2019/25

DP

Luc Bauwens and Yongdeng Xu

DCC-HEAVY: A multivariate GARCH model based on realized variances and correlations

CORE
CENTER FOR OPERATIONS RESEARCH AND ECONOMETRICS

CORE

Voie du Roman Pays 34, L1.03.01

B-1348 Louvain-la-Neuve

Tel (32 10) 47 43 04

Email: immaq-library@uclouvain.be

https://uclouvain.be/en/research-institutes/lidam/core/discussion-papers.html

CORE DISCUSSION PAPER
2019/25

DCC-HEAVY: A MULTIVARIATE GARCH MODEL BASED ON REALIZED VARIANCES AND CORRELATIONS

Luc Bauwens
Université catholique de Louvain, CORE,
B-1348 Louvain-La-Neuve, Belgium
E-mail: luc.bauwens@uclouvain.be
Yongdeng Xu[1]
Cardiff Business School, Cardiff University, CF103EU, UK
E-mail: xuy16@cardiff.ac.uk

December 9, 2019

Abstract

This paper introduces the DCC-HEAVY and DECO-HEAVY models, which are dynamic models for conditional variances and correlations for daily returns based on measures of realized variances and correlations built from intraday data. Formulas for multi-step forecasts of conditional variances and correlations are provided. Asymmetric versions of the models are developed. An empirical study shows that in terms of forecasts the new HEAVY models outperform the BEKK-HEAVY model based on realized covariances, and the BEKK, DCC and DECO multivariate GARCH models based exclusively on daily data.

Keywords: Dynamic conditional correlations, Forecasting, Multivariate HEAVY, Multivariate GARCH, Realized correlations.

JEL Classification: C32, C58, G17

[1]Corresponding author.

1 Introduction

The covariance matrix of daily asset returns is used in several applications of financial management. Therefore, modeling the temporal dependence in the elements of the covariance matrix, often with the ultimate aim of forecasting, is a key area of financial econometrics. The most widespread model class for this purpose is that of multivariate generalized autoregressive conditional heteroscedasticity (MGARCH) models, wherein the conditional covariance matrix of daily returns is specified as a deterministic function of past daily returns. A survey of MGARCH models is provided by Bauwens et al. (2006).

The increasing availability of intraday data has led to the development of so-called "High-frEquency-bAsed VolatilitY"(HEAVY) models, initially in the univariate case (Engle and Gallo, 2006; Shephard and Sheppard, 2010). In the multivariate case, such models have been introduced in the stochastic volatility framework by Jin and Maheu (2013), and in the MGARCH one by Noureldin et al. (2012). In the latter framework, the main difference is that the conditional covariance matrix of daily returns is specified in the HEAVY case as a function of lagged realized covariances, instead of lagged outer products of daily returns in the traditional MGARCH case. Thus, multi-step forecasts of daily conditional covariance matrices by HEAVY models require as input forecasts of realized covariances, which are obtained from a dynamic model for the realized covariance matrix. HEAVY models are based on more accurate measurements of covariances than GARCH models, and they improve forecasts of the conditional covariance matrix of daily returns, as illustrated by Noureldin et al. (2012). A different class of models that relate the daily return volatility to a realized volatility measure is the realized GARCH model of Hansen et al. (2012); this type of model has been extended to the multivariate setup by Hansen et al. (2014) and Gorgi et al. (2019).

A well-known difficulty of MGARCH and HEAVY models is the so-called "curse of dimensionality". This means that the number of parameters of such models tends to be large, being at least a quadratic function of the dimension of the return vector of which the conditional covariance matrix is the focus of the modeling. Therefore, in empirical applications where the dimension is more than a handful of assets, a scalar parameterization is adopted. For example, Noureldin et al. (2012), for ten assets, and Opschoor et al. (2018), for thirty assets, adopt the scalar BEKK parameterization both for the daily conditional covariance matrix and for the conditional expectation of the realized covariance matrix. This parameterization, coupled with a targeting procedure, which is a way to estimate the constant parameter matrices of the model, facilitates the quasi-maximum likelihood (QML) estimation of the remaining scalar parameters enormously, since their number is independent of the dimension. The scalar parmaterization implies that the conditional variances and covariances follow the same dynamic pattern, in particular having the same persistence and the same sensitivity to the past realized covariances, which may be unrealistic for a large number of assets. A Factor HEAVY model is proposed by Sheppard and Xu (2019). This model avoids the restrictions of the scalar BEKK without incurring the curse of dimensionality.

New models are developed and named DCC-HEAVY and DEC-HEAVY since they use the scalar dynamic conditional correlation (DCC) and dynamic equicorrelation (DECO) formulations of MGARCH models for the specification of the dynamics of the daily conditional covariance matrix and of the conditional expectation of the realized covariance matrix. Since its introduction by Engle (2002), the scalar DCC formulation of MGARCH models has become much more widely used than the scalar BEKK one. This is due to the fact that the scalar DCC allows to specify the dynamics of the conditional variances differently for the different asset returns, and differently from the dynamics of the conditional correlation matrix, without rendering QML estimation complicated for a large number of assets, thanks to a two-step estimation procedure coupled with a targeting procedure for the matrix constant term. The same holds for the scalar DECO model of Engle and Kelly (2012), which is very easy to estimate because the inverse of the conditional covariance matrix is available analytically. Likewise, each component of the DCC- or DECO-HEAVY model can be estimated in two steps: for the daily covariance matrix process, the quasi-likelihood function results from the normal distribution (as for DCC and DECO), and for the process of the conditional expectation of the realized covariance matrix, it results from the Wishart distribution. Although the Wishart density has been used in this context to define a (quasi-)likelihood function (Gourieroux et al., 2009; Chiriac and Voev, 2011; Golosnoy et al., 2012; Noureldin et al., 2012; Opschoor et al., 2018), the possibility to split the Wishart based estimation in two steps – the first for the parameters of the realized variance processes, the second for the parameters of the realized correlation process – was introduced by Bauwens at al. (2012); see also Bauwens et al. (2016).

Specifically, the conditional covariance matrix of daily returns of the DCC- and DECO-HEAVY is written as the conditional correlation matrix pre- and post-multiplied by the corresponding diagonal matrix of the square roots of the conditional variances. The dynamics of the latter is driven by the lagged realized variances (as in univariate HEAVY models), and the dynamics of the conditional correlation matrix is driven by the lagged realized correlation matrix, which is an original feature with respect to the HEAVY model of Braione (2016). A vector multiplicative error representation of the DCC-HEAVY model is provided, where the dynamics is driven both by lagged measures and outer products of past returns, thus encompassing both DCC-HEAVY and DCC-GARCH. Stationarity conditions and formulas for multi-step forecasts are derived. Furthermore, asymmetric impact and HAR-type terms are added to the DCC- and DECO-HEAVY models.

The models are applied to the stocks in the Dow Jones Industrial Average (DJIA) index. Like in Shephard and Sheppard (2010), the effects of the lagged squared returns are insignificant when lagged realized variances are included in the conditional variance equations. Likewise, the effect of the lagged outer product of degarched returns is insignificant when the lagged realized correlation matrix is included in the conditional correlation equation. Moreover, applying the model confidence set approach based on statistical and economic loss functions, the empirical results show that the DCC- and DECO-HEAVY models provide better out-of-sample forecasts than the DCC-GARCH, DECO-GARCH, BEKK-GARCH, and BEKK-HEAVY models. Including asymmetric and HAR terms further improve the DCC- and DECO-HEAVY model forecasts.

The remainder of the paper is organized as follows. Section 2 introduces the DCC- and DECO-HEAVY models. Section 3 provides the multiplicative error representation, the multi-step forecast formulas, and model extensions. Section 4 presents the estimation procedure. Section 5 outlines the results of an empirical analysis. Section 6 concludes. A supplementary appendix (SA) includes additional theoretical and empirical results, in particular for the dataset used by Noureldin et al. (2012).

2 DCC-HEAVY and DCC-DECO Models

In the first subsection, the multivariate HEAVY framework of Noureldin et al. (2012) is reminded, and in particular the scalar BEKK-HEAVY model. The DCC-HEAVY and DCC-DECO models are defined in the next subsections.

2.1 Multivariate HEAVY Framework

Let $r_t = (r_{1t}, r_{2t} \dots r_{kt})'$ denote the $k \times 1$ daily return vector of day t corresponding to k assets, and $r_{(j)t} = (r_{(j)1t}, r_{(j)2t} \dots r_{(j)kt})'$ the corresponding j-th intra-daily return vector at time j on day t, where $j = 1, 2, ..., m$. Assuming, for instance, six and hours of trading per day and five-minute returns, $m = 72$. The outer product of daily returns is the $k \times k$ matrix $r_t r_t'$. The most simple realized covariance measure for the k assets on day t is the $k \times k$ matrix defined as

$$RC_t = \sum_{j=1}^{m} r_{(j)t} r_{(j)t}'. \tag{1}$$

Assuming that $m > k$, RC_t is positive definite. Denote by v_t the $k \times 1$ realized variance vector of day t, consisting of the diagonal elements of RC_t, and by RL_t the realized correlation matrix of day t, defined as

$$RL_t = \{\text{diag}(RC_t)\}^{-1/2} RC_t \{\text{diag}(RC_t)\}^{-1/2}, \tag{2}$$

where $\text{diag}(RC_t)$ is the diagonal matrix obtained by setting the off-diagonal elements of RC_t equal to zero, and the exponent $-1/2$ transforms each diagonal element into the inverse of its square root. Thus, the off-diagonal elements of RL_t are the realized correlation coefficients for the asset pairs, and its diagonal elements are equal to unity.

A multivariate HEAVY model specifies a dynamic process for the conditional covariance matrix H_t of the daily return and another one for the conditional mean M_t of the realized covariance matrix of day t:

$$E(r_t r_t' | \mathcal{F}_{t-1}) := H_t, \tag{3}$$
$$E(RC_t | \mathcal{F}_{t-1}) := M_t, \tag{4}$$

where $\mathcal{F}_{t-1}$ is the information set generated by the daily and intra-daily observations, and $E(r_t|\mathcal{F}_{t-1}) = 0$ is assumed for simplicity (otherwise r_t denotes the demeaned return vector). The link between both moments comes from the dependence of H_t on past values of functions of RC_t.

For the specification of the dynamics of H_t and M_t, Noureldin et al. (2012) adopt the BEKK-type model to ensure that the conditional covariance matrix is positive semidefinite. The scalar version of the model, with the 'targeting' parameterization of the constant terms, is

$$\begin{aligned} H_t &= (1 - \beta_H)\bar{H} - \alpha_H \bar{M} + \alpha_H RC_{t-1} + \beta_H H_{t-1}, & (5)\\ M_t &= (1 - \alpha_M - \beta_M)\bar{M} + \alpha_M RC_{t-1} + \beta_M M_{t-1}, & (6) \end{aligned}$$

where $\alpha_H \geq 0$, $\beta_H \geq 0$, $\beta_H = 0$ if $\alpha_H = 0$, $\beta_H < 1$. Similar restrictions, with $\alpha_M + \beta_M < 1$, apply to β_M and α_M. The $k \times k$ matrices $\bar{H} = E(r_t r_t')$ and $\bar{M} = E(RC_t)$ are positive definite. These unknown parameter matrices can be estimated consistently by their empirical counterparts. and care must be taken that the constant term of H_t is also positive definite.

