HAL open science

ESTIMATING TIME-VARYING VARIANCES AND COVARIANCES VIA NEAREST NEIGHBOUR MULTIVARIATE PREDICTIONS: APPLICATIONS TO THE NYSE AND THE MADRID STOCK EXCHANGE INDEX

Eduardo Acosta-González, Fernando Fernández-Rodríguez, Julián Andrada-Félix

To cite this version:

Eduardo Acosta-González, Fernando Fernández-Rodríguez, Julián Andrada-Félix. ESTIMATING TIME-VARYING VARIANCES AND COVARIANCES VIA NEAREST NEIGHBOUR MULTIVARIATE PREDICTIONS: APPLICATIONS TO THE NYSE AND THE MADRID STOCK EXCHANGE INDEX. Applied Economics, 2009, 41 (26), pp.3437-3445. 10.1080/00036840701439371 . hal-00582118

HAL Id: hal-00582118

https://hal.science/hal-00582118

Submitted on 1 Apr 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

For Peer

ESTIMATING TIME-VARYING VARIANCES AND COVARIANCES VIA NEAREST NEIGHBOUR MULTIVARIATE PREDICTIONS: APPLICATIONS TO THE NYSE AND THE MADRID STOCK EXCHANGE INDEX

Journal:	*Applied Economics*
Manuscript ID:	APE-06-0258.R1
Journal Selection:	Applied Economics
JEL Code:	C13 - Estimation < C1 - Econometric and Statistical Methods: General < C - Mathematical and Quantitative Methods, G15 - International Financial Markets < G1 - General Financial Markets < G - Financial Economics, C14 - Semiparametric and Nonparametric Methods < C1 - Econometric and Statistical Methods: General < C - Mathematical and Quantitative Methods
Keywords:	Nonparametric estimation, Stock market indexes, Variance and covariance prediction

scholarONE™
Manuscript Central

Submitted Manuscript

ESTIMATING TIME-VARYING VARIANCES AND COVARIANCES VIA NEAREST NEIGHBOUR MULTIVARIATE PREDICTIONS: APPLICATIONS TO THE NYSE AND THE MADRID STOCK EXCHANGE INDEX

EDUARDO ACOSTA-GONZALEZ[a], JULIAN ANDRADA-FELIX[a] and FERNANDO FERNANDEZ-RODRIGUEZ[a*]

[a] Department of Quantitative Methods, University of Las Palmas de G.C. Spain
Campus de Tafira. Facultad de Ciencias Económicas y EE.
35017 Las Palmas de Gran Canaria. Spain

[*] E-mail corresponding author: ffernandez@dmc.ulpgc.es

RUNNING TITLE:
ESTIMATING VARIANCES AND COVARIANCES VIA NEAREST NEIGHBOUR MULTIVARIATE PREDICTIONS

Abstract

In this paper we present a technique to obtain the time-varying covariance matrix for several time series for nearest neighbour predictors. To illustrate the use of this technique, we analyse the time-varying variances and correlations between the daily returns on two equity stock market indexes, the New York Stock Exchange (NYSE) and the Madrid Stock Exchange Index (MSEI).

Keywords: Nonparametric estimation; Stock Market Indexes; Time-varying variance and covariance prediction.

JEL classification: C13; C14; G15

For Peer Review

I. Introduction

Cross-country covariances of asset returns play a key role in international risk management and portfolio diversification. So, potential benefits from international diversification from domestic investors are significant if cross-country covariances are small. However, investors may not be able to diversify away much domestic risk if the cross-country covariances are large.

Since the market crash of 1987, a wide range of literature has emerged investigating the contagion phenomenon and the fundamental factors which are affecting the cross-country stock return correlations. A determining paper was Roll (1989) which presents a survey on the link of international stock markets, signalling that international correlations of returns increased dramatically during crash periods. Several seminal papers focused their attention on analyzing how news from one international market influences other markets` volatility processes (see Engle et al., 1990; Hamao et al., 1990 and King and Wadhawani, 1990, among others).

Other papers have established that correlations across the indexes, like their unconditional sample counterparts, are typically nonnegative (e.g. Koch and Koch, 1991 and Engle and Susmel, 1993). This empirical fact has also been corroborated by papers such as King, Sentana and Wadhwani (1994) or Darbar and Deb (1997). More recently, literature asks if the spread contagion is due to economic and/or financial links between the base country and secondary country, determining which channels of financial

For Peer Review

Submitted Manuscript

contagion are the most significant in transmitting crises between countries (see Collins and Gavron, 2004 and other papers enclosed).

The methodology employed for estimating variances and covariances of asset returns are also crucial. As Darbar and Deb (1997) point out, if the correlations are time varying, examining sample correlations may be inadequate because the time path of conditional correlations reveals information that cannot be obtained by simply computing the sample correlation statistic. The sample correlation coefficient may also be a misleading measure of independence between two markets. For instance, a large value of a sample correlation may be driven by an episode of unusually high conditional correlation.

ARCH family models (see Engle, 1982) have been one of the most popular frameworks for modelling time-varying variances of asset returns. The ARMA-ARCH model rests on the assumption that the time series are linear in mean but nonlinear in variance. Nevertheless, several nonlinear and no parametrical alternative models have been considered recently. If we consider the hypothesis that the financial series have a nonlinear structure in mean, a very well-known model for predicting the returns are the Nearest Neighbour (NN from now on) prediction method.. In this paper we propose a way of estimating time-varying variances and covariances when the nonparametric NN multivariate estimator method is used to predict returns.

In ARCH approach literature it is very common to use a multivariate framework for modelling time-varying covariances among international financial markets. In this case, there is a high number of parameters which has to be estimated. With N series, N(N+1)/2 processes, which characterize

For Peer Review

the variance covariance matrix, need to be estimated. So, the imposition of restrictions is necessary to reduce the number of parameters of the model to be estimated (see Engle and Susmel, 1993).

