DISCUSSION PAPERS IN STATISTICS AND ECONOMETRICS

SEMINAR OF ECONOMIC AND SOCIAL STATISTICS
UNIVERSITY OF COLOGNE

No. 07/10

Forecasting international stock market correlations: Does anything beat a CCC?

by

Hans Manner

Olga Reznikova

DISKUSSIONSBEITRÄGE ZUR STATISTIK UND ÖKONOMETRIE

SEMINAR FÜR WIRTSCHAFTS- UND SOZIALSTATISTIK
UNIVERSITÄT ZU KÖLN

Albertus-Magnus-Platz, D-50923 Köln, Deutschland

DISCUSSION PAPERS IN STATISTICS AND ECONOMETRICS

SEMINAR OF ECONOMIC AND SOCIAL STATISTICS
UNIVERSITY OF COLOGNE

No. 07/10

Forecasting international stock market correlations: Does anything beat a CCC?

by

Hans Manner[1]
Olga Reznikova[2]

Abstract

It is well known that the correlation between financial series varies over time. Here, the forecasting performance of different time-varying correlation models is compared for cross-country correlations of weekly G5 and daily European stock market indices. In contrast to previous studies only the correlation and not the entire covariance matrix is forecasted and multi-step forecasts are considered. The forecast comparison is done by considering statistical and economic criteria. The results suggest that under a statistical criterion time-varying correlation models perform quite well for weekly data, but cannot outperform the constant correlation model for daily data. Considering economic criteria it is hard to beat a constant correlation model.

Keywords: dynamic conditional correlation, regime switching, stochastic correlation, smooth correlations, indirect model comparison, portfolio construction

JEL Classification: C53, G17.

[1]Seminar für Wirtschafts– und Sozialstatistik, Universität zu Köln, D-50923 Köln

[2]Institut de statistique, biostatistique et sciences actuarielles, Université catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium.

Forecasting international stock market correlations: Does anything beat a CCC?

Hans Manner[1] Olga Reznikova[2]

October 20, 2010

Abstract

It is well known that the correlation between financial series varies over time. Here, the forecasting performance of different time-varying correlation models is compared for cross-country correlations of weekly G5 and daily European stock market indices. In contrast to previous studies only the correlation and not the entire covariance matrix is forecasted and multi-step forecasts are considered. The forecast comparison is done by considering statistical and economic criteria. The results suggest that under a statistical criterion time-varying correlation models perform quite well for weekly data, but cannot outperform the constant correlation model for daily data. Considering economic criteria it is hard to beat a constant correlation model.

Keywords: Dynamic conditional correlation, regime switching, stochastic correlation, smooth correlations, indirect model comparison, portfolio construction

JEL Classification: C53, G17

1 Introduction

Modeling and forecasting volatilities and correlations of financial assets is of great interest to financial institutions. In particular the modeling of conditional variances has been studied very thoroughly in the literature on GARCH models starting with Engle (1982) and the competing class of stochastic volatility models as surveyed by, e.g., Andersen and Shephard (2009). The forecasting performance of competing GARCH models is studied from a theoretical perspective by Hansen and Lunde (2006) and empirically by, e.g., Hansen and Lunde (2005). More recently, multivariate volatility models have been studied and reviews can be found in Bauwens et al. (2006) and Asai et al. (2006). However, there are still only few studies on the forecasting comparison of multivariate volatility models. Notable exceptions are the methodological contributions by Patton and Sheppard (2009) and Laurent et al. (2009), and the empirical study by Caporin and McAleer (2010). A detailed overview of various volatility models and forecasting methods is provided in Andersen et al. (2006). Whereas usually the focus is on volatilities and covariances, in this paper we are interested in time-varying correlations. While correlations themselves are usually not as useful as covariances, they are an important part of the covariance matrix. Since it is common practice to model and estimate conditional variances and correlations separately, it is important to know whether some correlation models are more useful than others. Evidence of correlations changing over time is documented by, e.g., Longin and Solnik (1995), Erb et al. (1994), Engle (2002) and Pelletier (2006). Time-varying correlation models have proven to be an inherent part of financial management and a number of distinct models have been proposed recently. These models are based on quite different assumptions concerning

[1]Chair of Statistics and Econometrics, University of Cologne, 50937 Cologne, Germany, manner@statistik.uni-koeln.de

[2]Institut de statistique, biostatistique et sciences actuarielles, Université catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium.

the correlation dynamics and, to the best of our knowledge, a systematic model comparison does not exist.

The aim of this paper is to compare the forecasting ability of different time-varying correlation models. In contrast to previous studies we focus exclusively on forecasting correlations, not the whole covariance matrix, and we do not only consider one-step, but multi-step forecasts. The main problem in this context is that the object of interest, i.e. the conditional correlation coefficient, is unobservable. Therefore appropriate measures to compare and evaluate forecasts need to be used. As in one of our applications we are focusing on international stock market indices from countries in different time zones, the commonly applied strategy to compare the forecast to a proxy based on high frequency data cannot be applied. We circumvent this problem by using indirect measures of statistical and economic nature to compare the out-of-sample performance of our competing models.

The performance of the models is analyzed for pairs of the G5 weekly country stock market indices over the past two decades and daily European stock market indices over the last 6 years. The out-of-sample forecasting period is selected such that the recent subprime crisis is covered.

The rest of the paper is organized as follows. In Section 2 we explain the forecasting methodology, the criteria for forecast evaluation and the suggested correlation models. In Section 3 we describe our data and present the empirical results. Finally, Section 4 concludes.

2 Methodology

In this section we describe the methodology and the competing models that are used to study the forecasting abilities of correlation models.