The scalar BEKK-GARCH model corresponds to the equation (5) for H_t above, where RC_{t-1} is replaced by $r_{t-1}r_{t-1}'$ and $\bar{M}$ by $\bar{H}$, thus using only the daily information.

2.2 DCC-HEAVY Model

An alternative to the BEKK specification for H_t and M_t is the DCC model of Engle (2002). Any covariance matrix can be written as the product DRD, where D is the corresponding diagonal matrix of standard deviations and R is the correlation matrix.

2.2.1 DCC-HEAVY Specification of H_t

Let us denote by h_t the $k \times 1$ vector of conditional variances (that is, the diagonal elements of H_t), by $h_t^{1/2}$ the vector of conditional standard deviations (obtained by taking the square root of each entry of h_t), and by R_t the corresponding conditional correlation matrix. Then, the conditional covariance matrix, H_t can be written as

$$H_t = \mathrm{Diag}(h_t^{1/2}) R_t \, \mathrm{Diag}(h_t^{1/2}), \tag{7}$$

where, for x being a $k \times 1$ vector, Diag(x) is the $k \times k$ diagonal matrix with the entries of x as diagonal elements.

Assumption (3) for the HEAVY-BEKK is replaced by

$$\begin{aligned} E(\mathrm{diag}(r_t r_t')|\mathcal{F}_{t-1}) &:= \mathrm{Diag}(h_t), && (8)\\ E(u_t u_t'|\mathcal{F}_{t-1}) &:= R_t, \text{ where } u_t = r_t \odot h_t^{-1/2}. && (9) \end{aligned}$$

Thus, instead of specifying altogether the dynamics of the conditional variances and covariances of the returns as for instance in (5), the DCC-HEAVY model specifies the dynamics of the conditional variances of the returns and of the conditional correlation matrix (R_t) of the degarched returns (i.e. the observed returns divided by their conditional standard deviations). Notice indeed that

$$E(u_t u_t'|\mathcal{F}_{t-1}) = E(r_t r_t'|\mathcal{F}_{t-1}) \odot \left(h_t^{-1/2}(h_t^{-1/2})'\right) = H_t \odot \left(h_t^{-1/2}(h_t^{-1/2})'\right)$$

is a matrix with unit diagonal elements, and off-diagonal elements that are the conditional correlation coefficients, that is, the conditional covariances divided by the corresponding conditional standard deviations. This is the same setting as in the DCC-GARCH model of Engle (2002), and since a covariance is the product of two standard deviations and a correlation, the expectation of the covariance is not the corresponding function of the expected standard deviations and expected correlation.

The dynamics of the conditional variance vector is specified as

$$h_t = \omega_h + A_h v_{t-1} + B_h h_{t-1}, \tag{10}$$

where ω_h is a $k \times 1$ positive vector, and A_h and B_h are $k \times k$ matrices, such that each entry of h_t is positive. To ease the restrictions necessary for this and to avoid parameter proliferation, A_h and B_h are restricted to be diagonal matrices with positive entries on the diagonal, and the elements on the diagonal of B_h to be smaller than unity. The diagonality restrictions imply that each conditional variance depends on its own lag and the corresponding previous realized variance, as in the HEAVY-r model of Shephard and Sheppard (2010). More generally, A_h can be non-diagonal to allow spillover effects. If B_h is restricted to be diagonal, the model of the k variances can be estimated in k separate parts (see Section 4). The DCC-GARCH model of Engle (2002) for the conditional variances is similar to (10), with r_{t-1}^2 (the squared elements of r_{t-1}) replacing v_{t-1}.

The conditional correlation matrix is specified through a scalar dynamic equation:

$$\begin{aligned} R_t &= \tilde{R} + \alpha_r RL_{t-1} + \beta_r R_{t-1}, \quad (11) \\ \tilde{R} &= (1-\beta_r)\bar{R} - \alpha_r \bar{P}, \quad (12) \end{aligned}$$

where $\alpha_r \geq 0$, $\beta_r \geq 0$, $\beta_r = 0$ if $\alpha_r = 0$, $\beta_r < 1$, $\bar{R}$ is the $k \times k$ unconditional correlation matrix of u_t, and $\bar{P}$ is the $k \times k$ unconditional expectation of RL_t. The elements of $\bar{R}$ and $\bar{P}$ can be set to their empirical counterpart to simplify the estimation, so that only two parameters (α_r and β_r) remain to be estimated. By substituting (12) in (11), R_t is equal to $\bar{R} + \alpha_r(RL_{t-1} - \bar{P}) + \beta_r(R_{t-1} - \bar{R})$, and by taking the unconditional expectation on both sides, $E(R_t) = \bar{R}$ if $E(RL_t) = \bar{P}$.

Since $\bar{R}$, $\bar{P}$ and RL_{t-1} have unit diagonal elements, and assuming that the initial matrix R_0 is a correlation matrix, it is obvious that R_t has unit diagonal elements, but to be a well-defined correlation matrix it must be positive definite (PD) for all admissible values of (α_r, β_r). This is not necessarily the case. For example, for $k = 2$, the single correlation coefficient is equal to $\bar{r} + \alpha_r(rl_{t-1} - \bar{p})$ if $\beta_r = 0$, and setting $\bar{r} = 0.81$, $\alpha_r = 0.4$, $rl_{t-1} - \bar{p} = 0.5$ (this being a relatively large value), this is equal to the inadmissible value 1.01. Hence, when estimating the specification of R_t proposed above, one must ensure that R_t is PD. For the data used in the empirical application, no more restrictions on the space of (α_r, β_r) had to be imposed than stated below (12) in order to keep the positivity property of R_t during the iterative numerical procedure for maximizing the log-likelihood function. For this kind of data, the values of (α_r, β_r) for which R_t may become non-PD are quite different from the ranges of values typically estimated for these parameters, that is $0 < \alpha_r < 0.2$ and $0.6 < \beta_r < 1$. Figure 1 shows an example of the parameter values for which R_t becomes non-PD for $t > \tilde{t}$ in the sample period for the data used in the empirical application. The value of $\tilde{t}$ depends on (α_r, β_r), being smaller and smaller as (α_r, β_r) moves to the lower right corner of the figure.

Insert Figure 1 here

A scalar dynamic specification that avoids the non-positivity problem is

$$\begin{aligned} R_t &= (1-\alpha_r-\beta_r)R + \alpha_r RL_{t-1} + \beta_r R_{t-1}, \quad (13) \\ R &= w\bar{R} + (1-w)\bar{P}, \quad w \in [0,1], \quad (14) \end{aligned}$$

where $\bar{R}$ and $\bar{P}$ are defined as written below (12) . By imposing $\alpha_r + \beta_r < 1$, R_t is a convex combination of correlation matrices, hence it is a well-defined correlation matrix, guaranteed to be PD if at least one of the three matrices R, RL_{t-1}, R_{t-1} is PD (the other must be positive semidefinite). This is in the spirit of the TVC-GARCH specification of Tse and Tsui (2002), who use an empirical correlation matrix built from past daily returns instead of the realized correlation matrix.

Taking the unconditional expectation on both sides of (13) gives

$$E(R_t) = \frac{1 - \alpha_r - \beta_r}{1 - \beta_r} R + \frac{\alpha_r}{1 - \beta_r} \bar{P}, \tag{15}$$

which is a well-defined correlation matrix if R itself is a constant correlation matrix, not necessarily defined as (14). By assuming (14), if $w = 0$, $E(R_t) = \bar{P}$ (the unconditional expectation of RL_t), and if $w = 1$, it is is like (15) with $\bar{R}$ (the unconditional expectation of u_t) replacing R.

Notice that even if $w > 1$, $w\bar{R} + (1 - w)\bar{P}$ has unit diagonal elements, so that it looks like a correlation matrix, but it may not be positive definite. So when $w > 1$, R_t of (13) with $R = w\bar{R} + (1 - w)\bar{P}$ has the same problem of positivity as R_t of (11)-(12). Actually, the latter specification is a constrained version of (13)-(14) if w is *not* restricted to be in $[0, 1]$, the constraint being that $w = (1 - \beta_r)/(1 - \alpha_r - \beta_r)$. For this value of w, $E(R_t)$ by (15) is equal to $\bar{R}$, coherently with the result obtained for (11)-(12).

An advantage of the proposed specifications of R_t, based on the realized correlation matrix, is that there is no need to transform the covariance matrix of the degarched returns into a correlation one, as in the DCC-GARCH model of Engle (2002), which is specified by

$$R_t = \{\mathrm{diag}(Q_t)\}^{-1/2} Q_t \{\mathrm{diag}(Q_t)\}^{-1/2}, \tag{16}$$

with the $k \times k$ symmetric PD matrix

$$Q_t = (1 - \alpha_q - \beta_q)\bar{Q} + \alpha_q u_{t-1} u'_{t-1} + \beta_q Q_{t-1}, \tag{17}$$

where $\alpha_q \geq 0$, $\beta_q \geq 0$, $\beta_q = 0$ if $\alpha_q = 0$, $\alpha_q + \beta_q < 1$, and $\bar{Q}$ is a $k \times k$ PD matrix. Q_t is actually the conditional covariance matrix of u_t, with non-unit diagonal elements, and by (16), it is transformed into a correlation matrix. However, this parameterisation raises two issues, one about estimation, the other about forecasting.

First, because $E(u_t u'_t | \mathcal{F}_{t-1}) = R_t \neq Q_t$, $\bar{Q}$ is not equal to $E(Q_t)$ and therefore is not consistently estimated by $\sum_t u_t u'_t / T$. Thus using this average to estimate $\bar{Q}$ as proposed by Engle (2002), so that only α_q and β_q remain to be estimated by QML, introduces an asymptotic bias in their estimator. See Aielli (2013) for details and an alternative formulation of (17) that avoids this problem.

Second, at date t, the $(s+1)$-step ahead forecast of R_{t+s+1} requires $E_t(u_{t+s} u'_{t+s}) = E_t(R_{t+s})$, which is not available in closed from due to the nonlinear relation (16) between R_{t+s} and Q_{t+s}. By assuming $E_t(R_{t+s}) \approx E_t(Q_{t+s})$, Engle and Sheppard (2001) obtain the closed form forecast recurrence relation

$$E_t(R_{t+s+1}) = \bar{Q} + (\alpha_q + \beta_q)E_t(R_{t+s} - \bar{Q}) \tag{18}$$

that starts with $E_t(R_{t+1}) = R_t$. The correlation forecasts are thus approximate and biased.

The DCC-HEAVY model differs from DCC-GARCH model in three ways: 1) the dynamics of conditional variances h_t are driven by the lagged realized volatilities v_{t-1}; 2) the conditional correlation R_t is modelled directly rather than parameterised in a sandwich form as in (16); 3) the dynamics of the conditional correlation matrix R_t are driven by the lagged realized correlation matrix. The last two features allow us to obtain exact closed forms for s-step ahead correlation forecasts, as explained in Section 3.

2.2.2 DCC-HEAVY Specification of M_t

The specification of the DCC-HEAVY model requires to define the dynamics of

$$M_t = \mathrm{Diag}(m_t^{1/2}) P_t \, \mathrm{Diag}(m_t^{1/2}), \tag{19}$$

where m_t is the vector containing the main diagonal of M_t, that is, the conditional means of the realized variances, and P_t is the corresponding conditional mean of the realized correlation matrix.

The conditional expectation of the realized variances is specified as

$$m_t = \omega_m + A_m v_{t-1} + B_m m_{t-1}, \tag{20}$$

where ω_m is a positive $k \times 1$ vector, A_m and B_m are $k \times k$ matrices that are restricted to be diagonal matrices with positive entries, as discussed after (10).

