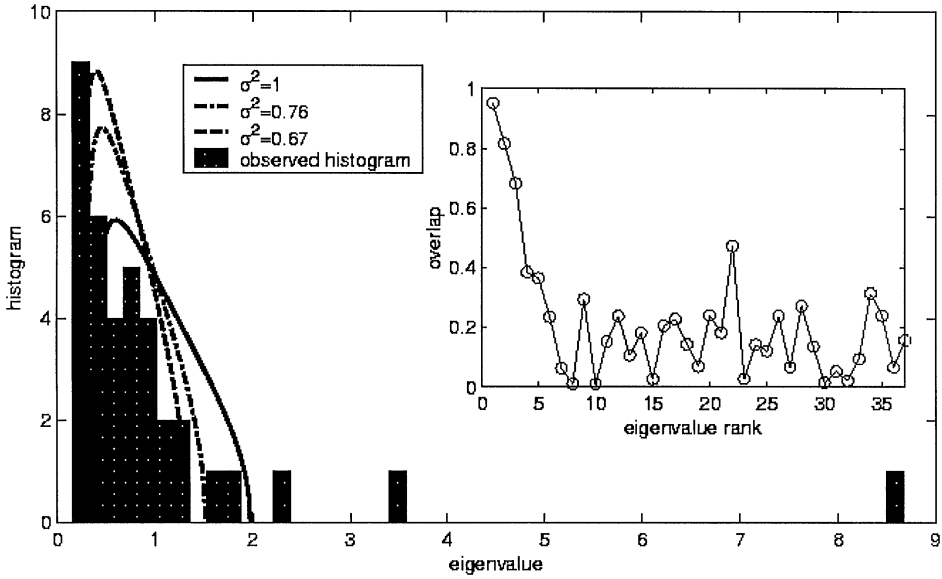

Fig. 1. The histogram of eigenvalues of the matrix C_{ij}. Lines are defined by Eq. (1) with $T = 226$, $N = 37$, and various values of σ, corresponding to the exclusion of fluctuations contained in the largest ($\sigma^2 = 0.76$), or two largest ($\sigma^2 = 0.67$) eigenvalues. The inset shows the overlap of eigenvectors computed for two consecutive 113-day intervals as a function of the rank of an eigenvalue. The three (perhaps even four) leading eigenvectors clearly have a higher than random overlap.

constructing the minimal spanning tree (MST) was described in Ref. [8], and recently applied to cross-correlations of world financial indices in Ref. [9]. Authors of this work also found a non-trivial correlation pattern manifested by a strong regional grouping of indices in the MST.

The raw data we had at our disposal consists of the daily open, high, low, and close prices of leading market indices (one per country) in 15 European, 14 Asian, and eight North and South American countries. We first calculated the daily open-to-close returns of each of these indices. We further selected from our set only those trading days for which we had a valid record for each and every country on our list. All data have gaps in them e.g. due to national holidays, when a particular market was closed. That left us with precisely 226 trading days approximately uniformly distributed between April 28, 1998 and December 20, 2000. Each of these remaining daily returns was normalized in such a way that $\sum_{t=1}^{226} \delta x_i(t) = 0$, and $\sum_{t=1}^{226} \delta x_i(t)^2 = 1$. The histogram of all 37 eigenvalues of the correlation matrix C_{ij}, shown in Fig. 1, revealed that the majority of eigenvalues are consistent with a null hypothesis of independent identically distributed Gaussian variables $\delta x_i(t)$. The prediction for the eigenvalue density

$$\rho_{\mathrm{RMT}}(\lambda) = \frac{T}{2\pi\sigma^2 N} \frac{\sqrt{(\lambda - \lambda_-)(\lambda_+ - \lambda)}}{\lambda} \tag{1}$$

given for this situation by (RMT) ([10], for the derivation of a more general formula using physics methods, see Ref. [11]) reasonably agrees with our data below $\lambda \simeq 1.3$.

