

Semi-Parametric Modelling of Correlation Dynamics

Christian M. Hafner, Dick van Dijk and Philip Hans Franses*

*Econometric Institute
Erasmus University Rotterdam*

July 2005

ECONOMETRIC INSTITUTE REPORT EI 2005-26

Abstract

In this paper we develop a new semi-parametric model for conditional correlations, which combines parametric univariate GARCH-type specifications for the individual conditional volatilities with nonparametric kernel regression for the conditional correlations. This approach not only avoids the proliferation of parameters as the number of assets becomes large, which typically happens in conventional multivariate conditional volatility models, but also the rigid structure imposed by more parsimonious models, such as the dynamic conditional correlation model. An empirical application to the 30 Dow Jones stocks demonstrates that the model is able to capture interesting asymmetries in correlations and that it is competitive with standard parametric models in terms of constructing minimum variance portfolios and minimum tracking error portfolios.

Keywords: Multivariate GARCH, dynamic conditional correlation, kernel regression, minimum variance portfolio, tracking error minimization.

JEL Classification: C14, C32, G11.

*Address for correspondence: Dick van Dijk, Econometric Institute, Erasmus University Rotterdam, P.O. Box 1738, NL-3000 DR Rotterdam, The Netherlands, e-mail: djvandijk@few.eur.nl

1 Introduction

Estimating and forecasting the (conditional) covariance structure of financial asset returns plays a crucial role in important areas of finance such as portfolio construction, asset pricing and risk management. It is therefore not surprising that modelling the dynamics of conditional covariance matrices has received ample attention. This research area was boosted by the introduction of the (Generalized) AutoRegressive Conditional Heteroskedasticity ((G)ARCH) model by Engle (1982) and Bollerslev (1986). Early-stage multivariate GARCH models, including the VEC-model of Bollerslev, Engle and Wooldridge (1988) and the BEKK-model of Engle and Kroner (1995), describe the dynamics of all elements of the covariance matrix in a flexible way using many parameters. As such, they suffer from the “curse of dimensionality”, meaning to say that the number of parameters to be estimated in these models increases very rapidly as the number of included assets increases.¹ This basically prevented their successful application in empirically relevant settings, where portfolios might consist of tens or even hundreds of assets. More parsimonious models, such as the Constant Conditional Correlation (CCC) model of Bollerslev (1990), are applied more frequently in practice. These models limit the possible dynamic patterns of the covariance matrix in important ways, however. In the CCC model, for example, the covariance between two assets changes over time only because of variation in their conditional volatilities, as their correlation is assumed to be constant.

Recently, several extensions of the CCC model have been proposed that allow for variation in the conditional correlations. On the one hand, Engle (2002) and Tse and Tsui (2002) developed Dynamic Conditional Correlation (DCC) models, where the conditional correlations evolve according to a GARCH-type structure. The attractive feature of this approach is that the number of parameters in the conditional correlation model can be limited by using the idea of “correlation targeting”, which means that the unconditional correlations implied by the model are restricted to be equal to the unconditional sample correlations. In that case, the basic DCC-model of Engle (2002), for example, involves only two unknown parameters. Cappiello, Engle and Sheppard (2003) and Hafner and Franses (2003) considered

¹The number of parameters in an unrestricted VEC-model for N assets equals $(N(N+1)/2) \times (1 + 2N(N+1)/2)$. The BEKK-model contains $2N^2 + N(N+1)/2$ parameters.

generalizations to allow for richer correlation dynamics, while still keeping the number of parameters within reasonable bounds even for large numbers of assets.² On the other hand, Pelletier (2005) and Silvennoinen and Teräsvirta (2005) proposed Regime-Switching Conditional Correlation models, where the correlations switch back and forth between a limited number of different values, according to an unobserved Markov-Switching process or according to the value of observed exogenous variables, respectively. The main disadvantage of these models is that correlation targeting can no longer be applied in a straightforward manner, such that the number of parameters to be estimated again grows rapidly (that is, quadratically) with the number of assets.

In this paper we put forward a new semi-parametric model for the conditional covariance matrix that is flexible yet easy to estimate even in high dimensions. The model is semi-parametric in the sense that the conditional variances are described parametrically, for example using standard univariate GARCH-type models, while the conditional correlations are estimated using nonparametric techniques. The conditional correlations are assumed to depend on an observable exogenous (or pre-determined) variable. The appropriate choice of this conditioning variable depends on the particular application. For example, in Section 4 below, we model the correlations between the 30 individual stocks included in the Dow Jones Industrial Average (DJIA) index, and allow these to depend on the market volatility and market return. These variables are motivated by both theoretical and empirical considerations. First, as shown by Andersen et al. (2001), in case asset returns exhibit a factor structure, the correlations are affected by the conditional volatility of the latent factor. In particular, under quite realistic assumptions concerning the factor loadings high factor volatility induces high correlations between individual asset returns. In the context of the DJIA stocks, a factor structure is a sensible possibility and, furthermore, the market portfolio is a natural candidate to use as a proxy for the latent common factor. Additional empirical evidence supporting the hypothesis that correlations increase in high volatility states can be found in Longin and Solnik (1995, 2001), and Ramchand and Susmel (1998), but see Loretan and English (2000) and Forbes and Rigobon (2002) for critical discussion. Second, modelling the

²As discussed in Section 2 below, the number of parameters in these models grows linearly with the number of assets.

correlations as a function of the lagged index return is motivated by the literature on “correlation breakdown”, which has documented that typically asset correlations increase in bear markets, but are not affected (or to a much lesser extent) in bull markets, see Longin and Solnik (2001), Ang and Chen (2002), Butler and Joaquin (2002), and Campbell, Koedijk and Kofman (2002), among others.

The idea of allowing the conditional correlations to depend on exogenous factors resembles the regime-switching model of Silvennoinen and Teräsvirta (2005). However, instead of *a priori* imposing a certain parametric specification for the correlation functions, we suggest a nonparametric estimator. Basically, a Nadaraya-Watson kernel regression is applied to each conditional correlation individually, using the same bandwidth to guarantee that the resulting estimator of the correlation matrix is positive definite. In order to obtain a genuine correlation matrix in finite samples, we apply the same transformation as in the DCC model of Engle (2002). Our nonparametric estimator only requires the assumption that correlations are smooth functions of the conditioning variable, while the data completely determines the appropriate shape of the functional relationship. An additional advantage of is that the individual correlation functions are allowed to exhibit quite different shapes, which is problematic in currently available parametric specifications as will become clear below.

The paper proceeds as follows. In Section 2, we briefly review the various conditional correlation models that have recently been proposed. We describe our new semi-parametric model in Section 3, including a detailed discussion of the issues involved in the nonparametric part of the estimation procedure. In Section 4, we consider daily returns of the 30 stocks included in the Dow Jones Industrial Average Index over the period 1989-2003, and apply several models to describe the dynamics in the conditional covariance matrix of these stocks. We find that our new semi-parametric model is competitive with rival specifications, in particular in terms of tracking error minimization. We conclude in Section 5, also pointing out interesting directions for future research. The Appendix contains a proof of the consistency and asymptotic normality of the semi-parametric estimator developed in Section 3.

2 Dynamic conditional correlation models

In this section we review existing models for describing the dynamics in conditional correlations of asset returns. These models are a specific sub-group within the general class of multivariate GARCH models. We refer to Bauwens, Laurent and Rombouts (2005) for elaborate discussion of other models in this class.

Let r_t denote an N -dimensional vector time series, such as daily returns on the stocks in the Dow Jones index. Suppose for simplicity (but without loss of generality) that the conditional mean of r_t is constant and equal to zero, but that its conditional covariance matrix is time-varying. That is, we have the generic model

$$\mathbb{E}[r_t|\mathcal{F}_{t-1}] = 0, \quad (1)$$

$$\mathbb{V}(r_t|\mathcal{F}_{t-1}) = \mathbb{E}[r_t r_t'|\mathcal{F}_{t-1}] = H_t, \quad (2)$$

where \mathcal{F}_t is the information set that includes all information up to and including time t . The conditional covariance matrix H_t can be decomposed as

$$H_t = D_t(\theta) R_t D_t(\theta), \quad (3)$$

with $D_t(\theta) = \text{diag}(\sqrt{h_{1t}}, \dots, \sqrt{h_{Nt}})$ a diagonal matrix with the square root of the conditional variances h_{it} , parameterized by the vector θ , on the diagonal. For example, h_{it} can be described by the univariate GARCH-model of Bollerslev (1986) or nonlinear extensions such as the GJR-GARCH model of Glosten, Jagannathan and Runkle (1993). The matrix R_t , with the (i, j) -th element denoted as ρ_{ijt} , is the possibly time-varying conditional correlation matrix. As noted by Engle (2002), R_t also is the conditional covariance matrix of the standardized disturbances $\varepsilon_t = D_t^{-1}(\theta)r_t$, that is

$$\mathbb{E}[\varepsilon_t \varepsilon_t'|\mathcal{F}_{t-1}] = R_t. \quad (4)$$

Different specifications of R_t give rise to different models. First, the constant conditional correlation (CCC) model of Bollerslev (1990) assumes that $R_t = R$ is constant. In that case, estimation is straightforward and can be performed in two steps. First, one estimates univariate GARCH models for the individual series r_{it} , $i = 1, \dots, N$, using (quasi) maximum likelihood to obtain estimates of the conditional variances h_{it} . Second, R can be consistently estimated by the sample covariance

matrix of the standardized residuals $\hat{\varepsilon}_t = \hat{D}_t^{-1} r_t$, that is $\hat{R} = T^{-1} \sum_{t=1}^T \hat{\varepsilon}_t \hat{\varepsilon}_t'$ with T denoting the sample size.

Second, Engle (2002) and Tse and Tsui (2002) introduced the class of dynamic conditional correlation (DCC) models, where R_t is allowed to vary according to a GARCH-type process, such as

$$Q_t = (1 - \alpha - \beta)\bar{Q} + \alpha\varepsilon_{t-1}\varepsilon_{t-1}' + \beta Q_{t-1}, \quad (5)$$

$$R_t = Q_t^*{}^{-1} Q_t Q_t^*{}^{-1}, \quad (6)$$

where $\bar{Q} = E[\varepsilon_t \varepsilon_t']$ is the unconditional covariance matrix of standardized returns, Q_t^* is a diagonal matrix containing the square roots of the diagonal elements of Q_t and α and β are scalars. The attractive feature of the DCC model is not only that two-step estimation is still feasible, but also that “correlation targeting” can be used in the second step. That is, after estimating univariate GARCH models for the conditional volatilities h_{it} in the first step, \bar{Q} can be replaced by the sample covariance matrix of the standardized residuals $\hat{\varepsilon}_t$. This imposes the restriction that the unconditional correlations as implied by the model are equal to the unconditional sample correlations, and reduces the number of parameters to be estimated in the second step to two, namely α and β . Engle and Sheppard (2001) prove consistency and asymptotic normality of this two-step estimator. Note that correlation targeting is the multivariate analogue of variance targeting in the univariate GARCH framework as introduced by Engle and Mezrich (1996).

The DCC model in (5) has two particular features that may render it too restrictive in practice. First, the model imposes all correlations to have the same dynamic pattern as governed by the parameters α and β . It is not difficult to come up with examples of situations where this restriction may be violated. For example, correlations between stocks from the same industry may behave rather differently from correlations between stocks from different industries. Second, the model implies symmetry, in the sense that a pair of positive standardized returns ε_{it} and ε_{jt} has the same effect on the conditional correlation as a pair of negative returns of the same magnitude. Recent empirical evidence in Ang and Chen (2002), among others, indicates the presence of asymmetries in correlations, suggesting for example that correlations increase for large downward stock price movements but not, or less, for large upward movements.

Several extensions of the basic model in (5) have been developed that alleviate these shortcomings. Hafner and Franses (2003) suggested the semi-generalized DCC (SGDCC) model that allows for asset-specific news impact parameters by replacing (5) with

$$Q_t = (1 - \bar{\alpha}^2 - \beta)\bar{Q} + \alpha\alpha' \odot \varepsilon_{t-1}\varepsilon_{t-1}' + \beta Q_{t-1}, \quad (7)$$

where \odot denotes the Hadamard product and α now is an $(N \times 1)$ vector, $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_N)'$, and $\bar{\alpha} = N^{-1} \sum_{i=1}^N \alpha_i$ is the average news impact parameter. The memory parameter β can be made asset-specific in identical fashion. Note that the number of parameters in (7) increases linearly with the number of assets N . Cappiello et al. (2003) allowed for nonlinear impact of shocks on the conditional correlations, in particular of simultaneous negative shocks in assets i and j , by considering the asymmetric DCC (ADCC) model

$$Q_t = (1 - \alpha - \beta)\bar{Q} - \gamma\bar{N} + \alpha\varepsilon_{t-1}\varepsilon_{t-1}' + \beta Q_{t-1} + \gamma n_{t-1}n_{t-1}', \quad (8)$$

where $n_t = \mathbb{I}[\varepsilon_t < 0] \odot \varepsilon_t$, with $\mathbb{I}[A]$ being the indicator function for the event A , and $\bar{N} = \mathbb{E}[n_t n_t']$. In estimation of the ADCC model, \bar{N} can be replaced by its sample analogue $T^{-1} \sum_{t=1}^T \hat{n}_t \hat{n}_t'$, such that the number of parameters is equal to three independent of the number of assets included. Combining the ideas in (7) and (8) is also possible, see Cappiello et al. (2003).

Third, Pelletier (2005) developed a Markov-Switching conditional correlation model that allows the conditional correlations to switch between k distinct values, by specifying the conditional correlation matrix in (3) as

$$R_t = R_{s_t}, \quad (9)$$

where s_t is an unobserved first-order Markov process with k states, with transition probabilities $P[s_t = j | s_{t-1} = i] = p_{ij}$, for $i, j = 1, \dots, k$. Note that in contrast to the DCC-type models discussed above, this model does not necessarily imply any direct dynamic dependence in the conditional correlation matrix, although there is indirect dependence through the Markov-character of s_t . Estimation of the parameters in this model can again be done using a two-step approach, where in the second step the EM-algorithm can be employed, see Pelletier (2005) for details. Note however that correlation targeting is not possible here. Hence, the elements in the R_{s_t} , $s_t =$

$1, \dots, k$ matrices have to be treated as unknown parameters such that the number of parameters to be estimated in the second step equals $k(N(N-1)/2 + (k-1))$.

Fourth, Silvennoinen and Teräsvirta (2005) consider a smooth transition conditional correlation (STCC) model that allows the conditional correlations to vary continuously between two extremes, according to the value of an observed variable x_t ,

$$R_t = R_1(1 - G(x_t; \gamma, c)) + R_2G(x_t; \gamma, c), \quad (10)$$

where the function $G(x_t; \gamma, c)$ is continuous and bounded between 0 and 1. For example, $G(x_t; \gamma, c)$ can be the logistic function

$$G(x_t; \gamma, c) = \frac{1}{1 + \exp(-\gamma(x_t - c))}, \quad (11)$$

which tends to 0 (1) for very small (large) values of the transition variable x_t , relative to the threshold parameter c . The parameter γ determines the smoothness of the change from 0 to 1 as x_t increases. The transition variable x_t should of course be included in the information set \mathcal{F}_{t-1} to make the model operational and useful for forecasting purposes. Although the parameters in this model can again be estimated using a two-step approach, correlation targeting is not possible, such that the number of parameters to be estimated in the second step equals $2N(N-1)/2+2$. An interesting special case of the STCC model (10) with (11) arises when the parameter $\gamma \rightarrow \infty$, such that the change of the correlation matrix from R_1 and R_2 occurs instantaneously at $x_t = c$. In that case, estimation of the parameters is simplified considerably, in the sense that correlation targeting becomes possible again. Note that, if $\gamma = \infty$ such that $G(x_t; \gamma, c) = \mathbb{I}[x_t \geq c]$, $R_1 = \mathbb{E}[n_{1t}n'_{1t}]$ with $n_{1t} = \mathbb{I}[x_t < c] \odot \varepsilon_t$ and R_2 can be similarly defined as $R_2 = \mathbb{E}[n_{2t}n'_{2t}]$ with $n_{2t} = \mathbb{I}[x_t \geq c] \odot \varepsilon_t$. For a fixed value of c , R_1 and R_2 can then be replaced by their sample analogues $(\sum_{t=1}^T \mathbb{I}[x_t < c])^{-1} \sum_{t=1}^T \hat{n}_{1t}\hat{n}'_{1t}$ and $(\sum_{t=1}^T \mathbb{I}[x_t \geq c])^{-1} \sum_{t=1}^T \hat{n}_{2t}\hat{n}'_{2t}$, respectively. This leaves only c to be estimated, which can be achieved by means of grid search.