The dynamic process for the conditional expectation of the realized correlation matrix is defined in the following way:

$$P_t = (1 - \alpha_p - \beta_p)\bar{P} + \alpha_p RL_{t-1} + \beta_p P_{t-1}, \tag{21}$$

where $\alpha_p \geq 0$, $\beta_p \geq 0$, $\beta_p = 0$ if $\alpha_p = 0$, $\alpha_p + \beta_p < 1$, and $\bar{P}$ is a correlation matrix that is the unconditional mean of RL_t. The elements of $\bar{P}$ can be set to their empirical counterpart to render the estimation simpler. $E(RL_t)$ is not equal to the unconditional correlation matrix $E(P_t)$, due to the nonlinearity of the the transformation from covariance to correlation. However, Bauwens et al. (2012) show that if RC_t is computed from a large enough number of high-frequency returns, $\bar{P}$ should be almost equal to $E(RL_t)$.

Equations (7)-(21) form the DCC-HEAVY model. By setting $\alpha_r = \beta_r = \alpha_p = \beta_p = 0$, the model simplifies into a constant conditional correlation HEAVY model. Estimation is discussed in Section 4.

2.3 DECO-HEAVY Model

The DECO-HEAVY model differs from the DCC-HEAVY in the specification of the conditional correlation matrix corresponding to H_t and of the conditional mean of the realized correlation matrix corresponding to M_t.

The specification of the conditional correlation matrix corresponding to H_t, denoted by R_t^E, is based on the assumption that all the conditional correlations are the same time-varying correlation $\rho_t \in (-1/(1+k), 1)$, chosen to be the average of the correlation coefficients of $R_t = (r_{t,ij})$, defined by (11)-(12):

$$\begin{aligned} R_t^E &= (1-\rho_t^E)I_k + \rho_t^E J_k, && (22) \\ \rho_t^E &= \frac{2}{k(k-1)} \sum_{i>j} r_{t,ij}. && (23) \end{aligned}$$

where J_k is $k \times k$ matrix of ones. The DECO-GARCH model of Engle and Kelly (2012) is using as dynamic equicorrelation coefficient the average of the DCC-GARCH correlations defined by (16) together with the modification of (17) proposed by Aielli (2013).

Likewise, the conditional mean of the realized correlation matrix corresponding to M_t, denoted by P_t^E, is specified as:

$$\begin{aligned} P_t^E &= (1-\rho_t^P)I_k + \rho_t^P J_k, && (24) \\ \rho_t^P &= \frac{2}{k(k-1)} \sum_{i>j} p_{t,ij}. && (25) \end{aligned}$$

The main advantage of DECO with respect to DCC, especially when k is very large, is the availability of analytical expressions of the inverse and determinant of the equicorrelated matrices, which are used in the computation of the likelihood function for estimation, and of economic loss functions for forecast evaluations. For more details, see Engle and Kelly (2012).

3 Representation, Forecasting and Extensions

In this section, the DCC-HEAVY and the closely related DCC-GARCH, and DCCX-GARCHX models are represented as multiplicative error models (MEM), from which stationarity conditions and closed-form formulas for multi-step forecasts follow directly. An extension of the the DCC-HEAVY model is proposed by adding asymmetric impact and HAR terms.

3.1 Multiplicative Error Representation

A MEM for a positive variable x_t specifies it as the product of a positive conditional mean and a positive error that follows some distribution with expectation equal to one. If x_t is a squared centred return, the conditional mean is the conditional variance. This can be extended to the elements of a vector. In this subsection, the focus is on the conditional expectation formulation, not on the distribution.

For the conditional and realized variance equations, define the vectors of $2k \times 1$ elements $x_t = [(r_t^2)', v_t']'$ and $\mu_t = [h_t', m_t']'$. The conditional expectation formulation of the conditional and realized variance equations (10) and (20) is

$$\begin{aligned} E(x_t|\mathcal{F}_{t-1}) &:= \mu_t, \\ \mu_t &= \omega + Ax_{t-1} + B\mu_{t-1}, \end{aligned} \tag{26}$$

where

$$\omega = \begin{bmatrix} \omega_h \\ \omega_m \end{bmatrix}, \; A = \begin{bmatrix} 0 & A_h \\ 0 & A_m \end{bmatrix}, \; B = \begin{bmatrix} B_h & 0 \\ 0 & B_m \end{bmatrix} \tag{27}$$

and the 0 symbol stands for a $k \times k$ matrix of zeros. If

$$A = \begin{bmatrix} A_h & 0 \\ 0 & A_m \end{bmatrix}, \tag{28}$$

the top part becomes the DCC-GARCH model, and the realized variance model is kept in the bottom part. If

$$A = \begin{bmatrix} A_h & A_{hm} \\ 0 & A_m \end{bmatrix}, \tag{29}$$

the model (referred to as DCC-GARCHX) includes in the top part both the lagged squared return and the lagged realized variance, encompassing the two previous models.

Defining the $2k \times k$ matrices $Y_t = [u_t u_t', RL_t]'$ and $\Phi_t = [R_t, P_t]'$, the conditional expectation formulation of the conditional and realized correlation matrix equations (11) and (21) is

$$\begin{aligned} E(Y_t|\mathcal{F}_{t-1}) &:= \Phi_t, \\ \Phi_t &= \Omega + (\alpha \otimes J_k)Y_{t-1} + (\beta \otimes J_k)\Phi_{t-1}, \end{aligned} \tag{30}$$

where

$$\Omega = \begin{bmatrix} \tilde{R} \\ \bar{P} \end{bmatrix}, \; \alpha = \begin{bmatrix} 0 & \alpha_r \\ 0 & \alpha_p \end{bmatrix}, \; \beta = \begin{bmatrix} \beta_r & 0 \\ 0 & \beta_p \end{bmatrix},$$

and the 0 symbol is scalar in this case. If $\alpha = \begin{bmatrix} \alpha_q & 0 \\ 0 & \alpha_p \end{bmatrix}$, $\beta = \begin{bmatrix} \beta_q & 0 \\ 0 & \beta_p \end{bmatrix}$, Q_t replaces R_t in the definition of Φ_t, and the constant term is adapted, the top part becomes the 'quasi-correlation' equation (17) of the DCC-GARCH model, the realized correlation equation being kept in the bottom part. If

$$\alpha = \begin{bmatrix} \alpha_q & \alpha_r \\ 0 & \alpha_p \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_r & 0 \\ 0 & \beta_p \end{bmatrix}, \tag{31}$$

both $u_{t-1}u'_{t-1}$ and RL'_{t-1} are included in the top part (referred to as DCCX-GARCH).

Processes such as defined by (26) and (30) can be written as VARMA(1,1) by defining appropriate error terms. From this representation, it follows that the covariance stationarity condition is that the largest eigenvalue of $A + B$ is smaller than unity for (26), and likewise for the largest eigenvalue of $\alpha + \beta$ for (30). For the matrices α and β defined after (30), the previous condition is equivalent to $\beta_r < 1$ and $\alpha_p + \beta_p < 1$, as written after (11) and (21). The unconditional first moment is then obtained by applying standard results for the VARMA representations. For example, $E(x_t) = (I_{2k} - A - B)^{-1}\omega$, which gives $E(v_t) = (I_k - A_m - B_m)^{-1}\omega_m$ and $E(r_t^2) = (I_k - B_h)^{-1}[\omega_h + A_h E(v_t)]$ for the variances of the DCC-HEAVY model. For the correlations, $E(RL_t) = \bar{P}$ and $E(u_t u'_t) = [\tilde{R} + \alpha_r E(RL_t)]/(1 - \beta_r) = \bar{R}$, the last equality resulting from (12).

3.2 Multiple-step Ahead Forecasting

Forecasts of the conditional covariance matrices of daily returns are used in several financial applications. The s-step ahead forecast of H_{t+s}, computed at date t, is defined in the case of DCC-type models as

$$H_{t+s|t} = \text{Diag}\left\{[E_t(h_{t+s})]^{1/2}\right\} E_t(R_{t+s}) \text{Diag}\left\{[E_t(h_{t+s})]^{1/2}\right\}, \tag{32}$$

where $E_t(.)$ is a short notation for $E(.|\mathcal{F}_{t-1})$. Notice that as in the DCC model, $H_{t+s|t}$ is not equal to $E_t(H_{t+s})$, due to the nonlinearity of the transformation of covariances into correlations and of the square root function.

To obtain $E_t(h_{t+s})$ and $E_t(R_{t+s})$, the conditional expectation expressions of the previous subsection are useful to compute $E_t(\mu_{t+s})$ and $E_t(\Phi_{t+s})$, denoted by $\mu_{t+s|t}$ and $\Phi_{t+s|t}$, respectively.

Starting from (26) leads to $E_t(\mu_{t+s})$. In moving more than one step ahead, $x_{t+s|t}$ is not known and needs to be substituted with its corresponding conditional expectation $\mu_{t+s|t}$, hence

$$\begin{aligned} \mu_{t+1|t} &= \omega + Ax_t + B\mu_t, \\ \mu_{t+s|t} &= \omega + (A+B)\mu_{t+s-1|t} \quad \text{for } s \geq 2, \end{aligned} \tag{33}$$

which can be solved recursively, giving the closed form forecast

$$\mu_{t+s|t} = \tilde{\omega} + C^{s-1}\mu_{t+1|t} \quad \text{for } s \geq 2,$$

where $\tilde{\omega} = (I - C)^{-1}(I - C^{s-1})\omega$ and $C = A + B$.

Proceeding in the same way for $E_t(\Phi_{t+s})$ from (30) gives

$$\begin{aligned} \Phi_{t+1|t} &= \Omega + (\alpha \otimes J_k)Y_t + (\beta \otimes J_k)\Phi_t, \\ \Phi_{t+s|t} &= \Omega + ((\alpha + \beta) \otimes J_k)\, \Phi_{t+s-1|t} \quad \text{for } s \geq 2. \end{aligned} \tag{34}$$

The closed form forecast is

$$\Phi_{t+s|t} = \tilde{\Omega} + \left(c^{s-1} \otimes J_k\right) \Phi_{t+1|t} \text{ for } s \geq 2,$$

where $\tilde{\Omega} = (I - c)^{-1}(I - c^{s-1})\Omega$ and $c = \alpha + \beta$.

For instance, the DCC-HEAVY model s-step ahead forecasts $\mu_{t+s|t}$ and $\Phi_{t+s|t}$ are derived from (33) and (34) by setting A, B, α, and β to the matrices defined after (26) and (30). The s-step ahead forecast $E_t[h_{t+s}]$ of the conditional variance vector corresponds to the first k elements of $\mu_{t+s|t}$, and the s-step forecast $E_t[R_{t+s}]$ of the conditional correlation corresponds to the $k \times k$ upper block of $\Phi_{t+s|t}$.

3.3 Model Extensions

It has long been discovered that stock markets react differently to positive and negative news. The asymmetric effect is now commonly used to refer to any volatility model, univariate and multivariate alike, in which the (co)variances respond asymmetrically to positive and negative shocks. The DCC-HEAVY model can be extended by incorporating the asymmetric effect into the variance and correlation equations. In the variance equation, the asymmetric effect implies that volatility tends to increase more following negative return shocks than equally-sized positive shocks. In the correlation equation, the asymmetric effect implies that the correlation between stock returns tends to increase when the market turns down. The extended model is called ADCC-HEAVY model.