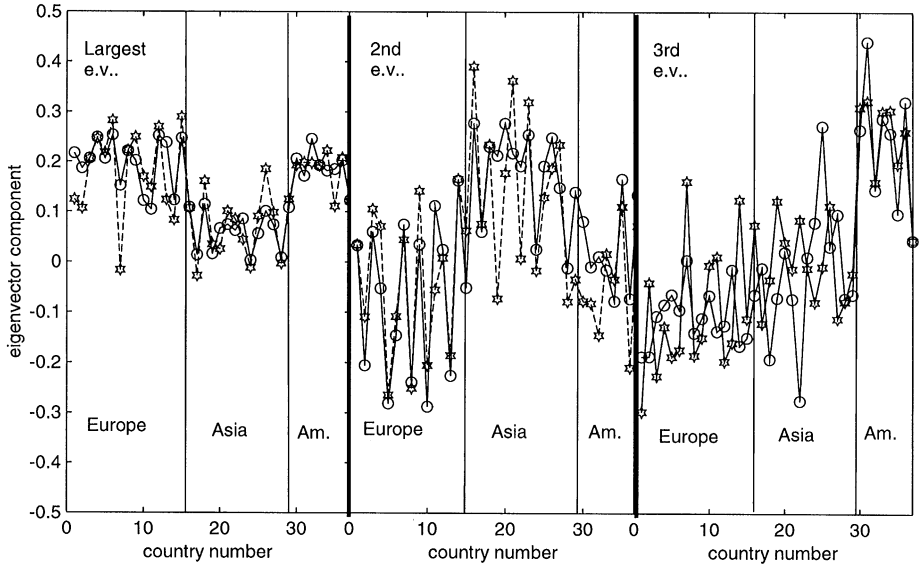

Fig. 2. Components of the three leading eigenvectors calculated for two consecutive 113-day intervals as a function of the country number (see Table 1).

This formula, derived in the limit of very large T and N, predicts sharp lower and upper cutoffs, $\lambda_{\pm} = \sigma^2(1 + N/T \pm 2\sqrt{N/T})$, in the eigenvalue density. This gives a strict quantitative test for deciding whether a particular eigenvalue reflects a real correlation signal present in the data, or is just a spurious noise effect caused by the finite length T of the data set. In principle, any eigenvalue significantly above the upper cutoff λ_+ should be treated as signal. The variance σ of $\delta x_i(t)$ can be renormalized from its starting value $\sigma = 1$ by the presence of correlations in the data. Indeed, we obtain the best fit of the noise-band part of the spectrum for $\tilde{\sigma}^2 = 0.67$, consistent with the empirically observed correlations. The three largest eigenvalues 8.7, 3.5, and 2.2 sufficiently exceed the theoretical upper limit $\lambda_+ = 1.97$ to be attributed to true correlation patterns. Indeed, in a control test we found that the probability for the largest eigenvalue of a correlation matrix, constructed from uncorrelated univariate Gaussian variables with N and T as our data, to exceed 2.2 is as low as 0.05%. This should be contrasted with a typical 5% to 1% confidence level of correlations between a pair of individual indices reported e.g. in Refs. [2,3]. In order to check the reproducibility of largest eigenvalues and their corresponding eigenvectors, we divided our sets into two consecutive 113-point subsets and repeated our analysis. The existence of 3 outlier eigenvalues did not change, however, their values were slightly modified. The largest eigenvalue was measured to be 9.8 during the first time interval, and 7.8 during the second. As can be concluded from the inset of Figs. 1 and 2, the corresponding eigenvectors are remarkably stable with overlaps between eigenvectors for the first and the second subintervals being 0.95, 0.82, and 0.68 for the largest, the second, and the third

eigenvalues, respectively. The largest possible overlap, realized when two eigenvectors are identical, is equal to 1. On the other hand, overlaps between eigenvectors from the noise-band between λ_- and λ_+ seem to be purely random (see inset of Fig. 1). Similar effects but with larger number of outliers above λ_+ (up to around 25) were observed [4,5] for individual stocks traded on US stock exchanges.