The aim of the paper is to propose a different approach for estimating the time-varying covariance matrix using nonparametric local linear models. In our methodology the volatility of an observation is associated with the risk of an NN in-sample prediction. In some extensions, the NN procedure involves regression by ordinary least squares (OLS) of the future evolution of the k nearest neighbours chosen on their preceding histories. However, OLS is efficient only when the contemporaneous correlation of the OLS residuals do not exist, otherwise the estimators are inefficient. In this paper we use the seemingly unrelated regression (SUR) estimation methods to estimate the model (Zellner, 1962). In this context we call this technique the locally-adjusted seemingly unrelated regression (LASUR) model. This will not only be able to obtain the predictions of the series but also the time-varying variances and covariances. So, future volatility and co-variability, one-step ahead of two series, is associated with the prediction error covariance matrix of an NN-LASUR prediction. To illustrate the use of this technique we analyse the time-varying variances and correlations between the daily returns on two equity stock market indices, the NYSE and the MSEI in the context of in-sample prediction.

Following Härdle (1990), the nonparametric approach to estimating a regression curve has four main purposes. First, it provides a versatile method of exploring a general relationship between variables. Second, it gives predictions of observations yet to be made without reference to a fixed

For Peer Review

Submitted Manuscript

parametric model. Third, it provides a tool for finding spurious observations by studying the influence of isolated points. Fourth, it constitutes a flexible method of substituting for missing values or interpolating between adjacent X-values.

Besides, we add three additional purposes related to the specific covariances of stocks correlation problems:

- Our non-parametric method is robust to extreme values estimating volatilities and correlations guaranteeing that these measures are not affected by the stock market slumps. Given the dramatic changes in the stocks co-movement patterns around market crashes we used a nonparametric method in order to develop a precise study of stock indexes correlation in these periods as well as guaranteeing that the volatility and correlation will be not affected by the extreme values.
- As opposed to parametric models for estimating the volatility (like ARCH family models), our non-parametric model produces normal standardized residues or standardized residues very close to normality
- Our non-parametric model is not affected by the possible existence of structural changes.

The plan of the paper is as follows. In section 2 the technique is developed. Section 3 describes an illustrative example of the LASUR procedure that uses real data sets, namely the New York Stock Exchange (NYSE) and the Madrid Stock Exchange Index (MSEI). In section 4 we show the conclusions.

For Peer Review

Editorial Office, Dept of Economics, Warwick University, Coventry CV4 7AL, UK

II. Multivariant Nearest Neighbour Predictions with SUR Estimation

Let $\{x_t\}_{t=1}^{T}$ be a time series of scalar observations. We introduce the method of local polynomial regression, where the underlying regression curve is locally and linearly approximated around a vector of d consecutive observations by a linear function, with the objective of getting an in-sample prediction of x_{t+1}. The out-of sample prediction is easily gotten just by taking the sub-sample up to the current t-moment and disregarding the rest of the sample.

The first step of this methodology consists of embedding the series in a d-dimensional space, called the phase space of the time series. Segments of d consecutive elements

$$x_t^d = (x_t, x_{t-1}, ..., x_{t-(d-1)}) \quad t = d, d+1, ..., T \tag{1}$$

of the series, frequently called d-histories, are considered vectors in the space $\square^{d}$, and recorded in the matrix Z^d.

$$Z^d = \begin{bmatrix} x_d & x_{d-1} & \dots & x_1 \\ x_{d+1} & x_d & \dots & x_2 \\ . & . & \dots & . \\ x_t & x_{t-1} & \dots & x_{t-(d-1)} \\ . & . & \dots & . \\ x_{T-1} & x_{T-2} & \dots & x_{T-d} \end{bmatrix} \tag{2}$$

The parameter d is referred to as the embedding dimension, and each row-vector of the matrix Z^d in (2) is called the d-history. For the out-of-sample prediction of x_{t+1} the last row of Z^d is the x_{t-1}^d vector.

For Peer Review

This embedding procedure is well-known in the literature of deterministic chaos, where the Takens (1981) theorem guarantees, under broad conditions, that a hidden deterministic dynamic may be mimicked in the d-dimensional space by the set of d-histories. With independence of the theory of deterministic chaos, nonparametric regression and local polynomial modelling methods propose similar procedures for modeling and predicting time series (see Härdle, 1990, Green and Silverman, 1994 and Hastie et al., 2001).

In order to predict an observation x_{t+1}, the local linear approximation considers the k nearest neighbors vectors $x_{t_\ell}^d$ for $\ell = 1,.....,k$ to the vector x_t^d taken from the row vectors of Z^d in (2). These nearest neighbors $x_{t_1}^d, x_{t_2}^d,....,x_{t_k}^d$ are the first k minima of the function $\left\| x_{t_\ell}^d - x_t^d \right\|$, where $d \le t_\ell \le T-1$ and $\|\cdot\|$ being the Euclidean distance.

Then we can model the data locally as

$$Y_t^d = X_t^d \beta_t^d + \varepsilon_t^d \tag{3}$$

where ε_t^d is a perturbations vector of size k, X_t^d is a matrix of size (k,d+1) formed from the nearest neighbors row-vectors selected from Z^d, plus a column of ones to take into account the constant in the model:

$$X_t^d = \begin{bmatrix} 1 & x_{t_1} & x_{t_1-1} & \dots & x_{t_1-(d-1)} \\ 1 & x_{t_2} & x_{t_2-1} & \dots & x_{t_2-(d-1)} \\ . & . & . & \dots & . \\ 1 & x_{t_k} & x_{t_k-1} & \dots & x_{t_k-(d-1)} \end{bmatrix}, \tag{4}$$

For Peer Review

and Y_t^d is a vector of size k formed from the one step ahead observation of every row-vector of X_t^d, that is

$$Y_t^d = \begin{bmatrix} x_{t_1+1} \\ x_{t_2+1} \\ . \\ x_{t_k+1} \end{bmatrix}. \tag{5}$$