2.1 Comparing correlation forecasts

Consider the stock market return $r_{i,t}$ of market i, for $i = 1, \ldots, N$ and $t = 1, \ldots, T$. Assume that $E(r_{i,t}) = 0$ for all i and t. The correlation coefficient between markets i and j at time t is defined as

$$\rho_{ij,t} = \frac{E_{\mathscr{F}}(r_{i,t} r_{j,t})}{\sqrt{E_{\mathscr{F}}(r_{i,t}^2) E_{\mathscr{F}}(r_{j,t}^2)}} = \frac{E_{\mathscr{F}}(r_{i,t} r_{j,t})}{\sigma_{i,t}\sigma_{j,t}}, \tag{1}$$

where $E_{\mathscr{F}}$ denotes the expectation conditional on some information set $\mathscr{F}$. We are interested in modeling and forecasting $\rho_{ij,t}$, which is assumed to be time varying. Given the stylized fact that stock market volatilities vary over time we must provide an appropriate model for the volatility of each market, before starting to estimate a model for the correlations. Furthermore, in order to be able to compare the forecasting performance of different models for the correlation dynamics one must actually provide a forecast of the covariance matrix. Due to non-linearity this cannot be simply achieved by multiplying the forecasts of the standard deviations with the correlation forecast. For these reasons and since the main goal of this paper is to compare correlation models exclusively, we decide to filter out the time-varying volatility prior to the analysis and continue to work with standardized data. Standardized returns are defined as

$$r_{i,t}^* = \frac{r_{i,t}}{\hat{\sigma}_{i,t}}, \tag{2}$$

where $\hat{\sigma}_{i,t}$ is an estimate of the conditional standard deviation of asset i at time t. In our application we first fit an AR(p) model and estimate $\sigma_{i,t}$ using the best fitting GARCH type model on the residuals, where the best fitting model is chosen to be the one minimizing the Bayesian information criterion (BIC) from a number of candidate specifications. We consider the standard GARCH(1,1), EGARCH(1,1) and GJR-GARCH(1,1) with Normal and Student t errors:

- GARCH(1,1)

$$\sigma_{i,t}^2 = \omega + \alpha r_{i,t-1}^2 + \beta \sigma_{i,t-1}^2;$$

- EGARCH(1,1)

$$\log\left(\sigma_{i,t}^2\right) = \omega + \alpha \frac{r_{i,t-1}^2}{\sqrt{\sigma_{i,t-1}^2}} + \gamma \left| \frac{r_{i,t-1}^2}{\sqrt{\sigma_{i,t-1}^2}} \right| + \beta \log\left(\sigma_{i,t-1}^2\right);$$

- GJR-GARCH(1,1)

$$\sigma_{i,t}^2 = \omega + \alpha r_{i,t-1}^2 + \gamma r_{i,t-1}^2 \mathbb{I}_{\{r_{i,t-1}<0\}} + \beta \sigma_{i,t-1}^2.$$

Using the pseudo-observations $r_{i,t}^*$ we estimate each candidate model for the time-varying correlation on the in-sample period $t = 1, \dots, T^*$. We then produce an h-step forecast using the information available at time T^*, $\mathscr{F}_{T^*}$. We denote this forecast $\rho_{T^*+h|T^*}$. Next, we update our information set to $\mathscr{F}_{T^*+1}$ to re-estimate the models using the observations for $t = 2, \dots, T^* + 1$, forecast $\rho_{T^*+1+h|T^*+1}$ and repeat this until we have $\rho_{T|T-h}$. Thus we perform a rolling window approach for forecasting correlations.

Comparing the forecasting performance of dynamic correlation models is much harder than in traditional forecasting exercises. First of all, correlation itself is unobserved and an appropriate proxy must be used to evaluate any loss function of interest. An appropriate proxy for correlation would be the realized correlation calculated from high frequency data. Unfortunately, since we are dealing with data for international stock markets, due to non-synchronous trading realized correlation often cannot be computed. Hansen and Lunde (2006) show that even when using a noisy proxy the appropriate choice of loss function leads to a consistent ranking of volatility models. However, whereas for variance (covariances) one can simply use squared returns (cross products of returns), correlations are bounded in $(-1, 1)$ so even the obvious choice of cross product of standardized returns cannot be used as a proxy as there is no guarantee that it stays within this range. This only leaves us with the possibility to use model based or indirect measures to compare the competing models.

For a given sequence of correlation forecasts $\rho_{ij,t+h|t}$ for $t = T^*, \ldots, T-h$, and the out-of-sample pseudo-observations $r^*_{i,T^*+h}, \ldots, r^*_{i,T}$ we compare the statistical out-of-sample fit of our models by computing the predictive log-likelihood (PLL) of a bivariate standard normal distribution. Higher PLL suggest a better statistical out-of-sample fit of a given model.

Next to this statistical measure, following Chan et al. (1999) we construct the global minimum variance portfolio (MVP). With the covariance matrix $H_{t+h|t}$ constructed using the correlation forecasts and the estimated GARCH volatilities, the portfolio weights are

$$w_{t+h} = \frac{H^{-1}_{t+h|t} \cdot \iota}{\iota' \cdot H^{-1}_{t+h|t} \cdot \iota}, \tag{3}$$

where ι is a (2×1) vector of ones. Denote the return of the MVP at time t by r_t^{MVP} and its sample variance for the out-of-sample period by σ^2_{MVP}. The model that minimizes σ^2_{MVP} is considered the best correlation forecasting model.

Next to comparing the PLL and σ^2_{MVP} we are interested in testing whether a benchmark model, in our case a constant conditional correlation model, is not inferior to any of the alternatives. This can be achieved by using the test for superior predictive ability (SPA) by Hansen (2005). Denote by $d_{k,t}$ the difference between the criterion, or loss function, for forecast evaluation of the benchmark model and model $k = 1, \ldots, m$, where a small value of the criterion corresponds to a good model performance. Then for $\mathbf{d}_t = (d_{1,t}, \ldots, d_{m,t})'$ the null hypothesis of interest is

$$H_0 : E(\mathbf{d}_t) \leq 0. \tag{4}$$

Denote $\bar{d}_k = n^{-1} \sum_{t=1}^{n} d_{k,t}$, with $n = T - T^*$ the number of out-of-sample observations. Then the test statistics for the SPA test is given by

$$T_n^{SPA} = max \left[\max_{k=1,\ldots,m} \frac{n^{1/2} \bar{d}_k}{\hat{\omega}_k}, 0 \right], \tag{5}$$

where $\hat{\omega}_k^2$ is a heteroscedasticity and autocorrelation consistent (HAC) estimator of $\omega_k^2 \equiv \text{Var}(n^{1/2}\bar{d}_k)$. P-values for the test are computed using the stationary bootstrap by Politis and Romano (1994) as explained in Hansen (2005).