A crucial ingredient of the STCC model is of course the transition variable x_t . Berben and Jansen (2005) take $x_t = t$ in an application to industry stock returns, effectively modelling a gradual structural change in correlations. Alternatively, x_t could be taken to be a lagged market return or volatility, motivated by the correlation breakdown effect and by the presence of a factor structure, respectively, as discussed in the Introduction. Silvennoinen and Teräsvirta (2005) develop an LM-type test

for the constancy of the correlation matrix against the STCC-alternative, which also can be used as a guiding tool in the choice of the appropriate transition variable x_t .

3 A semi-parametric conditional correlation model

In this section we put forward our semi-parametric conditional correlation model. The model can be considered a generalization of the STCC model of Silvennoinen and Teräsvirta (2005) given in (10) and (11). We are sympathetic to the idea that the conditional correlations depend on exogenous factors such as the market return or volatility. However, the STCC model is restrictive, in the sense that it imposes that all correlations depend on x_t in the same way, as determined by the transition function $G(x_t; \gamma, c)$. This assumption may be unrealistic in practice, as one can easily imagine that different correlations respond differently to changes in x_t ; see the factor model in Andersen et al. (2001) and the empirical application in Silvennoinen and Teräsvirta (2005) for examples. It might be possible to generalize the STCC model to allow for asset-specific smooth transition in correlations, similar to the SGDCC model of Hafner and Franses (2003) in (7). In that case, however, it would be difficult, if not impossible, to guarantee positive definiteness of the correlation matrix for all values of x_t . In addition, even if it is reasonable to assume that all correlations have the same functional relationship to x_t , the specific form of this dependence, such as the logistic function in (11), is not likely to be known *a priori*. Hence, we suggest to model the conditional correlation matrix in a nonparametric way.

In sum, we propose the following semi-parametric model:

$$r_t = D_t(\theta)\varepsilon_t, \quad (12)$$

with $D_t(\theta)$ as defined before, $\mathbb{E}[\varepsilon_t \mid \mathcal{F}_{t-1}] = 0$ and

$$\mathbb{E}[\varepsilon_t \varepsilon_t' \mid \mathcal{F}_{t-1}, x_t = x] = R(x), \quad (13)$$

where x_t is an observed variable, for example, the market return or volatility at time $t - 1$. Assuming that we have a \sqrt{T} -consistent estimator of θ , which we denote by $\hat{\theta}$, the standardized residuals are then defined by $\hat{\varepsilon}_t = D_t(\hat{\theta})^{-1}r_t$. These are used in the second stage to estimate $R(x)$ nonparametrically, as

$$\hat{R}(x) = \hat{Q}^*(x)^{-1} \hat{Q}(x) \hat{Q}^*(x)^{-1} \quad (14)$$

where

$$\widehat{Q}(x) = \frac{\sum_{t=1}^T \widehat{\varepsilon}_t \widehat{\varepsilon}_t' K_h(x_t - x)}{\sum_{t=1}^T K_h(x_t - x)} \quad (15)$$

where $K_h(\cdot) = (1/h)K(\cdot/h)$, K is a kernel function and h a bandwidth parameter. As before, $\widehat{Q}^*(x)$ is a diagonal matrix with the square roots of the diagonal elements of $\widehat{Q}(x)$ on its diagonal. Hence the (i, j) -th element of $\widehat{R}(x)$ in (14) can be written as

$$\widehat{\rho}_{ij}(x) = \frac{\widehat{q}_{ij}(x)}{\sqrt{\widehat{q}_{ii}(x)\widehat{q}_{jj}(x)}}.$$

The estimator $\widehat{R}(x)$ in (14) is essentially a transformed Nadaraya-Watson estimator applied to the elements of the matrix $\widehat{\varepsilon}_t \widehat{\varepsilon}_t'$. Of course, other nonparametric estimators such as local polynomials could be used as well, see e.g. Fan and Gijbels (1996), Härdle, Lütkepohl and Chen (1997) and Pagan and Ullah (1999) for reviews, and Fan and Yao (1998) and Ziegelmann (2002) for applications to volatility modelling. Recall that the conditional covariance matrix of r_t can be written as

$$H_t = \mathbf{V}(r_t \mid \mathcal{F}_{t-1}, x_t = x) = D_t(\theta)R(x)D_t(\theta),$$

where the semi-parametric character of the model for H_t now becomes obvious: $D_t(\theta)$ is modelled parametrically using standard univariate GARCH-type models, for example, while the correlation matrix $R(x)$ is treated in a nonparametric fashion.

Note that (15) essentially boils down to univariate kernel regression. Hence, unlike many of the early-stage multivariate GARCH models and unlike many estimation problems in nonparametric analysis, our semi-parametric model does not suffer from the curse of dimensionality. It is straightforward to see that (14) generalizes the STCC model in (10) as we do not impose that all correlations are related in the same way to x_t , nor do we assume any particular parametric form for this dependence. Also recall that one is likely to encounter difficulties in estimation when applying the STCC model to large systems given that correlation targeting is not possible. Using the non-parametric estimator in (14) and (15) this problem obviously is avoided. Note that this also suggests that our semi-parametric model might be used in exploratory data analysis to examine the shape of the dependence of ρ_{ijt} on x_t .

Apart from the choice of x_t , the crucial decision to be made when implementing the non-parametric estimator of the conditional correlation concerns the bandwidth

h . One option would be to use a constant bandwidth, which could be selected using cross-validation or the plug-in method, see e.g. Pagan and Ullah (1999). This has the drawback that the variance of $R(x)$ becomes very large near the boundaries of the support of x_t or more generally in areas of the support of x_t where the data is sparse. To alleviate the problem of data sparsity we suggest to use local bandwidths. In particular, we allow h to depend on the design density $f(x)$ such that $h(x) = bf(x)^{-a}$, where b is a positive constant and $0 \leq a \leq 1$. For positive values of a , this implies that the bandwidth becomes larger when relatively few observations around the point $x_t = x$ are available such that $f(x)$ is small, while $h(x)$ becomes smaller in high density regions. For $a = 0$ one has the standard kernel smoother with constant bandwidth $h(x) = b$. Jennen-Steinmetz and Gasser (1988) showed that $a = 1/4$ corresponds roughly to spline smoothing and $a = 1$ to generalized nearest-neighbor smoothing. In the empirical application in the next section, we will illustrate the sensitivity of the estimates of $R(x)$ to different values of a .

In Theorem 1 below we state the asymptotic properties of the nonparametric estimator of the conditional correlation matrix as given (14) and (15). Using the notation $\eta_t = \text{vech}(\varepsilon_t \varepsilon_t')$ and $r(x) = \text{vech}(R(x))$, we have

$$r(x) = \mathbb{E}[\eta_t \mid x_t = x] \quad (16)$$

$$\mathbf{V}(\eta_t \mid x_t = x) = \mathbb{E}[(\eta_t - r(x))(\eta_t - r(x))' \mid x_t = x] \quad (17)$$

Also, denote by $\hat{f}(x)$ an estimator of $f(x)$, the density of x_t . Under the assumptions stated in the appendix, the nonparametric estimator of $R(x)$ is consistent and asymptotically normal, as detailed in the following theorem.

Theorem 1 *Under Assumptions (A1) to (A6), and $b \rightarrow 0$ as T increases such that $Tb^5 \rightarrow 0$, it holds that*

1. *The diagonal elements of the matrix \hat{Q} in (15) converge in probability to 1:*

$$\hat{q}_{ii}(x) \xrightarrow{p} 1, \quad i = 1, \dots, N.$$

2. *The estimator $\hat{R}(x)$ in (14) is consistent and asymptotically normal:*

$$\sqrt{Tb}(\hat{r}(x) - r(x)) \xrightarrow{\mathcal{L}} N(0, \Sigma(x))$$

where

$$\Sigma(x) = \frac{\int K^2(u)du}{f(x)^{1-a}} V(\eta_t | x_t = x)$$

The proof of the theorem is given in the Appendix. Also, an explicit expression for $V(\eta_t | x_t = x)$ is given there, which depends on $R(x)$ and on the fourth moment characteristics of the innovations ε_t . If these are assumed to be normally distributed, then a consistent estimator of $V(\eta_t | x_t = x)$ is easily constructed by replacing $R(x)$ by $\hat{R}(x)$. If the distribution is unknown, then fourth moments of ε_t can be estimated using corresponding moments of the standardized residuals $\hat{\varepsilon}_t$, but for consistency we would need the assumption of finite eight moments.

4 Empirical application

In this section we explore the potential of our semi-parametric correlation model to describe and forecast the conditional covariance matrix of asset returns. We are mainly interested in empirically relevant situations where the number of assets N is fairly large, in which it is difficult to apply early-stage multivariate GARCH models and the regime-switching correlation models (9) and (10). We use daily returns of the 30 stocks that constituted the Dow Jones Industrial Average (DJIA) index between November 1999 and April 2004. Several conditional correlations models are estimated for these stock returns. We evaluate and compare the models not only by means of statistical criteria, but mainly by applying the models for constructing minimum variance portfolios and minimum tracking error volatility portfolios.

4.1 Data

The sample period covers 15 years, running from January 1, 1989 until December 31, 2003. Days on which the stock exchange was closed were removed from the sample, leaving a sample size of $T = 3784$ observations. We use the first decade 1989-1998 (2528 observations) for in-sample estimation and analysis of competing models, and set aside the final five years 1999-2003 (1256 observations) for out-of-sample forecasting.

Part of the challenge in this application stems from the fact that the general trend in the stock market was quite different during the initial estimation period and the forecasting period. The 1990's were characterized by a prolonged bull market, which

ended with the burst of the internet bubble and was followed by a bearish market around the turn of the millennium. Tables 1 and 2 show annualized means and standard deviations of the daily stock returns, which obviously reflect the overall market sentiment. All 30 stocks had large positive average returns during the first decade of our sample period, while the average return turned negative for 18 stocks during the final five years. Only the stocks of Alcoa Inc. (AA) and 3M Company (MMM) had higher returns during the period 1999-2003 than during the preceding 10 years. It is also seen that the volatility of stock returns was considerably higher during the last five years of the sample period. For most stocks the standard deviation increased by about 50%. Tables 1 and 2 also contain estimation results from the three-factor regression model developed by Fama and French (1993, 1996), given by

$$r_{it} - r_{f,t} = \alpha + \beta_{i,M}(r_{M,t} - r_{f,t}) + \beta_{i,SMB}r_{SMB,t} + \beta_{i,HML}r_{HML,t} + \varepsilon_t, \quad (18)$$

where r_{it} is the daily return for stock i , $r_{f,t}$ is the risk-free rate, $r_{M,t}$ is the market portfolio return, $r_{SMB,t}$ (Small-Minus-Big) and $r_{HML,t}$ (High-Minus-Low) are size and book-to-market factor portfolio returns. The size factor return $r_{SMB,t}$ is computed as the difference between the returns on portfolios of stocks with small and large market capitalization, while $r_{HML,t}$ is obtained similarly using portfolios of stocks with high and low book-to-market values. We employ the daily market and factor returns available on the website of Kenneth French.³ Most of the DJIA stocks move one-to-one with the market, given that the estimates of $\beta_{i,M}$ are fairly close to 1. Exceptions include Johnson and Johnson (JNJ) and JP Morgan (JPM), for which we find very low and high market betas, respectively. In general, the difference in $\beta_{i,M}$ during the two sub-periods is not very large. The negative estimates of $\beta_{i,SMB}$ are not surprising of course, given the size of the DJIA stocks. For the majority of stocks we find that $\beta_{i,HML}$ was substantially higher during the period 1999-2003 than during the years 1989-1998. We return to these estimates below.

– Tables 1 and 2 about here –

The increase in volatility mentioned above is also evident from Table 3, which shows summary statistics of the distribution of daily stock return variances, covariances and correlations. It is seen that the average daily stock return variance doubled

³http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

during the last five years of the sample period, from just over 3% to more than 6%. The same holds for the average covariance between the 30 DJIA stocks, which increased from 0.83 to 1.69. Interestingly, the standard deviation of the variances also doubled, while the standard deviation of the covariances even tripled. This leads to the conclusion that probably the average correlation did not change much, but the spread in the correlations increased considerably. Indeed, as shown in the final column of Table 3, the average correlation in both subperiods was equal to 0.28, but the standard deviation of the correlations jumped from 0.066 to 0.12.

– Table 3 about here –

4.2 Models

For the implementation of all conditional correlation models considered (except Riskmetrics as discussed below), we use the asymmetric GJR-GARCH model of Glosten et al. (1993) to describe the conditional volatilities of the individual daily stock returns. In this model, the conditional volatility of r_{it} evolves according to

$$h_{it} = \omega_i + \alpha_i r_{i,t-1}^2 + \gamma_i r_{i,t-1}^2 \mathbb{I}[r_{i,t-1} < 0] + \beta_i h_{i,t-1}, \quad (19)$$

allowing positive and negative lagged returns to have a different impact on current volatility. Typically, $\gamma_i > 0$ for stock returns, such that negative returns have a larger effect on volatility than positive returns of the same magnitude. Results obtained with symmetric GARCH(1,1) and with asymmetric power GARCH (AP-GARCH) volatility models are qualitatively similar to the ones reported below, and are available upon request.

We examine two choices for the conditioning variable in the nonparametric correlation estimator (15): the logarithm of the contemporaneous daily market volatility and the one-day lagged weekly market return, denoted as $\log(h_{M,t})$ and $r_{M,w,t-1}$, respectively. As explained in the Introduction, these variables are motivated by the empirical observations that correlations increase when market volatility is high and when market returns are low. Market returns are measured by the returns on a value-weighted portfolio of all NYSE, AMEX, and NASDAQ stocks, again taken from Kenneth French’s website. We are able to employ the logged contemporaneous market volatility $\log(h_{M,t})$ as conditioning variable x_t and still use the model

for forecasting purposes by estimating a GJR-GARCH model (19) for daily market returns and obtaining the fitted value $\hat{h}_{M,t}$. Note that this can be constructed using only information dated $t - 1$ and earlier.

We estimate our semi-parametric correlation model for different choices of the tuning parameter a , to examine the sensitivity of the correlation estimates with respect to the use of local bandwidths, as discussed in Section 3. To make the nonparametric estimates comparable for different values of a , we set the constant b equal to 1.50 for $a = 0$, and determine the value of b for other values of a such that the average bandwidth $T^{-1} \sum_{t=1}^T h(x_t)$ is the same for all values of a . Throughout we use the quartic kernel function $K(u) = \frac{15}{16}(1 - u^2)^2 I(|u| \leq 1)$ and a standard kernel density estimator with plug-in bandwidth for the density of x_t .

For comparison, we also estimate the CCC model with $R_t = R$, the standard DCC model (5) and (6), the asymmetric DCC model (8), and the STCC model (10) also using the logged contemporaneous daily market volatility and one-day lagged weekly market return as transition variable x_t in the logistic function (11). We also consider the RiskMetrics (RM) model which specifies the conditional covariance matrix as

$$H_t = (1 - \lambda)r_{t-1}r'_{t-1} + \lambda H_{t-1},$$

with $\lambda = 0.94$. Finally, following Chan, Karceski and Lakonishok (1999) we implement single-factor (SF) and three-factor (TF) models. The three-factor model uses the excess return on the value-weighted market portfolio and the size and book-to-market factor returns, and effectively boils down to implementing (18) with the raw stock return r_{it} as the dependent variable. The conditional covariance matrix of the stock returns is then given by

$$H_t = B\Omega_t B' + S_t, \tag{20}$$

where B is the matrix of factor loadings of the stocks, Ω_t is the conditional covariance matrix of the factors, and S_t is a diagonal matrix containing the idiosyncratic conditional variances. For convenience, we assume that the factors are uncorrelated, such that Ω_t also is diagonal, and that the conditional variance of each factor can be adequately described by means of a GJR-GARCH model (19). The same model is also used for the idiosyncratic variances in S_t . The single-factor model only includes the excess market return, that is β_{SMB} and β_{HML} in (18) are set equal to 0.

4.3 Estimation results

Figures 1 and 2 plot the average correlation between the DJIA stocks together with the 10th and 90th percentiles obtained from the for the semi-parametric models with the contemporaneous log-volatility $\log(h_{M,t})$ of the value-weighted market portfolio and with the lagged weekly market return $r_{M,w,t-1}$, respectively, estimated with daily returns over the period 1989-1998. From Figure 1 it is seen that for all values of a considered, the average correlation behaves similarly and increases with lagged volatility from 0.25 to 0.50, consistent with a one-factor model as discussed before. The increase in correlations for high volatility is more pronounced for small values of a . Turning to Figure 2, we observe that the average correlation appears to increase for negative lagged market returns, while it remains fairly constant for positive values of $r_{M,w,t-1}$, in particular for $a \leq 0.50$. This corresponds quite well with the patterns documented in Longin and Solnik (2001) and Ang and Chen (2002). For $a = 0.75$, the correlations increase for large positive index returns as well, albeit the effect of negative index returns is stronger.