Finally, β_t^d is a vector of local parameters of size d+1 associated with the segments of d consecutive observations x_t^d. This vector can be estimated by OLS:

$$\hat{\beta}_t^{d,(OLS)} = \left((X_t^d)'(X_t^d)\right)^{-1}(X_t^d)'(Y_t^d) \quad . \tag{6}$$

For the predictions of x_{t+1}, the observations of the vector x_t^d are used as explanatory variables, so

$$\hat{x}_{t+1} = x_t^d \hat{\beta}_t^{d,(OLS)}. \tag{7}$$

This procedure may be generalized to the case of several simultaneous time series. In this case, we use the cross information of several time series in forming a predictive vector. There have been several practical attempts in finance, such as Mizrach (1992) or Fernández-Rodríguez et al. (1999) (see also Fernández-Rodríguez et al., 2003 for a review) in the context of prediction.

For Peer Review

A serious problem arises in the generalization of univariate NN techniques to the multivariate case. It is the so called "curse of dimensionality", which was termed by Bellman (1961). Consider a p-dimensional unit hypercube as a neighbour of a target point. Following Hastie et al. (2001), if we wanted to capture a fraction r of the unit volume of the observations, the expected edge length would be $e_p(r) = r^{1/p}$. So, in ten dimensions $e_{10}(0.01) = 0.63$ and $e_{10}(0.1) = 0.8$. Therefore, to capture 1% and 10% of the data to form a local neighbour, we must cover 63% or 80% of the range of each input variable. Consequently, a local neighbourhood in a higher dimension is no longer local. As Fan and Gijbels (1997) point out, if a local neighbourhood contains 10 data points along each axis, then there are 10^d data points in the corresponding d-dimensional neighbourhood. Subsequently, even when d is moderate, a large data set is required to fit a k-NN multivariate model, and such large data sets are often not available in practical situations. Therefore, due to the "curse of dimensionality", k-NN multivariante techniques are hardly useful in the presence of many regressors.

Now we present an alternative approach in dealing with the "curse of dimensionality" which will be summarized as follows: For simplicity, let's consider a two-dimensional time series $\{x_{1,t}, x_{2,t}\}_{t=1}^{T}$, with two matrices of d-histories Z_i^d (for i=1,2) as in (2). In order to select the nearest neighbors we embed each of these series in the space $\mathbb{R}^d \times \mathbb{R}^d$, paying attention to the vectors $(x_{1,t}^d, x_{2,t}^d) \in \mathbb{R}^d \times \mathbb{R}^d$ which represent simultaneous histories in both series. So, we can look for the closest k simultaneous histories

For Peer Review

$(x_{1,t_1}^d, x_{2,t_1}^d), \ldots, (x_{1,t_k}^d, x_{2,t_k}^d)$ to the current one $(x_{1,t}^d, x_{2,t}^d)$. That is, we have to minimize the function

$$\left\| x_{1,t_\ell}^d - x_{1,t}^d \right\| + \left\| x_{2,t_\ell}^d - x_{2,t}^d \right\| \quad where \;\; d \le t_\ell \le T-1 . \tag{8}$$

for every current simultaneous history $(x_{1,t}^d, x_{2,t}^d)$, fixed the time t. Given the two-dimensional time series $\{x_{1,t}, x_{2,t}\}_{t=1}^T$, and a current day t, the out of sample prediction is obtained by minimizing the function (8) when only simultaneous histories $(x_{1,t_l}^d, x_{2,t_l}^d)$ with the restriction $t_l < t < T$ are considered.

Following (3), (4) and (5) we can model the data locally with two equations:

$$\begin{aligned} Y_{1,t}^d &= X_{1,t}^d \;\; \beta_{1,t}^d + \varepsilon_{1,t}^d \\ Y_{2,t}^d &= X_{2,t}^d \;\; \beta_{2,t}^d + \varepsilon_{2,t}^d \end{aligned} . \tag{9}$$

Observe that the equations in (9) are seemingly unrelated although the information for selecting $X_{1,t}^d$ and $X_{2,t}^d$ are obtained in (8) using the simultaneous histories. The purpose of model (9) is to collect the multivariant information of both series and, at the same time, to have a minimum number of parameters to be estimated with the end of avoiding the curse of dimensionality.

The OLS estimation of (9) is efficient only if the contemporaneous correlation of the OLS residuals do not exist, otherwise OLS is inefficient. When the contemporaneous correlation between $\varepsilon_{1,t}^d$ and $\varepsilon_{2,t}^d$ exists, it may

For Peer Review

be more efficient to estimate the two equations jointly, rather than to estimate each one separately using OLS. The appropriate joint estimation technique is known as seemingly unrelated regression estimation (SURE), (Zellner, 1962), that in this context we shall call the locally-adjusted seemingly unrelated regression (LASUR).

We can rewrite (9) as

$$Y = X\beta + \varepsilon \tag{10}$$

where

$$X = \begin{bmatrix} X_{1,t}^{d} & 0 \\ 0 & X_{2,t}^{d} \end{bmatrix}, \qquad Y = \begin{bmatrix} Y_{1,t}^{d} \\ Y_{2,t}^{d} \end{bmatrix}, \qquad \varepsilon = \begin{bmatrix} \varepsilon_{1,t}^{d} \\ \varepsilon_{2,t}^{d} \end{bmatrix}, \qquad \beta = \begin{bmatrix} \beta_{1,t}^{d} \\ \beta_{2,t}^{d} \end{bmatrix}. \tag{11}$$

