VALENTIN RAGEA

TESTING CORRELATION STABILITY DURING HECTIC FINANCIAL MARKETS

Valentin Ragea, University of Vienna and
Center for Central European Financial Markets
Email: a0108223@unet.univie.ac.at
Acknowledgments: Thomas Breuer, Ines Fortin, Alois Geyer, Stephen Taylor and an anonymous referee for *Financial Markets and Portfolio Management* have made valuable comments. I also thank the participants of the 6th Conference of Swiss Association for Financial Markets Research (April 2003) in Zurich, the participants at the Doctoral Seminar of European Finance Association (August 2003) in Glasgow, the BWZ Research Seminar (March 2003), the CCEFM Doctoral Seminar (May 2003) and Econometrics Research Seminar at Institute for Advanced Studies (May 2003) in Vienna. The work at this paper was partially funded by Gutmann Center for Portfolio Management.

1. Introduction

The correlation stability question could be asked in two contexts. Firstly, is correlation stable through time? A plot of a 1-year window correlation between major equity indices shows clearly that the correlation has a wide variation through time.[1] Secondly, do highly volatile markets produce disruptions in the joint move of the risk factors? This paper attempts an answer for the second question.

There are several situations where the correlation stability during hectic periods is of vital importance. Firstly, a common practice (see LAUBSCH (1999), BREUER and KRENN (1999), KIM and FINGER (2000)) in market risk management is to use, besides the historical stress scenarios, the so-called "predictive stress tests", in which, given current conditions, the large movements in the nominal level of risk factors are computed and the impact of these large movements is assessed. Although correlation between stressed risk factors is set to zero, the estimated values are still used when some of the risk factors are not included in the stress scenario. This technique, quite popular in practice (KUPIEC (1998)), could be theoretically justified if the underlying statistical properties of the financial time series hold not only during the normal times upon which they were derived, but also during periods of high market volatility.

Out of these statistical properties, the constant correlation (or in a larger sense, the dependence invariance) is critical. Starting with modern portfolio theory pioneered by Markowitz, the correlation has had the central role in reducing the portfolio risk and it is highly desirable to perform this task when it is most needed: during turbulent times. A widespread opinion on financial markets is that "during market events, correlation change dramatically" (BOOKSTABER (1997)), and the opinion is shared by major market participants (JP Morgan (1999)) and regulators (Bank of International Settlement (1999), Table A12). The same

idea is summarized below in a remark of Federal Reserve Board Chairman Alan Greenspan[2]:
"Furthermore, joint distributions estimated over periods without panics will misestimate the degree of correlation between asset returns during panics. Under these circumstances, fear and disengagement by investors often result in simultaneous declines in the values of private obligations, as investors no longer realistically differentiate among degrees of risk and liquidity, and increases in the values of riskless government securities. Consequently, the benefits of portfolio diversification will tend to be overestimated when the rare panic periods are not taken into account."
Asset management is a second area interested in the behavior of correlation, although the focus moves from the risk manager's perspective of the worst-case scenario to the medium and long term effect on the final value of the assets. A third issue in which the correlation behavior during hectic times is a key factor is the derivative pricing and hedging if the underlying asset is a basket. The last, and the most subtle, problem involving the stability of correlation is to be assigned to the so-called "model risk", and it has its root in the computational advantages of the Gaussian assumption. An example in this respect is a multivariate form of the widely used GARCH model. BOLLERSLEV (1990) showed that, if the correlations are considered to be time invariant, the maximum likelihood estimate of the correlation matrix is equal to the sample correlation matrix, which is always semi-positive definite, and the optimization runs smoothly. Otherwise, it could not be theoretically guaranteed a robust estimation of the parameters. Given this difficulty, it became a common practice, especially when a data set with a large number of time series is manipulated, to consider the correlation as invariant in time and to pursue further modeling tasks based on this assumption.
The present paper is organized as follows. Section 2 briefly reviews the literature and the difficulties posed by the described problem, Section 3 presents the model and the hypothesis testing, Section 4 describes the data and the empirical results and Section 5 concludes. As the estimation issues and partly the theoretical foundation of the hypothesis testing are relatively technical, they were completely transferred to the appendix. Appendix A introduces the EM algorithm as an appropriate estimation method and briefly describes how to obtain the standard errors (variance-covariance matrix) of the estimated parameters.

2. Literature Review

The first (intuitive) idea (see BOYER et al. (1997), LORETAN and ENGLISH (2000)) to test for changes in correlation is to split and assign the data to either a "hectic" or a "quiet" period and compute conditional correlation for each subsample. The technique is applied over a pair of returns, the splitting criterion usually refers to one component and is formulated as the returns are two/three/four standard deviations away from the mean or the standard deviation for a certain group (months, quartiles, deciles etc.) is larger than a certain multiplier of the sample standard deviation. All these comparisons support the changing correlation hypothesis. However, repeating the calculations with a simulated i.i.d. bivariate sample, the same conclusion is reached, even if the data has obviously (it was generated so!) constant correlations. BOYER et al. (1999) notice the same flaw for more realistic return data: a AR(1) process and a bivariate GARCH process with constant correlation of the type introduced by BOLLERSLEV (1990). Hence, their conclusion is compelling: truncating the data introduces strong biases in the estimated parameters. This is the most important fact concerning the issue under research, and this conclusion deserves to be outlined since it is non-intuitive and likely to be ignored in practice. Therefore, for a valid test, an explicit model of "normal" vs. "hectic" regime is needed, and the test should be constructed based on the statistical properties of the correlation in each of these two regimes.

Another approach (LONGIN and SOLNIK (2000)) to test for stability of correlation is to use the extreme value theory framework (EVT). It comes natural to apply the results developed for the study of rare events to the behavior of financial assets during turbulent times. EVT deals with previously explained pitfalls by providing theoretically-sound statistical properties for truncated data. One more advantage is that, in a straightforward application of EVT, no assumption about the return distribution is needed. At the core of the test is the modeling of the extreme return dependence through logistic copula, with two reasons behind. Firstly, it is a parsimonious specification that helps in the estimation stage. Secondly, there is a direct relationship between the parameter of the logistic copula and the correlation coefficient that greatly simplifies the task of constructing a correlation stability test. The empirical research uses monthly equity index returns and focuses on five developed markets: U.S, U.K., France, Germany and Japan. The tests were applied to both lower and upper tails (both positive and negative jumps). The conclusions support the idea of constant correlations for extreme positive returns, but this conclusion could not be maintained for large negative returns, where the results are mixed.

However, there are some drawbacks of this method that deserves some scrutiny. The dependence modeled through logistic function, which, although in the line with the theoretical foundations of the EVT, it is just an arbitrary choice to simplify estimation and hypothesis testing. Furthermore, the EVT is rather weak as the sample is small (there are just a few extreme events) and the estimated parameters are highly inefficient. An improvement in efficiency introduces bias in the estimated parameters and requests an explicit model for returns. Another approach is taken by KIM and FINGER (2000). They use a mixture of bivariate normal distribution to model explicitly the quiet and hectic regimes. As our approach also follows the same path, we will pay a closer attention to their work. The estimation is done in two steps. Firstly, the marginal distribution of one risk factor (so-called "core" asset, S&P 500 index) is fitted on a mixture of univariate normal distributions. Since the likelihood function has several local maxima, a grid search is employed and the hectic probability is restricted to the interval [0.01, 0.49]. In the second step, a bivariate model is considered; the parameters of the "core asset" are plugged in from the first step, and the parameters of the second risk factor as well as the correlations between the two assets in both quiet and hectic markets are computed. The estimation procedure for this second step is not theoretically explained, the authors noting only that the estimators provided are unbiased.

Two tests for stability of correlation are used. Firstly, a new model is estimated with returns being random draws from a (bivariate) normal distribution (rather than a mixture of two normal distributions). With the parameters from this model, a Monte Carlo simulation is run and a 90% confidence interval for correlation is computed. Afterwards, the hectic correlation is plotted against this confidence interval. The second test is motivated by the truncation flaws explained above. Though it does not truncate the data explicitly, the model, through the conditional event on a single marginal return, is subject to the same bias issues. To avoid this flaw, the authors suggest to simulate new pairs of returns $(\boldsymbol{x},\boldsymbol{y}')$, where $\boldsymbol{y}' = \boldsymbol{y}|\boldsymbol{x}$. $\boldsymbol{y}'$ is computed from original mixture of bivariate normal distribution, but with both quiet and hectic correlation set equal with unconditional cor-relations. A 90% confidence level is computed again, and the hectic correlation is researched to be in or outside this interval.