Components of the three highest ranking eigenvectors, measured for our data both in its entirety, and when divided in two equal subintervals, are given in Table 1 and plotted in Fig. 2. The first interesting result is that virtually all components of the largest eigenvector (e.v.) are positive, which means that there are no indices which are anti-correlated with others. Since eigenvectors corresponding to different eigenvalues have to be orthogonal to each other, other eigenvectors must contain negative components. The first (largest) eigenvector has strong support in European and American sectors, while its components in the Asian sector are somewhat smaller (yet still positive). The second eigenvector, on the other hand, is largely dominated by Asian stocks while the third one by American stocks.

Another interesting observation is that all three eigenvector components for some of the Asian emerging markets such as China, India, Pakistan, and Taiwan are too small to be detected. That means that in the first approximation, these indices are not influenced by the world index dynamics at all. In Europe, we saw no such correlation-free countries. However, the eigenvector components of Greece, Portugal, Russia, and Turkey were somewhat smaller than those of other European stock indices. North and South American stock indices have approximately equal components with, perhaps, only Peru and Venezuela somewhat falling behind.

It is interesting to compare our findings to those which were previously obtained for *weekly* returns using the minimal spanning tree (MST) technique [9]. In this work, it was also observed that indices are strongly grouped by the region with most indices of, say, Asian countries forming a separate branch of the MST. The market indices of Turkey, Greece, India, and Pakistan were found to be weakly correlated with the world index in both Ref. [9] and our study. We believe that the spectral analysis method, used in our work, gives a complimentary picture of reproducible correlations contained in the matrix C_{ij} to that of the MST method. One of the strong points about the spectral analysis method is that it gives a clear quantitative criterion for separating the signal from the noise. Also, with just a few large eigenvalues and their corresponding eigenvectors the output of the spectral analysis is easier to interpret and compare between, say, different time windows than that of the MST technique.

The leading eigenvector component of a stock market index of a given country can serve as a rough measure of the level of globalization of the financial sector of this country. This point of view is illustrated in Fig. 3, where the largest eigenvalue component is plotted as a function of the rank of the country in the A.T. Kearney/Foreign Policy Magazine Globalization Index^{TM} [1]. One can see a clear correlation between high globalization rank (1 being the highest and 50 the smallest) and the leading eigenvector component.

Table 1
Components of three leading eigenvectors computed for the whole 226-day time interval and its first and second halves. The last column is the rank of the globalization index of the country as defined in Ref. [1]