The assumptions are

$$E[\varepsilon] = 0 \quad \text{and} \quad E\left[\varepsilon\varepsilon'\right] = \Omega = \Sigma \otimes I_k \quad \text{where} \quad \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} \tag{12}$$

and $E\left[\varepsilon_{i,t}^{d}\left(\varepsilon_{j,t}^{d}\right)'\right] = \sigma_{ij} I_k$ for i,j=1,2. The identity matrix I_k is of dimension k and $\otimes$ is the Kronecker product. In this case Ω cannot be written as a scalar multiplied by a 2k-dimensional identity matrix. It follows that the locally adjusted generalized least squares estimator:

$$\hat{\beta}^{d,(LASUR)} = \left[X'\left(\hat{\Sigma}^{-1} \otimes I\right)X\right]^{-1} X'\left(\hat{\Sigma}^{-1} \otimes I\right) Y \tag{13}$$

For Peer Review

is the best linear unbiased estimator for β. $\hat{\beta}^{d,(LASUR)}$ has lower variance than $\hat{\beta}^{d,(OLS)}$ in (6) because it takes into account the contemporaneous correlation between the disturbances in different equations. The covariance matrix of $\hat{\beta}^{d,(LASUR)}$ is consistently estimated by

$$Cov\left(\hat{\beta}^{d,(LASUR)}\right) = \left[X'\left(\hat{\Sigma}^{-1} \otimes I\right)X\right]^{-1} \tag{14}$$

and the one step ahead prediction vector, which is denoted by $\hat{X}_{t+1}$, is obtained as:

$$\hat{X}_{t+1} = \begin{bmatrix} \hat{x}_{1,t+1} \\ \hat{x}_{2,t+1} \end{bmatrix} = P_t' \hat{\beta}^{d,(LASUR)} \tag{15}$$

where

$$P_t' = \begin{bmatrix} 1 & x_{1,t} & x_{1,t-1} \cdots & . & x_{1,t-(d-1)} & 0 & 0 & 0 & . & 0 \\ 0 & 0 & 0 & . & 0 & 1 & x_{2,t} & x_{2,t-1} \cdots & . & x_{2,t-(d-1)} \end{bmatrix}. \tag{16}$$

To obtain $\hat{\Sigma}$ in (13) we first estimate each equation by OLS and obtain the residual vector $\hat{\varepsilon}_i$. Consistent estimates of Σ are given by

$$\hat{\sigma}_{ij} = \frac{\left(\hat{\varepsilon}_{i,t}^{d}\right)'\left(\hat{\varepsilon}_{j,t}^{d}\right)}{k-d} \quad for \quad i,j = 1,2 \tag{17}$$

For Peer Review

that will be unbiased on finite samples when the same d-dimensional vectors are used in each series. Once $\hat{\beta}^{d,(LASUR)}$ is calculated it can be used to form a new estimator for σ_{ij} using (17) that can be used again to obtain new values of $\hat{\beta}^{d,(LASUR)}$ and so on, in an iterative procedure until a convergence criteria is achieved. Besides, if the disturbances, $\varepsilon_{1,t}^{d}\ \varepsilon_{2,t}^{d}$, follow a multivariate normal distribution this estimator is also a maximum likelihood estimator. There are two conditions under which $\hat{\beta}^{d,(OLS)}$ is identical to $\hat{\beta}^{d,(LASUR)}$. The first is when all contemporaneous correlations are zero. In this case, Σ is a diagonal matrix and (6) is then equal to (13). The second condition is when all the explanatory variables in each equation are identical in name, $\left(X_{1,t}^{d} = X_{2,t}^{d}\right)$ what it is no possible in the procedure described in this section.

The aim of the paper is to propose a different approach for estimating the conditional covariance matrix using nonparametric local linear models. In finance, volatility is associated with the risk of a specific prediction. In the ARCH (Engle, 1982) and GARCH (Bollerslev, 1986) approximations, time series volatility is measured by means of the conditional variance of its unexpected component, that is, with the risk of an ARMA-ARCH prediction. In this paper, the volatility of an observation is associated with the risk of an NN in-sample prediction. Future volatility and co-variability, one-step ahead of two series, is associated with the prediction error covariance matrix of a NN-LASUR prediction. Since the one-step ahead prediction error vector is

$$e_{t+1} = X_{t+1} - \hat{X}_{t+1} = P_t^{'}\beta + \varepsilon_{t+1} - P_t^{'}\hat{\beta}^{d,(LASUR)} = P_t^{'}\left(\beta - \hat{\beta}^{d,(LASUR)}\right) + \varepsilon_{t+1}, \qquad (18)$$

For Peer Review

where $X_{t+1} = \begin{bmatrix} x_{1,t+1} \\ x_{2,t+1} \end{bmatrix}, \varepsilon_{t+1} = \begin{bmatrix} \varepsilon_{1,t+1} \\ \varepsilon_{2,t+1} \end{bmatrix}$ and $e_{t+1} = \begin{bmatrix} e_{1,t+1} \\ e_{2,t+1} \end{bmatrix}$,

the prediction error covariance matrix of a NN prediction is defined as:

$$\operatorname{cov}(e_{t+1}) = E\left[\left(P_t^{'}\left(\beta - \hat{\beta}^{d,(LASUR)}\right) + \varepsilon_{t+1}\right)\left(P_t^{'}\left(\beta - \hat{\beta}^{d,(LASUR)}\right) + \varepsilon_{t+1}\right)^{'}\right] =$$

$$P_t^{'}\left(X^{'}\left(\hat{\Sigma}^{-1} \otimes I\right)X\right)^{-1} P_t + \hat{\Sigma}. \qquad (19)$$

Note that (19) is the covariance matrix $\hat{\Sigma}$ plus a positive definite matrix. Thus, the future volatility and co-variability one-step ahead of both series depends not only on the stochastic element in the system $\hat{\Sigma}$, but is also a function of P_t and matrix X. The P_t matrix contains the vectors $x_{1,t}^{d}$ and $x_{2,t}^{d}$ that have been used to find the nearest neighbours which are in the X matrix.