Finally, we compute the out-of-sample 5% Value-at-Risk (VaR) of an equally weighted portfolio and test the adequacy of the VaR forecasts using the out-of-sample dynamic quantile (DQ) test of Engle and Manganelli (2004). Let us define $\text{hit}_t = \mathbb{I}(r_t^* < \text{VaR}_t) - 0.05, t = T^* + h, \ldots, T$, where r_t^* is the portfolio return, $\mathbf{Hit} = [\text{hit}_{T^*+h}, \ldots, \text{hit}_T]'$ and let us construct $\mathbf{X}_{(N_R, 2+q)}$ with the typical row

$$\mathbf{X}_t = [1 \ \text{VaR}_t \ \text{hit}_{t-1} \ \ldots \ \text{hit}_{t-q}],$$

where q is the number of lags and $N_R = T - T^* - \max(h-1, q)$ is the number of out-of-sample observations used for the test. Then the DQ statistic

$$DQ = \frac{N_R^{-1}\mathbf{HitX}\left[\mathbf{X}'\mathbf{X}\right]^{-1}\mathbf{X}'\mathbf{Hit}'}{0.05(1-0.05)}$$

has a χ_q^2 distribution. Under the null hypothesis the model is considered to be correctly specified and the VaR estimates are adequate.

2.2 Models for correlation forecasting

Below we describe the correlation models that are compared in this study, and we explain how forecasts are obtained.

Constant conditional correlation (CCC): Correlations are treated as constant and are estimated using the sample correlation. The forecasts are obtained by assuming that correlations will not change.

Consistent dynamic conditional correlation (cDCC): The cDCC model was proposed by Aielli (2009) as a modification of a widely used DCC model of Engle (2002). The correlations are driven by lagged residuals $\tilde{r}_t$ and an autoregressive term

$$\begin{aligned} Q_t &= (1-\alpha-\beta)\Psi + \alpha\tilde{r}_{t-1}\tilde{r}'_{t-1} + \beta Q_{t-1}, & (6)\\ R_t &= \text{diag}\{Q_t\}^{-1/2} Q_t \text{diag}\{Q_t\}^{-1/2}, & (7)\end{aligned}$$

where $\tilde{r}_t = \text{diag}\{Q_t\}^{1/2} r_t^*$, so that $\text{Var}(\tilde{r}_t|\mathscr{F}_{t-1}) = Q_t$ and $\text{Var}(\tilde{r}_t) = \Psi$. The estimation is done via quasi maximum likelihood (QML) and is similar to the estimation of the DCC, with the only difference that $\widehat{\Psi}$ is a sample covariance matrix of $\tilde{r}_t$, so it depends on $\widehat{\alpha}$ and $\widehat{\beta}$ and is estimated within the second step of the QML and not prior to it. As recognized by Engle and Sheppard (2001) and Aielli (2009) the direct forecast of $\rho_{ij,t+h|t}$ is unfeasible with the DCC and cDCC as

$$E_t(\rho_{ij,t+h}) = E_t\left(\frac{q_{ij,t+h}}{\sqrt{q_{ii,t+h}\ q_{jj,t+h}}}\right) \tag{8}$$

is not linear. Engle and Sheppard (2001) suggest the approximation

$$\widehat{\rho}_{ij,t+h|t} = \frac{q_{ij,t+h|t}}{\sqrt{q_{ii,t+h|t}\ q_{jj,t+h|t}}}, \tag{9}$$

where $q_{ij,t+h|t}$ is the forecast of $q_{ij,t+h}$ and is obtained from

$$q_{ij,t+h|t} = \psi_{ij}(1-\alpha-\beta)\sum_{k=0}^{h-2}(\alpha+\beta)^k + (\alpha+\beta)^{h-1} q_{ij,t+1} \tag{10}$$

and $q_{ij,t}$ is an element of Q_t .

Semiparametric/smooth dynamic correlation (SDC): The SDC model of Hafner and Reznikova (2010) can be seen as a generalization of the model proposed by Hafner et al. (2006). Correlation is treated as a smooth function of time and is estimated nonparametrically by local maximum likelihood

$$\widehat{\rho}(\tau, T^*) = \arg\max_{\rho} \sum_{t=1}^{T^*} \ell\left(r_{1,t}^*, r_{2,t}^*, \rho\right) K_b\left(t/T^* - \tau\right), \tag{11}$$

where $\ell(\cdot,\cdot,\cdot)$ is the log-likelihood of a bivariate normal distribution with mean zero and variance one, $K_b(\cdot) = (1/b)K(\cdot/b)$ is a kernel function with bandwidth $b > 0$ and $\tau \in [0, 1]$. Prior to estimation a MSE-optimal bandwidth $\widehat{b}$ is selected. In order to avoid the problem of the boundary bias, we approximate ρ with the first order Taylor approximation

$$\rho(\tau, T) = \rho_0(\tau, T) + \rho_1(\tau, T)\left(\frac{t}{T} - \tau\right). \tag{12}$$

The forecast is based on the bandwidth selected on the in-sample period and on a linear approximation of $\widehat{\rho}(1, t)$

$$\rho_{t+h|t} = \widehat{\rho}_0(1, t) + \widehat{\rho}_1(1, t)h/t,\ t = T^*, \ldots, T^* - h, \tag{13}$$

where $\widehat{\rho}_0(1, t), \widehat{\rho}_1(1, t)$ are the estimates of the coefficients of the first order Taylor approximation of $\rho(\tau, t)$, where $\tau = 1$.

Regime switching correlation (RSC): This approach is based on Pelletier (2006) and it is assumed that there are two regimes governed by different correlations ρ_1 and ρ_2. Let k_t be a latent random variable that takes on the value $k = 1, 2$ when regime k is the current state. k_t is assumed to follow a Markov chain of order one with π_{ij} the probability of moving to regime j in period t conditional on being in state i at time $t - 1$. Estimation and forecasting of this type of models is based on the EM algorithm using the Kalman filter and a detailed description can be found in Hamilton (1994) Ch. 22. In particular, the forecast of ρ_t is the average of ρ_1 and ρ_2 weighted by the expected probabilities of being in state 1 and 2, respectively.