Index	Country	Index name, symbol	1st, e.v.			2nd, e.v.			3rd, e.v.			Rank in [1]
			All days	1st half	2nd half	All days	1st half	2nd half	All days	1st half	2nd half	
1	Austria	ATX, ˆATX	0.20	0.22	0.12	0.02	0.03	0.03	−0.19	−0.19	−0.30	7
2	Belgium	BEL-20, ˆBFX	0.16	0.19	0.11	−0.17	−0.21	−0.11	−0.12	−0.19	−0.04	N/A
3	Denmark	KFX, ˆKFX	0.21	0.21	0.21	0.07	0.06	0.11	−0.17	−0.11	−0.23	11
4	Finland	HELSINKI GENERAL, ˆHEX	0.22	0.25	0.25	0.02	−0.05	0.07	−0.14	−0.08	−0.13	5
5	France	CAC40, ˆFCHI	0.21	0.21	0.22	−0.28	−0.28	−0.26	−0.15	−0.06	−0.19	16
6	Germany	DAX, ˆGDAXI	0.27	0.25	0.28	−0.14	−0.15	−0.11	−0.14	−0.09	−0.17	14
7	Greece	GENERAL SHARE, ˆATG	0.10	0.15	−0.02	0.07	0.07	0.04	0.10	0.00	0.16	24
8	Netherlands	AEX GENERAL, ˆAEX	0.23	0.22	0.22	−0.25	−0.24	−0.25	−0.17	−0.14	−0.19	2
9	Norway	TOTAL SHARE, ˆNTOT	0.22	0.20	0.25	0.07	0.03	0.14	−0.12	−0.11	−0.15	9
10	Portugal	BVL30, ˆBVL30	0.14	0.12	0.17	−0.26	−0.29	−0.21	−0.07	−0.07	−0.01	15
11	Russia	MOSCOW TIMES, ˆMTMS	0.12	0.10	0.15	0.05	0.11	−0.06	−0.07	−0.14	0.01	45
12	Sweden	STOCKHOLM GENERAL, ˆSFOG	0.26	0.25	0.27	0.01	0.02	0.01	−0.17	−0.13	−0.20	3
13	Switzerland	SWISS MARKET, ˆSSMI	0.21	0.24	0.12	−0.23	−0.23	−0.19	−0.03	−0.01	−0.16	4
14	Turkey	ISE NATIONAL-100, ˆXU100	0.11	0.12	0.08	0.17	0.16	0.17	−0.08	−0.17	0.12	37
15	UK	FTSE100, ˆFTSE	0.26	0.25	0.29	0.01	−0.05	0.06	−0.17	−0.15	−0.11	8
16	Australia	ALL ORDINARIES, ˆAORD	0.11	0.11	0.11	0.35	0.28	0.39	−0.02	−0.06	0.07	23
17	China	SHANGHAI COMPOSITE, ˆSSEC	0.00	0.01	−0.03	0.07	0.06	0.08	−0.04	−0.01	−0.12	48
18	Honk Kong	HANG SENG, ˆHSE	0.13	0.11	0.16	0.25	0.23	0.23	−0.17	−0.19	−0.03	N/A
19	India	BSE30, ˆBSESN	0.02	0.02	0.03	0.07	0.21	−0.07	0.03	−0.07	0.12	49
20	Indonesia	JAKARTA COMPOSITE, ˆJKSE	0.06	0.07	0.03	0.24	0.28	0.18	0.05	0.02	0.04	38
21	Japan	NIKKEI 225, ˆN225	0.08	0.07	0.10	0.30	0.22	0.36	−0.07	−0.07	−0.01	29
22	Malaysia	KLSE COMPOSITE, ˆKLSE	0.07	0.06	0.09	0.15	0.19	0.01	−0.17	−0.28	0.08	20
23	New Zealand	NZSE40, ˆNZ40	0.07	0.09	0.04	0.29	0.25	0.32	0.00	0.01	−0.01	21
24	Pakistan	KARACHI100, ˆKSE	0.00	0.00	−0.01	0.00	0.03	−0.02	0.08	0.08	−0.08	N/A
25	Philippines	PSE COMPOSITE, ˆPSI	0.07	0.06	0.09	0.17	0.19	0.13	0.18	0.27	−0.01	34
26	Singapore	STRAITS TIMES, ˆSTI	0.13	0.10	0.19	0.24	0.25	0.19	0.03	0.03	0.11	1
27	South Korea	SEOUL COMPOSITE, ˆKS11	0.09	0.07	0.10	0.19	0.15	0.23	−0.03	0.10	−0.11	31
28	Taiwan	TAIWAN WEIGHTED, ˆTWII	0.00	0.01	0.00	−0.03	−0.01	−0.08	−0.09	−0.07	−0.08	N/A
29	Thailand	SET, ˆSETI	0.12	0.11	0.12	0.08	0.14	−0.03	−0.07	−0.06	−0.02	30
30	Argentina	MERVAL, ˆMERV	0.21	0.21	0.19	0.01	0.08	−0.08	0.31	0.27	0.31	39
31	Brazil	BOVESPA, ˆBVSP	0.19	0.17	0.20	−0.04	−0.01	−0.08	0.42	0.44	0.32	44
32	Canada	TSE 300 COMPOSITE, ˆTSE	0.23	0.25	0.20	−0.06	0.01	−0.14	0.17	0.14	0.16	10
33	Chile	IPSA, ˆIPSA	0.20	0.19	0.19	−0.01	−0.02	0.02	0.30	0.29	0.30	26
34	Mexico	IPC, ˆMXX	0.20	0.18	0.22	−0.06	−0.08	−0.03	0.29	0.26	0.30	41
35	Peru	LIMA GENERAL, ˆIGRA	0.17	0.19	0.11	0.14	0.17	0.11	0.17	0.10	0.20	43
36	USA	SP500, ˆSPC	0.21	0.20	0.21	−0.13	−0.07	−0.21	0.31	0.32	0.26	12
37	Venezuela	IBC, ˆIBC	0.13	0.12	0.12	0.10	0.13	0.07	0.09	0.04	0.04	35