Define $d_t = (d_{1t}, d_{2t}, ..., d_{kt})'$ where $d_{it} = 1$ if $r_{it} < 0$ and $d_{it} = 0$, if $r_{it} \geq 0$, for $i = 1, .., k$, and $D_t = I_k + \text{OFFDIAG}(d_t d_t')$, where $\text{OFFDIAG}(A)$ is the matrix obtained by keeping the off-diagonal elements of A and replacing its diagonal elements by zeros. The conditional variance and correlation equations of the ADCC-HEAVY model are then

$$h_t = \omega_h + A_h v_{t-1} + B_h h_{t-1} + \Gamma_h d_{t-1} \odot v_{t-1}, \tag{35}$$

$$R_t = \tilde{R} + \alpha_r RL_{t-1} + \beta_r R_{t-1} + \gamma_r D_{t-1} \odot RL_{t-1}, \tag{36}$$

$$\tilde{R} = (1 - \beta_r)\bar{R} - (\alpha_r \bar{P} + \gamma_r \bar{D} \odot \bar{P}), \tag{37}$$

where Γ_h is a $k \times k$ diagonal matrix and $\bar{D}$ is the sample mean of D_t. Equation (35) extends (10); asymmetric effects correspond to positive values on the diagonal of Γ_h. Equations (36)-(37) extend (11)-(12). The asymmetric effect corresponds to a positive value of γ_r: the impact of the lagged realized correlation between assets i and j on their current conditional correlation is equal to $\alpha_r + \gamma_r$ only if both $r_{i,t-1}$ and $r_{j,t-1}$ are negative, otherwise the impact is reduced to α_r if γ_r is positive. Notice that the diagonal elements of $D_{t-1} \odot RL_{t-1}$ and $\bar{D} \odot \bar{P}$ are equal to one, so that the same holds for R_t. Like in the simpler model where $\gamma_r = 0$, in estimation $(\alpha_r, \beta_r, \gamma_r)$ are constrained to values such that R_t is PD for all t. In the empirical applications, this did not create any difficulty.

The same asymmetric effects are included in the realized variance and correlation equations. Furthermore, the heterogeneous autoregressive (HAR) model of Corsi (2009) has emerged as a simple and powerful way to include the long-memory feature of realized volatilities. The model was extended to the multivariate setting by Chiriac and Voev (2011) and Oh and Patton (2016). Adding also HAR terms to the realized variance and correlation equations, respectively (20) and (21), results in the richer dynamic equations

$$m_t = \omega_m + A_m v_{t-1} + B_m m_{t-1} + \Gamma_m d_{t-1} \odot v_{t-1} + A_m^w v_{t-1}^w + A_m^m v_{t-1}^m, \tag{38}$$
$$P_t = \tilde{P} + \alpha_p RL_t + \beta_p P_{t-1} + \gamma_p D_{t-1} \odot RL_{t-1} + \alpha_p^w RL_{t-1}^w + \alpha_p^m RL_{t-1}^m, \tag{39}$$

where Γ_m, A_m^w, A_m^m are $k \times k$ diagonal matrices, $v_{t-1}^w = \frac{1}{5}\sum_{j=1}^{5} v_{t-1-j}$, $v_{t-1}^m = \frac{1}{22}\sum_{j=1}^{22} v_{t-1-j}$, $RL_{t-1}^w = \frac{1}{5}\sum_{j=1}^{5} RL_{t-1-j}$, $RL_{t-1}^m = \frac{1}{22}\sum_{j=1}^{22} RL_{t-1-j}$, and $\tilde{P} = (1 - \alpha_p - \beta_p - \gamma_r \bar{D}^r - \alpha_p^w - \alpha_p^m)\bar{P}$. Although P_t has unit diagonal elements, in estimation the parameter space must be constrained to ensure that this matrix is PD for all t. HAR terms are not added to the conditional covariance and correlation equations, since these effects are insignificant in the empirical application.

Other ways to define the asymmetric effect have been used. For example, one can define $d_{it} = 1$ when the stock volatility $RC_{ii,t}$ is 'high' (above some threshold). In the correlation equation, that implies that the correlation between two stock returns increases more when both stocks are highly volatile than if only one or none is highly volatile. Another asymmetric effect is to let all correlations increase if the market volatility (for instance measured by VIX) increases, see Bauwens and Otranto (2016).

Section A of the SA provides the formulas of multiple step forecasts of the extended models.

4 Estimation

The DCC- and DECO-HEAVY model are parameterized with a finite-dimensional $p \times 1$ parameter vector $\theta \in \Theta \subset \mathbb{R}^p$. Partitioning θ' into (θ'_H, θ'_M), where the $p_H \times 1$

vector θ_H is the parameter vector of the HEAVY model for H_t, and the $p_M \times 1$ vector θ_M is the parameter vector of the HEAVY model for M_t, θ_H and θ_M can be estimated separately, as they are variation free in the sense of Engle et al. (1983). Moreover, each of these estimations can be split into two steps, as explained below.

4.1 Estimation of θ_H

To get a quasi-likelihood function, we add to assumption (8)-(9) the hypothesis that the distribution of the innovation of the return vector is multivariate Gaussian:

$$u_t|\mathcal{F}_{t-1} \sim N(0, R_t), \tag{40}$$

implying that $r_t|\mathcal{F}_{t-1} \sim N(0, H_t)$, with H_t defined in (7).

Neglecting irrelevant constants, the quasi-log-likelihood function for T observations, given initial values, is

$$\begin{aligned} QL_H(\theta_H) &= -\frac{1}{2}\sum_{t=1}^{T}\left(\log|H_t| + r_t' H_t^{-1} r_t\right) && (41) \\ &= -\frac{1}{2}\sum_{t=1}^{T}\left(2\log|\mathrm{Diag}(h_t^{1/2})| + \log|R_t| + u_t' R_t^{-1} u_t\right). \end{aligned}$$

This function can be maximized numerically in a single step but for large k, the large dimension of the parameter space makes this difficult. A two-step estimation can be based on a partition of θ_H into θ_{H1}, the parameters of the variance equation (10), and θ_{H2}, the parameters of the correlation equation (11) or (13). The two-step procedure was proposed by Engle (2002) for the DCC-GARCH model.

The first step consists in estimating θ_{H1} by maximizing the quasi-log-likelihood obtained by replacing R_t in the second line of (41) by the identity matrix, so that the objective function does not depend on θ_{H2}:

$$QL_{H1}(\theta_{H1}) = -\frac{1}{2}\sum_{t=1}^{T}\left(2\log|\mathrm{Diag}(h_t^{1/2})| + u_t' u_t\right). \tag{42}$$

The matrix A_h of (10) can be non-diagonal to allow spillover effects, but the matrix B_h must be diagonal to allow separate estimation of the k equations. In our empirical analysis, both A_h and B_h are restricted to be diagonal.

The second step maximizes (41) with respect to θ_{H2}, fixing θ_{H1} to the value $\hat{\theta}_{H1}$ obtained at the first step. The second step objective function can be written as

$$QL_{H2}(\theta_{H2};\hat{\theta}_{H1}) = -\frac{1}{2}\sum_{t=1}^{T}\left(\log|R_t| + \hat{u}_t' R_t^{-1}\hat{u}_t\right), \tag{43}$$

where $\hat{u}_t = r_t \odot \hat{h}_t^{-1/2}$ and $\hat{h}_t$ means that h_t defined in (10) is evaluated at $\theta_{H1} = \hat{\theta}_{H1}$.

The two-step estimator is presumably a consistent but inefficient estimator of the parameter θ_H under conditions (such as asymptotic identification, strict stationarity and ergodicity of the time series processes) similar to those stated in Engle and Shephard (2001) for the DCC-GARCH model. The two-step estimation method is numerically tractable for large k, if $\bar{R}$ and $\bar{P}$ are targeted, i.e., estimated by their empirical counterparts, as proposed below (12). The impact of the targeting is a loss of efficiency. Noureldin et al., 2012) discuss this issue for the BEKK-HEAVY model, which is estimated in a single step. Further work to derive rigorously the asymptotic distribution of the single and two-step estimators of the DCC-HEAVY, with and without targeting, is needed and beyond the scope of this paper.

4.2 Estimation of θ_M

Since a realized covariance matrix is symmetric positive definite, a natural choice of distribution to form a likelihood function for the realized covariance process is the Wishart distribution. This assumption has been used in several papers, e.g. Gouriéroux et al. (2009), Golosnoy, Gribisch, and Liesenfeld (2012), Bauwens et al. (2012, 2016).

We assume that conditionally on the past information set, RC_t follows a central Wishart distribution of dimension k and denote this assumption by

$$RC_t|\mathcal{F}_{t-1} \sim W_k(\nu, M_t/\nu),$$

where ν is the degrees of freedom parameter restricted by $\nu > k-1$. The chosen parameterization implies that

$$E(RC_t|\mathcal{F}_{t-1}) := M_t.$$

Using the expression of a Wishart density function, and of M_t in (19), the quasi-log-likelihood function for a sample of T observations, given initial conditions, is

$$\begin{aligned} QL_M(\theta_M) &= -\frac{v}{2}\sum_{t=1}^{T}\left[\log|M_t| + \mathrm{tr}(M_t^{-1}RC_t)\right] \qquad (44) \\ &= -\frac{v}{2}\sum_{t=1}^{T}\left\{2\log|D_{m,t}| + \log|P_t| + \mathrm{tr}[(D_{m,t}P_tD_{m,t})^{-1}RC_t]\right\}, \end{aligned}$$

where θ_M is the vector of the parameters that appear in (20) and (21) and $D_{m,t}$ stands for $\mathrm{Diag}(m_t^{1/2})$. In the above expression, terms that depend on ν but do not depend on θ_M are not included. The parameter ν is considered as a nuisance parameter that can be neglected to estimate θ_M and practically it can be set to unity without loss of information since the score for θ_M is proportional to the value of this parameter.

As shown by Bauwens et al. (2012), the Wishart assumption provides a quasi-likelihood function, which can serve as objective function to get a single step estimator. They also show that the DCC-HEAVY model part for the realized covariance matrix can be estimated in two steps. The parameter space θ_M is split into θ_{M1} for the parameters in the realized volatility model and θ_{M2} for the parameters in the realized correlation model. We denote by QL_{M1} the quasi-log-likelihood where P_t in (44) is replaced by the identity matrix and ν set to unity:

$$QL_{M1}(\theta_{M1}) = -\frac{1}{2}\sum_{t=1}^{T}\left[2\log|D_{m,t}| + \mathrm{tr}(D_{m,t}^{-1}RC_tD_{m,t}^{-1})\right]. \tag{45}$$

This estimation is split into k separate estimations, when the matrices A_m and B_m of (20) are restricted to be diagonal.

We denote by QL_{M2} the quasi-log-likelihood where θ_{M1} is fixed at the value $\hat{\theta}_{M1}$ obtained at the first step:

$$QL_{M2}(\theta_{M2};\hat{\theta}_{M1}) = -\frac{1}{2}\sum_{t=1}^{T}\left\{\log|P_t| + \mathrm{tr}[(P_t^{-1} - I_k)\hat{D}_{m,t}^{-1}RC_t\hat{D}_{m,t}^{-1}]\right\}, \tag{46}$$

where $\hat{D}_{m,t}$ means that $\hat{D}_{m,t}$ is evaluated at $\theta_{M1} = \hat{\theta}_{M1}$. The parameter vector θ_{M2} includes α_p, β_p and the elements of $\bar{P}$. The latter can be targeted by the unconditional mean of the realized correlations, as discussed after (21), in which case the second step maximization is done with respect to two parameters and is therefore feasible for large k.