Stochastic autoregressive correlation (SCAR): Correlation is assumed to be the inverse Fisher transform of a stationary Gaussian autoregressive process of order one, $\rho_t = (\exp(2\lambda_t) - 1)/(\exp(2\lambda_t) + 1)$ with

$$\lambda_t = \alpha + \beta\lambda_{t-1} + \nu\varepsilon_t, \tag{14}$$

where ε_t is an i.i.d. $N(0,1)$ innovation. This model was first introduced by Yu and Meyer (2006). We estimate the model by a simulated maximum likelihood approach using importance sampling as explained in Hafner and Manner (2010). Forecasts are calculated by using standard time-series techniques to forecast λ_t:

$$\widehat{\lambda}_{t+h} = \frac{\alpha}{1-\beta} + \beta^h \left(\widehat{\lambda}_t - \frac{\alpha}{1-\beta} \right). \tag{15}$$

As suggested in Hafner and Manner (2010) the forecasts of ρ_t are obtained using a second order Taylor approximation of $\Psi(\lambda_t)$, which takes into account the nonlinearity of the inverse Fisher transformation:

$$\widehat{\rho}_{t+h} = \Psi(\widehat{\lambda}_{t+h}) + \frac{-4(\exp(2\widehat{\lambda}_{t+h}) - 1)\exp(2\widehat{\lambda}_{t+h})}{(\exp(2\widehat{\lambda}_{t+h}) + 1)^3} s^2_{t+h}, \tag{16}$$

where $s^2_{t+h} = \nu^2(1-\beta^{2h})/(1-\beta^2)$ is the forecast error for λ_t.

3 Empirical Study

In this section we present the results of our empirical study using the methodology explained above. We consider the returns of stock market indices on a weekly (Section 3.1) and daily (Section 3.2) frequency.

3.1 Weekly international stock market indices

The first dataset we use are weekly (Wednesday) returns of the MSCI stock market index for France, Germany, Japan, UK and US, meaning there are 10 country pairs. Our sample starts October 11, 1989 and ends March 3, 2010, resulting in a sample of 1064 observations. The last 5 years, or 260 observations, constitute the out-of-sample period. Weekly data have been chosen to avoid the non-synchronous trading effect of daily data.

The estimation results for the AR-GARCH models can be found in Tables 1 and 2. Table 3 reports the results for the predictive log-likelihood.

Table 1: Estimators of AR model for weekly international stock market returns

Country	**AR(p)**	
	a_0	a_1
France	0.0011 (0.0008)	−0.116 (0.031)
Germany	0.0010 (0.0010)	
Japan	−0.0005 (0.0010)	
UK	0.0010 (0.0008)	−0.089 (0.031)
USA	0.0008 (0.0007)	

Note: Table 1 reports the estimated parameters and standard errors of the $AR(p)$ model for the log-returns of MSCI index of G5 countries (weekly returns, October 11, 1989 to March 3, 2010), where order p is selected with the BIC.

Table 2: Estimators of GARCH model for weekly international stock market returns

Country	**Model**	**GARCH**				
		ω	α	γ	β	**d.o.f.**
France	EGARCH - St	−0.486 (0.121)	−0.083 (0.023)	0.190 (0.041)	0.953 (0.015)	10.40 (2.61)
Germany	GARCH - St	0.000015 (0.000007)	0.111 (0.022)		0.880 (0.021)	8.02 (1.61)
Japan	GARCH - St	0.000054 (0.000021)	0.129 (0.028)		0.820 (0.037)	9.52 (2.66)
UK	EGARCH - St	−0.458 (0.104)	−0.092 (0.021)	0.176 (0.037)	0.957 (0.012)	16.28 (6.77)
USA	GJR-Garch - St	0.000012 (0.000004)	−0.010 (0.022)	0.172 (0.034)	0.894 (0.024)	8.66 (2.15)

Note: Table 2 reports the estimated parameters and standard errors of the appropriate $GARCH$ model for the log-returns of MSCI index of G5 countries corrected for autocorrelation (weekly returns, October 11, 1989 to March 3, 2010), where the appropriate GARCH type model is selected with the BIC. Specifically, ω denotes the constant term, α stays for the coefficient of the error term, γ reflects the leverage effect, β is the coefficient of the autoregressive term and **d.o.f** are the degrees of freedom for models with Student t errors.

Table 3: Forecast comparison using predictive log-likelihood for weekly international stock market returns

Pair		h	CCC	cDCC	SDC	RSC	SCAR	SPA-pval
France	Germany	1	0.8687	1.1614	1.1147	**1.1665**	1.1126	0.0000
		4	0.8672	**1.1660**	1.0963	1.1569	1.1115	0.0000
		12	0.8706	**1.1639**	1.0500	1.1388	1.1033	0.0000
France	Japan	1	0.1721	**0.1972**	0.1436	0.1966	0.1926	0.0087
		4	0.1727	**0.1948**	0.1408	0.1897	0.1893	0.0000
		12	0.1747	0.1887	0.1194	0.1906	**0.1922**	0.0000
France	UK	1	0.6551	0.7742	0.6145	**0.7942**	0.7525	0.0000
		4	0.6529	0.7678	0.6339	**0.7886**	0.7546	0.0000
		12	0.6579	0.7624	0.4203	**0.7955**	0.7648	0.0000
France	US	1	0.3528	**0.4192**	0.3700	0.4180	0.4191	0.0000
		4	0.3529	0.4119	0.3639	0.4105	**0.4133**	0.0000
		12	0.3555	0.3968	0.2872	**0.4021**	0.4003	0.0000
Germany	Japan	1	0.1615	0.1786	0.1157	**0.1812**	0.1776	0.0513
		4	0.1618	0.1786	0.1254	**0.1811**	0.1754	0.0004
		12	0.162	0.1776	0.1067	**0.1816**	0.1762	0.0002
Germany	UK	1	0.5373	0.5867	0.4788	**0.5975**	0.5854	0.0006
		4	0.5332	**0.5999**	0.4604	0.5838	0.5780	0.0000
		12	0.5363	**0.6070**	0.3131	0.5785	0.5772	1.0000
Germany	US	1	0.3108	0.3509	0.2770	0.3437	**0.3571**	0.0425
		4	0.3102	0.3442	0.2594	0.3371	**0.3538**	0.0547
		12	0.3104	0.3351	0.0074	0.3308	**0.3360**	0.0605
Japan	UK	1	0.1532	0.1613	0.1510	**0.1742**	0.1737	0.0032
		4	0.1531	0.1587	0.1467	**0.1711**	0.1685	0.0002
		12	0.154	0.1564	0.1319	**0.1716**	0.1697	0.0001
Japan	US	1	0.0973	0.1092	0.0630	**0.1107**	0.1088	0.2188
		4	0.0981	0.1063	0.0713	**0.1144**	0.1108	0.0419
		12	0.098	0.1043	0.0690	**0.1112**	0.1107	0.0153
UK	US	1	0.3315	0.3636	0.3217	**0.3637**	**0.3637**	0.0005
		4	0.3318	0.3565	0.3130	**0.3621**	0.3616	0.0003
		12	0.3381	0.3548	0.2833	0.3647	**0.3649**	0.0008

Note: Table 3 compares the forecasting performance of the competing models by the predictive log-likelihood criterion and reports the SPA p-values for the null hypothesis that no model perform better than the CCC based on 10000 bootstrap replications. The data are weekly returns from October 11, 1989 to March 3, 2010 standardized by the volatilities estimated using an appropriate GARCH model. The last 5 years of data constitute the out-of-sample period.