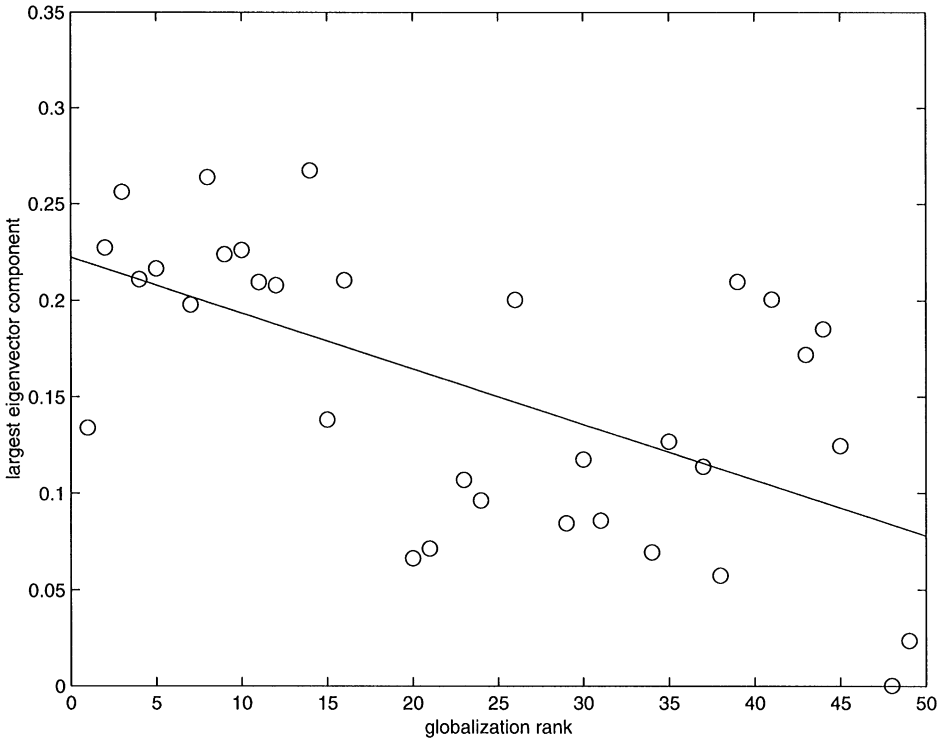

Fig. 3. The component of the highest ranking eigenvector as a function of the globalization rank of the country from Ref. [1]. The straight line is a linear fit to the data.