The data included 19 time series (with S&P 500 considered as the core asset), from broad classes: FX, short and long term bond prices, developed markets equity indexes and commodities. The first test shows the correlation for 13 out of 18 peripheral assets lies outside the confidence interval (the correlation is changing). However, in the second test, when the bias is removed, the number drops from 13 to 4 (most of them long term bond prices).

As a conclusion of this brief survey, the changing correlation belief seems to be more popular among practitioners rather than academics. However, the tests used till now were based on restrictive assumption and a clear-cut definition of "hectic" or "stressed" market is difficult to provide.

The model we propose, although following the same normal vs. hectic market specification as KIM and FINGER (2000), differs in the following points:

- A different estimation method. The procedure drops the differentiation between a core versus a marginal asset and the estimation technique uses the information from the both data series in computing probabilities of hectic/quiet regimes and asset correlation in each regime. The bias associated with a single-factor splitting criterion is thus avoided.
- A new test for correlation stability. The asymptotic properties of the correlation in each regime are carefully derived and a formal test introduced.
- The empirical research. Up to our knowledge, there is no academic research to test stability of stock indexes correlation between developed markets and emerging Europe.

3. Testing Correlation Stability using Mixture of Normal Distributions

3.1 The Model

In order to allow explicit computations of correlation coefficient in both normal and hectic periods, we model the pair of returns as a random draw from a bivariate normal distribution with either a low or a high variance:

$$x_i = \begin{cases} N\left[\mu_q, \Sigma_q\right] \\ \text{with probability} \quad p \quad (\text{quiet days}) \\ N\left[\mu_h, \Sigma_h\right] \\ \text{with probability} \quad 1-p \quad (\text{hectic days}) \end{cases} \tag{1}$$

where:

N	is a bivariate normal distribution
$\boldsymbol{x_i} = (x_1 \;\; x_2)$	is the (bivariate) return in day i
$\boldsymbol{\mu_q} = (\mu_{q1} \;\; \mu_{q2})$	is the mean vector for the quiet distribution/regime
$\boldsymbol{\mu_h} = (\mu_{h1} \;\; \mu_{h2})$	is the mean vector for the hectic distribution/regime

$$\Sigma_q = \begin{pmatrix} \sigma_{q1}^2 & \sigma_{q1}\sigma_{q2}\rho_q \\ \sigma_{q1}\sigma_{q2}\rho_q & \sigma_{q2}^2 \end{pmatrix}$$

is the covariance matrix for the quiet distribution/regime

$$\Sigma_h = \begin{pmatrix} \sigma_{h1}^2 & \sigma_{h1}\sigma_{h2}\rho_h \\ \sigma_{h1}\sigma_{h2}\rho_h & \sigma_{h2}^2 \end{pmatrix}$$

is the covariance matrix for the hectic distribution/regime.

The additional condition $p > 0.5$ is needed to assure model identification. Thus, the 11-parameter vector is:

$$\Psi = \left(\mu_{q1} \;\; \mu_{q2} \;\; \sigma_{q1} \;\; \sigma_{q2} \;\; \rho_q \;\; \mu_{h1} \;\; \mu_{h2} \;\; \sigma_{h1} \;\; \sigma_{h2} \;\; \rho_h \;\; p\right)$$

The model could also be interpreted as a regime-switching model. A day is assigned to either a quiet regime or to a hectic regime. Although in both cases, the return is (bivariate) normally distributed, the variance of the quiet regime is lower than the variance of the hectic regime. Returns from two consecutive days are assigned independently to a certain regime. This assumption ignores volatility clustering, one of the common artifacts of the financial time series. However, we will show in Section 4 that the model accounts for this feature.

Modeling returns with mixture of normal distributions was first proposed by KON (1984), and introduced in risk management by ZANGARI (1996). The main reasons behind their popularity is the elegant specification of a fat tail distribution (and, if it is the case, the right skew) for asset returns and, in the same time, the ability to use some of the computational advantages of the Gaussian distribution.

One more advantage of this model is the fully accounting of dependence using the correlation coefficient. Correlation is an exhaustive measure of dependence only for elliptical distributions (the most representative of this class being the normal and the Student-*t* distributions). For a lucid appraisal of correlation limitations see EMBRECHTS et al. (1998). However, correlation is widely used and intuitively understood on the industry side and hence the reluctance to use more advanced measures of dependence.

Before any further proceedings, it is worth to verify whether this more sophisticated model performs better than the single multivariate normal distribution. The test employed and the results are presented in Table 1 and they support the double mixture assumption.

One of the regularity conditions requested for a straightforward application of the likelihood ratio test is the true parameter value to be interior to the parameter space. This condition is violated in the present case, in which, under the null hypothesis p = 1. Therefore, we use a modified Likelihood Ratio (MLR) test suggested by WOLFE (1971). The null hypothesis is that the data follows a single bivariate normal distribution. The test statistic and its distribution are:

$$MLR = -2c(l - l_c) \sim \chi^2_{df} \qquad (2)$$

where:

- c = sample size – dimension of the data – 0.5 × no. of components in the mixture model – 1 (c = 1032)
- l is the log likelihood of the single multivariate normal model
- l_c is the log likelihood of the mixture of multivariate normal model
- df = 2 × (number of parameters in the mixture model – number of parameters in the single-distribution model – number ofmixing probabilities) (df = 10).

The null hypothesis is that data follows a single distribution. The critical values are: 15.99 (10% confidence level), 18.31 (5% confidence level), 23.21 (1% confidence level). The test values are presented in the table.

Table 1: Modified Likelihood Ratio Test to Discriminate between a Single and a Mixture of Bivariate Normal Distribution.

Index	MLR test	
	North America	Europe
Eastern Europe	102.76	111.76
Czech Republic	81.08	104.87
Hungary	168.04	179.88
Poland	72.56	79.42
Russia	132.18	138.77
Turkey	181.7	205.43

3.2 Estimation. EM Algorithm Solutions for a Mixture of Two Bivariate Normal Distributions

The direct estimation of the model is cumbersome as a global maximum does not exist for this function (HAMILTON (1991)). Consequently, the attempt to directly maximize the log likelihood function leads to instability, local solutions, and non-convergence. The problem gets even more complicate as the parameter space is rather large (11 unknowns).

If there are no restrictions on covariance matrices, the log likelihood function is unbounded. If the first mean vector is set equal with one of the return pair, the likelihood function goes to infinity as volatility goes to 0. Following the same reasoning, MCLACHLAN and KRISHNAN (1997) draw the attention of another potential peril in the estimation. A relative large local maximum could occur as a consequence of a fitted component with a very small (but nonzero) variance (in case of multivariate data we interpret variance as generalized variance, i.e. the determinant of the covariance matrix). This component comprises data very closed to each other. As a practical conclusion, there is a need to monitor the variance of the components to identify these spurious local maximizers.

However, once these difficulties are surmounted, under mild regularity conditions, there is a sequence of roots of the likelihood equation that is consistent and asymptotically efficient. These roots converge in probability to local maxima in the interior of the parameter space. To find these roots, we propose an alternative solution, based on the expectation-maximization (EM) algorithm. Although it doesn't overcome all the drawbacks previously mentioned (the most important, it does not guaranty a correct estimate), the method makes the estimation feasible. A primer of EM algorithm is given in Appendix A, and, in this section, we will only mention that it is suitable to the models whose estimation, although difficult to solve, becomes a trivial problem when one variable of the model is consider known. We will refer to this data-augmented model as the "complete-data model" and to the original model as "incomplete-data model".

The probability distribution function (p.d.f.) of the (incomplete-data) model introduced in the previous subsection is:

$$f(x_i) = p \cdot f_q(x_i) + (1-p) \cdot f_h(x_i) \tag{3}$$

where $f_q(x_i)$ and $f_h(x_i)$ are the p.d.f. of bivariate normal distributions $N(\mu_q, \Sigma_q)$ and, respectively, $N(\mu_h, \Sigma_h)$. The log likelihood function is:

$$l(x;\Psi) = \sum_{i=1}^{n} \ln\left[p \cdot f_q(x_i) + (1-p) \cdot f_h(x_i)\right] \tag{4}$$

The complete data model could be specified by adding (considering known) the missing information mathz. z_i tells whether the return pair $\boldsymbol{x}_i$ corresponds to a normal or hectic day.

$$\boldsymbol{y} = (\boldsymbol{z}, \boldsymbol{x}) \tag{5}$$

$$\boldsymbol{y}_i \,|\, (z_i=0) \sim N(\mu_q, \Sigma_q) \tag{6}$$

$$\boldsymbol{y}_i \,|\, (z_i=0) \sim N(\mu_h, \Sigma_h) \tag{7}$$

$$z_i = \begin{cases} 0 \, with\ probability \quad p \quad (quiet\ days) \\ 1 \, with\ probability \quad 1-p \quad (hectic\ days) \end{cases} \tag{8}$$

The probability density function for the complete data is:

$$f_c(z_i, x_i) = \left[p \cdot f_q(x_i)\right]^{1-z_i} \left[(1-p) \cdot f_h(x_i)\right]^{z_i} \tag{9}$$

where f_q and f_h are defined as above. The new log likelihood function is:

$$l_c(y;\Psi) = \\ = \sum_{i=1}^{n} (1-z_i)\left[\ln p + \ln f_q(x_i)\right] + z_i\left[\ln(1-p) + \ln f_h(x_i)\right] \tag{10}$$

which has obvious computational advantages over l.