Table 4: Forecast comparison using σ^2_{MVP} for weekly international stock market returns

Pair		h	CCC	cDCC	SDC	RSC	SCAR	SPA-pval
France	Germany	1	**0.1342**	0.1429	0.1491	0.1383	0.1476	1.0000
		4	**0.1353**	0.1425	0.1502	0.1387	0.1477	1.0000
		12	**0.1390**	0.1451	0.1553	0.1411	0.1486	1.0000
France	Japan	1	0.0866	0.0864	0.0871	**0.0861**	**0.0861**	0.1985
		4	0.0872	0.0872	0.0888	0.0871	**0.0870**	0.3802
		12	0.0893	**0.0885**	0.0904	0.0892	0.0892	0.2827
France	UK	1	**0.1181**	0.1231	0.1257	0.1202	0.1220	1.0000
		4	**0.1190**	0.1234	0.1261	0.1213	0.1221	1.0000
		12	**0.1224**	0.1275	0.1356	0.1247	0.1260	1.0000
France	US	1	0.0790	0.0785	0.0786	**0.0776**	**0.0776**	0.1576
		4	0.0797	0.0792	0.0787	0.0783	**0.0782**	0.0533
		12	0.0818	0.0810	0.0812	0.0806	**0.0805**	0.0266
Germany	Japan	1	0.0891	0.0897	0.0951	0.0892	**0.0890**	0.6227
		4	**0.0897**	0.0902	0.0953	**0.0897**	0.0898	1.0000
		12	0.0918	0.0920	0.0973	**0.0917**	**0.0917**	0.5841
Germany	UK	1	**0.1223**	0.1285	0.1343	0.1260	0.1259	1.0000
		4	**0.1232**	0.1285	0.1349	0.1248	0.1249	1.0000
		12	**0.1267**	0.1437	0.1402	0.1274	0.1274	1.0000
Germany	US	1	0.0778	0.0759	0.0752	0.0752	**0.0750**	0.0982
		4	0.0785	**0.0751**	0.0760	0.0757	0.0758	0.0503
		12	0.0805	0.0788	0.0805	0.0788	**0.0787**	0.0395
Japan	UK	1	0.0804	0.0808	0.0814	0.0803	**0.0802**	0.3650
		4	0.0809	0.0816	0.0822	**0.0808**	0.0809	1.0000
		12	**0.0829**	0.0832	0.0843	**0.0829**	**0.0829**	1.0000
Japan	US	1	**0.0578**	0.0587	0.0605	0.0583	0.0582	1.0000
		4	**0.0583**	0.0592	0.0613	0.0588	0.0588	1.0000
		12	**0.0594**	**0.0594**	0.0625	0.0597	0.0597	0.6855
UK	US	1	**0.0791**	0.0800	0.0801	0.0793	0.0794	1.0000
		4	**0.0797**	0.0803	0.0807	**0.0797**	**0.0797**	1.0000
		12	0.0819	0.0818	0.0822	**0.0816**	**0.0816**	0.2779

Note: Table 4 compares the forecasting performance of the competing models by the variance of the global minimum variance portfolio (multiplied by 100) and reports the SPA p-values for the null hypothesis that no model perform better than the CCC based on 10000 bootstrap replications. The data are weekly returns from October 11, 1989 to March 3, 2010 standardized by the volatilities estimated using an appropriate GARCH model. The last 5 years of data constitute the out-of-sample period.

Table 5: Dynamic quantile test for the adequacy of VaR forecasts for weekly international stock market returns

Pair		h	CCC	cDCC	SDC	RSC	SCAR
France	Germany	1	0.069	0.073	0.073	0.073	0.073
		4	0.243	0.331	0.330	0.330	0.330
		12	0.187	0.261	0.261	0.227	0.262
France	Japan	1	**0.030**	**0.044**	0.138	**0.029**	**0.030**
		4	**0.038**	**0.049**	0.146	0.053	**0.038**
		12	**0.025**	**0.033**	0.112	**0.033**	**0.035**
France	UK	1	**0.001**	0.107	0.107	0.106	0.107
		4	**0.001**	0.083	0.083	0.083	0.083
		12	**0.000**	0.060	0.060	0.061	0.060
France	US	1	0.189	0.188	0.617	0.368	0.367
		4	0.230	0.231	0.665	0.408	0.409
		12	0.162	0.275	0.273	0.275	0.275
Germany	Japan	1	**0.031**	0.089	**0.018**	**0.049**	0.103
		4	**0.031**	0.053	**0.013**	**0.018**	**0.018**
		12	**0.020**	**0.020**	**0.008**	**0.011**	**0.012**
Germany	UK	1	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
		4	**0.004**	**0.001**	**0.003**	**0.003**	**0.003**
		12	**0.002**	**0.001**	**0.000**	**0.001**	**0.001**
Germany	US	1	0.374	0.339	0.330	0.342	0.342
		4	0.393	0.339	0.305	0.340	0.339
		12	0.318	0.258	0.462	0.259	0.259
Japan	UK	1	0.222	0.263	0.221	0.412	0.412
		4	0.183	0.200	0.191	0.366	0.220
		12	0.163	0.191	0.272	0.204	0.204
Japan	US	1	0.545	0.565	0.656	0.700	0.577
		4	0.627	0.612	0.586	0.638	0.634
		12	0.789	0.520	0.568	0.532	0.532
UK	US	1	0.208	0.207	0.591	0.207	0.207
		4	0.232	0.232	0.226	0.232	0.232
		12	0.203	0.203	0.106	0.203	0.203

Note: Table 5 reports the p-values of the dynamic quantile test by Engle and Manganelli (2004) for the null hypothesis of a correct out-of-sample Value-at-Risk. The data are weekly returns from October 11, 1989 to March 3, 2010 standardized by the volatilities estimated using an appropriate GARCH model. The last 5 years of data constitute the out-of-sample period.