We further decided to explore how the outcome of the above eigenvector/eigenvalue analysis dependent on the time horizon Δt over which one computes the returns of an index. When instead of daily open-to-close returns we repeated our analysis for one day close-to-close returns, the largest eigenvalues have changed from 2.2, 3.5, and 8.7 to 2.1, 3.7, and 11.0, respectively. A noticeable 25% increase in the largest eigenvalue perhaps can be attributed to markets having more time to respond to the news. Also, while daily open-to-close returns in Asia, Europe, and Americas have almost no overlap (i.e., time when two markets are simultaneously open), this situation is improved when one considers daily close-to-close fluctuations. As shown in Fig. 4, the largest eigenvalue continued to grow (albeit slowly), as the time horizon of close-to-close returns was changed from 1 to 10 trading days (weekly return usually corresponds to just five trading days), reaching the value of 16.2 for the longest time horizon. However, the largest eigenvectors computed for these very different time horizons remained remarkably stable. For example, the average overlap between the highest rank eigenvectors computed for these 10 different time horizons turned out to be 0.99, i.e., these eigenvectors on average are only 1% different from each other! Overlaps were somewhat smaller for lower ranking eigenvectors with average values of 0.77 for the second and 0.61 for the third largest eigenvalues. Still, as can be seen in the inset

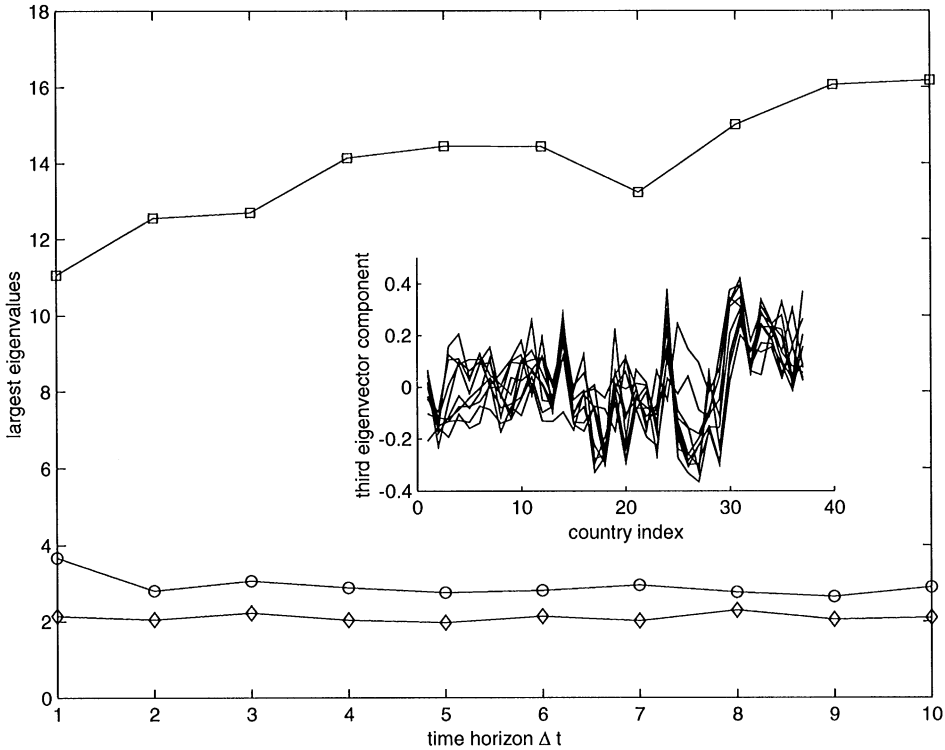

Fig. 4. Three largest eigenvalues of the correlation matrix of close-to-close returns as a function of the time horizon, i.e., the number of days used to calculate the return. The inset shows the components of the third eigenvector for all 10 time horizons.

of Fig. 4, even in the third eigenvector many of the main features are very robust with respect to changes in the time horizon.