Following the procedure outlined in Appendix A, the parameters are computed using an iterative algorithm, starting from arbitrary values. For the E-step, $E_{\Psi^{(k)}}\left[l_c|x\right]$ has to be computed.

$$\begin{aligned} E_{\Psi^{(k)}}\left[l_c|x\right] &= \left[\sum_{i=1}^{n} (1-z_i)\left[\ln p + \ln f_q(x_i)\right] + z_i\left[\ln(1-p) + \ln f_h(x_i)\right] \Big| x\right] \\ &= \left[\sum_{i=1}^{n} \left[\ln p + \ln f_q(x_i)\right] E_{\Psi^{(k)}}\left[(1-z_i)|x\right] + \left[\ln(1-p) + \ln f_h(x_i)\right] E_{\Psi^{(k)}}\left[z_i|x\right]\right] \end{aligned} \tag{11}$$

Using the fact that z_i is a Bernoulli random variable and Bayes' formula:

$$
\begin{aligned}
& E_{\Psi^{(k)}}\left[(1-z_i)\middle|x\right] = \\
& = P\left(z_i = 0)\middle|x\right) \\
& = \frac{P(z_i=0)P(x_i,z_i=0)}{P(z_i=0)P(x_i,z_i=0)+P(z_i=1)P(x_i,z_i=1)} \quad (12) \\
& = \frac{p\cdot f_q(x_i)}{p\cdot f_q(x_i)+(1-p)\ln f_h(x_i)} \\
& = \alpha_i
\end{aligned}
$$

Combining previous two relations, we obtain:

$$
\begin{aligned}
& E_{\Psi^{(k)}}\left[l_c\middle|x\right] = \\
& = \sum_{i=1}^{n}\left\{\begin{array}{l}\alpha_i\left[\ln p+\ln f_q(x_i)\right]+ \\ (1-\alpha_i)\left[\ln(1-p)+\ln f_h(x_i)\right]\end{array}\right\} \quad (13)
\end{aligned}
$$

The M-step optimization is:

$$
\Psi^{(k+1)} = \arg\max_{\Psi} E_{\Psi^{(k)}}\left[l_c\middle|x\right] \quad (14)
$$

and has the following closed-form solution:

$$
p^{(k+1)} = \frac{\sum_{i=1}^{n}\alpha_i}{n} \quad (15)
$$

$$
\mu_q^{(k+1)} = \frac{\sum_{i=1}^{n}\alpha_i x_i}{\sum_{i=1}^{n}\alpha_i} \quad (16)
$$

$$
\Sigma_q^{(k+1)} = \frac{\sum_{i=1}^{n}\alpha_i\left(x_i-\mu_q\right)^T\left(x_i-\mu_q\right)}{\sum_{i=1}^{n}\alpha_i} \quad (17)
$$

$$
\mu_h^{(k+1)} = \frac{\sum_{i=1}^{n}(1-\alpha_i)x_i}{\sum_{i=1}^{n}(1-\alpha_i)} \quad (18)
$$

$$
\Sigma_h^{(k+1)} = \frac{\sum_{i=1}^{n}(1-\alpha_i)(x_i-\mu_h)^T(x_i-\mu_h)}{\sum_{i=1}^{n}(1-\alpha_i)} \quad (19)
$$

It should be mentioned that the formulas we obtain as the first step of an iterative procedure are the same with those given by KIM and FINGER (2000) for their final estimates and this fact explains the bias of their parameters.

3.3 The Tests for Correlation Stability

As the estimation is still based on maximum likelihood (ML) method, the results inherit the properties of the ML estimates, i.e. asymptotic normality and unbiasedness. These properties support the use of a t-test. The null hypothesis is:

$$
H_0: \quad \rho_q = \rho_h
$$

and the t-statistics is:

$$
t = \frac{\rho_q-\rho_h}{\sqrt{\sigma_{\rho_q}^2+\sigma_{\rho_h}^2-2\sigma_{\rho_q,\rho_h}}} \quad (20)
$$

where $\sigma_{\rho_q}^2$ and $\sigma_{\rho_h}^2$ are the standard errors of estimated parameters ρ_q and, respectively, ρ_h. Since the statistics is *t*-distributed with 1000 degrees of freedom, one could reject H_0 at 95% confidence level if *t* is outside [–1.96, 1.96] interval.

An alternate way is to employ the classical Wald test. This choice is also supported by the known failure of the likelihood ratio test in the case of mixture distributions. The Wald statistic, with one restriction as applied in this context, is:

$$
W = n\left[h(\hat{\Psi})\right]^T\left[H_{\hat{\Psi}}^T I_{\hat{\Psi}}^{-1} H_{\hat{\Psi}}\right]^{-1}\left[h(\hat{\Psi})\right] \sim \chi_1^2 \quad (21)
$$

where:

$h(\Psi) = 0$	is the restriction ($\rho_q - \rho_h = 0$)
n	is the sample size
$I^{-1}{}_{\Psi}$	is the inverse of the observed-data information matrix and
H_{Ψ}	is the partial derivative of $h(\Psi)$ with respect to the parameter vector, that reduces to ($0\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ -1\ 0$).

After making the necessary calculations, the W-statistics becomes:

$$W = \frac{(\rho_q - \rho_h)^2}{\sigma^2_{\rho_q} + \sigma^2_{\rho_h}} \tag{22}$$

4. Data and Empirical Results

The empirical application tests the correlation stability of the emerging Eastern Europe stock indices with respect to their Western/American counterparts. The MSCI index family is used, due to the consistent methodology for both the developed and emerging markets. This is particularly useful, as a comparison between the two and a research into the joint move is attempted. Table 2 lists the stock indices used for empirical research, briefly explains the country coverage and presents basic statistics.

All indices are capitalization-weighted. The sample covers the period from 4 January 1999 to 30 November 2002, with n = 1012 data points. There were eliminated periods of concomitant lack of activity on one or two markets and low transaction levels on other markets (religious holidays).

The mean (μ) and the standard deviation (σ) are sample estimates. The skewness (S) and the ex cess kurtosis (K) are computed upon the following formulas:

Table 2: Empirical Data. Description and Basic Statistics

Panel A: Description

Region/Country	Index	Observations
North America	MSCI North America	Covers U.S. and Canada
West. Europe (EU)	MSCI Europe	Covers E.U. area (15 member countries)
Eastern Europe	MSCI Emerging Europe	Covers Czech Republic,Hungary, Poland and Russia
Czech Rep.	MSCI Czech Rep.	
Hungary	MSCI Hungary	
Poland	MSCI Poland	
Russia	MSCI Russia	
Turkey	MSCI Turkey	

Panel B: Basic Statistics

	μ		σ		S	K	JB	AC(1)	AC(2)	AC(3)
	Daily	Annual	Daily	Annual						
North America	–0.02%	–6.05%	1.37%	21.76%	0.13	1.34	78.34	0.01	–0.03	–0.05
Europe	–0.03%	–7.58%	1.31%	20.74%	–0.13	1.79	137.59	0.06	–0.07	–0.06
East. Europe	0.02%	4.20%	1.71%	27.16%	–0.16	1.71	127.44	0.06	0.01	–0.02
Czech Rep.	0.04%	10.48%	1.74%	27.60%	0.08	1.14	55.54	0.06	–0.04	–0.03
Hungary	–0.01%	–2.09%	1.85%	29.41%	0.13	3.91	648.11	0.08	–0.01	–0.07
Poland	–0.01%	–1.61%	2.00%	31.75%	0.10	1.29	71.26	0.08	–0.02	–0.02
Russia	0.14%	42.78%	3.23%	51.24%	0.09	3.92	648.31	0.05	0	–0.05
Turkey	0.01%	2.33%	4.04%	64.06%	0.09	5.46	1259.17	0.06	–0.01	–0.02

$$S = \frac{1}{\sigma^3 n} \sum_{t=1}^{n} (R_t - \mu)^3$$

$$K = \frac{1}{\sigma^4 n} \sum_{t=1}^{n} (R_t - \mu)^4 - 3$$

To test for normality, the widely-used Jarque-Bera (JB) test is employed. The statistic follows a χ^2_2 distribution and is given by:

$$JB = \frac{T}{6}\left(S^2 + \frac{K^2}{4} \right)$$

Although the normality hypothesis is rejected with JB test, the skewness and autocorrelation have little contribution, and the kurtosis is the main factor to explain the non-normality. This observation particularly encourages a mixture of normal distribution model.