III. Empirical Results for NYSE and Madrid Stock Exchange Indexes

To illustrate the use of this technique we analyse the time-varying variances and correlations between the daily returns on two equity stock market indexes, the New York Stock Exchange (NYSE) and the Madrid Stock Exchange Index (MSEI). As our task is to describe the relations in the past between these two indexes throughout the sample period, therefore the in-sample prediction of variances and correlations are appropriated. Nevertheless, in other cases the out-of sample prediction could be carried out.

For Peer Review

Submitted Manuscript

The data consist of daily close-to-close returns for indexes of the two equity markets from January 2, 1991 to December 30, 2002, which provide a total of 2916 observations. A graph of the data is provided in figure 1.

[Figure 1]

[Table 1]

In table 1 we present a statistical summary based on the first differences of the logarithms of the stock market indexes (multiplied by 100). The daily average return for both series is less than 1 per cent. The standard deviation of the returns varies from 0.94 per cent per day for the NYSE, to 1.27 for the MSEI. In both cases, the returns are skewed to the left and the unconditional distribution of returns has kurtosis coefficients in excess of 3, implying fatter tails than the normal distribution. According to the Jarque-Bera test, we strongly reject the null hypothesis of normality in both cases. Finally, we report the sample correlation coefficient with a value of 0.4056.

Following Casdagli and Eubank (1992), the number of neighbours k=28 and the embedding dimension d=6 have been estimated by minimizing the sample prediction error.

[Table 2]

In table 2 we report statistics for the estimated time-varying covariance and correlation matrices. An important use of the time-varying

For Peer Review

variance is the evaluation of the accuracy of the NN forecast. In standard time series methodology, which uses constant variance ARMA processes, the variance of the forecast error does not depend on the current information set. If the series being forecasted by NN displays time-varying variances, the current information set can indicate the accuracy by which the series can be forecasted by NN. To ensure that the computed value of time-varying variance is reasonable, in figure 2 we present the first differences in logarithms of the NYSE and its prediction intervals for the last 50 observations. The prediction intervals are obtained by

$$\hat{x}_{NYSE,t+1} \pm 2 * SE(e_{NYSE,t+1}) . \tag{20}$$

where SE(·) represents the time-varying standard error.

[Figure 2]

For the full sample, the first differences in logarithms of the NYSE lie outside the prediction intervals at 6.8 %. For the differences in logarithms of the MSEI at 6.5%, it implies that LASUR estimation of the time-varying variance is able to collect correctly the risk associated with the predictions since similar results are expected for any measure of dispersion if the distribution of the data is bell-shaped, and anyway these percentages are clearly smaller than those corresponding to Chebyshev‘s inequality.

The upper part of Figure 3 shows the centred moving averages for 101 observations to eliminate irregular fluctuations from the time-varying variances of both indexes. The behaviour of these two series seems similar, but MSEI variances are around the double of those corresponding to NYSE.

For Peer Review

The sample correlation between these two series is 0.38 indicating a positive relation. In order to investigate causal relations between these series we estimate the next bivariate regressions to apply Granger's causality test

$$\begin{aligned} V_{N,t} &= \alpha_0 + \alpha_1 V_{N,t-1} + \ldots + \alpha_5 V_{N,t-5} + \beta_1 V_{M,t-1} + \ldots + \beta_5 V_{M,t-5} + u_{N,t} \\ V_{M,t} &= \alpha_0 + \alpha_1 V_{M,t-1} + \ldots + \alpha_5 V_{M,t-5} + \beta_1 V_{N,t-1} + \ldots + \beta_5 V_{N,t-5} + u_{M,t} \end{aligned} \tag{21}$$

where $V_{N,t}$ and $V_{M,t}$ are the time-varying variances of the first differences in the logarithms of the NYSE and MSEI, respectively. The Wald statistics for the null joint hypothesis: $\beta_1 = \ldots = \beta_5 = 0$ is rejected for each equation. This means that the process is a feedback system where $V_{N,t}$ causes $V_{M,t}$ and $V_{N,t}$ causes $V_{M,t}$.

[Figure 3]

Finally, in Figure 3 we have signalled the period from September to November 1998. In this time interval the largest slump took place in the sample period which corresponds to the Russian financial crisis. Our estimations during the index fall reflect an important increase in the variances in both markets, which diminishes until reaching a trough.

Previous studies, some of which have been mentioned in the introduction, suggest that international correlation is much higher in periods of volatile markets and that market trends affect international correlations. In our case, the correlation seems to be influenced by the volatility of both

For Peer Review

indexes but not by the market trend. Taking the standard deviation as a measure of volatility we have estimated the next two reciprocal regressions

$$\hat{C}_t = 0.89 - 0.006\, D_{N,t}^{-1} \qquad R^2 = 0.15 \tag{22}$$
$$(0.00)\ \ (0.00)$$

$$\hat{C}_t = 0.89 - 0.005\, D_{M,t}^{-1} \qquad R^2 = 0.13 \tag{23}$$
$$(0.00)\ \ (0.00)$$

where $\hat{C}_t$ is the fitted time-varying correlation between NYSE and MSEI. $D_{N,t}$, and $D_{N,t}$ are the time-varying standard deviations of NYSE and MSEI, respectively. P-values are given in brackets. Other functional forms have been considered but the reciprocal gets a good fitting. As Figure 4 shows, for both markets when volatility increases, correlations increase approaching asymptotically to the limit of 0.89. Figure 4 illustrates the relation represented in equations (22) and (23).

[Figure 4]

In Figure 1 we have signalled two periods with different trends. The first one from 11/22/1994 to 08/25/2000 represents a bull market, and the second one from 08/25/2000 to the end of the sample represents a bear market. We have run the next regression to test whether there is any relation between trend market and correlations

For Peer Review

$$\hat{C}_t = \underset{(0.00)}{0.42} + \underset{(0.43)}{0.001}\, d_1 + \underset{(0.00)}{0.052}\, d_2 \qquad R^2 = 0.004$$

$$d_1 = \begin{cases} 1 & \textit{from}\ 11/22/1994\ \textit{to}\ 08/25/2000 \\ 0 & \textit{in other case} \end{cases}$$

$$d_2 = \begin{cases} 1 & \textit{from}\ 08/25/2000\ \textit{to}\ 12/30/2002 \\ 0 & \textit{in other case} \end{cases}$$

P-values are given in brackets. The coefficient of the bear market is significantly different from zero but the trend market explains correlations poorly.