The detailed asymptotic distribution theory for the single step and the two-step estimators, with or without targeting, is not yet available.

5 Empirical Application

5.1 Data Description

High-frequency data for 29 stocks belonging to the Dow Jones Industrial Average (DJIA) index are used; the 30th stock was dropped since it is not permanently in the index during the sample period. The sample period is 3 January 2001 to 16 April 2018 with a total of 4318 trading days, and the data source is the TAQ database.

The stock names and tickers are: Apple Inc.(AAPL), American Express Company (AXP), The Boeing Company (BA), Caterpillar Inc. (CAT), Cisco Systems, Inc. (CSCO), Chevron Corporation (CVX), The Walt Disney Company (DIS), DowDuPont Inc. (DWDP), General Electric Company (GE), The Goldman Sachs Group, Inc. (GS), Home Depot Inc.(HD), International Business Machines Corporation (IBM), Intel Corporation(INTC), Johnson & Johnson (JNJ), JPMorgan Chase & Co. (JPM), The Coca-Cola Company (KO), McDonald's Corporation (MCD), 3M Company (MMM), Merck &Co., Inc. (MRK), Microsoft Corporation (MSFT), NIKE, Inc. (NKE), Pfizer Inc.(PFE), The Procter & Gamble Company (PG), The Travelers Companies, Inc.(TRV), United Health Group Incorporated (UNH), United Technologies Corporation (UTX), Verizon Communications Inc. (VZ), Walmart Inc. (WMT), Exxon Mobil Corporation (XOM).

The daily realized covariance matrices are computed as explained in the beginning of Section 2, using five minute returns. The synchronization of intra-day prices of the 29 stocks was done using five minute intervals, the price closest (from the left) to the respective sampling point was taken; the first and last 15 minutes of the day (9:30-16:00) were excluded.

Descriptive statistics are provided in Table B1 of the SA. For each stock, it reports the time-series averages and standard deviations of its squared returns and realized variances, and the means and standard deviations of the time series averages of its realized covariances and correlations with the other 28 stocks. Each average realized variance does not account for the overnight variation and is therefore a fraction (in most cases 50 to 60 percent) of the corresponding average squared return.

Figure 2 shows a representative example of time series plots of the realized variances of two stocks (JPM and XOM) and the corresponding realized covariances and correlations.

Insert Figure 2 here

The focus of the empirical application is a forecasting comparison of the conditional covariance and correlation matrices, and the conditional variances of the 29 stocks using a set of models. Before reporting the results of the comparisons, estimation results are reported for the DCC and DECO models.

5.2 Estimation Results for the Full Period

Table 1 presents summary statistics (median, minimum, maximum) of the first step parameter estimates (except the constant terms) of the 29 variance equations of the GARCH, GARCHX, HEAVY-h and HEAVY-m models. The estimates for each stock, and the associated robust t-statistics, are given in Table B2 of the SA.

Insert Table 1 here

The estimates of the B_h parameters in the HEAVY-h model are smaller than in the GARCH model (median of 0.481 versus 0.916), while the estimates of the A_h parameters are much larger (medians of 0.781 versus 0.072). The influence of these differences is visible on the bottom panel of Figure 3, which shows (in log) the time-series of the average (over the 29 stocks) of the corresponding fitted conditional variances. The GARCH path is smoother than the HEAVY-h path, and the latter fluctuates locally more strongly than the former, responding faster to recent changes of volatility. Similar differences occur for each stock. The GARCH parameter estimates are much more homogenous across the different stocks than the HEAVY-h estimates. The values of the HEAVY-m parameter estimates are more similar to HEAVY-h than to GARCH, though they are much more homogenous than for HEAVY-h.

In the nesting GARCHX model, the coefficient (in A_h) of the lagged squared return is set to zero for eight stocks out of 29 because a non-negativity constraint is imposed and is binding; for these stocks, GARCHX estimates are identical to HEAVY-h. For the 21 other stocks, the estimate is positive (between 0.002 and 0.068), with t-statistics below 1.5 in 12 cases (out of 21), and 4 larger than two. Even though the t-statistic has a non-standard distribution since the null hypothesis of zero is on the boundary of the parameter space, these results suggest that for almost all stocks, the estimate is not significant. On the contrary, the estimate of the lagged realized variance coefficient (in A_{mh}) is positive (between 0.103 and 1.748, with 0.698 as median value) and the associated t-statistics are usually large enough to suggest they are significant (only three are below 2.5). The log-likelihood gain of GARCHX over HEAVY-h is equal to 5 for the 21 additional parameters, hence it appears to be minor. On the contrary, the gains of GARCHX and HEAVY-h over GARCH are substantial (41 and 36 respectively). Notice that GARCH and HEAVY-h are not nested, but they have the same number of parameters, so choosing between them using their log-likelihood values is equivalent to a choice based on model choice criteria. In brief, these results suggest that the conditional variance dynamics is better captured by the lagged realized variance than by the lagged squared return, confirming the findings of Shephard and Sheppard (2010).

Insert Figure 3 here

Table 2 presents the second step parameter estimates of the correlation models: DCC-GARCH, DCCX-GARCH, DCC-HEAVY-R (eq. (11)-(12)), DCC-HEAVY-P (eq. (21)), and the corresponding DECO versions. For the DCC models, these estimates are broadly in line with those of the variance equations. The estimate of β_r in DCC-HEAVY-R model is smaller than that of β_q in DCC-GARCH model (0.869

versus 0.988), and the estimate of α_r is larger than that of α_q (0.068 versus 0.003), implying less smooth and more reactive fitted correlations. The paths of average fitted correlations and covariances of the three models are shown on Figure 3. The paths of the DECO-HEAVY-R model are much less smooth and more reactive to recent information than for DCC-HEAVY-R, due to a smaller estimate of β_r (0.552 versus 0.869) and a larger one of α_r (0.447 versus 0.068). The paths for the DCC-GARCH are smoother and less reactive than in the HEAVY models. The described path differences are stronger for each pair of stocks, since averaging reduces the variability.

Insert Table 2 here

In the nesting DCCX- and DECOX-GARCH models, the coefficient estimates of the lagged realized correlation (0.068 and 0.447) are of the same magnitude as in the DCC- and DECO-HEAVY-R (0.061 and 0.369), with large t-statistics (9.40 for DCCX, 5.09 for DECOX). The coefficient estimate of the lagged return cross-product is very close to zero (with t-statistic 10.59) in DCCX, while it is equal to 0.029 (with t-statistic 1.94) in DECOX. The maximized second step log-likelihood values of DCCX-GARCH (-5115) and DCC-HEAVY (-5119) are slightly different, but they are much larger than the value of DCC-GARCH (-5143). For the DECO models: DECO-HEAVY (-5307) and DECOX-GARCH (-5306) are very close, but DECO-GARCH (-5319) is lower. Thus lagged realized correlations can be considered as more important drivers of the conditional correlations rather than return cross-products.

The maximized log-likelihood values and their decomposition into the variance and correlation parts are reported in Table 2. The decompositions suggest that the DCC-HEAVY and the DECO-HEAVY dominate the DCC-GARCH model in both the variance and correlation parts. However, most of the in-sample gain is coming from the variance part. The overall improvement is substantial. DCC-HEAVY improves slightly more than DECO-HEAVY. Notice that these models have the same number of parameters, so comparisons using log-likelihood values are equivalent to comparisons using model choice criteria

5.3 Forecasting Comparisons

A comparison of models can be made by evaluating the in-sample and out-of-sample forecasting performances of the models using the Model Confidence Set (MCS) of Hansen et al. (2011). A MCS identifies a set of models having the best forecasting performance at a chosen confidence level, based on a loss function. Six models are compared: DCC-GARCH, DCC-HEAVY, DECO-GARCH, DECO-HEAVY, BEKK-GARCH, BEKK-HEAVY. Out-of-sample s-step-ahead forecasts of the 29-dimensional covariance and correlation matrices are computed, for s=1, 5 and 22; for horizons 5 and 22, they are iterated forecasts. For DECO models, the correlations are computed from DECO itself, not from the underlying DCC.

5.3.1 Loss Functions

Statistical and economic loss functions are adopted along the lines proposed by Becker et al. (2015).

For the covariance matrix forecasts, two statistical loss functions are used, which compare the covariance matrix forecasts with respect to the actual (unobserved) covariance matrix Σ_{t+s}. The first one is based on the negative of the Wishart log-density function:

$$QLIK^a_{t,s}(\Sigma_{t+s}, H^a_{t+s|t}) = \mathrm{tr}[(H^a_{t+s|t})^{-1}\Sigma_{t+s}] + \log|H^a_{t+s|t}|, \tag{47}$$

where $H^a_{t+s|t}$ denotes the s-step forecast using model a conditional on time t information. The second loss function is based on the Frobenius norm of the difference between the forecast and benchmark matrices (see e.g. Golosnoy et al., 2012), defined by

$$FN^a_{t,s} = ||\Sigma_{t+s} - H^a_{t+s|t}|| = \left[\sum_{i,j}(\sigma_{ij,t+s} - h^a_{ij,t+s})^2\right]^{1/2}. \tag{48}$$

Since Σ_{t+s} is unobservable, the observed realized covariance matrix RC_{t+s} is used as proxy for it. These statistical loss functions provide a consistent ranking of volatility models in the sense of Patton (2011) and Patton and Sheppard (2009) as they are robust to noise in the proxy; see also Laurent et al. (2013).

For the correlation matrix forecasts, the QLIK and FN losses are computed from the same formulas as for covariances, that is (47) and (48). The only difference is that forecasted correlations and realized correlations replace forecasted covariances and realized covariances, respectively.

Once a time series of $T_{h,s}$ (covariance or correlation) forecasts is obtained for a model, the corresponding losses and their time series average are computed, i.e.

$$QLIK^a_s = \frac{1}{T_{h,s}}\sum_{t=1}^{T_{h,s}} QLIK^a_{t,s}, \text{ and } FN^a_s = \frac{1}{T_{h,s}}\sum_{t=1}^{T_{h,s}} FN^a_{i,t,s}. \tag{49}$$

This is performed for each model and each forecast horizon, so that models can be ranked by the MSC procedure and a MCS at a chosen confidence level can be identified.

For the variance forecasts of different models, we use the univariate loss functions

$$UQLIK^a_{i,t,s} = \frac{v_{i,t+s}}{h^a_{i,t+s|t}} - \log\left(\frac{v_{i,t+s}}{h^a_{i,t+s|t}}\right) - 1, \tag{50}$$

and

$$MSE^a_{i,t,s} = (v_{i,t+s} - h^a_{i,t+s|t})^2, \tag{51}$$

where $v_{i,t+s}$ is the observed realized variance of stock i at date $t+s$ and $h^a_{i,t+s|t}$ is the corresponding s-step forecast of model a, based on information available at date t. Once the time series average of each loss function has been computed for each stock, the mean across stocks is taken and used in the MCS procedure, i.e.

$$UQLIK^a_s = \frac{1}{T_{h,s}k}\sum_{t=1}^{T_{h,s}}\sum_{i=1}^{k} QLIK^a_{i,t,s}, \text{ and } MSE^a_s = \frac{1}{T_{h,s}k}\sum_{t=1}^{T_{h,s}}\sum_{i=1}^{k} MSE^a_{i,t,s}. \tag{52}$$

The economic loss functions are relevant for the covariance matrix forecasts. They are based on forecasted portfolio performances. The same economic loss functions as Engle and Kelly (2012) are used: global minimum variance portfolio (GMV), and minimum variance portfolio (MV); see also Engle and Colacito (2006). These loss functions are the variances of the forecasts of portfolio returns. Given a covariance matrix forecast $H^a_{t+s|t}$, the GMV portfolio weight vector w^a_{t+s} is computed as the minimizer of the portfolio variance $(w^a_{t+s})'H^a_{t+s|t}w^a_{t+s}$ subject to the constraint that the weights add to unity. Next the portfolio return is $w^a_{t+s}r_{t+s}$ is computed. Once this is done for each forecast date, the variance of the time series of optimal portfolio returns is the loss function used by the MCS procedure, since a a superior model produces optimal portfolios with lower forecast variance.