All indices are total return indices, including the tax-adjusted dividend income to the market (price) performance, as it better reflects the total return and removes the dividend effects. The dividend is reinvested after the deduction of withholding tax, applying the rate to non-resident individuals who do not benefit from double taxation treaties. To construct a country index, MSCI screens the individual stocks for size and liquidity, and select securities to cover 85% of free float-adjusted market capitalization for each country. MSCI regional indices are aggregated to form regional and global indices, where each country's weight in the composite index is proportional to its weight in the total investable universe.

Table 3 presents the results of the estimation. Table 4 shows the value of the *t* and *W*-test statistics. There are some findings that deserve commenting. Firstly, the probability for a return to be assigned to the hectic regime ranges, for almost all pairs, between 0.1 and 0.2, numbers that are in line with those obtained by KIM and FINGER (2000). Secondly, the correlations from two different regimes are close to each other, only with (North America, Turkey) and (North America, Hungary) behaving slightly different. Therefore, even without a formal test, the constant correlation hypothesis is hard to reject. Thirdly, the correlation is considerably higher among Western – Eastern Europe

Table 3: Estimated Parameters for Mixture of Bivariate Normal Distribution Model

Panel A: Estimated parameters for the pair (Europe, *)**

***	μ_{q1}	μ_{q2}	σ_{q1}	σ_{q2}	ρ_q	μ_{h1}	μ_{h2}	σ_{h1}	σ_{h2}	ρ_h	*p*	*LogLik*
East.Europe	0.01%	0.10%	1.05%	1.42%	**0.43**	–0.27%	–0.44%	2.22%	2.77%	**0.42**	0.84	5800
Hungary	–0.02%	–0.06%	1.04%	1.43%	**0.46**	–0.09%	0.30%	2.29%	3.35%	**0.44**	0.85	5768
Czech Rep.	–0.01%	0.06%	1.05%	1.46%	**0.45**	–0.17%	–0.09%	2.18%	2.74%	**0.37**	0.83	5769
Poland	0.00%	0.02%	1.05%	1.74%	**0.34**	–0.24%	–0.19%	2.29%	3.11%	**0.34**	0.85	5589
Russia	–0.03%	0.11%	1.36%	2.43%	**0.36**	–0.06%	0.31%	0.98%	5.85%	**0.29**	0.84	5110
Turkey	–0.02%	–0.11%	1.34%	2.90%	**0.21**	–0.09%	0.84%	1.06%	8.23%	**0.26**	0.86	4879

Panel B: Estimated parameters for the pair (North America, *)**

***	μ_{q1}	μ_{q2}	σ_{q1}	σ_{q2}	ρ_q	μ_{h1}	μ_{h2}	σ_{h1}	σ_{h2}	ρ_h	*p*	*LogLik*
East.Europe	–0.06%	0.07%	1.18%	1.43%	**0.19**	0.23%	–0.35%	2.30%	3.00%	**0.2**	0.87	5657
Hungary	–0.06%	–0.05%	1.19%	1.44%	**0.23**	0.26%	0.30%	2.27%	3.65%	**0.11**	0.88	5613
Czech Rep.	–0.05%	0.06%	1.16%	1.47%	**0.13**	0.15%	–0.10%	2.24%	2.87%	**0.12**	0.85	5617
Poland	–0.09%	–0.03%	1.18%	1.75%	**0.15**	0.43%	0.17%	2.25%	3.24%	**0.12**	0.87	5474
Russia	–0.07%	0.11%	1.21%	2.58%	**0.21**	0.29%	0.42%	2.21%	6.20%	**0.2**	0.88	5036
Turkey	–0.06%	–0.10%	1.19%	2.88%	**0.01**	0.15%	0.61%	2.10%	7.84%	**0.15**	0.85	4823
Europe	0.01%	0.05%	1.00%	1.13%	**0.41**	–0.16%	0.10%	2.31%	2.21%	**0.46**	0.84	5001

Table 4: Results of the Tests for Constant Correlation

Panel A: Statistics for the pair (Europe, *)**

***	Variance(ρ_q)	Variance(ρ_h)	Covariance(ρ_h, ρ_q)	t	Prob(t)	W	Prob(W)
Eastern Europe	0.12%	0.50%	–0.05%	0.14	0.8887	0.02	0.8875
Hungary	0.11%	0.47%	–0.04%	0.27	0.7872	0.08	0.7773
Czech Rep.	0.13%	0.52%	–0.06%	0.93	0.3526	0.94	0.3323
Poland	0.13%	0.62%	–0.06%	0.04	0.9681	0.01	0.9203
Russia	0.13%	0.74%	–0.04%	0.71	0.4779	0.52	0.4708
Turkey	0.13%	0.86%	–0.07%	–0.46	0.6456	0.23	0.6315

Panel B: Statistics for the pair (North America, *)**

***	Variance(ρ_q)	Variance(ρ_h)	Covariance(ρ_h, ρ_q)	t	Prob(t)	W	Prob(W)
Eastern Europe	0.16%	0.92%	–0.09%	–0.08	0.9363	0.01	0.9203
Hungary	0.09%	0.40%	–0.07%	1.63	0.1034	3.02	0.0822
Czech	0.18%	0.86%	–0.11%	0.09	0.9283	0.01	0.9203
Poland	0.16%	0.99%	–0.10%	0.28	0.7795	0.09	0.7642
Russia	0.14%	0.88%	–0.07%	0.13	0.8966	0.02	0.8875
Turkey	0.19%	0.70%	–0.06%	–1.49	0.1365	2.35	0.1253

pairs than North American – Eastern European ones, which is explainable by the larger degree of economic integration within Europe. Finally, using a formal test, only correlation between MSCI Turkey – MSCI North America and MSCI Hungary – MSCI North America have a tendency to change during volatile periods, but the null hypothesis could not be rejected with a high probability.

A probability of 0.2 for a return to come from the hectic regime should be interpreted with care. It is not the case that every two days in ten one will notice an unusual high/low return. Even it is drawn from the "hectic" distribution, there is a relative high probability that the return will be in a moderate range (a normal distribution with higher standard deviation still keeps a great deal of its "mass" close to the average). Given this observations, one should make a clear distinction between hectic/normal regimes and hectic/normal returns.

At the beginning of the paper, there were cited a few sources supporting the idea of structural breaks during stress markets, and these structural breaks include the dependence pattern. Although our results contribute against this thesis, it should be mentioned that the joint market movements during extreme market downturns (like the one on October 19, 1987) remain outside the scope of our analysis. The immediate reason, related to the model used in this paper, is that those extreme returns are "poured" in a distribution with more moderate ones, which attenuate the overall effect. However, there are other drawbacks that impede the research towards a clearer verdict. The extreme events are through their nature very rare, making impossible a sound test using econometric tools. Moreover, looking closely to an extreme event in our sample (September 11, 2001), other reasons become apparent. There is a liquidity effect, as turnover in selected markets was very low, and this questions the reliability of prices. In addition, it is very common to take emerging market exposure through, for example, exchange-traded emerging market mutual funds, available to investors in major financial centers and subject to the market mood and regulations of those particular locations. This fact propagates a shock from a market to another reinforcing the dependence, but,