The average log-likelihood of the best performing model is shown in bold. The results suggest that the models that allow for time-varying correlations perform much better than the constant correlation model. In particular, the RSC model shows the best overall performance followed by the cDCC and the SCAR models. The SDC models is the worst model and is even outperformed by the CCC in many cases. The p-values of the SPA test, applied to the negative log-likelihood, indicate that the difference in PLL between the CCC and the best competitor is statistically significant for almost all cases. The forecast horizon h does not appear to play role for the relative performance of the models.

In Table 4 the results for the variance of the global minimum variance portfolio are shown. The results are in strong contrast to the results for the PLL. Overall, the CCC is the best performing model followed by the SCAR model. Furthermore, the SPA test suggests that in cases where the CCC is not the best performing model by the point estimate of σ^2_{MVP} the difference in model performance is hardly ever significant (only in 2 out of 30 cases). This means that for portfolio selection a simple model is to be preferred over more complex specifications. This is probably due to the fact the the estimation error has a strong impact that is amplified when computing portfolio weights. This is in line with the findings of Caporin and McAleer (2010), who also find a good performance of simpler model for indirect comparison based on portfolio construction when estimating covariances.

Finally, Table 5 presents the p-values of the DQ test by Engle and Manganelli (2004) for the null hypothesis of correctly forecasted VaR. Except for four country pairs all models pass the test. For the pairs France-Japan, France-UK, Germany-Japan and Germany-UK the CCC model is always rejected, while the cDCC and SDC models are not rejected in 5 and 6 cases, respectively. Overall these results indicate that using a more sophisticated model does not results in a much better model performance than simply assuming constant correlations, but as these models never perform worse

Table 6: Estimators of AR model for daily European stock market returns

Index	AR(p)	
	a_0	a_1
DAX30	$2.9E-04$ $(3.6E{-}04)$	
CAC40	$-2.2E-05$ $(3.5E{-}04)$	-0.077 (0.025)
IBEX35	$1.7E-04$ $(3.8E{-}04)$	
MIB30	$-2.0E-04$ $(3.8E{-}04)$	

Note: Table 6 reports the estimated parameters and standard errors of the $AR(p)$ model for the log-returns of DAX30, CAC40, IBEX35 and MIB30 (daily returns, September 1, 2004 to August 31, 2010), where order p is selected with the BIC.

and sometime better than the CCC it is recommendable to use a dynamic correlation model for forecasting VaR.

3.2 Daily European stock market indices

The second dataset we use are daily returns of the CAC40, the DAX30, the IBEX35 and the MIB30 indices, resulting in 6 country pairs. Our sample ranges from September 1, 2004 until August 31, 2010 with a total of 1564 observations. The last 2 years, or 522 observations, constitute the out-of-sample period.

The results for the AR-GARCH model are presented in Tables 6 and 7. The results for the PLL that can be found in Table 8 show a different picture than the results for weekly data in Section 3.1. Although in most instance the RSC and SCAR models perform better than the CCC this difference is rarely statistically significant. In six cases the CCC is even the best performing forecasting model. The results for the construction of the MVP in Table 9 are in line with the ones for the weekly data. The CCC is clearly the best fitting model followed by RSC and SCAR. The SPA test shows that no model

Table 7: Estimators of GARCH model for daily European stock market returns

Index	Model	GARCH				
		ω	α	γ	β	**d.o.f.**
DAX30	EGARCH - Student	-0.316 (0.046)	-0.147 (0.018)	0.127 (0.024)	0.975 (0.004)	8.15 (1.68)
CAC40	EGARCH - Student	-0.299 (0.040)	-0.160 (0.017)	0.116 (0.023)	0.976 (0.004)	12.02 (3.40)
IBEX35	EGARCH - Student	-0.248 (0.038)	-0.126 (0.015)	0.127 (0.023)	0.983 (0.003)	7.00 (1.24)
MIB30	EGARCH - Student	-0.256 (0.035)	-0.126 (0.015)	0.131 (0.023)	0.982 (0.003)	8.81 (1.93)

Note: Table 7 reports the estimated parameters and standard errors of the appropriate *GARCH* model for the log-returns of DAX30, CAC40, IBEX35 and MIB30 corrected for autocorrelation (daily returns, September 1, 2004 to August 31, 2010), where the appropriate GARCH type model is selected with the BIC. Specifically, ω denotes the constant term, α stays for the coefficient of the error term, γ reflects the leverage effect, β is the coefficient of the autoregressive term and **d.o.f** are the degrees of freedom for models with Student t errors.

beats the CCC for any stock market pair or forecast horizon. Finally, the adequacy of the VaR forecasts as represented by the results of the DQ test in Table 10 seems to depend more on the data than on the model. Thus the VaR forecasts do not seem to be very precise, but this may also be a result of insufficiently modeled volatility and not only of the correlation forecasts.

Overall the results for the daily European stock index data indicate that it is very hard to beat the constant conditional correlation model and that dynamic correlation models do not in general outperform this benchmark.

4 Summary and Conclusion

In this paper we conducted a comparison of the forecasting performance of five time-varying correlation models. Unlike previous studies we focused on forecasting only correlation and not the covariance matrix, for which, prior to the analysis, we standardized the complete data set by filtering

Table 8: Forecast comparison using predictive log-likelihood for daily European stock market returns