– **Figures 1 and 2 about here** –

Panels (a) and (b) of Figure 3 display the average correlation and the 10th and 90th percentiles over time for the semi-parametric models with the two choices of x_t and $a = 0.50$ in both cases. It appears that using market volatility as conditioning variable leads to more sizable changes in the conditional correlations than using the market return. At the same time however, the average conditional correlation based on the model with $x_t = \log(h_{M,t})$ evolves more smoothly, with the average correlation based on the model with $x_t = r_{M,w,t-1}$ showing more erratic short-run behavior.

– **Figure 3 about here** –

Panels (c) and (d) of Figure 3 show the average correlation and the 10th and 90th percentiles based on the DCC model and the single-factor model, respectively. For the DCC model, the average correlation varies very little over time. In contrast, the conditional correlations implied by the SF model vary considerably. This is a direct consequence of the single-factor structure together with the use of a GJR-GARCH model for the conditional variance of the factor. In the SF model with the excess

market return as the single factor, B in (20) reduces to a vector consisting of the stocks' betas $\beta_{j,M}$, while $\Omega_t = h_{M,t}$ is a scalar. The conditional correlation between stocks i and j then is equal to

$$\rho_{ijt} = \frac{\beta_{i,M}\beta_{j,M}h_{M,t}}{\sqrt{\beta_{i,M}^2h_{M,t} + s_{it}^2}\sqrt{\beta_{j,M}^2h_{M,t} + s_{jt}^2}},$$

where s_{it}^2 is the idiosyncratic conditional variance of stock i . As shown by Andersen et al. (2001), it follows that $\partial\rho_{ijt}/\partial h_{M,t} > 0$ if $\beta_{i,M}, \beta_{j,M}, h_{M,t} > 0$. Hence, as long as the market betas of two stocks are both positive, an increase in market volatility will lead to an increase in their conditional correlation. Given that all betas for the DJIA stocks are positive and not too different from each other, see Table 1, all pairwise conditional correlations increase when $h_{M,t}$ becomes higher. Additionally, the GJR-GARCH model for the market portfolio returns implies considerable variation in $h_{M,t}$, leading to the substantial changes in the conditional correlations as seen in Figure 3.

Panel (b) of Figure 4 explicitly shows how the conditional correlations are related to the (logged) market volatility. Note that the increase in the average correlation is much more pronounced than in the corresponding semi-parametric model, compare Figure 1. Panel (a) of Figure 4 shows that the average correlation from the DCC model is not related to the market volatility at all. Similarly, in panel (c) of this figure a rather weak quadratic relation between the average correlation from the DCC model and the lagged weekly market return is visible, but much less pronounced than in the semi-parametric model. For the single-factor model we do find a strong relationship between the conditional correlations and the lagged weekly market return, of the same form as found for the semi-parametric model in Figure 2 and supporting the ‘correlation breakdown’ effect with a much larger increase in correlation for large negative market returns than for large positive ones. At first sight this may seem surprising but in fact it is a direct consequence of the use of a GJR-GARCH model for the market volatility. In that model, large past returns lead to high current volatility, while the parameter estimates are such that negative returns increase volatility more than positive returns of the same magnitude. The higher volatility in turn increases the correlations between the stocks.

– Figure 4 about here –

Finally, Figures 5 and 6 summarize the estimation results for the STCC model with $x_t = \log(h_{M,t})$ and $r_{M,w,t-1}$, respectively. The STCC models also imply that conditional correlations are higher in case of high market volatility and in case of low market returns. However, note that by construction the model implies that the conditional correlations only change monotonically from one level to another as the logistic transition function (11) changes from 0 to 1. Panel (c) of these figures plot the value of the transition functions, suggesting that the correlations spend most of the time in the low-volatility and positive market return regimes, where the conditional correlations are at their lower level.

– Figures 5 and 6 about here –

4.4 Minimizing portfolio variance using correlation forecasts

We follow the recommendation of Engle and Sheppard (2001) to judge the adequacy of the dynamic correlation models by examining certain characteristics of stock portfolios that are constructed based on covariance matrix forecasts from the models. In particular, we consider the global minimum variance portfolio (MVP), which is often used for judging the goodness of fit of multivariate volatility models, see e.g. Chan, Karceski and Lakonishok (1999). The MVP weights are given by $w_t = \hat{H}_t^{-1}\iota/(\iota'\hat{H}_t^{-1}\iota)$, where \hat{H}_t is the one-step ahead forecast of the conditional covariance matrix constructed at time $t - 1$, and ι is a $(N \times 1)$ vector of ones. Note that the MVP weights only depend on forecasts of the conditional covariance, such that forecasting expected returns, which is known to be notoriously difficult, is avoided. For each MVP based on the different covariance models we consider the average return, standard deviation, Sharpe ratio and tracking error relative to the S&P 500 index. The latter is defined as the square root of the mean squared difference between the portfolio's return and the S&P 500 return. For comparison purposes, we also construct equally-weighted and value-weighted portfolios and the minimum variance portfolio. For the equally-weighted portfolio (EWP), the portfolio weights are constant and equal to $w_t = N^{-1}\iota$, where . The time-varying weights for the value-weighted portfolio (VWP) are obtained as $w_t = \frac{w_{t-1}(1+r_{t-1})}{(1+r_t)\iota}$, starting with an equally-weighted portfolio at $t = 0$, cf. Engle and Sheppard (2001).

The EWP and VWP provide reasonable benchmarks to assess the extent to

which optimization actually helps in reducing portfolio variance. In addition, together with the MVP they can be used to further evaluate the covariance models by considering additional portfolio characteristics. First, we compute the variance of the standardized portfolio returns $r_{p,t} = w_t' r_t / \sqrt{w_t' \hat{H}_t w_t}$. If the multivariate conditional covariance model is correctly specified, the variance of $r_{p,t}$ should be equal to 1, for any choice of weights w_t . Second, we compute one-day Value-at-Risk (VaR) forecasts as $\hat{\sigma}_p z_q$, where $\hat{\sigma}_p$ is the one-step ahead forecast of the portfolio standard deviation, $\hat{\sigma}_p = \sqrt{w_t' \hat{H}_t w_t}$. In contrast to Engle and Shephard (2001), for example, we do not use the q -th quantile of the standard normal distribution for z_q . Rather, we employ the quantiles of the standardized in-sample portfolio returns obtained using the relevant weights w_t . We consider one-day VaR forecasts at $100 \times q = 1\%$, 5% and 10% . The accuracy of the VaR forecasts is assessed by means of the Dynamic Quantile test of Engle and Manganelli (2004). For that purpose, define the binary variable HIT_t such that $HIT_t = 1$ if the portfolio return is below the VaR forecast and 0 otherwise. Under the null hypothesis of a correctly specified model for the conditional covariance matrix and hence for the VaR forecasts, the HIT_t variable should have mean q and should be independent from all information available at $t - 1$, including lagged HIT_t 's and the VaR forecast for time t . This can be tested by constructing an F -statistic for the null hypothesis $\delta_0 = \dots = \delta_{l+1} = 0$ in the auxiliary regression

$$HIT_t - q = \delta_0 + \delta_1 HIT_{t-1} + \dots + \delta_l HIT_{t-l} + \delta_{l+1} VaR_t + e_t,$$

where we set $l = 5$.

Table 4 summarizes the in-sample MVP results over the period January 1, 1989-December 31, 1998, which suggest several conclusions. First, compared to the EWP and VWP, all models except Riskmetrics achieve in producing an MVP with lower standard deviation, although the reduction in volatility is quite modest. The best performing model is our semi-parametric model using the lagged weekly market return as conditioning variable, with an annualized standard deviation of 12.54% compared with 15.16% for the EWP. Second, minimizing portfolio variance comes at the cost of a sharp reduction in the portfolio's average return, such that the Sharpe ratios of the MVP's is considerably lower than the Sharpe ratios of the EWP and VWP. Third, the conditional correlation models outperform the single- and three-

factor models. Hence, it seems worthwhile to model the dynamics in the covariance structure of the individual stocks explicitly, rather than indirectly by means of a factor approach. Fourth, the performance of the different conditional correlations is very similar, with the largest difference in standard deviation of the MVP returns being less than 0.3%.

– **Table 4 about here** –

The conclusion that it is difficult to distinguish between the competing models based on in-sample measures of fit also emerges from Table 5. The variance of the standardized portfolio returns shows that actually all models perform quite badly in terms of the volatility of the MVP returns, which is consistently underestimated. In contrast, the variances of the standardized EWP and VWP returns are quite close to the theoretical value of one for all models except perhaps RiskMetrics. In terms of VaR forecasts, results also are similar across models (and by construction, violation frequencies are identical for the EWP). The $\text{VaR}(q)$ forecasts are violated too frequently for the MVP for all three levels of q , which aligns with the underestimation of the MVP volatility. The VaR violation frequencies are approximately correct for the EWP and VWP portfolios. The results of the Dynamic Quantile test confirm these observations, except that the small p -values for the test applied to the 10% VaR forecasts for the EWP are perhaps surprising, given that the violation frequency is close to perfect.

– **Table 5 about here** –

Next, we examine the out-of-sample forecasting performance of the models over the period January 1, 1999-December 31, 2003. For this purpose, we re-estimate all models after every 20 observations using a ten-year moving window. We examine the same portfolio characteristics as before, except that now we employ genuine out-of-sample forecasts of the conditional covariance matrix.

The MVP results in Table 6 resemble the results in Table 4, in the sense that most conclusions drawn for the in-sample MVP results continue to hold out-of-sample. The only exception is that, in terms of the standard deviation of MVP returns the DCC-type models now perform slightly better than the STCC and semi-parametric models. In addition, we note that the variance of the MVP during the out-of-sample

period is almost double the variance during the in-sample period. From Table 7 we observe that all models continue to underestimate the MVP variance in the out-of-sample period, while the corresponding Value-at-Risk forecasts are (thus) violated much too frequently. The conditional correlation models perform much better for the EWP and MVP, although the 5% and 10% VaR forecasts for these portfolios are violated too frequently as well.

– **Tables 6 and 7 about here** –

The most conspicuous finding in the MVP analysis presented above is the similarity in performance of all covariance models for the portfolio’s variance. The characteristics of the different MVP’s are examined further by estimating the Fama-French three-factor model as given in (18) using the portfolio’s excess return as dependent variable. The resulting estimates, shown in Tables 4 and 6, confirm that the portfolios are comparable in terms of sensitivities to the market return, size and book-to-market factors. In addition, all MVP’s (except the ones based on the single-factor and Riskmetrics models) must select stocks with low market betas, given that their market betas are close 0.8 and 0.7 for the in-sample and out-of-sample periods, respectively, compared with market betas close to one for the EWP and VWP. The emphasis on stocks with low market betas is also seen in Table 8, which shows the mean and standard deviation of the weights in the MVP based on the semi-parametric conditional correlation model with x_t equal to the daily log-volatility for the market return. During the in-sample period, the MVP is dominated by Exxon Mobil (XOM) with a mean portfolio weight close to 20%. The estimates of $\beta_{i,M}$ in Table 1 show that indeed this stock had the lowest market beta during this period. Also note that the stocks of 3M (MMM) and SBC Communications (SBC) have large mean portfolio weights during the in-sample period although their betas are not particularly low compared to the other stocks. The opposite is observed for the stock of Merck & Co (MRK), which has the second-lowest market beta but only an average mean portfolio weight. This illustrates the fact that the portfolio variance cannot only be reduced by putting much weight on low beta stocks, but also by tilting the portfolio towards stocks with low return volatilities. 3M and SBC are typical examples of the latter. Finally, we remark that the regression results of the three-factor model also shed light on the reason why the MVP variance is so much higher

during the out-of-sample period compared to the in-sample period. This appears largely due to the portfolio’s sensitivity to the HML factor, which is close to zero during the 1989-1998 period but substantially different from zero during 1999-2003.

– **Table 8 about here** –

In the context of factor models, Chan et al. (1999) attribute the similarity in performance of different covariance models in terms of MVP construction to the presence of a dominant factor, the market, which is much more important than the other influences on returns. Constructing the minimum variance portfolio then effectively boils down to minimizing the portfolio’s sensitivity to the market. This is achieved by tilting the portfolio toward stocks with low market betas and explains why all covariance models yield such similar results. Chan et al. (1999) suggest that any differences between the different covariance models may be brought to light more clearly when the dominant factor is removed. This turns out to be equivalent to tracking a benchmark portfolio that resembles the dominant factor, and the relevant portfolio to consider becomes the portfolio that minimizes the tracking error.

4.5 Minimizing tracking error using correlation forecasts

As the benchmark portfolio for the tracking error optimization problem we choose the S&P 500 index. The minimum tracking error portfolio (MTEP) is equivalent to the minimum variance portfolio for the stock returns in excess of the return on the benchmark portfolio. Hence we apply the various models to obtain forecasts of the conditional covariance matrix of the excess DJIA stock returns relative to the S&P 500 return, denoted \hat{H}_t^e , and compute the MTEP weights as $w_t = \hat{H}_t^{e^{-1}} \iota / (\iota' \hat{H}_t^{e^{-1}} \iota)$. In this case we also apply the STCC and semi-parametric models with conditioning variable x_t being the contemporaneous log-volatility of the size and book-to-market factor returns. This is motivated by the fact that these factors might be of considerable importance for the covariance structure of the excess returns as the dominant market factor is removed. Tables 9 and 10 display the characteristics of the resulting portfolios for the in-sample period 1989-1998 and out-of-sample period 1999-2003, respectively. We find that the semi-parametric model with $x_t = \log(h_{\text{HML},t})$ renders the smallest tracking error, both in-sample and out-of-sample, although the differences with other conditional correlation models are not large. Note that the single-

and three-factor models perform quite disappointing for the out-of-sample period, with tracking errors that are comparable to the tracking errors of the EWP and VWP. Finally, we observe that the out-of-sample tracking error is twice as large as the in-sample tracking error. This again is due to the increased sensitivity to the book-to-market factor.

– Tables 9 and 10 about here –

5 Concluding remarks

We have developed a new semi-parametric model for conditional correlations, which combines parametric univariate GARCH-type specifications for the individual conditional volatilities with nonparametric kernel regression for the conditional correlations. This approach not only avoids the proliferation of parameters as the number of assets becomes large, which typically happens in conventional multivariate conditional volatility models, but also the rigid structure imposed by more parsimonious models, such as the DCC model. Hence, our model is applicable in empirically relevant settings, where portfolios might consist of tens or even hundreds of assets.

The application to the DJIA stocks illustrates the potential of the semi-parametric model. The estimation results suggest that correlation functions are asymmetric when the conditioning variable is the lagged weekly market return. That is, correlations increase stronger in bear markets than they do in bull markets, in agreement with the ‘correlation breakdown’ effect. We also find support for the idea that correlations increase with market volatility, as would be implied by a factor structure for the stock returns. The semi-parametric model performs quite well when compared with standard parametric DCC-type models in terms of constructing minimum variance portfolios and minimum tracking error portfolios.

A point to note, finally, is that the time required to estimate the semi-parametric model turned out to be far below that of even the simple DCC models with two parameters. The reason is that no ill-conditioned likelihood needs to be maximized, but rather a simple data smoother is used. Future research should develop more guidance with respect to bandwidth selection, and find ways to avoid the curse of dimensionality when more than one conditioning variable is considered.

Acknowledgements

We thank participants of the 3rd Annual Advances in Econometrics Conference, Baton Rouge (November 2004), and the 55th Session of the International Statistical Institute (Sydney, April 2005), as well as an anonymous referee for useful comments and suggestions. Any remaining errors are ours.

Appendix: Assumptions and proof of Theorem 1

We use the following notation. First and second derivatives of a function $f : \mathbb{R} \rightarrow \mathbb{R}$ are denoted by $f'(\cdot)$ and $f''(\cdot)$, respectively. Let x be an N -vector and X an $N \times N$ symmetric matrix. The Euclidean norm of x is denoted as $|x|$. The operator **diag** builds a diagonal matrix from a vector argument, i.e.,

$$\text{diag}(x) = \begin{pmatrix} x_1 & 0 & \cdots & 0 \\ 0 & x_2 & & \vdots \\ \vdots & & \ddots & \\ 0 & \cdots & & x_N \end{pmatrix}$$

The operator **dg** stacks the diagonal of a matrix into a vector, i.e., $\text{dg}(X) = (X_{11}, X_{22}, \dots, X_{NN})'$. The operator **vec** stacks the columns of a matrix into a column vector, and **vech** stacks only the lower triangular part including the diagonal into a vector. Denote by D_N the $N^2 \times (N(N+1)/2)$ duplication matrix and by D_N^+ its generalized inverse, of size $(N(N+1)/2 \times N^2)$, with the properties $\text{vech}(X) = D_N^+ \text{vec}(X)$ and $\text{vec}(X) = D_N \text{vech}(X)$. Further, denote by L_N the $N \times N^2$ matrix with the property $L_N \text{vec}(X) = \text{dg}(X)$.