As table 2 shows, although the series sometimes experience negative correlations, such events are rare. They occur less than 10 per cent of the time, suggesting the conditional correlations across the indexes, like their unconditional sample counterparts, are typically nonnegative. This is a typical empirical finding (e.g. Koch and Koch, 1991 and Engle and Susmel, 1993). However, the range of the estimated correlations is considerable. This fact has been found in other research as in King, Sentana and Wadhwani (1994) or Darbar and Deb (1997). In figure 5, we present the centred moving averages for 101 observations from the time-varying correlations. This figure shows that the time-varying correlations appear to increase with time.

[Figure 5]

For Peer Review

In figure 5 we have signalled three periods to describe the time-varying correlation behaviour. The first period is from September to November 1998, which corresponds to the Russian financial crisis, where the effects are shown clearly in the index. During the index slump, an increase in the correlations in both markets is produced which diminishes until reaching a trough. Subsequently, when both indexes increase the correlation also increases. The second period is from the end of December, 1999 until the middle of May, 2000. This period is characterized by internal events in the Spanish and European markets, according to the fusion of technological firms, which brought about the so-called speculative bubble. This phenomenon, although not alien to the American market, was experienced intensely in the Spanish market, as it is possible to see in the NYSE and MSEI indexes. As shown in figure (3), this phenomenon entails an important decrease in the correlations between both markets. When this phenomenon begins to disappear in the MSEI, it is possible to observe that the correlations between both markets increase again. The last period considered is from the day before the terrorist attacks on the World Trade Centre and the Pentagon on September, 2001. Until this moment the evolution was characterized by a constant decrease in correlations. However, the terrorist attacks produced a change of tendency as we show in figure 5.

IV. Conclusions

For Peer Review

In Economics and financial time series it is important to estimate the time-varying variances and covariances of asset returns. These magnitudes are crucial for risk management and diversification. The degree to which investors can reduce their risk by diversifying their portfolio depends on the correlations among assets. The lower is the correlation between two assets, the higher is the potential benefit to be obtained by diversification. Typical empirical findings indicate that stock returns are positively correlated across countries, but these correlations are usually smaller than between domestic assets, which justify international diversification.

In this paper we present a technique for obtaining the time-varying covariance matrix for several time series for nearest neighbour predictors. To illustrate the use of this technique, we analyse the time-varying variances and correlations between the daily returns on two equity stock market indexes, the New York Stock Exchange (NYSE) and the Madrid Stock Exchange Index (MSEI).

We highlight four empirical conclusions of our work. (1) There is a positive time-varying relation between both indexes. (2) Consistent with previous empirical results, we find evidence of a positive time-varying relation between the returns volatility of NYSE and MSEI and its correlations respectively. This provided an interesting use of the technique presented in this paper since correlations asset returns are a critical parameter in risk management. The use of constant correlations between asset returns in portfolio selection might lead to the use of suboptimal portfolios. (3) The capacity of trend markets to explain correlations is very poor. If any positive relation existed, it would be in the bear market but not

For Peer Review

in the bull market. And finally, (4) the important shocks produced in the markets such as those described above, seem to be well collected by our variances and correlations estimation technique.

Acknowledgment

This research is supported by the Spanish Ministry of Science and Education through project SEJ2006-07701.

References

Bellman, R. E. (1961). *Adaptive Control Processes*. in Princeton University Press.

Casdagli M. and Eubank, S. (1992). *Nonlinear Modeling and Forecasting*. Addison-Wesley, Redwood City.

Collins, D. and Gavron, S. (2004) Channels of financial market contagion. *Applied Economics*, **36**, pp. 2461-2469.

Darbar, S.M. and Deb, P. (1997) Co-Movements in international equity markets, *Jorunal of Financial Research*, **20**, pp. 305-322.

Engle, R.F. (1982) Autoregressive conditional heterocesdasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, **50**, pp. 987-1007.

Engle, R.F. and Susmel, R. (1993) Common volatility in international equity markets, *Journal of Business and Economic Statistics*, **11**, pp. 167-176.

For Peer Review

Submitted Manuscript

Engle, R. F., Ng, V., and Rothschild, M. (1990) Asset Pricing with a Factor ARCH Covariance Structure: Empirical Estimates for Treasury Bills. *Journal of Econometrics*, **45**, pp. 213-239.

Fan, J and Gijbels, I. (1997). *Local Polynomial Modelling and its Applications*. Chapman & Hall, London.

Fernández-Rodríguez, Sosvilla-Rivero, F, S. and Andrada-Félix, J. (1999) Exchange Rate Forecast with Nearest-Neighbour Methods: Evidence from the EMS, *International Journal of Forecasting*, **15**, pp. 383-392.

Fernández-Rodríguez, F., Sosvilla-Rivero, S. and Andrada-Félix, J. (2003) Nearest Neighbour Predictions in Foreign Exchange Markets in *Computational Intelligence in Economics and Finance*. (Ed.) Shu-Heng Chen and Paul Wang, Springer-Verlag.

Green, P. J. and Silverman, B. W. (1994). *Nonparametric Regression and Generalized Linear Models*. Chapman & Hall, London.

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

Härdle, W. (1990). *Applied Nonparametric Regression*. Cambridge University Press, Cambridge.

Hastie, T., Tibshirani, R. and Friedman, J. (2001). *The Elements of Statistical Learning*. Springer, New York.

King, M., Sentana, E. and Wadhwani, S. (1994) Volatility and links between national stock markets, *Econometrica*, **92**, pp. 901-933.