Pair		h	CCC	cDCC	SDC	RSC	SCAR	SPA-pval
DAX30	CAC40	1	1.2105	1.1926	1.1440	1.2356	**1.2543**	0.0848
		4	1.2147	1.0934	1.0848	**1.2284**	1.2154	0.8008
		12	1.2084	1.0766	1.0222	1.2119	**1.2187**	0.6337
DAX30	IBEX35	1	**0.7001**	0.6789	0.5893	0.6760	0.6912	1.0000
		4	**0.6986**	0.6366	0.5413	0.6867	0.6807	1.0000
		12	**0.6901**	0.6423	0.4280	0.6702	0.6726	1.0000
DAX30	MIB30	1	0.8277	0.8170	0.6823	**0.8541**	0.8466	0.1873
		4	**0.8264**	0.7193	0.4386	0.8126	0.8057	1.0000
		12	**0.817**	0.6271	-0.0107	0.8143	0.8027	1.0000
CAC40	IBEX35	1	**0.8493**	0.8450	0.7828	0.8422	0.8488	1.0000
		4	0.8540	0.8462	0.7794	0.8662	**0.8672**	0.1392
		12	0.8434	0.8443	0.5680	**0.852**	0.8485	0.3628
CAC40	MIB30	1	0.9780	0.9846	0.9100	**0.9989**	0.9977	0.2359
		4	0.9823	0.9802	0.8888	**0.9908**	0.9882	0.3850
		12	0.9708	0.9839	0.6826	0.9868	**0.9907**	0.0128
IBEX35	MIB30	1	0.7704	0.7750	0.6584	**0.8063**	0.7964	0.0165
		4	0.7702	0.7802	0.5486	**0.7992**	0.7928	0.0122
		12	0.7572	0.7517	-0.0889	**0.7721**	0.7691	0.1322

Note: Table 8 compares the forecasting performance of the competing models by the predictive log-likelihood criterion and reports the SPA p-values for the null hypothesis that no model perform better than the CCC based on 10000 bootstrap replications. The data are daily returns from September 1, 2004 to August 31, 2010 standardized by the volatilities estimated using an appropriate GARCH model. The last 2 years of data constitute the out-of-sample period.

out time-varying volatility. Correlations were forecasted for both daily and weekly stock market returns over a horizon of 1, 4 and 12 periods using a rolling window approach. The models we considered were constant conditional correlation (CCC) of Bollerslev (1990), consistent dynamic conditional correlation (cDCC) of Aielli (2009), smooth dynamic correlation (SDC) of Hafner and Reznikova (2010), regime switching correlation (RSC) of Pelletier

Table 9: Forecast comparison using σ^2_{MVP} for daily European stock market returns

Pair		h	CCC	cDCC	SDC	RSC	SCAR	SPA-pval
DAX30	CAC40	1	0.0390	0.0388	0.0403	0.0382	**0.0378**	0.2735
		4	0.0392	0.0472	0.0447	**0.0383**	0.0407	0.6903
		12	0.0394	0.0511	0.0435	**0.0382**	0.0405	0.7307
DAX30	IBEX35	1	0.0393	0.0413	0.0431	**0.0378**	0.0397	0.6869
		4	0.0395	0.0451	0.0458	**0.0378**	0.0411	0.6027
		12	0.0396	0.0452	0.0469	**0.0372**	0.0410	0.4358
DAX30	MIB30	1	**0.0401**	0.0424	0.0465	0.0404	0.0410	1.0000
		4	**0.0403**	0.0499	0.0679	0.0421	0.0428	1.0000
		12	**0.0403**	0.0571	0.0774	0.0414	0.0421	1.0000
CAC40	IBEX35	1	**0.0430**	0.0435	0.0441	0.0434	0.0434	1.0000
		4	**0.0431**	0.0432	0.0443	0.0433	**0.0431**	1.0000
		12	**0.0428**	0.0433	0.0453	0.043	0.0430	1.0000
CAC40	MIB30	1	**0.0441**	0.0448	0.0453	0.0444	0.0443	1.0000
		4	0.0443	0.0443	0.0454	0.0442	**0.0441**	0.3583
		12	0.0443	**0.0441**	0.0483	0.0442	0.0442	0.5728
IBEX35	MIB30	1	**0.0439**	0.0458	0.0468	0.0441	0.0445	1.0000
		4	**0.0441**	0.0453	0.0474	**0.0441**	0.0442	0.5208
		12	**0.0441**	0.0452	0.0487	**0.0441**	0.0440	0.5787

Note: Table 9 compares the forecasting performance of the competing models by the variance of the global minimum variance portfolio (multiplied by 100) and reports the SPA p-values for the null hypothesis that no model perform better than the CCC based on 10000 bootstrap replications. The data are daily returns from September 1, 2004 to August 31, 2010 standardized by the volatilities estimated using an appropriate GARCH model. The last 2 years of data constitute the out-of-sample period.

(2006), and stochastic autoregressive correlation (SCAR) of Yu and Meyer (2006) and Hafner and Manner (2010). The out-of-sample performance of our candidate models was compared using the predictive log-likelihood, the variance of the global minimum variance portfolio and the 5% Value-at-Risk of an equally weighted portfolio. The significance of the difference in forecasting performance is tested using the test for superior predictive ability

Table 10: Dynamic quantile test for the adequacy of VaR forecasts for daily European stock market returns

Pair		h	CCC	cDCC	SDC	RSC	SCAR
DAX30	CAC40	1	**0.004**	**0.004**	**0.004**	**0.005**	**0.004**
		4	**0.004**	**0.004**	**0.004**	**0.004**	**0.004**
		12	**0.003**	**0.003**	**0.003**	**0.003**	**0.003**
DAX30	IBEX35	1	0.112	0.084	0.116	**0.017**	0.112
		4	0.116	0.085	0.115	**0.018**	0.085
		12	0.141	0.140	0.140	**0.008**	0.141
DAX30	MIB30	1	**0.011**	**0.004**	**0.010**	**0.004**	**0.010**
		4	**0.008**	**0.008**	**0.008**	**0.008**	**0.008**
		12	**0.008**	**0.008**	**0.008**	**0.008**	**0.008**
CAC40	IBEX35	1	0.126	0.125	0.204	0.094	0.126
		4	0.130	0.129	0.198	0.093	0.126
		12	0.157	0.231	0.231	0.157	0.157
CAC40	MIB30	1	**0.023**	**0.023**	**0.023**	**0.023**	**0.023**
		4	**0.022**	**0.022**	**0.034**	**0.022**	**0.022**
		12	**0.022**	**0.018**	**0.018**	**0.018**	**0.018**
IBEX35	MIB30	1	**0.035**	**0.043**	0.192	**0.043**	**0.043**
		4	**0.032**	**0.028**	0.176	**0.039**	**0.046**
		12	**0.032**	**0.029**	0.106	**0.038**	**0.038**

Note: Table 10 reports the p-values of the dynamic quantile test by Engle and Manganelli (2004) for the null hypothesis of a correct out-of-sample Value-at-Risk. The data are daily returns from September 1, 2004 to August 31, 2010 standardized by the volatilities estimated using an appropriate GARCH model. The last 2 years of data constitute the out-of-sample period.