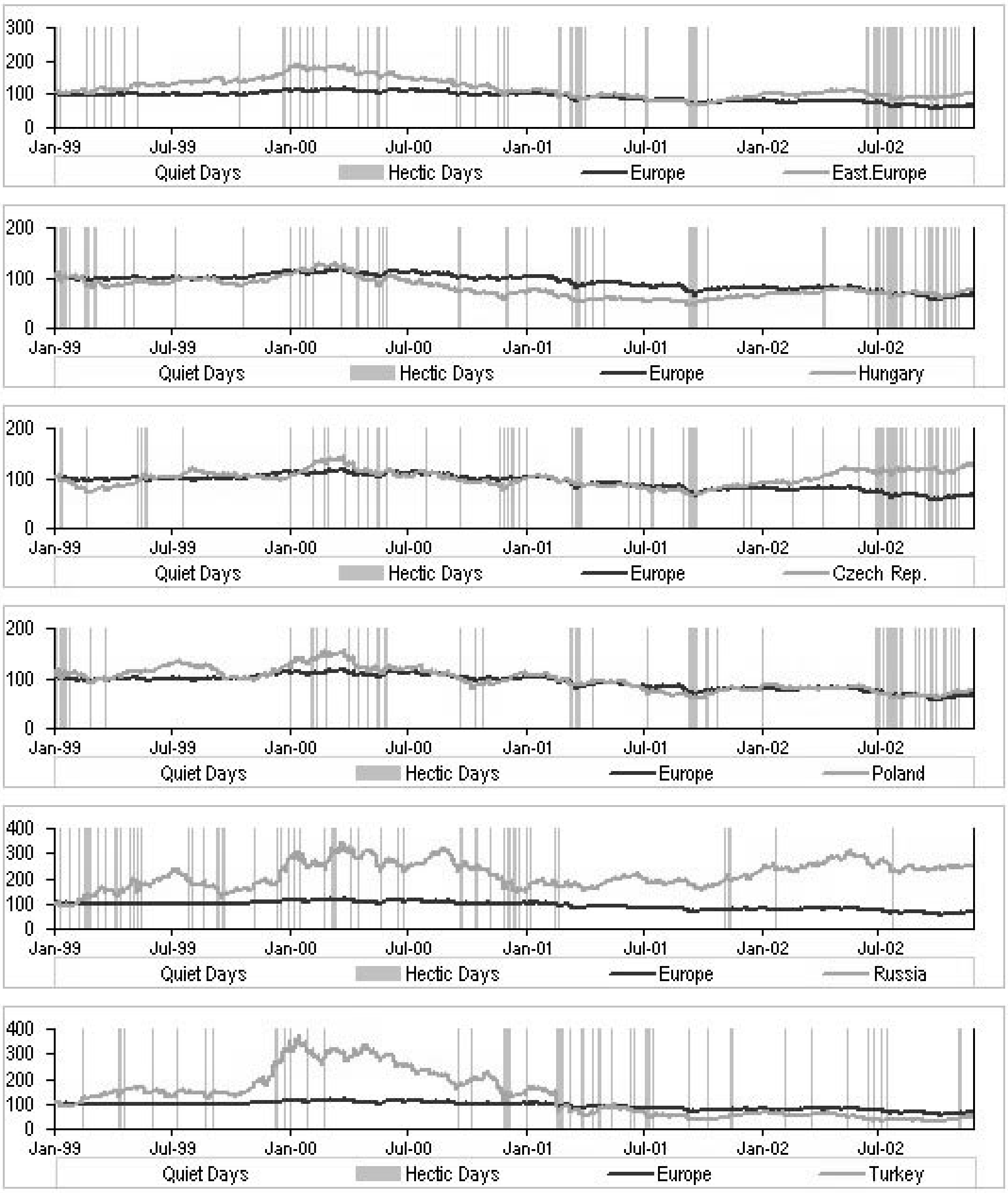
dependence, but, in the same time, it is ignored in the model, hindering the obtained results.
Additional insights are offered by the way the model discriminate between "quiet" vs. "hectic" days and "quiet" vs. "hectic" returns. 1 and 2 plot the accumulated performance of the index pairs under research and also shows the days which are likely to be considered "hectic". The first observation is that, although the model does not make efforts to take into consideration the volatility clustering, the "hectic" days are grouped in certain intervals. Secondly, these intervals coincide for Hungary, Czech Republic, Poland and their aggregate index, MSCI Eastern Europe.
Tables 3 and Table 4 shows the returns which are likely to be considered "quiet" and, respectively, "hectic". The chart confirms that the model correctly identifies the "hectic" returns as those points at the edge of the cloud. However, the labels "quiet" and "hectic" are attached in a probabilistic way ("this return is likely to come from a quiet day") in order to avoid truncation, which was shown at the beginning of the paper to produce biased estimates.

5. Conclusions

Testing stability of correlation during hectic financial markets is far from a trivial issue. For specifying a statistically-sound test, a complete specification of the return generating process is needed. Using a mixture of normal distribution framework, the conclusion tends to support the idea of unchanging correlation, which is in disagreement with what current market practice assumes.

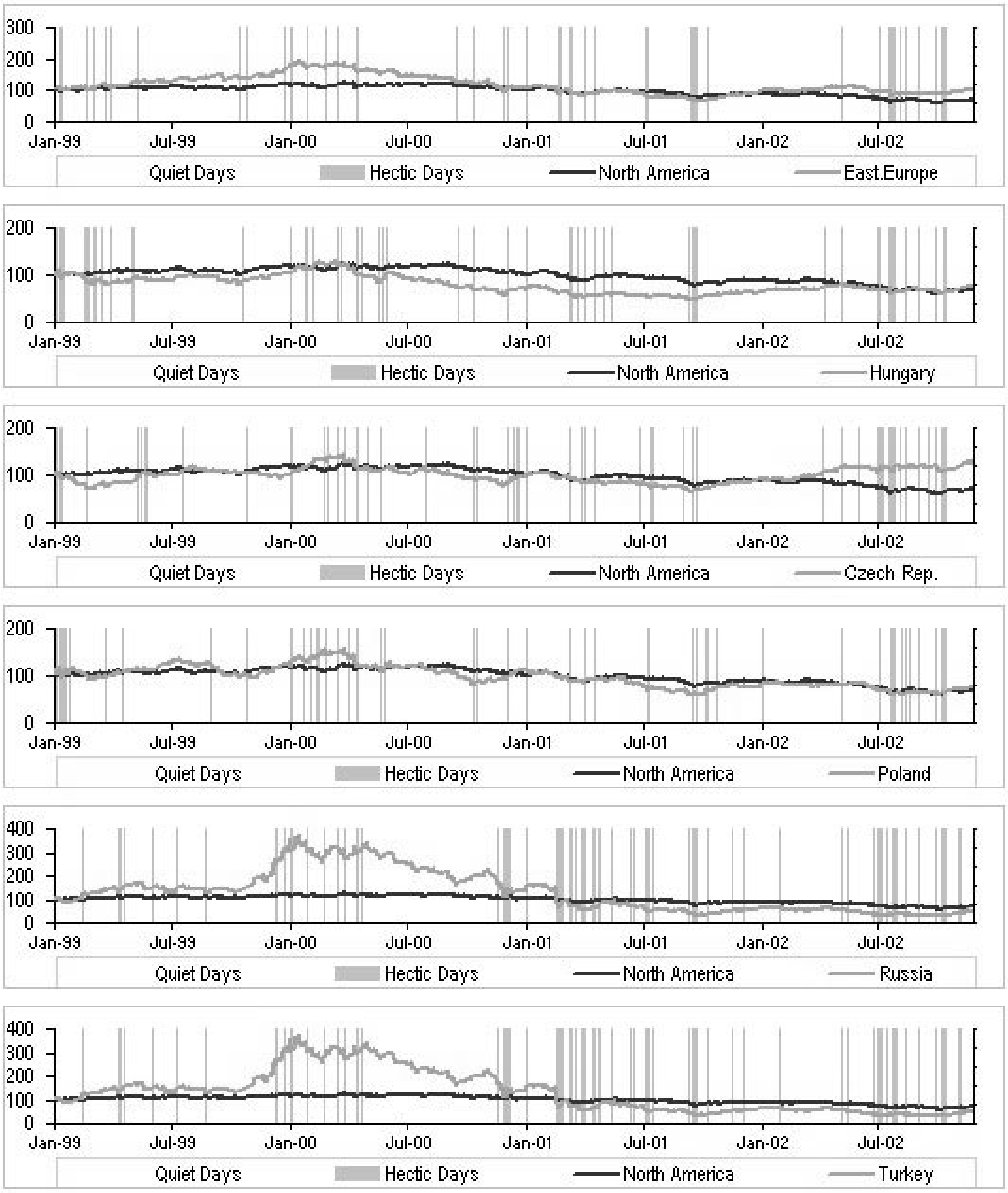
Using EVT, similar tests could be constructed if the underlying data generating process is a multivariate GARCH or normal mixture, as the extreme value behavior is known for both these processes. Boyer et al. (1999) suggest a Markov-chain framework to identify the regime switching and testing correlation changes between regimes. Extending the framework, there are other measures of dependence (rank correlation, copulas) that are known to better describe the joint moves when the Gaussian assumption is dropped and whose properties could be potential targets for further research.