In an attempt to establish how relevant are magnitudes (as opposed to signs) of price fluctuations to the observed correlation patterns, we have repeated the above analysis using $\delta X_i(t) = \mathrm{sign}(S_i(t + \Delta t) - S_i(t))$. The observed eigenvectors for different time horizons had 0.99 average overlap with those computed using $\delta X_i(t) = \ln S_i(t + \Delta t) - \ln S_i(t)$. The largest eigenvalue again grew with the time horizon from 7.4 for signs of one day close-to-close returns to 10.4 for signs of 5-day close-to-close returns. Lower rank eigenvectors of the correlation matrix of signs also had substantial overlaps of 0.93 and 0.86 with normal ones. This allows us to conclude that it is signs and not magnitude of returns which are mostly relevant for the observed correlation patterns.

Since the psychology and consequently actions of market participants during market crashes differs from those during market rallies, one expects to see the difference in correlations between positive and negative changes in indices. To this end, we have calculated matrices C_{ij}^{+} and C_{ij}^{-} using only positive and negative parts of signals $x_i(t)$ correspondingly. The largest eigenvalues of these two matrices for the whole 226-day interval were measured to be 13.7 and 16.8 correspondingly. From this one can infer

that negative changes (market "crashes") appear to be more correlated in different countries than the positive ones (market "rallies"). This effect is reproducible. Indeed, when we subdivided our data set in two 113-day intervals, the largest eigenvalues were measured to be $\lambda_+ = 14.3$, $\lambda_- = 18.2$ for the first subinterval and $\lambda_+ = 13.8$, $\lambda_- = 15.9$ for the second one. As one can see in both cases $\lambda_- > \lambda_+$ indicating stronger correlations between crashes.

In what was described above, we always computed (nearly) synchronous correlations of different market returns on the same day (or the same week for longer time horizons). One has to take into account that due to the time-zone difference, daily open-to-close returns computed on the same trading day (say, February 13) are not actually synchronous with Asian stock markets in the lead, followed by European and later with a small overlap by American markets. However, the significance of this time-zone difference for say weekly returns is much less pronounced. To check if the predictability of daily returns is restricted to the same trading day or survives for several days, we have investigated the correlations in daily close-to-close returns with all Asian indices shifted by S days. The negative values of the shift S corresponds to correlations of daily returns of Asian indices $|S|$ days after an observed pattern of returns of the European and American stocks, while positive S corresponds to Asian indices preceding the rest of the world by S days. The simplest way to detect the presence of correlations in this case is by calculating the average correlation coefficient connecting any of the 14 Asian indices with $15 + 8 = 23$ European and American ones. The size of the sample over which this average is taken is $14 \times 23 = 322$. To check for reproducibility of the observed patterns, we divided our data set consisting of 226 trading days into three equal length segments and calculated this average for each segment independently. The results are shown in the main panel of Fig. 5. Reproducible correlations seem to survive for up to 3 days on the negative part of the axis, corresponding to the reaction of Asian stocks to changes in European and American indices. At least the sign of the average correlation coefficient was found to be consistently positive in all three of our segments. $S = -1$ has the largest magnitude of correlations. These (positive) correlations represent the next trading day reaction of Asian stocks to changes in European and American ones. Due to time-zone differences, a somewhat smaller (yet still positive) correlation coefficient observed for $S = 0$ corresponds to the opposite effect, i.e., the response of European and American stocks to the events in Asia on the day before. Our results for $S = 0$ and -1 are in agreement with positive correlations between open-to-close returns at Tokyo and New York stock exchanges that were previously reported in the economics literature [2,3] as well as in recent econophysics papers [12]. On the positive S side, there seems to be a reproducible negative correlation at $S = 1$ followed by a noisy signal for larger values of S. The inset to Fig. 5 show the results of the same analysis for the whole 226-day interval when either Asian, or European and Asian stocks are shifted by S days. It is interesting that in this graph which has 3 times better statistics than the main panel of the figure, there seem to be oscillations of correlation coefficients with period of two trading days for negative values of S.