We work under the following assumptions:

Assumption 1

- (A1) The true parameter vector θ_0 governing the conditional variance matrix D_t exists uniquely and lies in the interior of a compact set.
- (A2) The first-stage estimator of θ is \sqrt{T} -consistent and asymptotically normal.
- (A3) The process $\{r_t\}$ is strictly stationary.
- (A4) The second-stage kernel function satisfies $\int uK(u)du = 0$, $\int K^2(u)du = \|K\|_2^2 < \infty$, $\mu_2(K) = \int u^2K(u)du < \infty$, $\int K(u)du = 1$, $\sup |K(u)| = B < \infty$, and $|u||K(u)| \rightarrow 0$ as $|u| \rightarrow \infty$.
- (A5) The off-diagonal elements of $R(x)$ are strictly between -1 and 1, uniformly in x , and are at least twice continuously differentiable on the support of x_t .
- (A6) The process $\xi_t = R^{-1/2}(x_t)\varepsilon_t$ is i.i.d. The density $p(\cdot)$ of ξ_1 exists and is positive on any compact subset of \mathbb{R}^N . Furthermore $\mathbb{E}[\xi_1] = 0$, $\mathbb{E}[\xi_1\xi_1'] = I_N$, and there exist constants $\delta, B > 0$ such that $\mathbb{E}[\xi_{it}^{4+\delta}] \leq B < \infty$, $i = 1 \dots, N$.

Note that (A6) implies that fourth moments of ε_t are finite, because $R(\cdot)$ is bounded. For notational convenience, define $\tilde{\xi}_t = \text{vec}(\xi_t\xi_t' - I_N)$, which by (A6) is i.i.d. with mean zero and $\mathbb{E}[\tilde{\xi}_t\tilde{\xi}_t'] = M_4$, say.

In the proof of Theorem 1, we make use of the following lemmata.

Lemma 1 *Let*

$$\tilde{R}(x) = \tilde{Q}^*(x)^{-1}\tilde{Q}(x)\tilde{Q}^*(x)^{-1}$$

where

$$\tilde{Q}(x) = \frac{\sum_{t=1}^T \tilde{\varepsilon}_t \tilde{\varepsilon}_t' K_h(x_t - x)}{\sum_{t=1}^T K_h(x_t - x)},$$

$\tilde{Q}^*(x)$ is a diagonal matrix with the square roots of the diagonal elements of $\tilde{Q}(x)$ on its diagonal, and $\tilde{\varepsilon}_t = D_t(\theta_0)^{-1}r_t$. Then, under assumptions (A1) to (A6)

$$\tilde{r}(x) - \hat{r}(x) = O_p\left(\frac{1}{Tb}\right)$$

uniformly in x , where $\tilde{r}(x) = \text{vech}(\tilde{R}(x))$.

Proof: Very similar to the proof of Lemma A.2 of Rodriguez-Poo and Linton (2001). \square

Lemma 2 Assume that the diagonal elements of D_t^2 are specified as

$$h_{it} = \omega_i + \alpha_i r_{i,t-1}^2 + \beta_i h_{i,t-1},$$

with $\alpha_i + \beta_i < 1$, $i = 1, \dots, N$. Under Assumptions (A1) to (A6) and $R_t = R(r_{t-1})$, the process $(r_t)_{t=0}^\infty$ is geometrically ergodic. Furthermore, if r_0 is drawn from the stationary distribution, then $(r_t)_{t=0}^\infty$ is geometrically α -mixing.

Proof: Let $\omega = (\omega_1, \dots, \omega_N)'$, $\alpha = (\alpha_1, \dots, \alpha_N)'$, and $\beta = (\beta_1, \dots, \beta_N)'$. Define the processes $V_t = (V_{1t}, \dots, V_{Nt})'$ with $V_{it} = \sum_{j=0}^\infty \beta_i^j r_{i,t-j}^2$, and the $2N$ -vector process $Z_t = (r_t', V_t')'$. By construction, Z_t is a Markov process which can be represented as $Z_t = m(Z_{t-1}) + g(Z_{t-1})\zeta_t$ with $\zeta_t = (\xi_t', \tilde{\xi}_t' L_N')'$ and where $m : \mathbb{R}^{2N} \rightarrow \mathbb{R}^{2N}$ and $g : \mathbb{R}^{2N} \rightarrow \mathbb{R}^{2N} \times \mathbb{R}^{2N}$, given by

$$m(Z_{t-1}) = \begin{pmatrix} 0_{N \times 1} \\ \text{diag}(\beta)V_{t-1} + \text{diag}(\tilde{V}_{t-1})R_t \end{pmatrix}$$

and with $\tilde{V}_t = (I_N - \text{diag}(\beta))^{-1}\omega + \text{diag}(\alpha)V_t$,

$$g(Z_{t-1}) = \begin{pmatrix} \text{diag}(\tilde{V}_{t-1})^{1/2} R_t^{1/2} & 0_{N \times N} \\ 0_{N \times N} & \text{diag}(\tilde{V}_{t-1}) L_N (R_t^{1/2} \otimes R_t^{1/2}) \end{pmatrix}.$$

Using assumption (A6), $E[\zeta_t \zeta_t']$ is finite and given by W , say. The condition of Doukhan (1994),

$$\limsup_{|z| \rightarrow \infty} \frac{|m(z)|^2 + \text{Tr}(W g'(z) g(z))}{|z|^2} < 1$$

is easily checked, as $R(\cdot)$ is bounded. By proposition 6 of Doukhan (1994, p. 107), geometric ergodicity of (Z_t) follows. Finally, by Davydov (1973), a geometrically ergodic Markov chain with initial value drawn from the stationary distribution is geometrically α -mixing. This proves the lemma. \square

Proof of Theorem 1:

Let us only consider the case of constant bandwidths, $a = 0$, so that $h = b$. The proof for the general case $a > 0$ is analogous to Jennen-Steinmetz and Gasser (1988). Also, we only sketch the proof for the factor $x_t = w' r_{t-1}$, where w is a fixed N -vector. First, by definition $\rho_{ii}(x) = E[\varepsilon_{it}^2 \mid x_{t-1} = x] = 1$, $i = 1, \dots, N$. The geometric mixing property of $(x_t)_{t=0}^\infty$ as implied by Lemma 2 ensures consistency and asymptotic normality of the standard kernel smoother as shown by Robinson (1983) under weaker conditions. Thus, $\hat{Q}(x)$ is a consistent estimator for $R(x)$. Because $\hat{Q}^*(x) \xrightarrow{p} I_N$ and using Slutsky's

Theorem, $\hat{R}(x)$ is also a consistent estimator for $R(x)$. (It has the additional advantage of being a correlation matrix for finite samples.)

Turning to the asymptotic distribution, note that

$$\sqrt{Th}(\hat{r}(x) - r(x)) = \sqrt{Th}(\hat{r}(x) - \tilde{r}(x)) + \sqrt{Th}(\tilde{r}(x) - r(x)) \quad (\text{A.1})$$

where the first term on the right hand side of (A.1) is $O_p((Th)^{-1/2})$ using Lemma 1, and therefore converges to zero in probability under the assumption $Th \rightarrow \infty$. Thus, the asymptotic distribution of the left hand side of (A.1) is the same as that of the second term on the right hand side of (A.1). This term however has the same asymptotic distribution as $\sqrt{Th}(\tilde{q}(x) - r(x))$, where $\tilde{q}(x) = \text{vech}(\tilde{Q}(x))$. We can write

$$\sqrt{Th}(\tilde{q}(x) - r(x)) = \frac{1}{\hat{f}(x)}(Th)^{-1/2} \sum_{t=1}^T K\left(\frac{x_t - x}{h}\right) \{\eta_t - r(x)\} \quad (\text{A.2})$$

where $\hat{f}(x) = (Th)^{-1} \sum_{t=1}^T K\left(\frac{x_t - x}{h}\right)$ is the kernel estimator of $f(x)$, the density of x_t , and $\eta_t = \text{vech}(\varepsilon_t \varepsilon_t')$. As \hat{f} converges to f in probability, the asymptotic distribution is determined by that of $(Th)^{-1/2} \sum_{t=1}^T K\left(\frac{x_t - x}{h}\right) \{\eta_t - r(x)\}$, which we can decompose as

$$(Th)^{-1/2} \sum_{t=1}^T K\left(\frac{x_t - x}{h}\right) \{\eta_t - r(x)\} = I_1 + I_2,$$

where, denoting $u_t = \eta_t - r(x_t)$,

$$I_1 = (Th)^{-1/2} \sum_{t=1}^T K\left(\frac{x_t - x}{h}\right) u_t$$

and

$$I_2 = (Th)^{-1/2} \sum_{t=1}^T K\left(\frac{x_t - x}{h}\right) \{r(x_t) - r(x)\}.$$

Since $r(x_t) = \mathbb{E}[\eta_t | x_t]$, u_t is a martingale difference sequence. By Lemma 2 it is a function of geometrically α -mixing processes and thus itself geometrically α -mixing. Note that it can be written as $u_t = D_N^+(R_t^{1/2} \otimes R_t^{1/2})\tilde{\xi}_t$ with $R_t = R(x_t)$ and

$$\mathbb{V}(u_t | x_t) = D_N^+(R_t^{1/2} \otimes R_t^{1/2}) M_4(R_t^{1/2} \otimes R_t^{1/2}) D_N < \infty$$

by (A6). Using the Cramér-Wold device, $I_1 \xrightarrow{\mathcal{L}} N(0, f(x)^2 \Sigma(x)) \iff c' I_1 \xrightarrow{\mathcal{L}} N(0, f(x)^2 c' \Sigma(x) c)$ for any fixed vector c of length $N(N+1)/2$. Thus, it suffices to show that a central limit theorem holds for $c' I_1$ with the required asymptotic variance. Using Lemma 5.2 of Härdle, Tsybakov and Yang (1998), one obtains $\mathbb{V}(c' I_1) = f(x)^2 c' \Sigma(x) c + o(1)$ with $\Sigma(x) = \|K\|_2^2 \mathbb{V}(u_t | x_t = x) / f(x)$. This shows that the asymptotic variance of $\hat{f}(x)^{-1} I_1$ is given by $\Sigma(x)$. To prove that the asymptotic distribution is normal, the conditions of the central limit theorem for square integrable martingale differences of Liptser and Shirjaev (1980, Corollary 6) can be shown to hold along the lines of Härdle, Tsybakov and Yang (1998).

For the second term, note first that

$$\begin{aligned}
\mathbb{E}[I_2] &= (Th)^{-1/2} \sum_{t=1}^T \mathbb{E} \left[K\left(\frac{x_t - x}{h}\right) \{r(x_t) - r(x)\} \right] \\
&= (Th)^{1/2} \int K(\psi) \{r(x + h\psi) - r(x)\} f(x + h\psi) d\psi \\
&= (Th)^{1/2} \left\{ h^2 \mu_2(K) \left(\frac{f(x)r''(x)}{2} + f'(x)r'(x) \right) + o(h^2) \right\} \\
&= (Th)^{1/2} \{O(h^2) + o(h^2)\} = O \left\{ (Th)^{1/2} h^2 \right\},
\end{aligned}$$

which converges to zero under the assumption $h = o(T^{-1/5})$. Now $Y_t = K(\psi)\{r(x_t) - r(x)\}$, $\psi = (x_t - x)/h$, is a mixing process as it is a measurable function of a mixing process. Hence, by Theorem 6.3 of Pötscher and Prucha (1997) a weak law of large numbers holds for Y_t if $\sup_T T^{-1} \sum_t^T \mathbb{E}[|Y_t|^{1+\varepsilon}] < \infty$ for some $\varepsilon > 0$. Using (A5), $|r(x_t) - r(x)| \leq 2|r(x_t)| \leq 2$ a.s. Furthermore, $K(\cdot)$ is bounded by (A4). Thus, $\mathbb{E}[|Y_t|^{1+\varepsilon}] \leq \mathbb{E}[|K(\psi)|^{2(1+\varepsilon)}]^{1/2} \mathbb{E}[|r(x_t) - r(x)|^{2(1+\varepsilon)}]^{1/2} \leq |2B|^{1+\varepsilon} < \infty$. Thus, $I_2 = o_p(1)$, which completes the proof. \square

References

- Andersen, T.G., Bollerslev, T., Diebold, F.X., Ebens, H., 2001. The distribution of realized stock return volatility. *Journal of Financial Economics* 61, 43–76.
- Ang, A., Chen, J., 2002. Asymmetric correlations of equity portfolios. *Journal of Financial Economics* 63, 443–494.
- Bauwens, L., Laurent, S., Rombouts, J.V.K., 2005. Multivariate GARCH models: A survey. *Journal of Applied Econometrics*, forthcoming.
- Berben, R.-P., Jansen, W.J., 2005. Comovement in international equity markets: A sectoral view. *Journal of International Money and Finance* 24, 832–857.
- Bollerslev, T., 1986. Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics* 31, 307–327.
- Bollerslev, T., 1990. Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH approach. *Review of Economics and Statistics* 72, 498–505.
- Bollerslev, T., Engle, R.F., Wooldridge, J.M., 1988. A capital asset pricing model with time varying covariances. *Journal of Political Economy* 96, 116–131.
- Butler, K.C., Joaquin, D.C., 2002. Are the gains from international portfolio diversification exaggerated? The influence of downside risk in bear markets. *Journal of International Money and Finance* 21, 981–1011.
- Campbell, R., Koedijk, K., Kofman, P., 2002. Increased correlation in bear markets. *Financial Analyst Journal* 58, 87–94.
- Cappiello, L., Engle, R.F., Sheppard, K., 2003. Asymmetric dynamics in the correlations of global equity and bond returns. *European Central Bank Working paper* No. 204.
- Chan, L.K.C., Karceski, J., Lakonishok, J., 1999. On portfolio optimization: Forecasting covariances and choosing the risk model. *Review of Financial Studies* 12, 937–674.
- Davydov, Y.A., 1973. Mixing conditions for Markov chains. *Theory of Probability and its Applications* 18, 312–328.
- Doukhan, P., 1994. *Mixing: Properties and Examples*, Lecture Notes in Statistics 85, Berlin, Springer-Verlag.
- Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* 50, 987–1007.
- Engle, R.F., 2002. Dynamic conditional correlation - A simple class of multivariate GARCH models. *Journal of Business and Economic Statistics* 20, 339–350.
- Engle, R.F., Kroner, K.F., 1995. Multivariate simultaneous generalized ARCH. *Econometric Theory* 11, 122–150.
- Engle, R.F., Manganelli, S., 2004. CaViaR: Conditional autoregressive Value at Risk by regression quantiles. *Journal of Business and Economic Statistics* 22, 367–381.
- Engle, R.F., Sheppard, K., 2001. Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH. *UCSD Discussion Paper* No. 2001-15.

- Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33, 3–56.
- Fama, E.F., French, K.R., 1996. Multifactor explanations of asset pricing anomalies. *Journal of Finance* 51, 55–84.
- Fan, J., Gijbels, I., 1996. *Local Polynomial Modelling and Its Applications*. New York, Chapman and Hall.
- Fan, J., Yao, Q., 1998. Efficient estimation of conditional variance functions in stochastic regression, *Biometrika* 85, 645–660.
- Forbes, K.J., Rigobon, R., 2002. No contagion, only interdependence: Measuring stock market comovements. *Journal of Finance* 57, 2223–2261.
- Glosten, L.R., Jagannathan, R., Runkle, D.E., 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance* 48, 1779–1801.
- Hafner, C.M., Franses, P.H., 2003. A generalized dynamic conditional correlation model for many assets. *Econometric Institute Report No. 2003-18*, Erasmus University Rotterdam.
- Härdle, W., Lütkepohl, H., Chen, R., 1997. A review of nonparametric time series analysis. *International Statistical Review* 65, 49–72.
- Härdle, W., Tsybakov, A., Yang, L., 1998. Nonparametric vector autoregression. *Journal of Statistical Planning and Inference* 68, 221–245.
- Jennen-Steinmetz, C., Gasser, T., 1988. A unifying approach to nonparametric regression estimation. *Journal of the American Statistical Association* 83, 1084–1089.
- Liptser, R.S., Shirjaev, A.N., 1980. A functional central limit theorem for martingales. *Theory of Probability and its Applications* 25, 667–688.
- Longin, F., Solnik, B., 1995. Is the correlation in international equity returns constant: 1960-1990?. *Journal of International Money and Finance* 14, 3–26.
- Longin, F., Solnik, B., 2001. Extreme correlation of international equity markets. *Journal of Finance* 56, 651–676.
- Loretan, M., English, W.B., 2000. Evaluating changes in correlations during periods of high market volatility. *Bank of International Settlements Quarterly Review*, 29–36.
- Pagan, A.R., Ullah, A., 1999. *Nonparametric Econometrics*. Cambridge, Cambridge University Press.
- Pelletier, D., 2005. Regime switching for dynamic correlations. *Journal of Econometrics*, forthcoming.
- Pötscher, B.M., Prucha, I.R., 1997. *Dynamic Nonlinear Econometric Modelling: Asymptotic Theory*. Berlin, Springer-Verlag.
- Ramchand, L., Susmel, R., 1998. Volatility and cross-correlation across major stock markets. *Journal of Empirical Finance* 5, 397–416.
- Robinson, P.M., 1983. Nonparametric estimators for time series. *Journal of Time Series*

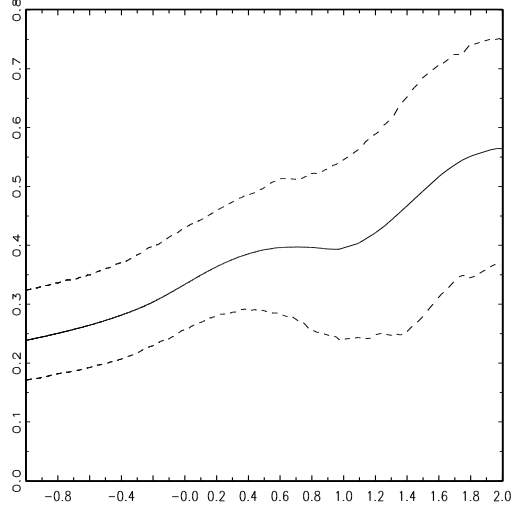
Analysis 4, 185–207.

Rodriguez-Poo, J.M., Linton, O., 2001. Nonparametric factor analysis of residual time series. *TEST* 10, 161–182.

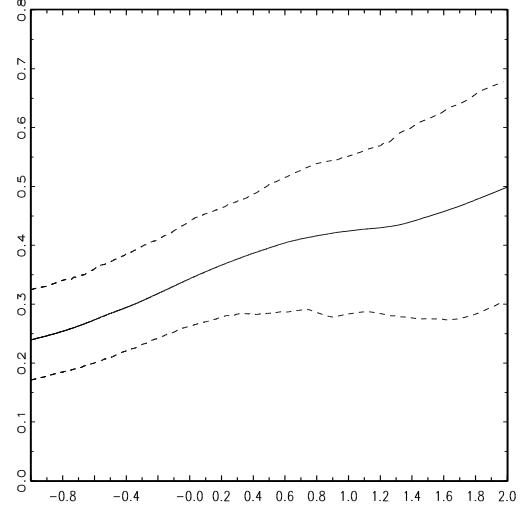
Silvennoinen, A., Teräsvirta, T., 2005. Multivariate autoregressive conditional heteroskedasticity with smooth transitions in conditional correlations. *SSE/EFI Working Paper Series in Economics and Finance* No. 577, Stockholm School of Economics.

Tse, Y.K., Tsui, A.K.C., 2002. A multivariate GARCH model with time-varying correlations. *Journal of Business and Economic Statistics* 20, 351–362.

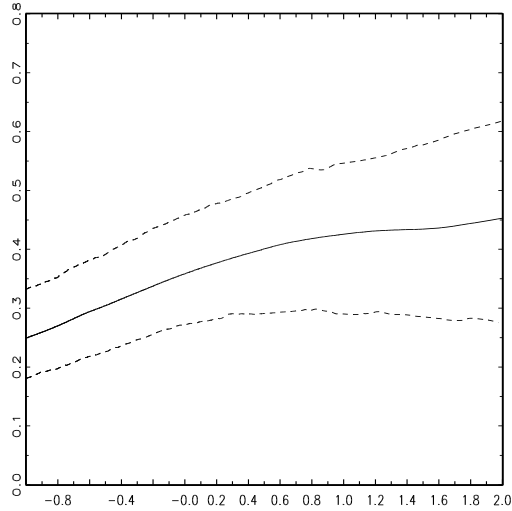
Ziegelmann, F.A., 2002. Nonparametric estimation of volatility functions: The local exponential estimator. *Econometric Theory* 18, 985–991.



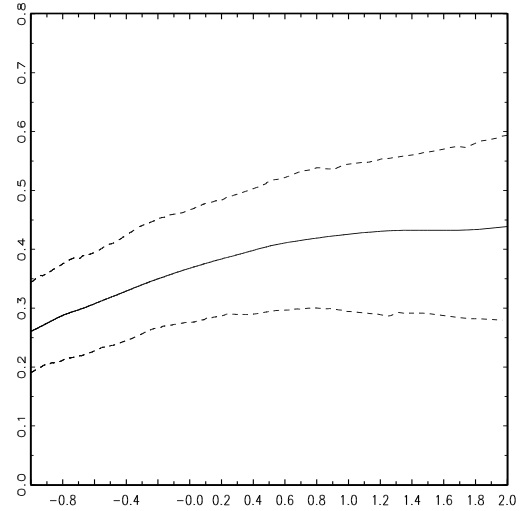
(a) $a = 0$



(b) $a = 0.25$

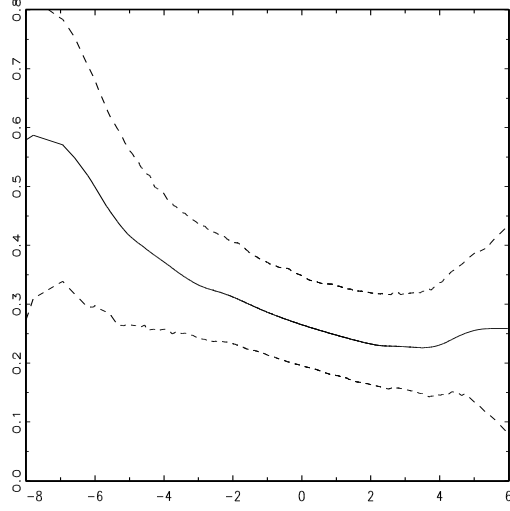


(c) $a = 0.50$

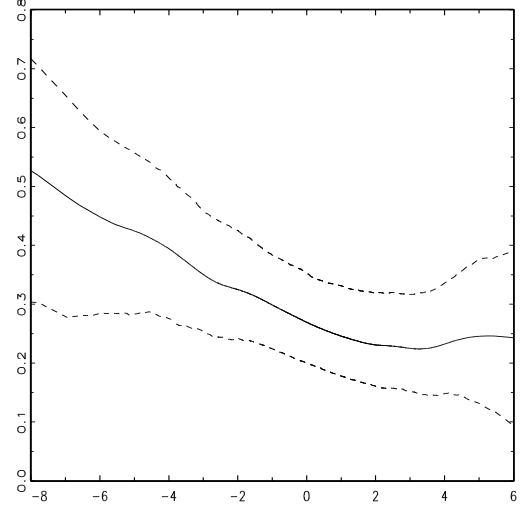


(d) $a = 0.75$

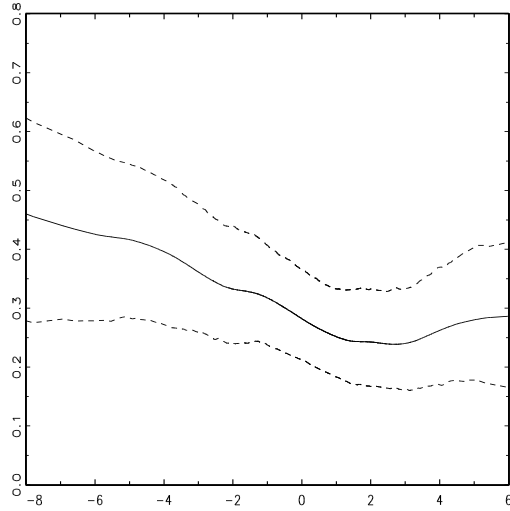
Figure 1: Average correlation between Dow Jones stocks (solid line) and 10th and 90th percentiles (dashed lines) obtained from the nonparametric estimator (14) with x_t being the contemporaneous daily log-volatility on the value-weighted market portfolio $\log(h_{M,t})$, and bandwidth $h(x) = bf(x)^{-a}$, estimated using daily returns over the period January 1, 1989-December 31, 1998, plotted as a function of $\log(h_{M,t})$.



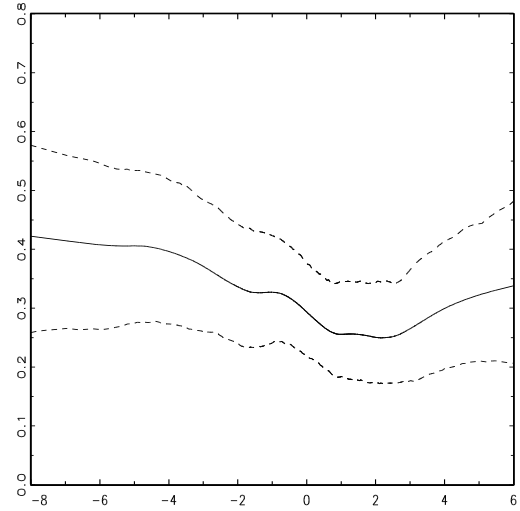
(a) $a = 0$



(b) $a = 0.25$

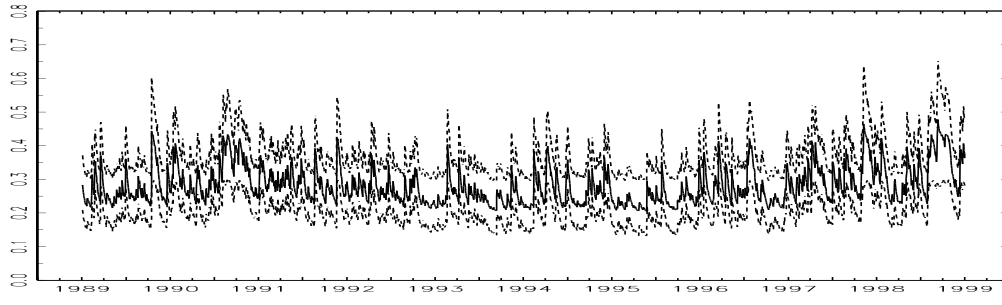


(c) $a = 0.50$

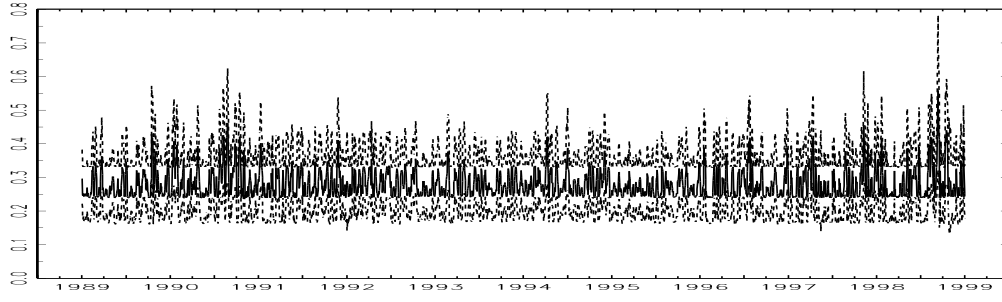


(d) $a = 0.75$

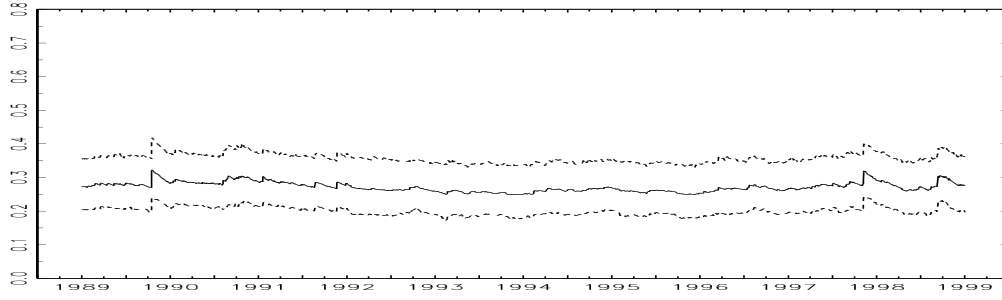
Figure 2: Average correlation between Dow Jones stocks (solid line) and 10th and 90th percentiles (dashed lines) obtained from the nonparametric estimator (14) with x_t being the one-day lagged weekly return on the value-weighted market portfolio $r_{M,w,t-1}$ and bandwidth $h(x) = bf(x)^{-a}$, estimated using daily returns over the period January 1, 1989-December 31, 1998, plotted as a function of $r_{M,w,t-1}$.



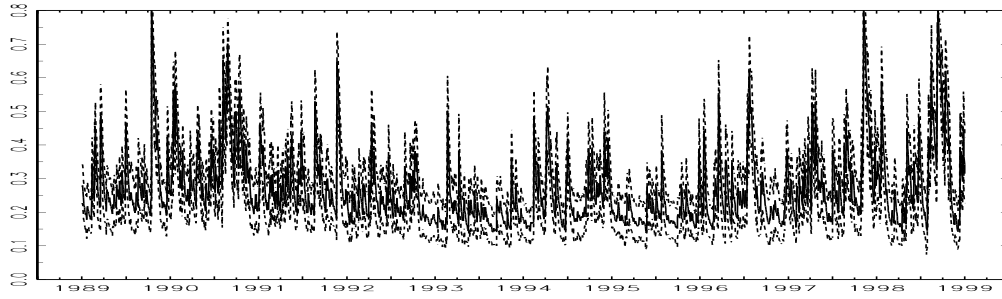
(a) SEMI - $\log(h_{M,t})$



(b) SEMI - $r_{M,w,t-1}$

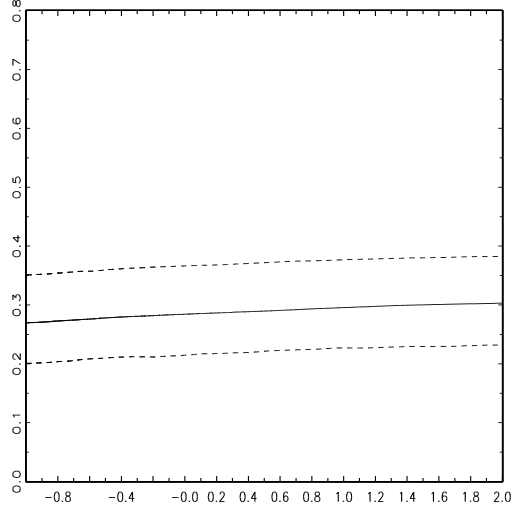


(c) DCC

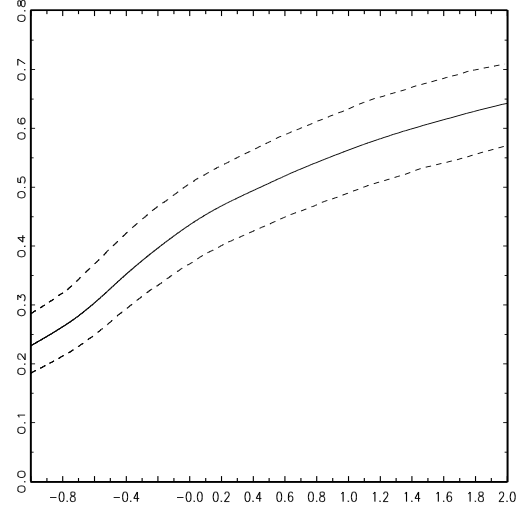


(d) SF

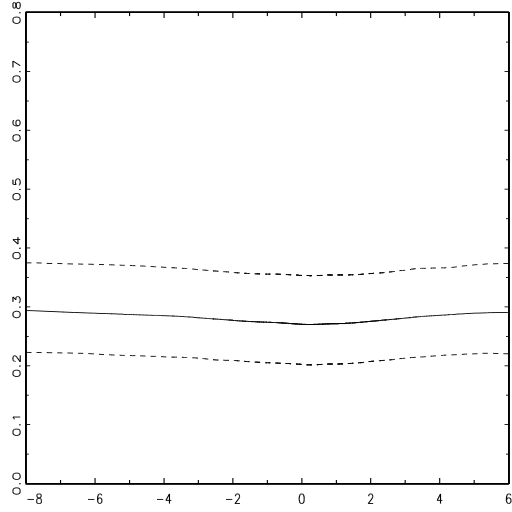
Figure 3: Average correlation between Dow Jones stocks (solid line) and 10th and 90th percentiles (dashed lines) obtained from the nonparametric estimator (14) with x_t being the contemporaneous daily log-volatility on the value-weighted market portfolio $\log(h_{M,t})$ and the one-day lagged weekly return on the value-weighted market portfolio $r_{M,w,t-1}$, from the DCC model (5) (5), and from the SF model, estimated using daily returns over the period January 1, 1989-December 31, 1998.