Figure 1: MSCI Indexes: Performance During Quiet/Hectic Periods



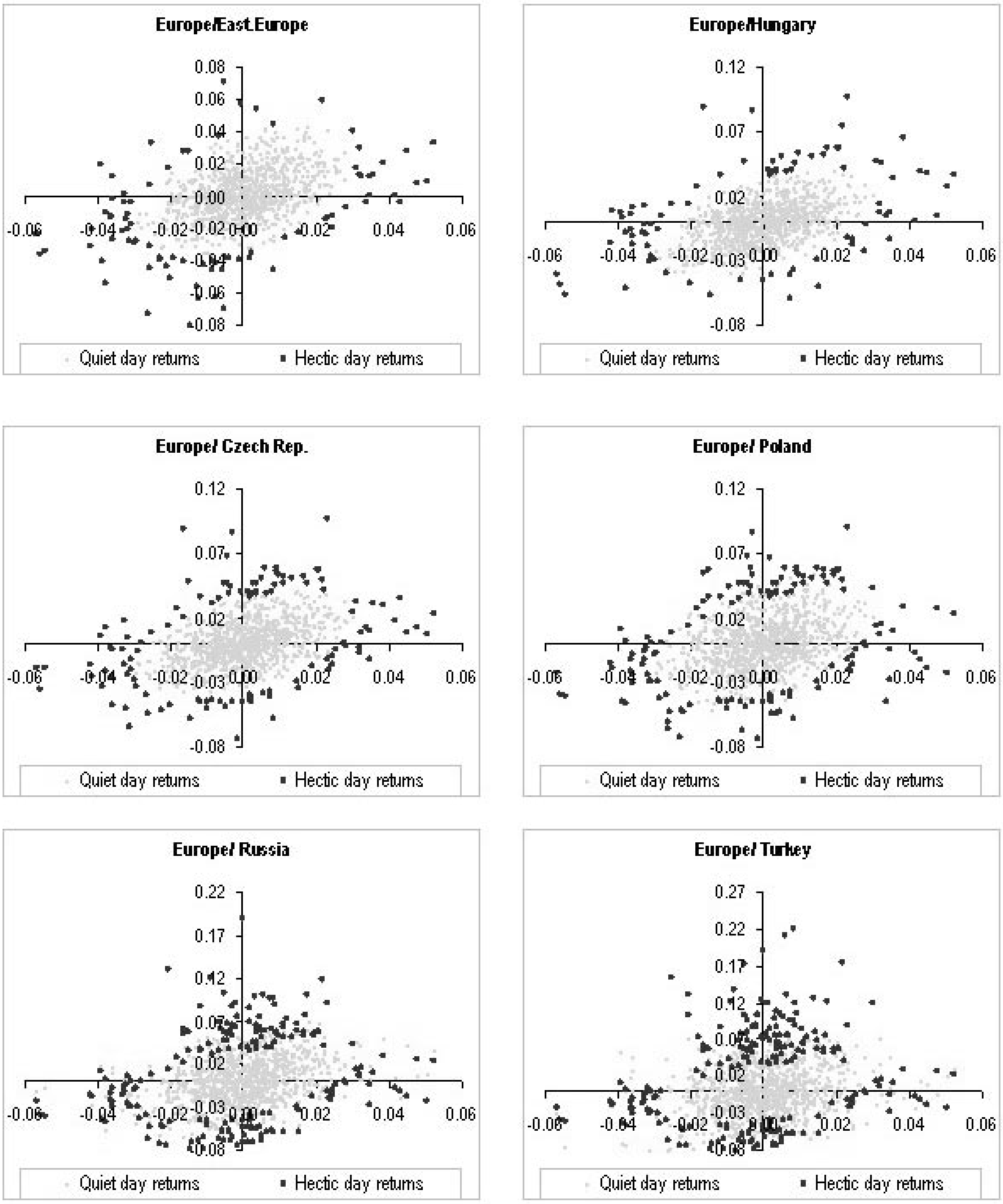
The performance is described by the accumulated returns between 1 Jan. 1999 and 30 Nov. 2002, 1 Jan. 1999 = 100. A particular day was labeled quiet if the conditional probability of its return to come from the quiet distribution given the return value was larger than 0.5.

Figure 2: MSCI Indexes: Performance During Quiet/Hectic Periods



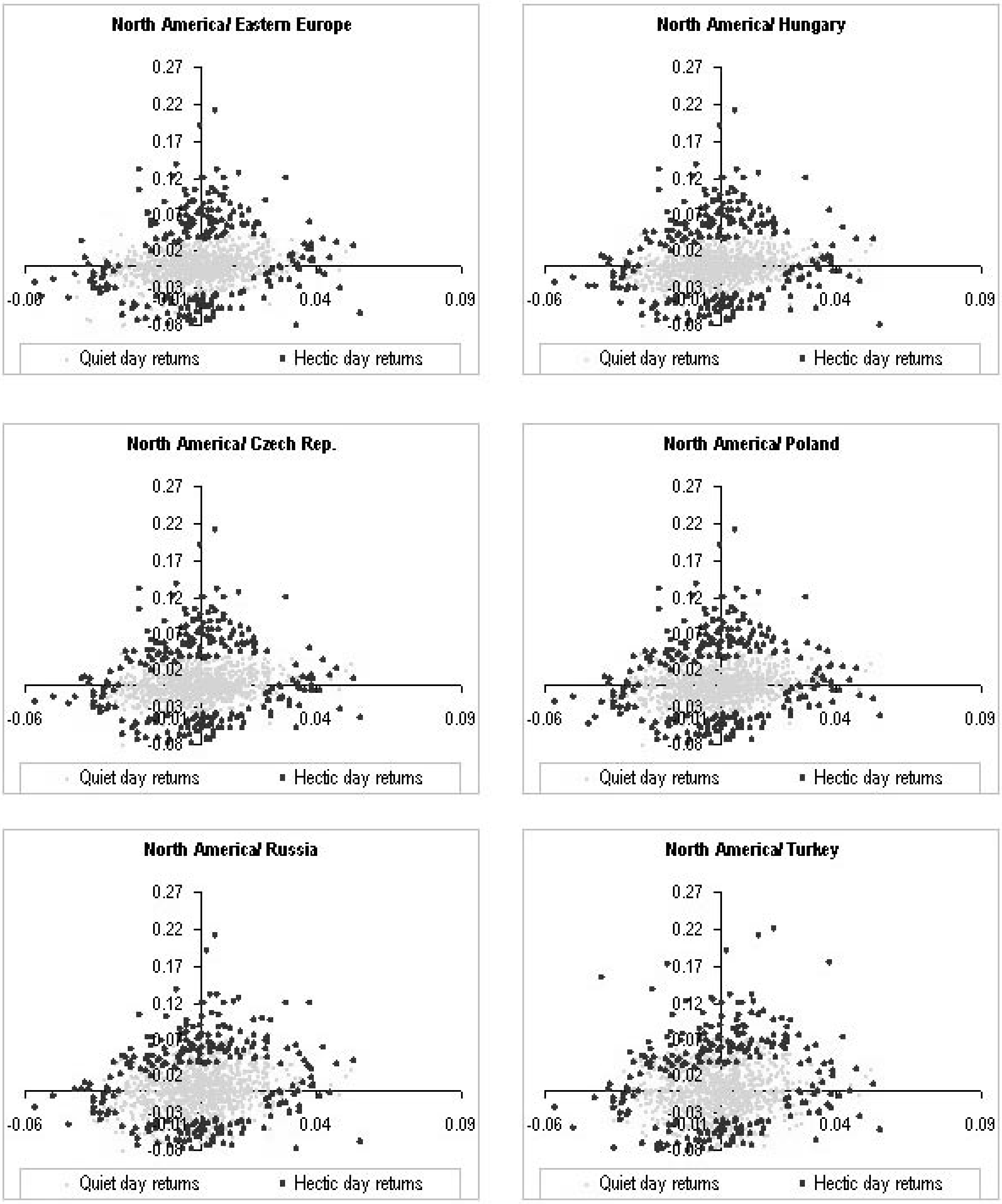
The performance is described by the accumulated returns between 1 Jan. 1999 and 30 Nov. 2002, 1 Jan. 1999 = 100. A particular day was labeled quiet if the conditional probability of its return to come from the quiet distribution given the return value was larger than 0.5.