For Peer Review

King, M. and Wadhawani, S. (1990) Transmission of Volatility Between Stock Markets. *The Review of Financial Studies*, **3**, pp. 5-33.

Koch, P. and Koch, R. (1991) Evolution in dynamic linkages across daily national stock indexes, *Journal of International Money and Finance*, **10**, pp. 231-251.

Mizrach, B. (1992) Multivariate Nearest-Neighbor Forecasts of the EMS exchange Rates, *Journal of Applied Econometrics*, **7**, pp. S151-S163.

Roll, R. (1989) Price volatility, International Market Links and their Implications for Regulatory Policies. *Journal of Financial Services Research*, **3**, pp. 211-246.

Takens, F. (1981) Detecting strange attractors in turbulence, in *Lect. Notes Math*. 898. Raud and L.S. Young editors. Springer-Verlag, New York.

Zellner, A. (1962) An efficient method of estimating seemingly unrelated regression equations and tests of aggregation bias, *Journal of American Statistical Association*, **57**, pp. 348-368.

For Peer Review

Submitted Manuscript

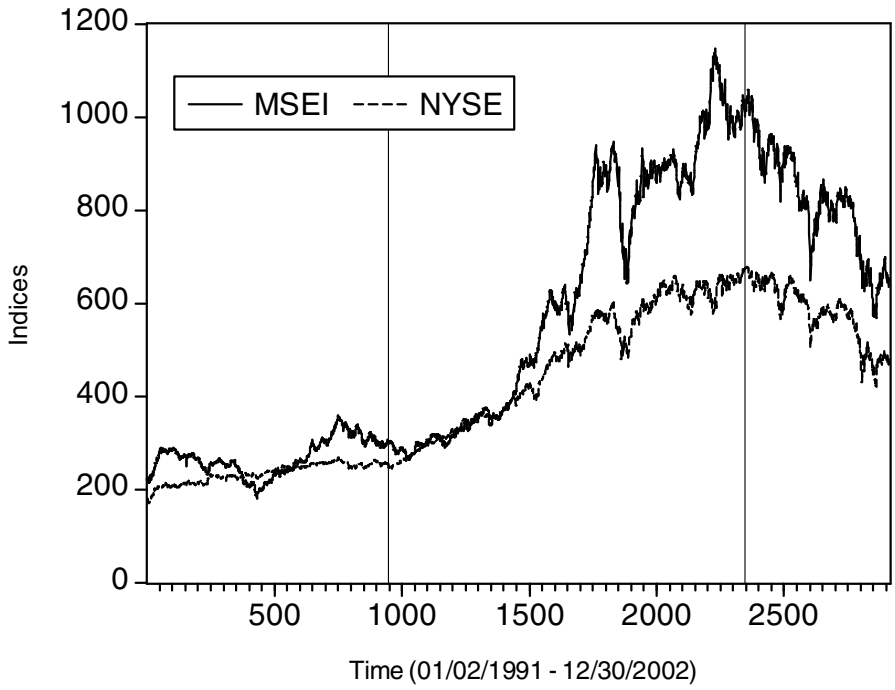

Figure 1. Stock Market Indexes: New York Stock Exchange (NYSE) and the Madrid Stock Exchange Index (MSEI)

Review

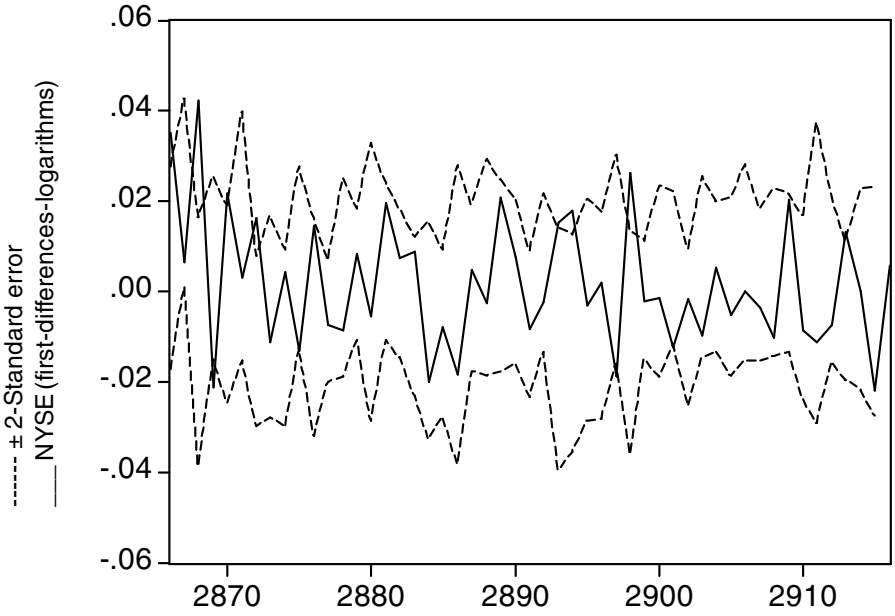

Figure 2. First differences in logarithms of the NYSE and the prediction interval