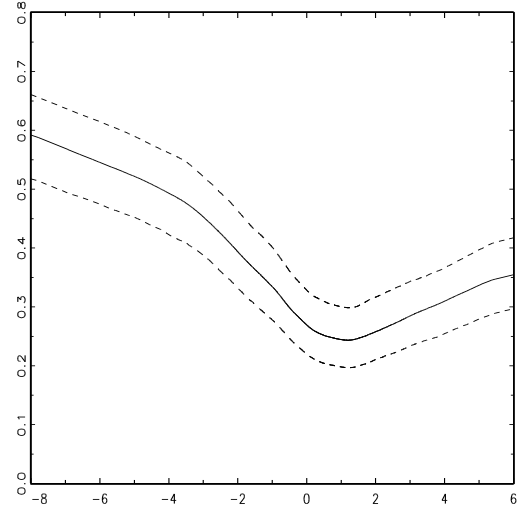
(a) DCC - $\log(h_{M,t})$



(b) SF - $\log(h_{M,t})$

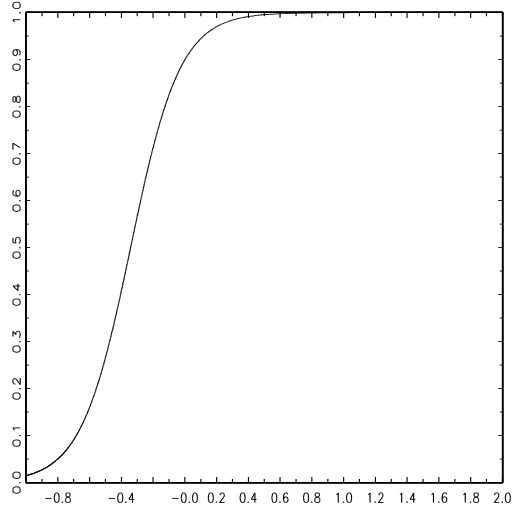


(c) DCC - $r_{M,w,t-1}$

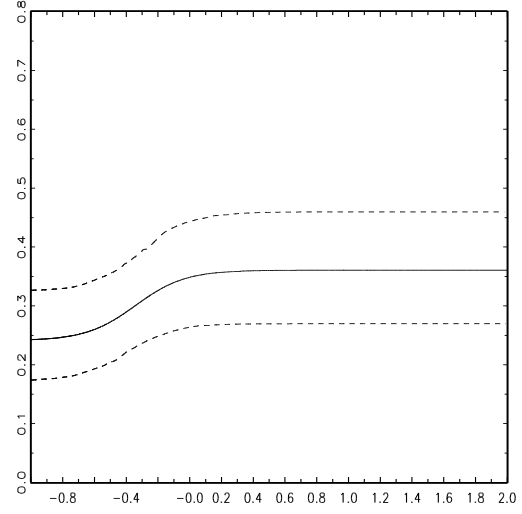


(d) SF - $r_{M,w,t-1}$

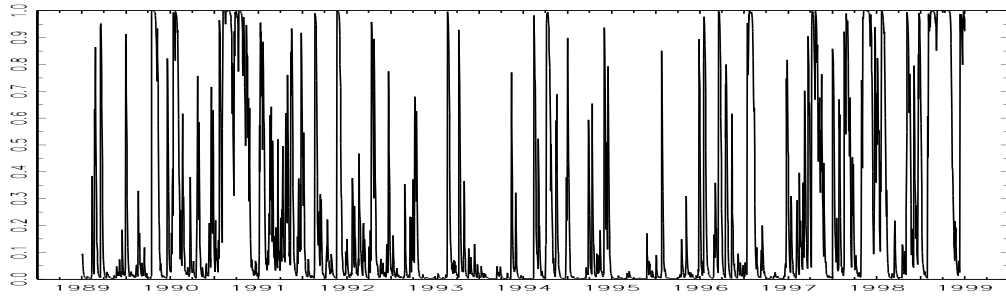
Figure 4: (Smoothed) Average correlation between Dow Jones stocks (solid line) and 10th and 90th percentiles (dashed lines) obtained from the DCC model (5) and the SF model, estimated using daily returns over the period January 1, 1989-December 31, 1998, plotted as a function of the contemporaneous daily log-volatility on the value-weighted market portfolio $\log(h_{M,t})$ and the one-day lagged weekly return on the value-weighted market portfolio $r_{M,w,t-1}$.



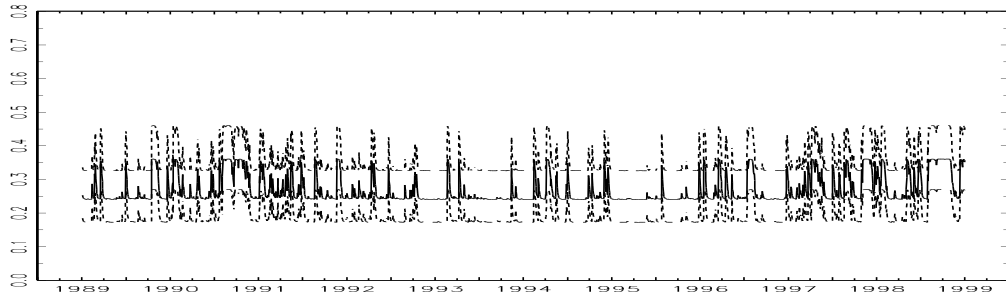
(a) Transition function against x_t



(b) Correlation quantiles against x_t

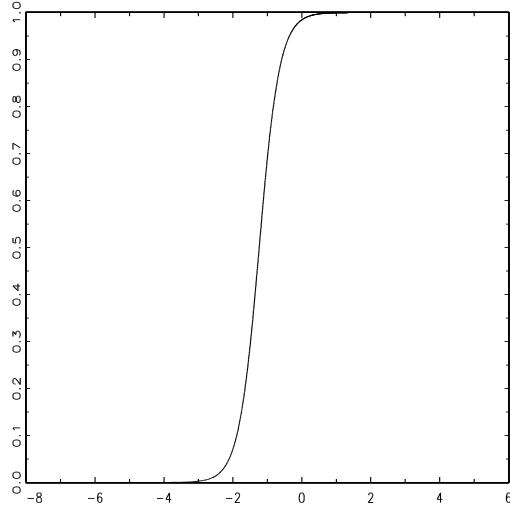


(c) Transition function over time

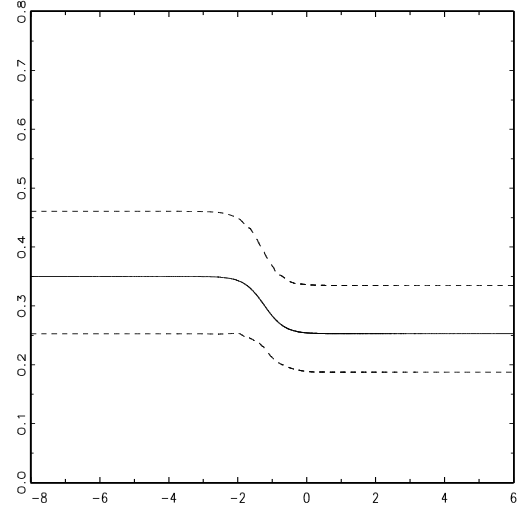


(d) Correlation quantiles over time

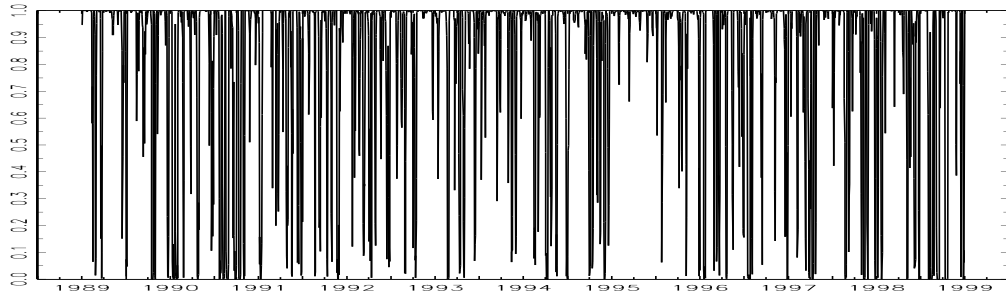
Figure 5: Logistic transition function $G(x_t; \gamma, c)$ as in (11), and average correlation between Dow Jones stocks (solid line) and 10th and 90th percentiles (dashed lines) obtained from the STCC model (10) with x_t being the contemporaneous daily log-volatility on the value-weighted market portfolio $\log(h_{M,t})$, estimated using daily returns over the period January 1, 1989-December 31, 1998.



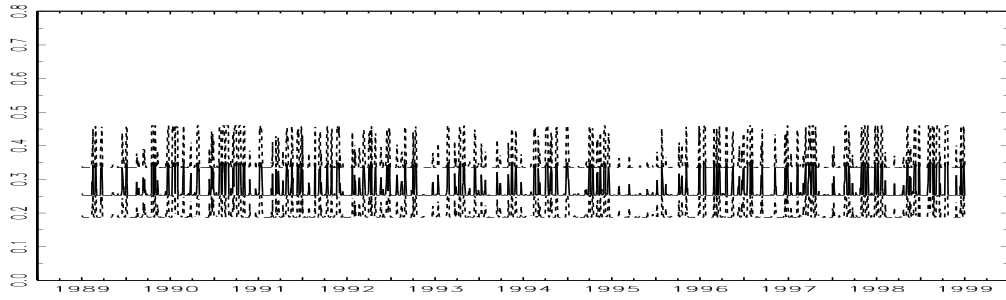
(a) Transition function against x_t



(b) Correlation quantiles against x_t



(c) Transition function over time



(d) Correlation quantiles over time

Figure 6: Logistic transition function $G(x_t; \gamma, c)$ as in (11), and average correlation between Dow Jones stocks (solid line) and 10th and 90th percentiles (dashed lines) obtained from the STCC model (10) with x_t being the one-day lagged weekly return on the value-weighted market portfolio $r_{M,w,t-1}$, estimated using daily returns over the period January 1, 1989-December 31, 1998.

Table 1: Characteristics of daily returns on DJIA stocks - 1989–1998

| | Mean | St.Dev | $\beta_{i,M}$ | $\beta_{i,SMB}$ | $\beta_{i,HML}$ |
|------|--------|--------|---------------|-----------------|-----------------|
| AA | 9.834 | 26.273 | 1.248(0.072) | 0.157(0.082) | 0.739(0.116) |
| AXP | 14.803 | 31.448 | 1.566(0.070) | −0.232(0.090) | 0.419(0.111) |
| BA | 9.023 | 28.416 | 1.191(0.077) | 0.108(0.094) | 0.022(0.119) |
| C | 26.579 | 34.322 | 1.840(0.077) | −0.168(0.112) | 0.489(0.137) |
| CAT | 10.611 | 28.838 | 1.292(0.068) | 0.030(0.091) | 0.597(0.122) |
| DD | 13.064 | 25.196 | 1.115(0.065) | −0.419(0.082) | 0.425(0.097) |
| DIS | 17.021 | 26.772 | 1.101(0.062) | −0.067(0.080) | −0.102(0.105) |
| EK | 7.139 | 27.269 | 0.881(0.058) | −0.065(0.086) | 0.061(0.107) |
| GE | 22.288 | 21.788 | 1.058(0.040) | −0.593(0.060) | −0.180(0.070) |
| GM | 6.112 | 28.822 | 1.637(0.066) | −0.116(0.083) | 1.498(0.119) |
| HD | 36.930 | 30.892 | 1.317(0.061) | −0.095(0.085) | −0.603(0.107) |
| HON | 16.721 | 27.152 | 1.240(0.085) | −0.060(0.109) | 0.524(0.120) |
| HPQ | 16.566 | 35.701 | 1.058(0.080) | −0.293(0.112) | −0.856(0.151) |
| IBM | 11.084 | 27.831 | 0.826(0.057) | −0.336(0.078) | −0.592(0.111) |
| INTC | 36.946 | 38.068 | 1.076(0.094) | −0.343(0.106) | −1.297(0.173) |
| IP | 6.569 | 25.008 | 1.261(0.063) | 0.042(0.074) | 0.847(0.099) |
| JNJ | 20.843 | 24.487 | 0.593(0.052) | −0.840(0.073) | −0.958(0.099) |
| JPM | 15.317 | 33.808 | 1.999(0.094) | 0.018(0.108) | 1.383(0.137) |
| KO | 25.139 | 24.056 | 0.763(0.056) | −0.964(0.070) | −0.619(0.086) |
| MCD | 18.858 | 24.841 | 0.801(0.062) | −0.436(0.079) | −0.289(0.097) |
| MMM | 8.797 | 20.664 | 0.799(0.047) | −0.360(0.063) | 0.139(0.080) |
| MO | 18.525 | 27.321 | 0.719(0.065) | −0.593(0.074) | −0.545(0.107) |
| MRK | 20.475 | 25.262 | 0.620(0.055) | −0.790(0.065) | −0.992(0.092) |
| MSFT | 45.455 | 35.043 | 0.889(0.070) | −0.430(0.087) | −1.603(0.129) |
| PG | 21.597 | 23.243 | 0.765(0.051) | −0.723(0.069) | −0.413(0.082) |
| SBC | 16.977 | 23.078 | 0.840(0.049) | −0.669(0.073) | 0.410(0.098) |
| T | 13.573 | 24.522 | 0.842(0.061) | −0.466(0.079) | −0.162(0.097) |
| UTX | 16.922 | 23.731 | 1.190(0.054) | 0.092(0.073) | 0.447(0.095) |
| WMT | 23.745 | 28.419 | 1.033(0.059) | −0.686(0.082) | −0.496(0.101) |
| XOM | 12.199 | 19.795 | 0.589(0.054) | −0.695(0.069) | 0.132(0.081) |

Note: The table reports the mean and standard deviation (in annualized percentage points) of daily DJIA stock returns over the period January 1, 1989 – December 31, 1998, together with coefficient estimates and heteroskedasticity-consistent standard errors from the three-factor model given in (18).

Table 2: Characteristics of daily returns on DJIA stocks - 1999–2003

| | Mean | St.Dev | $\beta_{i,M}$ | $\beta_{i,SMB}$ | $\beta_{i,HML}$ |
|------|---------|--------|---------------|-----------------|-----------------|
| AA | 10.981 | 42.868 | 1.513(0.072) | 0.073(0.113) | 1.426(0.147) |
| AXP | 6.821 | 41.112 | 1.424(0.072) | −0.635(0.100) | 0.343(0.123) |
| BA | 3.245 | 38.566 | 1.183(0.110) | 0.107(0.123) | 0.967(0.176) |
| C | 14.649 | 39.029 | 1.477(0.086) | −0.571(0.100) | 0.401(0.146) |
| CAT | 9.088 | 37.245 | 1.169(0.068) | −0.243(0.114) | 0.870(0.130) |
| DD | −6.524 | 34.531 | 1.112(0.070) | −0.362(0.105) | 1.036(0.116) |
| DIS | −5.558 | 42.299 | 1.290(0.112) | 0.158(0.122) | 0.564(0.175) |
| EK | −22.289 | 38.188 | 0.964(0.088) | 0.078(0.145) | 0.624(0.155) |
| GE | −3.011 | 35.971 | 1.212(0.066) | −0.556(0.078) | 0.069(0.108) |
| GM | −6.686 | 38.084 | 1.471(0.081) | −0.041(0.097) | 1.147(0.139) |
| HD | −1.170 | 45.755 | 1.269(0.082) | −0.425(0.125) | 0.271(0.154) |
| HON | −7.353 | 46.291 | 1.456(0.103) | −0.232(0.123) | 0.815(0.169) |
| HPQ | −3.961 | 54.846 | 1.340(0.088) | 0.362(0.130) | −0.509(0.159) |
| IBM | −0.465 | 40.505 | 0.896(0.062) | −0.330(0.099) | −0.496(0.112) |
| INTC | 1.816 | 58.322 | 1.419(0.092) | −0.240(0.137) | −0.888(0.165) |
| IP | −1.423 | 37.741 | 1.204(0.069) | −0.167(0.107) | 1.211(0.124) |
| JNJ | 4.101 | 28.501 | 0.458(0.071) | −0.629(0.088) | 0.132(0.119) |
| JPM | −6.019 | 44.804 | 1.678(0.102) | −0.407(0.121) | 0.550(0.180) |
| KO | −7.716 | 30.837 | 0.404(0.062) | −0.685(0.104) | 0.244(0.107) |
| MCD | −8.982 | 33.925 | 0.575(0.068) | −0.404(0.099) | 0.303(0.113) |
| MMM | 15.600 | 29.008 | 0.854(0.058) | −0.365(0.082) | 0.620(0.115) |
| MO | −2.553 | 39.104 | 0.512(0.089) | −0.355(0.132) | 0.603(0.166) |
| MRK | −9.816 | 32.048 | 0.584(0.064) | −0.722(0.095) | 0.161(0.121) |
| MSFT | −6.231 | 43.732 | 1.108(0.068) | −0.067(0.136) | −0.584(0.128) |
| PG | 1.997 | 33.769 | 0.320(0.077) | −0.647(0.100) | 0.120(0.113) |
| SBC | −16.382 | 39.224 | 0.935(0.085) | −0.504(0.115) | 0.487(0.139) |
| T | −28.409 | 47.886 | 1.192(0.093) | −0.189(0.142) | 0.537(0.171) |
| UTX | 9.378 | 38.708 | 1.389(0.174) | −0.083(0.147) | 1.087(0.258) |
| WMT | 7.921 | 37.369 | 0.811(0.062) | −0.813(0.111) | −0.102(0.120) |
| XOM | 0.154 | 27.661 | 0.776(0.059) | −0.304(0.091) | 0.757(0.108) |

Note: The table reports the mean and standard deviation (in annualized percentage points) of daily DJIA stock returns over the period January 1, 1999 – December 31, 2003, together with coefficient estimates and heteroskedasticity-consistent standard errors from the three-factor model given in (18).