Figure 3: MSCI Indexes: Return/Return Plot Conditioned on Hectic/Quiet Days. Europe



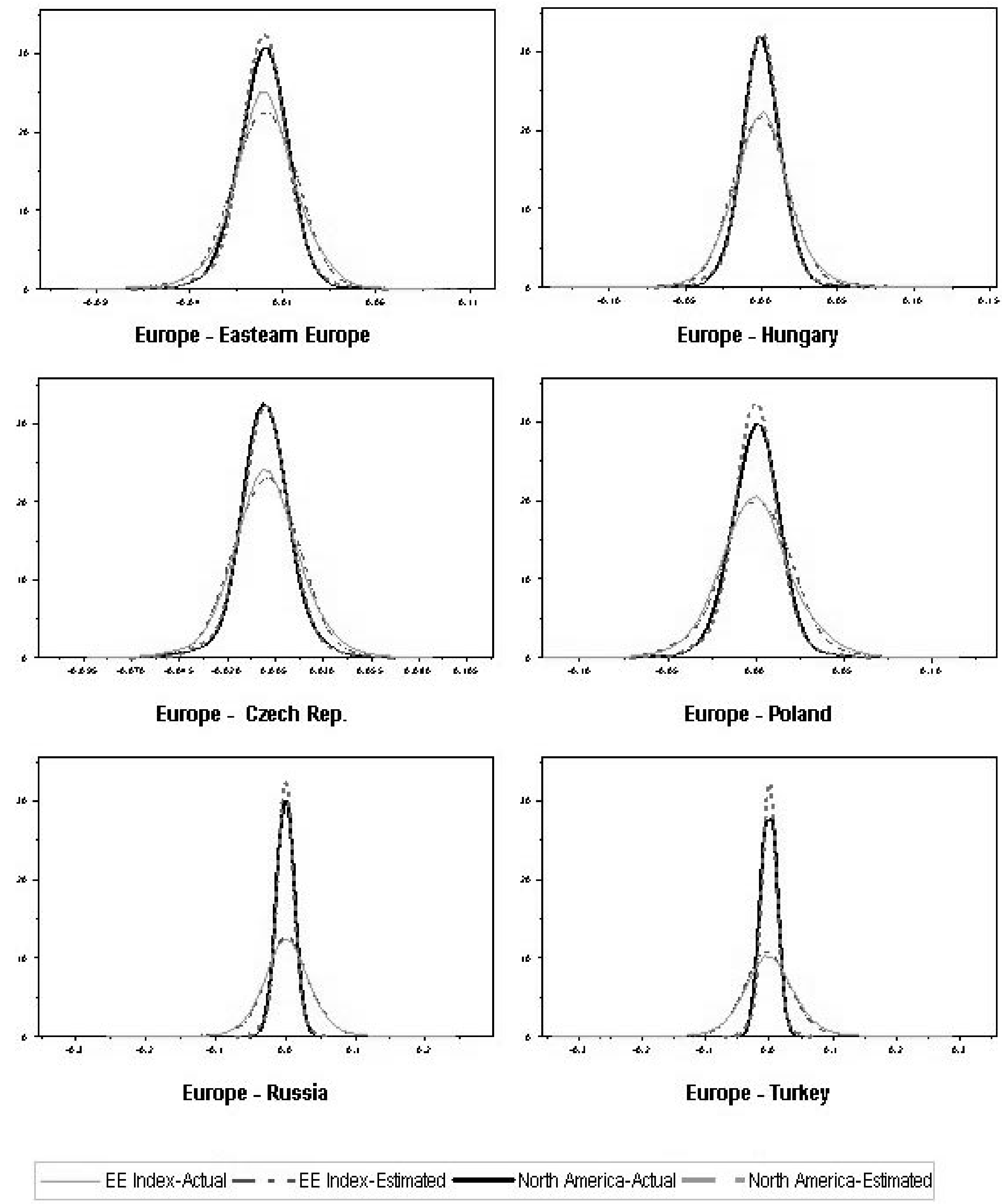
The chart presents the return/return plot of the pairs MSCI Europe – MSCI Eastern Europe / Hungary / Czech Rep. / Poland / Russia / Turkey between 1 Jan. 1999 and 30 Nov. 2002, conditioned on quiet day/hectic day. A particular day was labeled quiet if the conditional probability of its return to come from the quiet distribution given the return value was larger than 0.5.

Figure 4: MSCI Indexes: Return/Return Plot Conditioned on Hectic/Quiet Days. North America



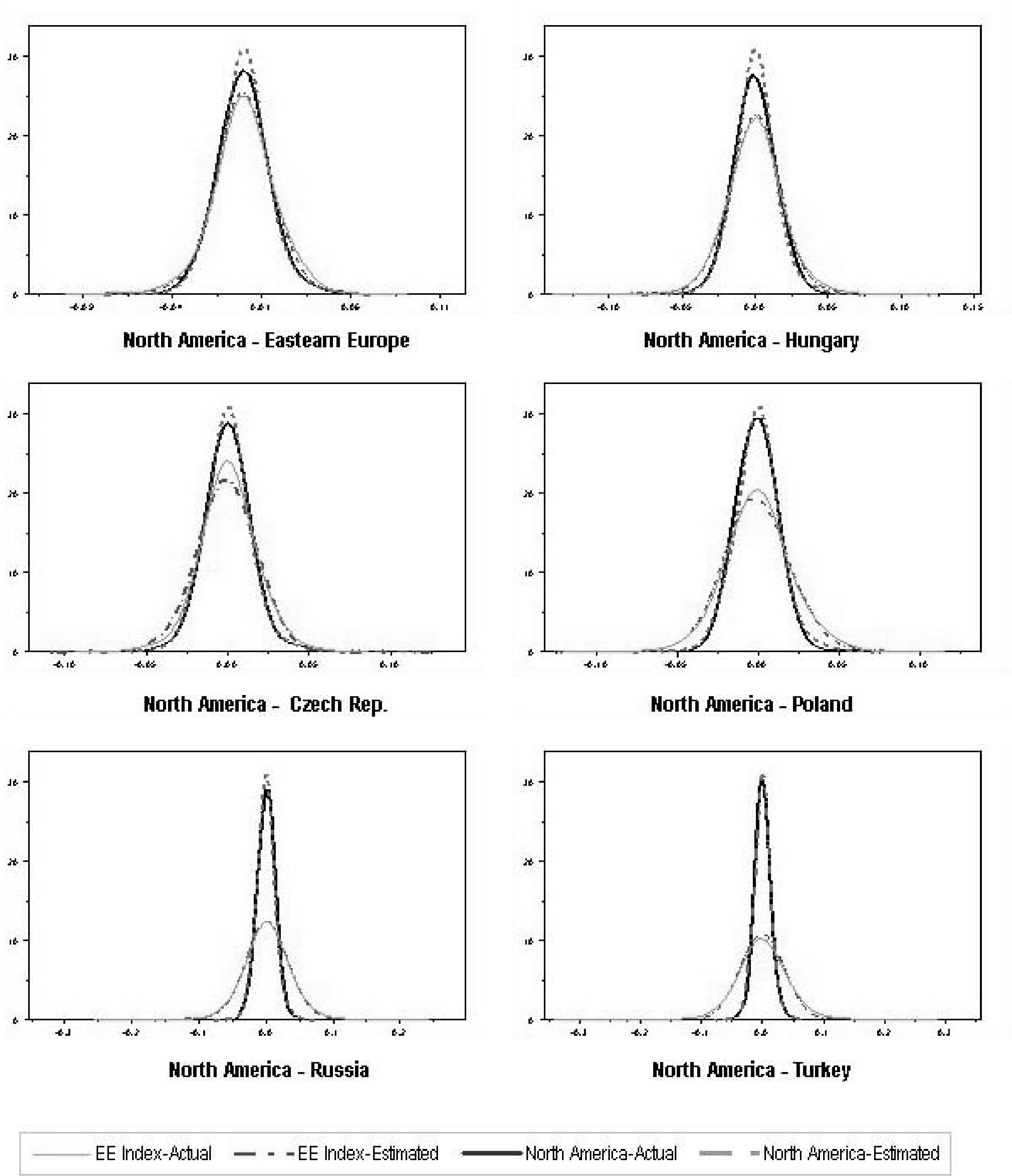
The chart presents the return/return plot of the pairs MSCI North America – MSCI Eastern Europe / Hungary / Czech Rep. / Poland / Russia / Turkey between 1 Jan. 1999 and 30 Nov. 2002, conditioned on quiet day/hectic day. A particular day was labeled quiet if the conditional probability of its return to come from the quiet distribution given the return value was larger than 0.5.

Figure 5: Density Plot: Actual Data vs. Estimated Model. Europe - East European Index



The charts present the density plot of the actual data (empirical kernel) and the density determined by the estimated parameters. A bivariate normal distribution is plotted as a pair of univariate probability distribution functions.

Figure 6: Density Plot: Actual Data vs. Estimated Model. North America - East European Index



The charts present the density plot of the actual data (empirical kernel) and the density determined by the estimated parameters. A bivariate normal distribution is plotted as a pair of univariate probability distribution functions.