The MV portfolio is obtained by minimizing the portfolio variance subject to the additional constraint that the expected portfolio return be larger than a chosen value. Following Engle and Kelly (2012), this value is fixed at 10% and the expected portfolio return at the mean of the data.

5.3.2 Results

To compute out-of-sample forecasts, each model is re-estimated every 5th observation based on rolling sample windows of 3,000 observations, resulting in a total of $T_h = 1318$ out-of-sample forecasts for $s = 1$, 1314 for $s = 5$, and 1297 for $s = 22$.

Table 3 reports the out-of-sample forecast losses defined by (49) and the economic loss functions, for the different models. The boldface values identify the models that belong to the 99% model confidence set (MCS99 hereafter) for each loss function, when the comparison is limited to the six symmetric models mentioned in the first paragraph of this subsection. The comparisons including the four asymmetric models included in the table are presented in subsection 5.4, and the MCS ranks in the table pertain to the comparisons of all models.

At forecast horizons 1 and 5, DCC-HEAVY belongs to the MCS99 for all loss functions, DECO-HEAVY belongs to it for a subset of loss functions, and all the other models are out of each MCS99. At horizon 22, no HEAVY model is in MCS99 of FN and QLIK; MCS99 includes DCC-GARCH and DECO-GARCH for FN and DCC-GARCH and BEKK-GARCH for QLIK For the GMV loss function, MCS99 consists of DCC-HEAVY, DECO-HEAVY and BEKK-GARCH, and for MV loss, it contains DCC-GARCH, DECO-HEAVY and BEKK-GARCH.

Insert Table 3 and 4 here

Table 4 reports the MCS99 sets for correlation and variance forecasts separately, excluding the BEKK models for these comparisons, and using only the statistical loss functions since the economic loss functions use the covariance matrix.

For correlations, DCC-HEAVY is in MCS99 of both loss functions at the three horizons. DECO-HEAVY is also in the MCS99 of FN at the three horizons, and DECO-GARCH in the QLIK MCS99 set at horizon 22.

For variances, the loss values defined by (50) are reported in the 'Variance' part of Table 4. Notice that DECO models are irrelevant (being the same as DCC in the first step of estimation). The results reveal that DCC-HEAVY is alone in MCS99 for both MSE and QLIK at horizons 1 and 5. At horizon 22, only DCC-GARCH is in both MCS99 sets. Shephard and Sheppard (2010) report that the performance of HEAVY with respect to GARCH deteriorates as the forecast horizon increases.

To trace where the forecast gains occur, Table 5 reports the ratios between the losses of the DCC-HEAVY and DCC-GARCH models, and likewise for the DECO models. One can see that the DCC/DECO-HEAVY models outperform the DCC/DECO-GARCH models both in covariance, correlation and variance forecast losses at the forecast horizons 1 and 5 (with a single exception for DECO). The improvements are larger for horizon 1 than 5, and smaller for DECO than for DCC (except one case). They can be important, e.g. DCC-HEAVY reduces the covariance and variance QLIK losses by 10% at least and up to 25%. At horizon 22, the GARCH models have the smallest losses, except for correlations where the differences are less than 2.5%. The loss improvements of GARCH with respect to HEAVY are between 7 and 17% when they occur.

Insert Table 5 here

In brief, the forecast comparisons are clearly favouring the DCC-HEAVY model at the short forecast horizons for all loss functions, and to a lesser extent the DECO-HEAVY model, at the expense of the BEKK-HEAVY and the three GARCH models.

At the longest forecast horizon, the results depend on the loss function; moreover there is a clear worse performance of DCC-HEAVY relative to DCC-GARCH for variance and covariance forecasts at this horizon.

The in-sample forecasting results are reported in the SA (Tables B3-B4). They do not differ much from the out-of-sample results.

5.4 Asymmetric and HAR Terms in DCC/DECO-HEAVY

The out-of-sample forecast performance of the DCC-HEAVY model deteriorates with respect to DCC-GARCH when the forecast horizon increases, switching from better to worse at some horizon between 5 and 22. DCC-HEAVY makes use of forecasts of realized variances and correlations. If the realized variance or correlation equations are incorrectly specified, the forecast error is brought to the conditional covariance and correlation forecasts. These forecast errors get larger when the forecast horizon is longer. To improve the forecast performance of DCC-HEAVY in multiple step ahead forecasts, the ADCC-HEAVY model defined by (35)-(37)-(38)-(39) is worth trying. The HAR terms are useful to capture the long memory feature of realized variances and correlations.

5.4.1 Estimation Results

The parameter estimates of the asymmetric variance equations are reported in the SA (Tables B5 and B6). For the conditional variances, the coefficient estimates of the asymmetric term (the diagonal elements of Γ_h in (35)) are positive for 22 stocks, with t-statistics above 2 for eight of them. For the realized variances, the coefficient estimates of the asymmetric term (the diagonal elements of Γ_m in (38)) are all positive, with t-statistics larger than 2.5 for 28 stocks. The coefficient estimates of the weekly HAR term are positive (with a single, insignificant, exception) but only fourteen of them have t-statistics above 2, indicating a moderate impact of this term. On the contrary, the monthly HAR term is positive for all stocks and appears to be strongly significant (with one exception).

The estimates of the asymmetric correlation equations are reported in Table 6. For the conditional correlations, the estimate of γ_r is positive in the ADCC (0.025) and very significant (t-statistic 6.51). This means that the impact of each lagged realized correlation on the next conditional correlation is stronger (being equal to 0.062+0.025) when both lagged returns of the corresponding assets are negative than otherwise (being then equal to 0.062). Nevertheless, the additional statistically significant impact of 0.025 is not very important on the next period correlation, since it increases it by 2.5% of the previous period correlation (if both lagged returns are negative): that means an increase of 0.0125 if the previous correlation is 0.5. For the ADECO model, the impact is of 2% and statistically insignificant.

Insert Table 6 here

For the realized correlations, the estimate of γ_p is positive and significant in both models (HAR-ADCC and HAR-ADECO), which means that the impact of each lagged realized correlation on the next conditional mean of the realized correlation is stronger (being equal to 0.063+0.010 in the HAR-ADCC) when both lagged returns of the corresponding assets are negative than otherwise (being then equal to 0.063). The same remark applies as above, about the limited value of the expected correlation change this implies between two consecutive days.

Concerning the HAR terms, the weekly term impact is negative (-0.033) and significant in HAR-ADCC, positive (0.075) and insignificant in HAR-ADECO. The monthly term impact is negative in HAR-ADCC but positive in HAR-ADECO, being significant in both models. Anyway, the effective daily changes of expected correlations these terms imply are small in HAR-ADCC and moderate in HAR-ADECO.

5.4.2 Forecast Comparisons

For covariance forecasts, Table 3 reports the loss function values for the four asymmetric models (ADCC-HEAVY, ADECO-HEAVY, ADCC-GARCH and ADECO-GARCH) in addition to the six symmetric models. The MCS99 for each loss function and horizon is identified by the underlined values; a bold underlined value is thus in the MCS99 for the ten models, and in that of the first six models, while a bold value, not underlined, is in the MCS99 of the first six models but is removed from the MCS99 when the ten models are compared.

The changes of the MCS99 due to the inclusion of the four asymmetric models in the comparisons are in favour of HEAVY models: the asymmetric HEAVY models are added to the MCS99 of some loss functions and horizons, but no asymmetric GARCH is model added for any loss function and horizon. For example, at horizons 5 and 22, ADCC-HEAVY is in the MCS99 sets for all loss functions; only the ADECO-HEAVY model also belongs to these sets in the case of the FN loss. The asymmetric HEAVY models attenuate or even reverse the worse performance of their symmetric counterparts relative to DCC at the three horizons, as can be seen by comparing the covariance loss ratios reported in Tables 7 and 8 and the corresponding values in Table 5.

For correlation forecasting, Table 4 also includes the asymmetric models in the comparisons. The changes of the MCS99 due to this are mixed: at horizon 1, only the asymmetric versions of the HEAVY models that were in the initial MCS99 are added;

at horizon 5, ADCC- and ADECO-HEAVY are added for both losses; at horizon 22, ADECO-HEAVY is added for FN loss, being alone in MCS99; for QLIK loss, ADCC-HEAVY, ADECO-GARCH and ADECO-HEAVY are added, DCC-HEAVY being removed and DECO-GARCH kept. The correlation loss ratios reported in Tables 7 and 8 differ slightly from those of Table 5, indicating that the asymmetric versions of the models do not improve much the symmetric versions.

For variance forecasting (Table 4), the main change in the MCS99 due to the inclusion of asymmetric models is that ADCC-HEAVY is added for both loss functions and all horizons, being even the single model in the sets at horizons 5 and 22. This model has the rank 1 in all comparisons. Clearly, ADCC-HEAVY improves variance forecasting, especially at horizons larger than 1. Loss ratios show that this improvement is important at horizon 22; for example, for MSE loss, it goes from -16.6% in Table 5 to +14.7% in Table 7 and +22.7% in Table 8.

Insert Tables 7 and 8 here

6 Conclusions

Multivariate volatility models that specify the dynamics of the daily conditional covariance matrix as a function of realized covariances have emerged in the literature since 2012. They are a valuable alternative to multivariate GARCH models wherein the dynamics depend on lagged squared returns and their cross-products, because realized variances and covariances are more precise measures of daily volatility.

Perhaps surprisingly, with the partial exception of Braione (2016), no dynamic conditional correlation formulation of a HEAVY model has been proposed in the literature, where BEKK-type formulations are used in the papers of Noureldin et al. (2012) and Opschoor et al. (2018). Our contribution fills this gap by developing DCC-type HEAVY models. Such models have the advantage, with respect to BEKK models, of separating the specification of the conditional variances from the specification of the conditional correlations. The same advantage occurs in the specifications of the expected realized variances and correlations. As for GARCH DCC models, this results in more flexible models, in the sense that the dynamics of variances is different between assets and from the dynamics of correlations. Sticking to scalar models for correlations, the models remain parsimonious in parameters.

An illustrative empirical application for twenty-nine assets illustrates the value of the flexibility of DCC and DECO versions of HEAVY models. These models, including extensions to include asymmetric effects, have superior forecasting performance with respect to BEKK formulations. As always in this type of empirical exercise, this finding is contingent to the dataset used and cannot be claimed to be valid in general. A robust conclusion is that HEAVY models dominate GARCH versions in term of forecasting performance.

This research has not developed the properties of the QML estimators of the parameters of the DCC- and DECO-HEAVY models. Such developments and additional empirical studies will come out of future research.