Table 3: Distributions of variances, covariances and correlations of daily returns on DJIA stocks

| | Variances | Covariances | Correlations |
|---------------------------|-----------|-------------|--------------|
| <u>Panel A: 1989–1998</u> | | | |
| Mean | 3.045 | 0.834 | 0.282 |
| St.Dev | 1.041 | 0.279 | 0.066 |
| Minimum | 1.549 | 0.353 | 0.157 |
| 10th percentile | 1.876 | 0.547 | 0.209 |
| 25th percentile | 2.370 | 0.646 | 0.235 |
| Median | 2.833 | 0.806 | 0.271 |
| 75th percentile | 3.287 | 0.959 | 0.316 |
| 90th percentile | 4.656 | 1.145 | 0.369 |
| Maximum | 5.728 | 2.610 | 0.569 |
| <u>Panel B: 1999–2003</u> | | | |
| Mean | 6.127 | 1.686 | 0.280 |
| St.Dev | 2.271 | 0.908 | 0.119 |
| Minimum | 2.947 | 0.230 | 0.032 |
| 10th percentile | 3.247 | 0.729 | 0.139 |
| 25th percentile | 4.591 | 0.971 | 0.191 |
| Median | 5.744 | 1.512 | 0.268 |
| 75th percentile | 7.139 | 2.247 | 0.353 |
| 90th percentile | 8.260 | 2.879 | 0.436 |
| Maximum | 13.111 | 6.228 | 0.751 |

Table 4: In-sample performance and characteristics of minimum variance portfolios based on forecasting models - 1989–1998

| | Standard | | Sharpe | Tracking | | | |
|-------------|----------|-----------|--------|----------|--------------|---------------|---------------|
| | Mean | Deviation | Ratio | Error | β_M | β_{SMB} | β_{HML} |
| SF | 9.24 | 13.42 | 0.306 | 10.89 | 0.675(0.037) | −0.420(0.051) | 0.287(0.057) |
| TF | 8.51 | 13.80 | 0.245 | 11.20 | 0.806(0.041) | −0.155(0.057) | 0.374(0.059) |
| RM | 13.81 | 17.82 | 0.487 | 15.53 | 0.806(0.044) | −0.155(0.069) | 0.256(0.077) |
| CCC | 11.78 | 12.82 | 0.519 | 8.17 | 0.791(0.023) | −0.308(0.028) | 0.170(0.037) |
| DCC | 11.57 | 12.73 | 0.506 | 8.18 | 0.786(0.023) | −0.301(0.029) | 0.172(0.037) |
| ADCC | 11.19 | 12.72 | 0.476 | 8.25 | 0.785(0.023) | −0.299(0.029) | 0.176(0.037) |
| STCC(h) | 11.40 | 12.71 | 0.492 | 8.42 | 0.780(0.024) | −0.305(0.030) | 0.195(0.039) |
| STCC(r) | 10.19 | 12.69 | 0.398 | 8.39 | 0.775(0.023) | −0.310(0.030) | 0.187(0.039) |
| SEMI(h) | 10.53 | 12.63 | 0.427 | 8.71 | 0.765(0.024) | −0.300(0.030) | 0.212(0.039) |
| SEMI(r) | 10.11 | 12.54 | 0.397 | 8.57 | 0.761(0.025) | −0.310(0.030) | 0.206(0.038) |
| EWP | 17.99 | 15.16 | 0.848 | 4.16 | 1.072(0.013) | −0.333(0.017) | −0.053(0.021) |
| VWP | 21.52 | 17.95 | 0.913 | 6.83 | 1.114(0.017) | −0.410(0.021) | −0.458(0.027) |

Note: The table reports in-sample (January 1, 1989 – December 31, 1998) summary statistics for the minimum variance portfolios based on the single factor (SF) model, the three factor (TF) model, the RiskMetrics (RM) model, the Constant Conditional Correlation (CCC) model, the Dynamic Conditional Correlation (DCC) model, the Asymmetric DCC (ADCC) model, the Smooth Transition Conditional Correlation (STCC) model with x_t equal to the daily log-volatility for the market return (h) or the one-day lagged weekly index return (r), and the semi-parametric conditional correlation (SEMI) model with the same choices of x_t . In addition, summary statistics are reported for the equally-weighted and value-weighted portfolios (EWP and VWP). For each MVP, the table shows the mean and standard deviation (in annualized percentage points) of the portfolio return, the annualized Sharpe ratio, the annualized tracking error relative to the S&P 500 index, and coefficient estimates and standard errors from the Fama-French three-factor model (18).

Table 5: Correlation models for Dow Jones stocks: In-sample results

| | SF | TF | RM | CCC | DCC | ADCC | STCC, $x_t =$ | | $a =$ | Semi-parametric model | | | | | | | |
|---|-------|-------|-------|-------|-------|-------|-----------------|---------------|-------|-----------------------|-------|-------|-------|---------------------|-------|-------|-------|
| | | | | | | | $\log(h_{M,t})$ | | | $x_t = \log(h_{M,t})$ | | | | $x_t = r_{M,w,t-1}$ | | | |
| | | | | | | | $\log(h_{M,t})$ | $r_{M,w,t-1}$ | | 0 | 0.25 | 0.50 | 0.75 | 0 | 0.25 | 0.50 | 0.75 |
| Portfolio standard deviations | | | | | | | | | | | | | | | | | |
| MVP | 1.142 | 1.078 | 2.678 | 1.097 | 1.086 | 1.089 | 1.096 | 1.097 | | 1.096 | 1.091 | 1.086 | 1.081 | 1.087 | 1.091 | 1.086 | 1.073 |
| EWP | 1.017 | 0.938 | 1.073 | 0.993 | 0.989 | 0.988 | 0.990 | 0.992 | | 1.000 | 0.991 | 0.970 | 0.952 | 1.003 | 0.999 | 0.982 | 0.956 |
| VWP | 1.024 | 0.905 | 1.070 | 1.007 | 1.001 | 1.001 | 1.005 | 1.005 | | 1.014 | 1.005 | 0.985 | 0.968 | 1.017 | 1.012 | 0.994 | 0.968 |
| MVP - Value-at-Risk violations | | | | | | | | | | | | | | | | | |
| 1% | 1.19 | 1.15 | 12.78 | 1.27 | 1.23 | 1.31 | 1.35 | 1.43 | | 1.35 | 1.31 | 1.35 | 1.42 | 1.54 | 1.58 | 1.54 | 1.62 |
| 5% | 5.22 | 5.26 | 22.71 | 5.86 | 5.82 | 5.94 | 6.02 | 5.82 | | 5.94 | 5.86 | 6.13 | 6.02 | 5.78 | 5.86 | 6.21 | 6.41 |
| 10% | 10.29 | 10.25 | 27.58 | 11.24 | 11.28 | 11.32 | 11.36 | 11.40 | | 11.59 | 11.59 | 11.36 | 11.32 | 11.52 | 11.52 | 11.52 | 11.32 |
| MVP: p -value of Dynamic Quantile statistic | | | | | | | | | | | | | | | | | |
| 1% | 0.267 | 0.037 | 0.000 | 0.041 | 0.034 | 0.046 | 0.037 | 0.018 | | 0.035 | 0.044 | 0.044 | 0.043 | 0.048 | 0.045 | 0.045 | 0.019 |
| 5% | 0.759 | 0.410 | 0.000 | 0.038 | 0.054 | 0.061 | 0.059 | 0.137 | | 0.084 | 0.089 | 0.050 | 0.047 | 0.056 | 0.065 | 0.041 | 0.037 |
| 10% | 0.676 | 0.878 | 0.000 | 0.030 | 0.048 | 0.029 | 0.068 | 0.038 | | 0.045 | 0.045 | 0.065 | 0.046 | 0.030 | 0.039 | 0.033 | 0.055 |
| EWP - Value-at-Risk violations | | | | | | | | | | | | | | | | | |
| 1% | 0.95 | 0.95 | 0.95 | 0.95 | 0.95 | 0.95 | 0.95 | 0.95 | | 0.95 | 0.95 | 0.95 | 0.95 | 0.95 | 0.95 | 0.95 | 0.95 |
| 5% | 4.94 | 4.94 | 4.94 | 4.94 | 4.94 | 4.94 | 4.94 | 4.94 | | 4.94 | 4.94 | 4.94 | 4.94 | 4.94 | 4.94 | 4.94 | 4.94 |
| 10% | 9.97 | 9.97 | 9.97 | 9.97 | 9.97 | 9.97 | 9.97 | 9.97 | | 9.97 | 9.97 | 9.97 | 9.97 | 9.97 | 9.97 | 9.97 | 9.97 |
| EWP: p -value of Dynamic Quantile statistic | | | | | | | | | | | | | | | | | |
| 1% | 0.918 | 0.918 | 0.085 | 0.795 | 0.799 | 0.799 | 0.708 | 0.418 | | 0.769 | 0.699 | 0.735 | 0.758 | 0.461 | 0.494 | 0.484 | 0.379 |
| 5% | 0.585 | 0.585 | 0.153 | 0.073 | 0.072 | 0.076 | 0.309 | 0.098 | | 0.291 | 0.735 | 0.601 | 0.640 | 0.096 | 0.131 | 0.186 | 0.148 |
| 10% | 0.672 | 0.803 | 0.109 | 0.024 | 0.027 | 0.035 | 0.069 | 0.027 | | 0.040 | 0.053 | 0.093 | 0.098 | 0.020 | 0.030 | 0.040 | 0.046 |
| VWP - Value-at-Risk violations | | | | | | | | | | | | | | | | | |
| 1% | 0.87 | 0.87 | 1.03 | 1.19 | 1.07 | 1.07 | 0.99 | 0.95 | | 0.99 | 0.95 | 0.95 | 0.95 | 0.87 | 0.91 | 0.91 | 0.91 |
| 5% | 4.59 | 4.67 | 5.14 | 5.03 | 5.03 | 4.99 | 4.95 | 5.03 | | 4.91 | 4.79 | 4.87 | 4.83 | 5.07 | 5.07 | 5.10 | 5.07 |
| 10% | 10.57 | 10.61 | 10.09 | 10.21 | 10.37 | 10.37 | 10.53 | 10.41 | | 10.45 | 10.49 | 10.37 | 10.37 | 10.33 | 10.37 | 10.33 | 10.41 |
| VWP: p -value of Dynamic Quantile statistic | | | | | | | | | | | | | | | | | |
| 1% | 0.093 | 0.089 | 0.136 | 0.162 | 0.766 | 0.774 | 0.819 | 0.658 | | 0.927 | 0.885 | 0.660 | 0.689 | 0.721 | 0.844 | 0.501 | 0.743 |
| 5% | 0.528 | 0.715 | 0.319 | 0.241 | 0.122 | 0.171 | 0.500 | 0.177 | | 0.368 | 0.459 | 0.516 | 0.570 | 0.176 | 0.137 | 0.149 | 0.123 |
| 10% | 0.478 | 0.421 | 0.146 | 0.103 | 0.122 | 0.122 | 0.154 | 0.190 | | 0.196 | 0.183 | 0.206 | 0.211 | 0.186 | 0.244 | 0.192 | 0.190 |

Note: The table reports in-sample results for dynamic conditional correlation models for Dow Jones stocks, estimated using daily returns over the period January 1, 1989-December 31, 1998. Results are based on the single factor (SF) model, the three factor (TF) model, the RiskMetrics (RM) model, the Constant Conditional Correlation (CCC) model, the Dynamic Conditional Correlation (DCC) model, the Asymmetric DCC (ADCC) model, the Smooth Transition Conditional Correlation (STCC) model with x_t equal to the logged contemporaneous market volatility ($\log(h_{M,t})$) or the one-day lagged weekly market portfolio return ($r_{M,w,t-1}$), and the semi-parametric conditional correlation (SEMI) model with the same choices of x_t . MVP denotes minimum variance portfolio, EWP and VWP indicate equally-weighted and value-weighted portfolios, respectively. Portfolio standard deviations refer to the in-sample standard deviation of standardized portfolio returns. Value-at-Risk violation percentages and Dynamic Quantile statistics for the corresponding HIT-sequences concern one-day ahead in-sample forecasts.

Table 6: Out-of-sample performance and characteristics of minimum variance portfolios based on forecasting models - 1999–2003

| | Mean | Standard Deviation | Sharpe Ratio | Tracking Error | β_M | β_{SMB} | β_{HML} |
|-------------|--------|-----------------------|-----------------|-------------------|--------------|---------------|---------------|
| SF | −8.78 | 19.46 | −0.628 | 23.50 | 0.509(0.061) | −0.329(0.074) | 0.671(0.108) |
| TF | −12.42 | 20.46 | −0.775 | 21.75 | 0.632(0.081) | −0.129(0.086) | 0.572(0.137) |
| RM | −16.39 | 22.52 | −0.881 | 24.13 | 0.450(0.056) | −0.243(0.067) | 0.170(0.088) |
| CCC | −7.48 | 17.51 | −0.624 | 15.97 | 0.705(0.035) | −0.269(0.048) | 0.507(0.064) |
| DCC | −8.66 | 17.20 | −0.704 | 15.97 | 0.677(0.036) | −0.261(0.047) | 0.459(0.064) |
| ADCC | −9.26 | 17.20 | −0.739 | 16.05 | 0.674(0.037) | −0.262(0.047) | 0.459(0.064) |
| STCC(h) | −8.29 | 17.69 | −0.664 | 17.09 | 0.675(0.038) | −0.260(0.049) | 0.518(0.067) |
| STCC(r) | −8.87 | 18.23 | −0.712 | 17.43 | 0.676(0.037) | −0.261(0.048) | 0.487(0.066) |
| SEMI(h) | −8.26 | 18.11 | −0.646 | 17.38 | 0.678(0.040) | −0.260(0.051) | 0.511(0.070) |
| SEMI(r) | −8.63 | 18.54 | −0.671 | 16.97 | 0.677(0.036) | −0.260(0.050) | 0.506(0.069) |
| EWP | −1.98 | 21.82 | −0.249 | 7.98 | 1.066(0.024) | −0.307(0.034) | 0.425(0.039) |
| VWP | −4.42 | 19.67 | −0.400 | 7.53 | 0.938(0.023) | −0.298(0.029) | 0.313(0.037) |

Note: The table reports out-of-sample (January 1, 1999 – December 31, 2003) summary statistics for the minimum variance portfolios based on various forecasting models. See Table 4 for models definitions and description of the summary statistics shown.