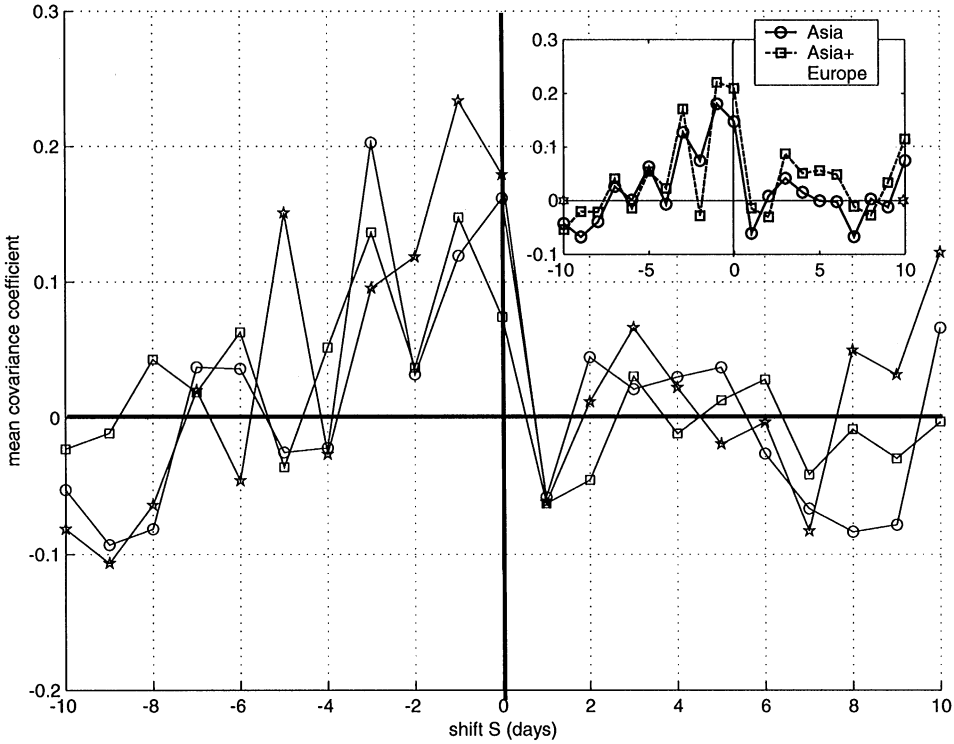

Fig. 5. The mean value of the correlation coefficient connecting Asian indices to the rest of the world as a function of the shift S. Negative values of S correspond to Asian indices taken $|S|$ days later than the rest of the world. Three data sets were taking in three equal length subintervals of our data set. The inset shows the same analysis repeated including all data points with Asian indices (circles) and Asian and European indices (squares) shifted. Note oscillations for negative S.

Work at Brookhaven National Laboratory was carried out under Contract No. DE-AC02-98CH10886, Division of Material Science, U.S. Department of Energy.

References

[1] Measuring Globalization, Foreign Policy Magazine, January/February 2001, pp. 56–65, also available online at http://www.atkearney.com/-pdf/eng/Globalization_index_S.pdf.
[2] K.G. Becker, et al., J. Finance XLV (1990) 1297–1306.
[3] W.-L. Lin, et al., Rev. Finan. Stud. 7 (1994) 507.
[4] L. Laloux, et al., Phys. Rev. Lett. 83, 1467.
[5] V. Plerou, et al., Phys. Rev. Lett. 83 (1999) 1471.
[6] S. Drozdz, et al., Physica A (2001) to appear, (cond-mat/0011488)
[7] S. Drozdz, et al., at cond-mat/0102402.
[8] R.N. Mantegna, Eur. Phys. J B 11 (1999) 193.
[9] G. Bonanno, et al., Phys. Rev. E 62 (2000) R7615.
[10] K.M. Wachter, Ann. Probab. 6 (1978) 1.
[11] A.M. Sengupta, P.P. Mitra, Phys. Rev. E 60 (1999) 3389.
[12] N. Vandewalle, et al., cond-mat/0001293.