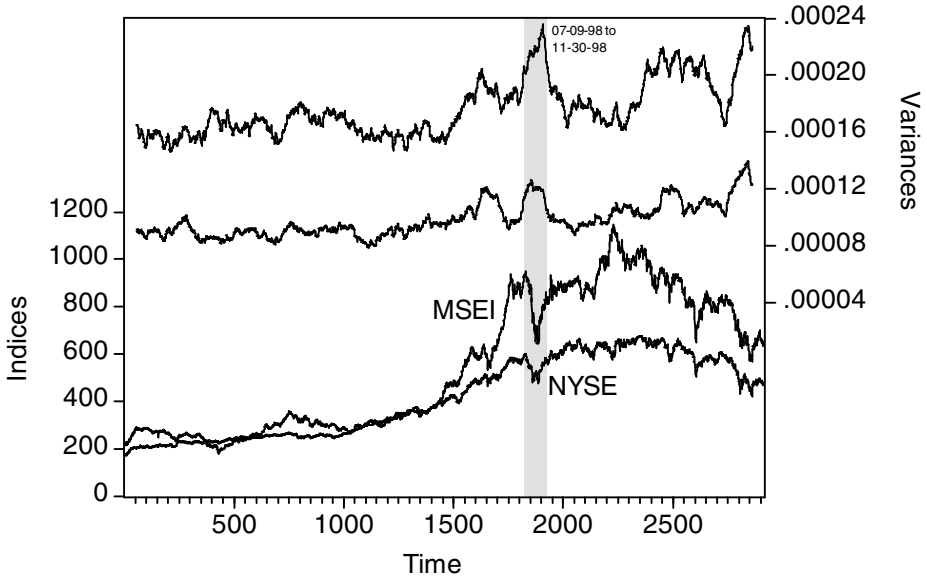

Figure 3. Centred moving averages for time-varying variances

Peer Review

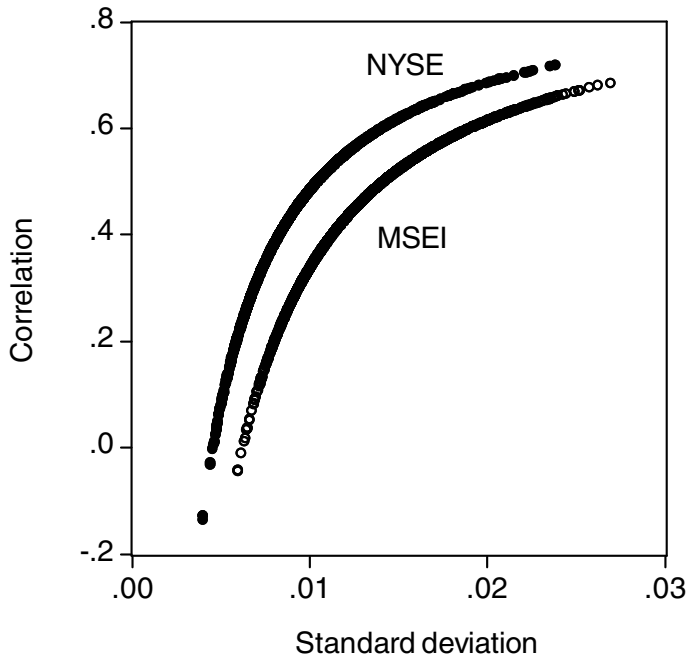

Figure 4. Relation between standard deviation and correlation

eer Review

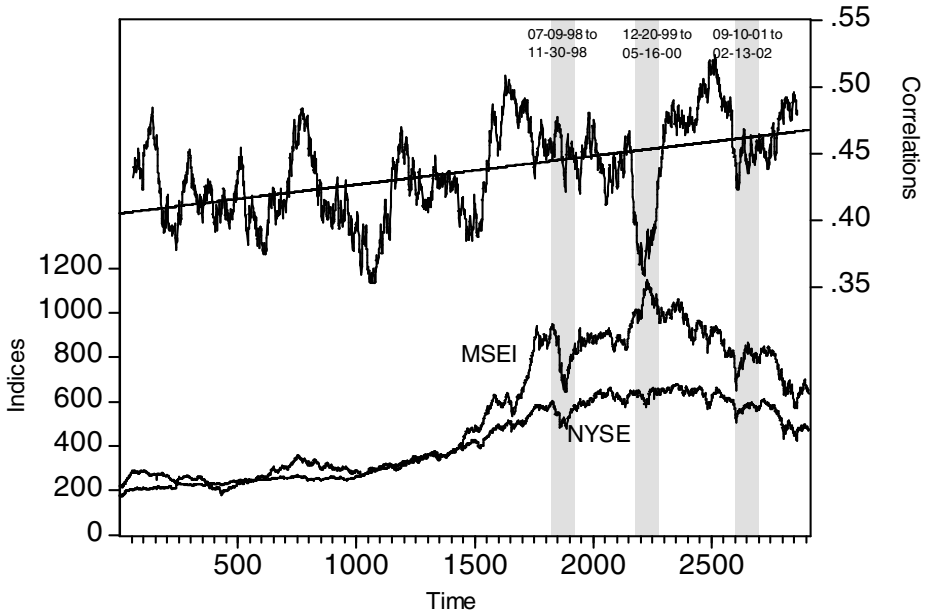

Figure 5. Centred moving averages for time-varying correlations

Table 1. Summary statistics for daily returns (multiplied by 100)

	NYSE	MSEI
Mean	0.033	0.036
Median	0.034	0.051
Maximum	5.179	6.362
Minimum	-6.791	-8.611
Std. Dev.	0.940	1.267
Skewness	-0.254	-0.249
Kurtosis	7.617	6.105
Jarque-Bera	2620.3	1201.3
(Probability)	(0.00)	(0.00)
Observations	2915	2915
Sample correlation = 0.4056		

For Peer Review

Table 2. Summary statistics for the estimates daily time-varying variances, covariances and correlations

	10th Per.	30th Per.	Median	70th Per.	90th Per.
Variance (NYSE)	$0.04*10^{-3}$	$0.07*10^{-3}$	$0.08*10^{-3}$	$0.11*10^{-3}$	$0.16*10^{-3}$
Variance (MSEI)	$0.09*10^{-3}$	$0.12*10^{-3}$	$0.15*10^{-3}$	$0.20*10^{-3}$	$0.30*10^{-3}$
Covariance	$0.002*10^{-3}$	$0.030*10^{-3}$	$0.052*10^{-3}$	$0.082*10^{-3}$	$0.143*10^{-3}$
Correlation	0.0137	0.3201	0.4919	0.6264	0.7839

For Peer Review