Acknowledgments

Thanks to Professors Lyudmila Grigoryeva (University of Konstanz) and Juan-Pablo Ortega (University of St.Gallen) for providing the datasets for the DJIA companies and to Oleksandra Kukharenko (University of Konstanz) for handling the data and computing the realized covariance matrices from the TAQ data.

References

[1] Aielli, G. (2013). Dynamic conditional correlation: on properties and estimation. Journal of Business & Economic Statistics 31, 282-299.

[2] Bauwens, L., Laurent, S. and Rombouts, J. V. K. (2006). Multivariate GARCH models: a survey. Journal of Applied Econometrics, 21, 79-109.

[3] Bauwens, L., Storti, G. and Violante, F. (2012). Dynamic conditional correlation models for realized covariance matrices. CORE DP 2012/60.

[4] Bauwens, L., Braione, M. and Storti, G. (2016). Forecasting Comparison of Long Term Component Dynamic Models for Realized Covariance Matrices. Annals of Economics and Statistics, 123/124, 103-134.

[5] Braione, M. (2016). A time-varying long run HEAVY model. Statistics and Probability Letters, 119, 36-44.

[6] Chiriac, R. and Voev, V. (2011). Modelling and forecasting multivariate realized volatility. Journal of Applied Econometrics, 26, 922-947.

[7] Corsi, F. (2009). A simple approximate long-memory model of realized volatility. Journal of Financial Econometrics, 7, 174-196.

[8] Engle, R. F., Hendry, D. F. and Richard, J. F. (1983). Exogeneity. Econometrica, 51, 277-304.

[9] Engle, R.F. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. Journal of Business & Economic Statistics, 20, 339-350.

[10] Engle, R.F. and Kelly, B. (2012). Dynamic equicorrelation. Journal of Business & Economic Statistics, 30, 212-228.

[11] Engle, R. F. and Gallo, G. M. (2006). A multiple indicator model for volatility using intra-daily data. Journal of Econometrics, 131, 3-27.

[12] Engle, R. F. and Sheppard, K. (2001). Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH. WP 8554, National Bureau of Economic Research.

[13] Gouriéroux, C., Jasiak, J. and Sufana, R. (2009). The Wishart autoregressive process of multivariate stochastic volatility. Journal of Econometrics, 150, 167-181.

[14] Golosnoy, V., Gribisch, B. and Liesenfeld, R. (2012). The conditional autoregressive Wishart model for multivariate stock market volatility. Journal of Econometrics, 167, 211-223.

[15] Gorgi, P., Hansen, P.R., Janus, P. and Koopman, S.J. (2019). Realized Wishart-GARCH: A Score-driven Multi-Asset Volatility Model. Journal of Financial Econometrics, 17, 1-32.

[16] Hansen, P. R., Lunde, A. and Nason, J.M. (2011). The model confidence set. Econometrica, 79, 453-497.

[17] Hansen, P. R., Huang, Z. and Shek H. (2012). Realized GARCH: A Joint Model of Returns and Realized Measures of Volatility. Journal of Applied Econometrics 27, 877-906.

[18] Hansen, P. R., Lunde, A. and Voev, V. (2014). Realized Beta GARCH: A Multivariate GARCH Model with Realized Measures of Volatility. Journal of Applied Econometrics 29, 774-799.

[19] Jin, X. and Maheu, J. M. (2012). Modeling realized covariances and returns. Journal of Financial Econometrics, 11, 335-369.

[20] Laurent, S., Rombouts, J. V. and Violante, F. (2013). On loss functions and ranking forecasting performances of multivariate volatility models. Journal of Econometrics, 173, 1-10.

[21] Noureldin, D., Shephard, N. and Sheppard, K. (2012). Multivariate high-frequency-based volatility (HEAVY) models. Journal of Applied Econometrics, 27, 907-933.

[22] Oh, D. H. and Patton, A. J. (2016). High-dimensional copula-based distributions with mixed frequency data. Journal of Econometrics, 193, 349-366.

[23] Opschoor, A., Janus, P., Lucas, A. and Van Dijk, D. (2018). New HEAVY models for fat-tailed realized covariances and returns. Journal of Business & Economic Statistics, 1-15.

[24] Patton A.J. (2011). Volatility forecast comparison using imperfect volatility proxies. Journal of Econometrics 160, 246-256.

[25] Patton A.J and Sheppard K.K. (2009). Evaluating volatility and correlation forecasts. In Handbook of Financial Time Series, Andersen TG, Davis RA, Kreiss JP, Mikosch T (eds). Springer: Berlin; 801-838.

[26] Shephard, N. and Sheppard, K. (2010). Realising the future: forecasting with high-frequency-based volatility (HEAVY) models. Journal of Applied Econometrics, 25, 197-231.

[27] Sheppard, K. and Xu, W. (2019). Factor High-Frequency-Based Volatility (HEAVY) Models. Journal of Financial Econometrics, 17, 33-65.

[28] Tse, Y. K. and Tsui, A. K. C. (2002). A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. Journal of Business & Economic Statistics, 20, 351-362.

Table 1: Summary of parameter estimates of the variance equations

	GARCH (26)-(28)		GARCHX (26)-(29)			HEAVY-h (10)		HEAVY-m (20)	
	A_h	B_h	A_h	A_{hm}	B_h	A_h	B_h	A_m	B_m
Min	0.024	0.830	0.002	0.103	0.199	0.353	0.199	0.297	0.514
Med	0.072	0.916	0.028	0.698	0.507	0.781	0.481	0.369	0.606
Max	0.117	0.971	0.068	1.748	0.885	1.748	0.775	0.470	0.695

Minimum (Min), median (Med), and maximum (Max) are the summary statistics of the estimates for the 29 stocks. All estimates are provided in Table B2 of the supplementary appendix. For A_h of GARCHX, the statistics are for the 21 non-zero estimates (8 estimates are equal to 0).

Table 2: Parameter estimates (and robust t-statistics) of the correlation equations

DCC-GARCH (17)		DCCX-GARCH (30)-(31)			DCC-HEAVY-R (11)-(12)		DCC-HEAVY-P (21)	
α_q	β_q	α_q	β_q	α_r	α_r	β_r	α_m	β_m
0.003	0.988	0.006	0.876	0.061	0.068	0.869	0.041	0.948
(8.56)	(457.39)	(10.59)	(56.76)	(9.40)	(10.40)	(54.65)	(30.59)	(506.6)

Log-likelihood decomposition (DCC-HEAVY vs DCC-GARCH)

	DCC-GARCH	DCCX-GARCHX	DCC-HEAVY	HEAVY Gains
Variance	-7785	-7644	-7649	136
Correlation	-5143	-5115	-5119	24
Total	-12927	-12759	-12768	159

DECO-GARCH		DECOX-GARCH			DECO-HEAVY-R		DECO-HEAVY-P	
α_q	β_q	α_q	β_q	α_r	α_r	β_r	α_m	β_m
0.046	0.937	0.029	0.615	0.369	0.447	0.552	0.347	0.607
(3.25)	(42.66)	(1.94)	(10.02)	(5.09)	(7.97)	(9.86)	(16.33)	(23.80)

Log-likelihood decomposition (DECO-HEAVY vs DECO-GARCH)

	DECO-GARCH	DECOX-GARCHX	DECO-HEAVY	HEAVY Gains
Variance	-7785	-7644	-7649	136
Correlation	-5319	-5306	-5307	12
Total	-13104	-12950	-12956	148

The HEAVY Gains are the differences of log-likelihood between DCC/DECO-HEAVY model to the DCC/DECO GARCH model.

Table 3: MCS for loss functions of out-of-sample covariance forecasting

	FN	MCS Rank	QLIK	MCS Rank	GMV	MCS Rank	MV	MCS Rank
$s = 1$								
DCC-GARCH	16.792	5	12.007	6	0.226	7	0.309	7
DCC-HEAVY	**13.292**	1	**8.821**	1	**0.193**	1	**0.272**	1
DECO-GARCH	16.513	6	14.182	9	0.237	10	0.330	10
DECO-HEAVY	**13.667**	4	11.059	4	0.214	3	0.308	4
BEKK-GARCH	23.140	9	15.016	8	0.213	4	0.302	3
BEKK-HEAVY	28.312	10	13.559	7	0.227	8	0.329	8
ADCC-GARCH	16.768	8	12.077	5	0.228	6	0.311	5
ADCC-HEAVY	13.304	2	8.842	2	0.196	2	0.276	2
ADECO-GARCH	16.780	7	14.125	10	0.239	9	0.336	9
ADECO-HEAVY	13.697	3	11.084	3	0.215	5	0.312	6
$s = 5$								
DCC-GARCH	18.008	5	13.268	6	0.228	7	0.314	7
DCC-HEAVY	**15.907**	4	**11.703**	2	**0.203**	2	**0.286**	2
DECO-GARCH	17.824	6	15.242	9	0.237	10	0.332	9
DECO-HEAVY	**16.135**	3	**11.850**	3	**0.206**	3	0.292	3
BEKK-GARCH	23.637	9	15.298	7	0.216	5	0.307	5
BEKK-HEAVY	29.791	10	14.711	8	0.234	9	0.337	10
ADCC-GARCH	18.126	8	13.395	4	0.229	6	0.316	6
ADCC-HEAVY	15.299	2	11.384	1	0.203	1	0.286	1
ADECO-GARCH	18.209	7	15.255	10	0.239	8	0.337	8
ADECO-HEAVY	15.062	1	12.937	5	0.211	4	0.302	4
$s = 22$								
DCC-GARCH	**20.720**	5	**15.798**	4	0.230	6	**0.319**	5
DCC-HEAVY	21.863	8	17.310	7	**0.220**	3	**0.313**	3
DECO-GARCH	**20.426**	4	17.327	9	0.235	9	0.329	8
DECO-HEAVY	22.029	6	17.217	8	**0.221**	4	**0.315**	4
BEKK-GARCH	25.201	9	**16.109**	3	**0.224**	5	**0.320**	6
BEKK-HEAVY	32.171	10	16.463	5	0.243	10	0.351	10
ADCC-GARCH	21.204	7	16.327	6	0.233	7	0.328	7
ADCC-HEAVY	18.487	2	14.890	1	0.215	1	0.305	1
ADECO-GARCH	20.879	3	18.081	10	0.236	8	0.337	9
ADECO-HEAVY	18.418	1	15.251	2	0.219	2	0.313	2

Values of loss functions in bold identify the models in the 99% level MCS when the comparison is limited to the first six models. Underlined values identify the models in the 99% level MCS when the comparison is done for all models. A value in bold but not underlined is thus in the MCS of the first six models, but is excluded when considering all models. The MCS rankings are for the global comparison.