Table 7: Correlation models for Dow Jones stocks: Out-of-sample results

| | | | | | | | | | | Semi-parametric model | | | | | | | |
|---|-------|-------|-------|-------|-------|-----------------|---------------|-------|-----------------------|-----------------------|-------|-------|---------------------|-------|-------|-------|-------|
| SF | TF | RM | CCC | DCC | ADCC | STCC, $x_t =$ | | $a =$ | $x_t = \log(h_{M,t})$ | | | | $x_t = r_{M,w,t-1}$ | | | | |
| | | | | | | $\log(h_{M,t})$ | $r_{M,w,t-1}$ | | 0 | 0.25 | 0.50 | 0.75 | 0 | 0.25 | 0.50 | 0.75 | |
| Portfolio standard deviations | | | | | | | | | | | | | | | | | |
| MVP | 1.381 | 1.195 | 2.645 | 1.240 | 1.208 | 1.208 | 1.238 | 1.554 | | 1.254 | 1.252 | 1.252 | 1.253 | 1.323 | 1.308 | 1.266 | 1.248 |
| EWP | 1.054 | 0.944 | 1.034 | 1.124 | 1.097 | 1.095 | 1.065 | 1.092 | | 1.060 | 1.052 | 1.043 | 1.037 | 1.132 | 1.123 | 1.107 | 1.101 |
| VWP | 1.054 | 0.932 | 1.038 | 1.135 | 1.106 | 1.103 | 1.072 | 1.097 | | 1.067 | 1.059 | 1.050 | 1.043 | 1.107 | 1.101 | 1.088 | 1.081 |
| MVP - Value-at-Risk violations | | | | | | | | | | | | | | | | | |
| 1% | 1.32 | 0.74 | 14.32 | 2.06 | 2.14 | 2.06 | 2.22 | 2.47 | | 2.39 | 2.47 | 2.39 | 2.39 | 2.63 | 2.47 | 2.31 | 2.55 |
| 5% | 6.42 | 6.58 | 25.27 | 9.55 | 9.63 | 9.71 | 9.79 | 9.79 | | 10.29 | 10.45 | 10.45 | 10.70 | 9.49 | 9.79 | 9.79 | 10.19 |
| 10% | 12.10 | 13.74 | 30.95 | 16.71 | 17.28 | 17.37 | 16.95 | 17.28 | | 17.28 | 17.04 | 17.20 | 17.45 | 17.28 | 17.27 | 17.04 | 17.68 |
| MVP: p -value of Dynamic Quantile statistic | | | | | | | | | | | | | | | | | |
| 1% | 0.517 | 0.151 | 0.001 | 0.015 | 0.009 | 0.007 | 0.018 | 0.047 | | 0.028 | 0.049 | 0.046 | 0.046 | 0.014 | 0.022 | 0.055 | 0.067 |
| 5% | 0.111 | 0.071 | 0.000 | 0.005 | 0.002 | 0.002 | 0.004 | 0.014 | | 0.004 | 0.003 | 0.002 | 0.002 | 0.020 | 0.012 | 0.013 | 0.008 |
| 10% | 0.061 | 0.053 | 0.000 | 0.006 | 0.004 | 0.004 | 0.005 | 0.009 | | 0.005 | 0.006 | 0.006 | 0.005 | 0.006 | 0.005 | 0.005 | 0.004 |
| EWP - Value-at-Risk violations | | | | | | | | | | | | | | | | | |
| 1% | 0.82 | 0.82 | 0.74 | 1.15 | 0.99 | 0.99 | 0.99 | 1.19 | | 0.91 | 0.91 | 0.91 | 0.91 | 1.04 | 1.04 | 1.11 | 1.11 |
| 5% | 4.61 | 4.61 | 5.60 | 7.57 | 7.49 | 7.49 | 7.16 | 5.81 | | 7.08 | 6.91 | 6.75 | 6.83 | 5.89 | 5.97 | 5.81 | 5.57 |
| 10% | 12.35 | 12.35 | 11.36 | 14.98 | 14.73 | 14.81 | 13.91 | 12.26 | | 14.07 | 13.99 | 13.91 | 14.07 | 12.58 | 12.58 | 12.58 | 12.42 |
| EWP: p -value of Dynamic Quantile statistic | | | | | | | | | | | | | | | | | |
| 1% | 0.560 | 0.546 | 0.398 | 0.272 | 0.498 | 0.497 | 0.568 | 0.565 | | 0.390 | 0.398 | 0.408 | 0.407 | 0.546 | 0.573 | 0.584 | 0.488 |
| 5% | 0.375 | 0.466 | 0.109 | 0.111 | 0.135 | 0.135 | 0.193 | 0.785 | | 0.320 | 0.332 | 0.334 | 0.337 | 0.518 | 0.518 | 0.582 | 0.676 |
| 10% | 0.173 | 0.122 | 0.107 | 0.021 | 0.033 | 0.026 | 0.051 | 0.319 | | 0.039 | 0.045 | 0.047 | 0.041 | 0.164 | 0.190 | 0.208 | 0.281 |
| VWP - Value-at-Risk violations | | | | | | | | | | | | | | | | | |
| 1% | 0.91 | 0.82 | 0.58 | 1.32 | 1.15 | 0.99 | 0.99 | 1.19 | | 0.99 | 0.99 | 0.99 | 0.99 | 1.11 | 1.11 | 1.11 | 1.19 |
| 5% | 5.10 | 4.86 | 5.76 | 7.90 | 7.41 | 7.49 | 7.49 | 6.05 | | 7.16 | 7.00 | 6.83 | 7.00 | 6.37 | 6.29 | 6.13 | 5.97 |
| 10% | 12.51 | 11.93 | 11.60 | 15.47 | 14.90 | 14.98 | 14.07 | 12.66 | | 14.07 | 13.99 | 14.16 | 14.16 | 12.66 | 12.74 | 12.81 | 12.50 |
| VWP: p -value of Dynamic Quantile statistic | | | | | | | | | | | | | | | | | |
| 1% | 0.809 | 0.657 | 0.258 | 0.316 | 0.173 | 0.145 | 0.540 | 0.772 | | 0.568 | 0.567 | 0.563 | 0.561 | 0.703 | 0.713 | 0.741 | 0.810 |
| 5% | 0.266 | 0.167 | 0.046 | 0.051 | 0.184 | 0.183 | 0.132 | 0.807 | | 0.232 | 0.249 | 0.307 | 0.212 | 0.559 | 0.461 | 0.585 | 0.680 |
| 10% | 0.075 | 0.073 | 0.087 | 0.014 | 0.026 | 0.028 | 0.031 | 0.123 | | 0.043 | 0.044 | 0.036 | 0.036 | 0.081 | 0.110 | 0.116 | 0.162 |

Note: The table reports out-of-sample results for dynamic correlation models for Dow Jones stocks over the period January 1, 1999-December 31, 2003. Results are based on the single factor (SF) model, the three factor (TF) model, the RiskMetrics (RM) model, the Constant Conditional Correlation (CCC) model, the Dynamic Conditional Correlation (DCC) model, the Asymmetric DCC (ADCC) model, the Smooth Transition Conditional Correlation (STCC) model with x_t equal to the logged contemporaneous market volatility ($\log(h_{M,t})$) or the one-day lagged weekly market portfolio return ($r_{M,w,t-1}$), and the semi-parametric conditional correlation (SEMI) model with the same choices of x_t . All models are re-estimated every 20 observations using a ten-year moving window of daily return observations. MVP denotes minimum variance portfolio, EWP and VWP indicate equally-weighted and value-weighted portfolios, respectively. Portfolio standard deviations are the standard deviation of standardized portfolio return forecasts. Value-at-Risk violation percentages and Dynamic Quantile statistics for the corresponding HIT-sequences concern one-day ahead out-of-sample forecasts.

Table 8: Summary statistics for portfolio weights

| | Minimum variance | | Minimum tracking error | |
|------|------------------|--------------|------------------------|------------|
| | 1989-1998 | 1999-2003 | 1989-1998 | 1999-2003 |
| AA | 4.28(3.57) | 0.34(3.67) | 2.43(0.79) | 1.31(0.74) |
| AXP | -0.98(3.17) | -1.74(3.83) | 1.84(0.69) | 2.64(1.20) |
| BA | 3.95(4.43) | 3.91(3.94) | 2.31(0.63) | 1.92(0.70) |
| CAT | 2.94(3.40) | 0.51(2.69) | 1.96(0.62) | 0.86(0.87) |
| C | -5.15(1.71) | -1.04(4.45) | 2.26(0.46) | 4.55(1.43) |
| DD | -0.59(3.82) | -1.12(3.38) | 3.01(0.79) | 0.93(1.32) |
| DIS | 3.08(3.23) | 0.55(3.04) | 2.23(0.58) | 1.91(0.76) |
| EK | 6.06(3.72) | 11.88(8.38) | 2.30(0.49) | 2.37(1.54) |
| GE | 0.42(8.34) | -6.03(6.29) | 6.08(1.61) | 5.81(1.74) |
| GM | 1.38(3.05) | 4.62(4.64) | 2.90(0.77) | 3.34(1.44) |
| HD | -0.03(3.38) | -0.31(4.75) | 2.72(0.71) | 2.12(1.50) |
| HON | 3.88(4.29) | 1.18(5.52) | 2.39(0.84) | 3.08(2.02) |
| HPQ | -0.42(1.89) | 0.35(2.32) | 1.68(0.33) | 2.69(0.94) |
| IBM | 7.26(5.77) | 6.22(5.24) | 4.15(1.56) | 4.92(1.41) |
| INTC | -1.56(1.67) | -2.26(2.06) | 2.15(0.37) | 4.57(1.38) |
| IP | 4.33(4.26) | 4.36(4.97) | 2.35(0.66) | 2.49(1.39) |
| JNJ | 1.95(4.96) | 11.84(8.41) | 3.22(1.03) | 5.00(1.87) |
| JPM | 2.03(3.93) | -0.28(3.21) | 3.91(1.35) | 4.57(1.28) |
| KO | -1.60(4.34) | 4.52(6.77) | 3.46(1.19) | 1.83(1.54) |
| MCD | 5.67(4.37) | 5.83(5.04) | 2.33(0.82) | 1.85(0.96) |
| MMM | 12.46(7.59) | 6.30(4.76) | 3.94(1.16) | 2.65(1.46) |
| MO | 3.12(3.51) | 3.97(4.52) | 2.94(1.13) | 1.69(0.71) |
| MRK | 3.73(4.17) | 1.31(5.11) | 4.75(1.14) | 4.03(1.49) |
| MSFT | 0.09(2.49) | 3.31(4.20) | 1.88(0.58) | 5.36(1.96) |
| PG | 4.19(5.00) | 12.45(10.56) | 4.07(0.94) | 3.68(1.93) |
| SBC | 8.38(6.77) | 3.58(5.38) | 5.91(1.34) | 4.42(1.29) |
| T | 6.41(5.42) | 1.85(3.08) | 4.58(1.22) | 3.46(0.88) |
| UTX | 6.69(4.80) | 4.85(7.47) | 3.51(0.97) | 2.59(1.92) |
| WMT | -1.48(3.17) | 2.11(5.83) | 2.80(0.82) | 4.90(1.60) |
| XOM | 19.50(6.93) | 16.97(7.96) | 9.95(1.81) | 8.46(2.44) |

Note: The table contains means and standard deviations of in-sample (1989–1998) and out-of-sample (1999–2003) portfolio weights in the minimum variance portfolio and minimum tracking error portfolio based on the semi-parametric conditional correlation model with x_t equal to the daily log-volatility for the market return.

Table 9: In-sample performance and characteristics of minimum tracking error portfolios based on forecasting models - 1989–1998

| | Mean | Standard Deviation | Sharpe Ratio | Tracking Error | β_M | β_{SMB} | β_{HML} |
|-------------------|-------|-----------------------|-----------------|-------------------|--------------|---------------|---------------|
| SF | 17.39 | 14.36 | 0.853 | 3.64 | 0.998(0.012) | −0.387(0.015) | −0.030(0.017) |
| TF | 16.82 | 14.39 | 0.812 | 3.75 | 1.020(0.012) | −0.346(0.015) | −0.010(0.018) |
| RM | 17.52 | 14.86 | 0.832 | 4.61 | 1.010(0.015) | −0.397(0.018) | −0.064(0.023) |
| CCC | 17.57 | 14.40 | 0.864 | 3.37 | 0.997(0.011) | −0.403(0.014) | −0.045(0.016) |
| DCC | 17.53 | 14.40 | 0.861 | 3.35 | 1.000(0.011) | −0.402(0.014) | −0.047(0.016) |
| ADCC | 17.49 | 14.37 | 0.862 | 3.32 | 1.007(0.011) | −0.401(0.014) | −0.046(0.016) |
| STCC(h_M) | 17.55 | 14.39 | 0.863 | 3.34 | 1.000(0.011) | −0.398(0.014) | −0.042(0.016) |
| STCC(h_{SMB}) | 17.48 | 14.39 | 0.857 | 3.35 | 0.996(0.011) | −0.403(0.014) | −0.049(0.016) |
| STCC(h_{HML}) | 17.58 | 14.39 | 0.865 | 3.32 | 0.998(0.010) | −0.402(0.014) | −0.046(0.016) |
| SEMI(h_M) | 17.34 | 14.37 | 0.849 | 3.28 | 1.004(0.010) | −0.390(0.013) | −0.035(0.015) |
| SEMI(h_{SMB}) | 17.37 | 14.44 | 0.848 | 3.28 | 1.006(0.010) | −0.395(0.014) | −0.044(0.016) |
| SEMI(h_{HML}) | 17.65 | 14.36 | 0.871 | 3.26 | 0.998(0.010) | −0.398(0.014) | −0.044(0.015) |
| EWP | 17.99 | 15.16 | 0.848 | 4.16 | 1.072(0.013) | −0.333(0.017) | −0.053(0.021) |
| VWP | 21.52 | 17.95 | 0.913 | 6.83 | 1.114(0.017) | −0.410(0.021) | −0.458(0.027) |

Note: The table reports in-sample (January 1, 1989 – December 31, 1998) summary statistics for the minimum tracking error portfolios based on various forecasting models. h_M , h_{SMB} and h_{HML} indicate that the variable x_t in the STCC and semi-parametric models is taken to be the (contemporaneous) daily log-volatility for the market, SMB and HML factor returns, respectively. See Table 4 for definitions of the other models and description of the summary statistics shown.

Table 10: In-sample performance and characteristics of minimum tracking error portfolios based on forecasting models - 1999–2003

| | Mean | Standard Deviation | Sharpe Ratio | Tracking Error | β_M | β_{SMB} | β_{HML} |
|-------------------|-------|-----------------------|-----------------|-------------------|--------------|---------------|---------------|
| SF | −3.16 | 21.96 | −0.301 | 7.91 | 1.062(0.025) | −0.330(0.034) | 0.406(0.040) |
| TF | −4.80 | 22.42 | −0.368 | 7.70 | 1.070(0.030) | −0.293(0.036) | 0.344(0.047) |
| RM | −5.25 | 23.25 | −0.374 | 7.88 | 1.019(0.018) | −0.296(0.028) | 0.098(0.035) |
| CCC | −3.14 | 21.61 | −0.305 | 6.91 | 1.032(0.020) | −0.346(0.030) | 0.332(0.034) |
| DCC | −3.69 | 21.70 | −0.329 | 6.52 | 1.025(0.020) | −0.342(0.029) | 0.284(0.033) |
| ADCC | −3.87 | 21.76 | −0.339 | 6.78 | 1.029(0.020) | −0.344(0.030) | 0.315(0.034) |
| STCC(h_M) | −3.18 | 22.03 | −0.301 | 6.60 | 1.036(0.020) | −0.343(0.029) | 0.275(0.035) |
| STCC(h_{SMB}) | −3.23 | 22.16 | −0.316 | 6.56 | 1.041(0.020) | −0.314(0.026) | 0.285(0.032) |
| STCC(h_{HML}) | −3.31 | 21.90 | −0.321 | 6.61 | 1.037(0.020) | −0.341(0.030) | 0.223(0.031) |
| SEMI(h_M) | −2.40 | 22.31 | −0.262 | 6.65 | 1.045(0.020) | −0.343(0.030) | 0.265(0.034) |
| SEMI(h_{SMB}) | −3.30 | 22.06 | −0.306 | 6.50 | 1.051(0.019) | −0.304(0.026) | 0.288(0.032) |
| SEMI(h_{HML}) | −3.47 | 22.30 | −0.310 | 6.36 | 1.032(0.019) | −0.347(0.30) | 0.219(0.033) |
| EWP | −1.98 | 21.82 | −0.249 | 7.98 | 1.066(0.024) | −0.307(0.034) | 0.425(0.039) |
| VWP | −4.42 | 19.67 | −0.400 | 7.53 | 0.938(0.023) | −0.298(0.029) | 0.313(0.037) |

Note: The table reports out-of-sample (January 1, 1999 – December 31, 2003) summary statistics for the minimum tracking error portfolios based on various forecasting models. See Table 9 for models definitions and description of the summary statistics shown.