(SPA) by Hansen (2005).

Our results show that when considering predictive log-likelihood (PLL) as the evaluation criterion dynamic correlation models only outperform the CCC for weekly international index returns. In particular, the RSC model performs best on average. Further, the SPA test by Hansen (2005) indicates that the difference in PLL between the CCC and the best performing models is statistically significant in almost all cases. For the daily European index returns, on the other hand, no model is able to systematically outperform the CCC in terms of PLL. The results for portfolio construction provide evidence in favor of the constant correlation model for both data sets. Although in some cases the competing models perform slightly better, the difference is not statistically significant. The forecasts of the Value-at-Risk do not depend so much on the correlation model, but on the data set, although the SDC model seem to perform slightly better than the other models. Finally, the forecast horizon did not have a clear impact on the results.

Concluding, it is difficult to beat a CCC model out-of-sample using an economic evaluation criterion, whereas most dynamic correlation models may offer significant improvements when considering statistical criteria for some data sets. Future research should consider data sets that allow the construction of realized correlation in order to have a good proxy for correlations. This would allow for direct comparison of the models under consideration and would allow for a wider range of loss functions. Furthermore, it should be of interest to look at different financial data such as exchange rate returns or commodity prices, as well as data for different sample periods and at different frequencies. Finally, it is worthwhile investigating whether exogenous variables that explain conditional correlations can be found and whether they can be used to improve forecast performance.

Acknowledgments

We thank Christian Hafner, Sébastien Laurent, Robert Vermeulen and Dennis Tuerk for helpful discussions and suggestions. Financial support from the contract "Projet d'Actions de Recherche Concertées" nr. 07/12/002 of the "Communauté française de Belgique", granted by the "Académie universitaire Louvain", is gratefully acknowledged.

References

G. P. Aielli. Dynamic conditional correlations: On properties and estimation. Available at ssrn: http://ssrn.com/abstract=1507743, 2009.

T. G. Andersen and N. Shephard. Stochastic volatility: Origins and overview. In T. G. Andersen, R. Davis, J.-P. Kreiss, and T. Mikosch, editors, *Handbook of Financial Time Series*. 2009. Forthcoming.

T. G. Andersen, T. Bollerslev, P. F. Christoffersen, and F. X. Diebold. Volatility and correlation forecasting. In C. G. G. Elliot and A. Timmermann, editors, *Handbook of Economic Forecasting*, pages 778–878, Amsterdam: North-Holland, 2006.

M. Asai, M. McAleer, and J. Yu. Multivariate stochastic volatility: a review. *Econometric Reviews*, 25:145–175, 2006.

L. Bauwens, S. Laurent, and J. V. K. Rombouts. Multivariate garch models: a survey. *Journal of Applied Econometrics*, 21:79–109, 2006.

T. Bollerslev. Modelling the coherence in short-run nominal exchange rates: a multivariate generalized arch approach. *Review of Economics and Statistics*, 72:498–505, 1990.

M. Caporin and M. McAleer. Ranking multivariate garch models by problem dimension. Available at ssrn: http://ssrn.com/abstract=1601236, 2010.

L. K. C. Chan, J. Karceski, and J. Lakonishok. On portfolio optimisation: Forecasting covariances and choosing the risk model. *The Review of Financial Studies*, 12:937–974, 1999.

R. F. Engle. Autoregressive conditional heteroskedasticity with estimates of the variance of u.k. inflation. *Econometrica*, 50:987–1008, 1982.

R. F. Engle. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business and Economic Statistics*, 20:339–350, 2002.

R. F. Engle and S. Manganelli. CAViaR: Conditional autoregressive value at risk by regression quantiles. *Journal of Business and Economic Statistics*, 22:367–381, 2004.

R. F. Engle and K. Sheppard. Theoretical and empirical properties of dynamic conditional correlation multivariate garch. NBER Working Paper 8554, National Bureau of Economic Research, 2001.

C. B. Erb, C. R. Harvey, and T. E. Viskante. Forecasting international equity correlations. *Financial Analysts Journal*, 50:32–45, 1994.

C. M. Hafner and H. Manner. Dynamic stochastic copula models: Estimation, inference and applications. *Journal of Applied Econometrics*, 2010. forthcoming.

C. M. Hafner and O. Reznikova. Efficient estimation of a semiparametric dynamic copula model. *Computational Statistics and Data Analysis*, 54: 2609–2627, 2010.

C. M. Hafner, D. van Dijk, and P. H. Franses. Semiparametric modelling of correlation dynamics. In T. Fomby and C. Hill, editors, *Advances in Econometrics*, volume 20, pages 59–103, Department of Economics, Louisiana State University, 2006.

J. Hamilton. *Time Series Analysis*. Princeton University Press, 1994.

P. R. Hansen. A test for superior predictive ability. *Journal of Business and Economic Statistics*, 23:365–380, 2005.

P. R. Hansen and A. Lunde. A forecast comparison of volatility models: Does anything beat a garch(1,1)? *Journal of Applied Econometrics*, 20: 873–889, 2005.

P. R. Hansen and A. Lunde. Consistent ranking of volatility model. *Journal of Econometrics*, 131:97–121, 2006.

S. Laurent, J. V. K. Rombouts, and F. Violante. Consistent ranking of multivariate volatility models. CORE Discussion Papers 2009002, Universit catholique de Louvain, 2009. URL `http://ideas.repec.org/p/cor/louvco/2009002.html`.

F. Longin and B. Solnik. Is the correlation in international equity returns constant: 1960-1990. *Journal of International Money and Finance*, 14: 3–26, 1995.

A. Patton and K. Sheppard. Dependence properties in risk management: Properties and pitfalls. In T. G. Andersen, R. A. Davis, J.-P. Kreiss, and T. Mikosch, editors, *Handbook of Financial Time Series*, Springer Verlag, 2009.

D. Pelletier. Regime switching for dynamic correlations. *Journal of Econometrics*, 131:445–473, 2006.

D. N. Politis and J. P. Romano. The stationary bootstrap. *Journal of the American Statistical Association*, 89:1303–1313, 1994.

J. Yu and R. Meyer. Multivariate stochastic volatility models: Bayesian estimation and model comparison. *Econometric Reviews*, 25:361–384, 2006.