APPENDIX A:

A Primer on EM Algorithm

The roots of the EM algorithm could be traced well in the past, but the seminal reference that formalized EM and provided a proof of convergence is DEMPSTER et al. (1977). A more recent book devoted entirely to EM and applications is MCLACHLAN and KRISHNAN (1997). The algorithm is applied for the maximum likelihood estimation of the so-called "missing data problems" and is remarkable because of the simplicity and generality of the associated theory, as well as because of the variety of the examples that falls under its reach.

The missing data problem refers to the situation when, given a sample $\boldsymbol{x}$ with a probability distribution function $f(\boldsymbol{x};\boldsymbol{\Psi})$ and a log likelihood function $l(\boldsymbol{x};\boldsymbol{\Psi}) = \ln f(\boldsymbol{x};\boldsymbol{\Psi})$, a straightforward estimation of parameter vector $\boldsymbol{\Psi}$ is difficult to perform. However, if the sample $\boldsymbol{x}$ is augmented with some missing data $\boldsymbol{z}$, the estimation becomes trivial. Let $\boldsymbol{y}=(\boldsymbol{x}, \boldsymbol{z})$ be the complete data distributed with the probability distribution function $g_{\boldsymbol{\Psi}}(\boldsymbol{y})$ and the new log likelihood function $l_c(\boldsymbol{y};\boldsymbol{\Psi}) = \ln g(\boldsymbol{y};\boldsymbol{\Psi})$.

EM algorithm consists in two steps. The E-step (expectation step) starts by considering an arbitrary parameter vector $\boldsymbol{\Psi}^{(0)}$. As part of the data is missing, the complete-data log likelihood function $l_c(\boldsymbol{y};\boldsymbol{\Psi})$, given $\boldsymbol{\Psi}^{(0)}$, is a random variable. However, we can make it more "precise" by conditioning it with the known data $\boldsymbol{x}$ and approximate it through its expectation:

$$E_{\Psi^{(0)}}\left[\ln g(y)\middle|x\right] \tag{23}$$

The M-step (maximization step) continues by maximizing the previously computed quantity over the parameter space and updating the old parameter vector with the result of this optimization problem:

$$\Psi^{(1)} = \arg\max_{\Psi} E_{\Psi^{(0)}}\left[\ln g(y)\middle|x\right] \tag{24}$$

The procedure is repeated as long as the differences between old and new parameters are arbitrary small.

The parameter vector $\boldsymbol{\Psi}^{(k)}$, computed through the iterative procedure explained above, converges, under some regularity conditions, to the maximum likelihood parameters of the incomplete problem and shares their properties: efficiency, unbiasedness and consistency. A full proof is beyond the scope of this paper and we direct the interested reader to the references at the beginning of this section.

It should be mentioned that the solution of the EM algorithm could converge to a local maximum of the likelihood function. In general, the problem is fixed by careful initial parameterization.

The standard errors of EM-computed MLE estimates are determined in classical way using Fisher's information matrix. Let us introduce the following notations:

$$S(x,\Psi) = \frac{\partial l(\Psi)}{\partial \Psi} \tag{25}$$

be the incomplete-data log likelihood function gradient vector (or score statistic);

$$I(\Psi,x) = \frac{\partial^2 l(\Psi)}{\partial \Psi \partial \Psi^T} \tag{26}$$

be the matrix of the negative of the second-order partial derivatives of the incomplete-data log likelihood function with respect to the elements of $\boldsymbol{\Psi}$. Assuming all regularity conditions, the expected Fisher information matrix is:

$$\begin{aligned} \Im(\Psi,x) &= E_{\Psi}\left[S(x,\Psi)S^T(x,\Psi)\right] \\ &= -E_{\Psi}\left[I(\Psi,x)\right] \end{aligned} \tag{27}$$

The ML estimates converge asymptotically to the following distribution:

$$\hat{\Psi} \sim N(\Psi_0, \frac{1}{n}\Im^{-1}) \tag{28}$$

where $\boldsymbol{\Psi}_0$ denotes the true parameter vector.

Although the computational challenges associated with this procedure are often the reasons that instigated the use of the EM algorithm in the first instance, it is more tractable to compute the value of the log likelihood function second derivative at $\boldsymbol{\Psi}$ than to find its global maximum.

ENDNOTES

[1] The different exposure to technology stocks is an often cited explanation for this behavior in the last decade.

[2] Cited from KIM and FINGER (2000). Original remark appeared in GREENSPAN (1999).

REFERENCES

BANK FOR INTERNATIONAL SETTLEMENTS, Committee on the Global Financial System (1999): “A Review of Financial Market Events in Autumn 1998”, Basel.

BOLLERSLEV, T. (1990): “Modeling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Model”, Review of Economics and Statistics 72, pp. 498–505.

BOOKSTABER, R. (1997): “Global Risk Management: Are We Missing the Point?”, Journal of Portfolio Management 23:3, pp. 102–107.

BOYER, B. H., M. S. GIBSON and M. LORETAN (1999): “Pitfalls in Tests for Changes in Correlations”, International Finance Papers No. 597, Board of Governors of the Federal Reserve System.

BREUER, T. and G. KRENN (1999): “Stress Testing”. Guidelines for Market Risk Vol. 5, Oesterreichische National Bank.

CAMPBELL, J. Y., A. W. LO and A. MACKINLAY (1997): “The Econometrics of Financial Markets”, Princeton University Press, Princeton.

DEMPSTER, A. P., N. M. LAIRD and D.B. RUBIN (1977): “Maximum Likelihood from Incomplete Data vie EM Algorithm”, Journal of the Royal Statistical Society 39, pp. 1–38.

EMBRECHTS, P., A. MCNEIL and D. STRAUMANN (1998): “Correlation and Dependency in Risk Management: Properties and Pitfalls”, Working Paper, ETH, Zurich.

ENGLE, R. (1984): “Wald, Likelihood Ratio, and Lagrange Multiplier Tests in Econometrics”, in “Handbook of Econometrics" (eds. Z. Griliches and M. D. Intriligator), Elsevier Science.

FERGUSON, S. T. (1996): “A Course in Large Sample Theory”, Chapman&Hall, London.

GODFREY, L. G. (1988): “Misspecification Tests in Econometrics”, Cambridge University Press.

GREENSPAN, A. (1999): “Measuring Financial Risk in the Twenty-first Century”, Conference sponsored by the Office of the Comptroller of the Currency, Washington, D.C., October 14.

HAMILTON, J. (1991): “A quasi-Bayesian Approach to Estimating Parameters for Mixtures of Normal Distributions”, Journal of Business and Economic Statistics, Vol. 9, No. 1.

J.P. MORGAN and Co. (2000): “Annual Report 1999”, New York.

KIM, J. and C. C. FINGER (2000): “A Stress Test to Incorporate Correlation Breakdown”, Journal of Risk 2, No. 3.

KON, S. J. (1984): “Models of Stock Returns. A Comparison”, Journal of Finance 39, 1, pp. 147–165.

LAUBSCH, A. J. (1999): “Risk Management: A Practical Guide”, RiskMetrics Group, New York.

LONGIN, F. and B. SOLNIK (2001): “Correlation Structure of International Equity Markets during Extremely Volatile Periods”, Journal of Finance, Vol. LVI, No. 2.

LORETAN, M. and W. B. ENGLISH (2000): “Evaluating 'Correlation Breakdowns' during Periods of Market Volatility”, Board of Governors of the Federal Reserve System, International Finance Discussion Papers.

MCLACHLAN, G. and K. BASFORD (1998): “Mixture Models: Inference and Applications to Clustering”, Marcel Dekker Inc., New York and Basle.

MCLACHLAN, G. and T. KRISHNAN (1997): “The EM Algorithm and Extensions”, Wiley series in probability and statistics, John Wiley & Sons.

MCNEIL, A. and T. SALADIN (1997): “The Peaks over Thresholds Method for Estimating High Quantiles of Loss Distributions”, Working Paper, ETH Zurich.

SILVEY, S. D. (1995): “Statistical Inference”, Chapman & Hall.

TSE, Y. K. (2000): “A Test for Constant Correlations in a Multivariate GARCH Model”, Journal of Econometrics 98, pp. 107–127.

WHITE, H. (1994): “Estimation, Inference and Specification Analysis”, Cambridge University Press.

WOLFE, J. H. (1971): “A Monte Carlo study of the sampling distribution of the likelihood ratio for mixtures of multinormal distributions” (NPTRL-STB-72-2), San Diego: Naval Personnel and Training Research Laboratory.

ZANGARI, P. (1996): “An Improved Methodology for Measuring VaR”, RiskMetrics Monitor 2nd Quarter, pp. 7–25.