Table 4: MCS for loss functions of out-of-sample correlation and variance forecasting

	Correlation				Variance			
	FN	MCS Rank	QLIK	MCS Rank	MSE	MCS Rank	UQLIK	MCS Rank
$s = 1$								
DCC-GARCH	6.119	8	20.833	3	2.544	3	0.509	4
DCC-HEAVY	**5.505**	4	**20.450**	2	**2.210**	2	**0.386**	2
DECO-GARCH	5.787	6	21.283	8				
DECO-HEAVY	**5.364**	1	21.081	5				
ADCC-GARCH	6.037	7	20.817	4	2.752	4	0.504	3
ADCC-HEAVY	5.478	3	20.417	1	2.205	1	0.385	1
ADECO-GARCH	5.719	5	21.266	7				
ADECO-HEAVY	5.379	2	21.121	6				
$s = 5$								
DCC-GARCH	6.161	8	20.852	4	2.838	3	0.561	3
DCC-HEAVY	**5.741**	4	**20.527**	3	**2.659**	2	**0.501**	2
DECO-GARCH	5.959	7	21.023	7				
DECO-HEAVY	**5.574**	1	21.713	8				
ADCC-GARCH	6.100	6	20.844	5	3.097	4	0.562	4
ADCC-HEAVY	5.728	3	20.480	2	2.347	1	0.488	1
ADECO-GARCH	5.915	5	20.982	6				
ADECO-HEAVY	5.609	2	20.383	1				
$s = 5$								
DCC-GARCH	6.258	8	20.851	7	**3.359**	2	**0.661**	2
DCC-HEAVY	**6.061**	2	**20.623**	5	3.916	4	0.710	4
DECO-GARCH	6.237	7	**20.536**	1				
DECO-HEAVY	**6.112**	4	21.111	8				
ADCC-GARCH	6.224	6	20.830	6	3.706	3	0.676	3
ADCC-HEAVY	6.025	1	20.560	2	2.865	1	0.620	1
ADECO-GARCH	6.112	3	20.564	4				
ADECO-HEAVY	6.154	5	20.564	3				

Values of loss functions in bold identify the models in the 99% level MCS when the comparison is limited to the first six models. Underlined values identify the models in the 99% level MCS when the comparison is done for all models. A value in bold but not underlined is thus in the MCS of the first six models, but is excluded when considering all models. The MCS rankings are for the global comparison.

Table 5: Loss ratios between DCC-HEAVY and DCC-GARCH: out-of-sample forecasting

		DCC-HEAVY vs DCC-GARCH		DECO-HEAVY vs DECO-GARCH	
		FN	QLIK	FN	QLIK
Covariance	s=1	0.792	0.735	0.828	0.780
	s=5	0.883	0.882	0.905	0.777
	s=22	1.055	1.096	1.078	0.994
Correlation	s=1	0.900	0.982	0.927	0.991
	s=5	0.932	0.984	0.935	1.033
	s=22	0.968	0.989	0.980	1.028
		MSE	UQLIK		
Variance	s=1	0.869	0.759		
	s=5	0.937	0.892		
	s=22	1.166	1.074		

For the Variance panel, results of DCC and DECO are identical.

Table 6: Parameter estimates (and robust t-statistics) of the asymmetric correlation equations

ADCC-HEAVY-R (36)-(37)			HAR-ADCC-HEAVY-M (39)				
α_r	β_r	γ_r	α_p	β_p	γ_p	α_p^w	α_p^m
0.062	0.857	0.025	0.063	0.964	0.010	-0.033	-0.005
(11.05)	(52.60)	(6.51)	(40.27)	(353.68)	(12.05)	(-18.36)	(-3.06)
ADECO-HEAVY-R			**HAR-ADECO-HEAVY-P**				
α_r	β_r	γ_r	α_p	β_p	γ_p	α_p^w	α_p^m
0.438	0.573	0.020	0.195	0.511	0.173	0.075	0.100
(6.78)	(9.70)	(0.54)	(9.40)	(8.53)	(13.53)	(1.55)	(5.32)

Table 7: Loss ratios between ADCC-HEAVY and DCC-GARCH: out-of-sample forecasting

		ADCC-HEAVY vs DCC-GARCH		ADECO-HEAVY vs DECO-GARCH	
		FN	QLIK	FN	QLIK
Covariance	s=1	0.792	0.736	0.830	0.782
	s=5	0.850	0.858	0.845	0.849
	s=22	0.892	0.943	0.902	0.880
Correlation	s=1	0.895	0.980	0.929	0.992
	s=5	0.930	0.982	0.941	0.970
	s=22	0.963	0.986	0.987	1.001
		MSE	UQLIK		
Variance	s=1	0.867	0.758		
	s=5	0.827	0.869		
	s=22	0.853	0.937		

For the Variance panel, results of DCC and DECO are identical.

Table 8: Loss ratios between ADCC-HEAVY and ADCC-GARCH: out-of-sample forecasting

		ADCC-HEAVY vs ADCC-GARCH		ADECO-HEAVY vs ADECO-GARCH	
		FN	QLIK	FN	QLIK
Covariance	s=1	0.793	0.732	0.816	0.785
	s=5	0.844	0.850	0.827	0.848
	s=22	0.872	0.912	0.882	0.843
Correlation	s=1	0.907	0.981	0.940	0.993
	s=5	0.939	0.983	0.948	0.971
	s=22	0.968	0.987	1.007	1.000
		MSE	UQLIK		
Variance	s=1	0.801	0.765		
	s=5	0.758	0.868		
	s=22	0.773	0.916		

For the Variance panel, results of DCC and DECO are identical.

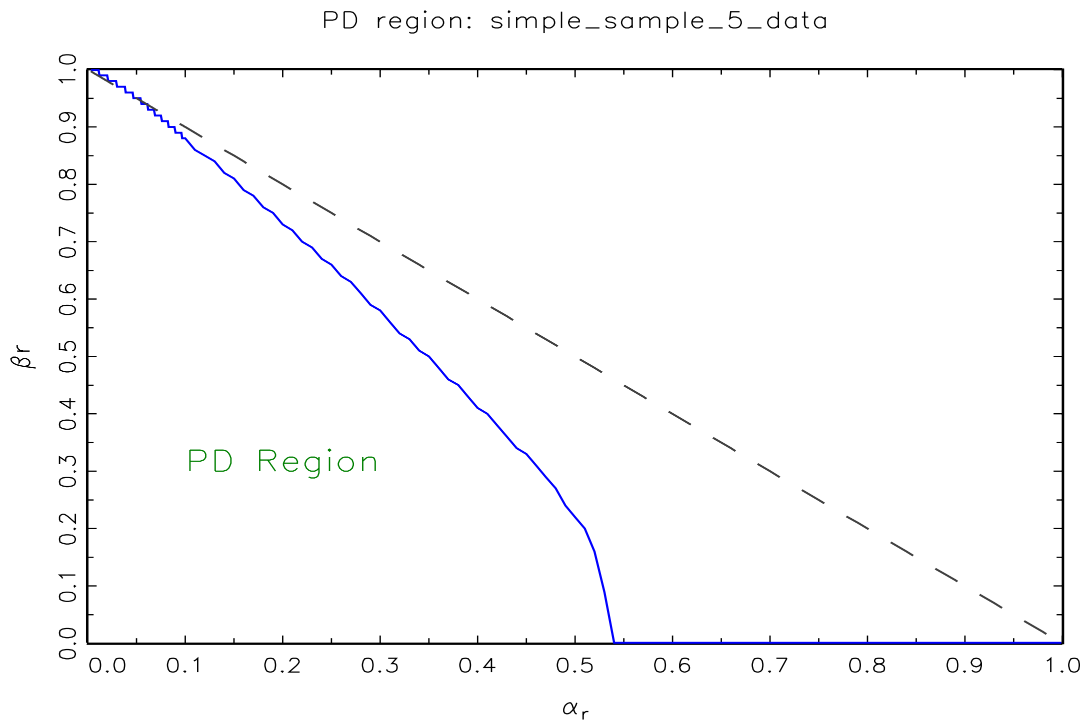

Figure 1: The figure shows the parameter values for which R_t defined by (11)-(12) becomes PD for $t > \tilde{t}$ in the sample period, for the data described in Section 5.1. Below the continuous line, R_t remains PD throughout the full sample period.

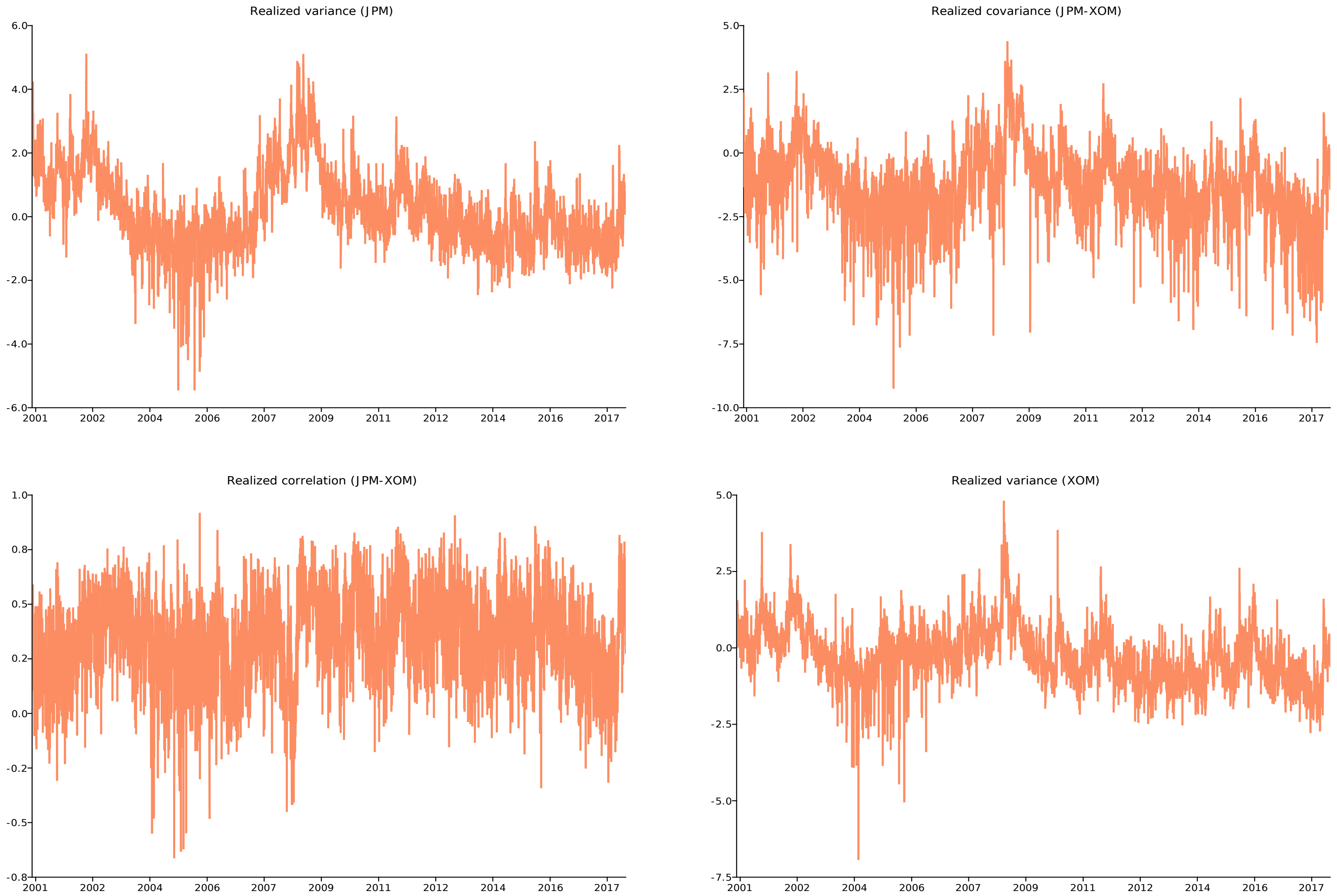

Figure 2: The figure shows the log realized volatilities of JP Morgan (JPM) and Exxon Mobil (XOM), the log realized covariance and the realized correlation between these stocks, over the sample period 03/01/2001-16/04/2018.

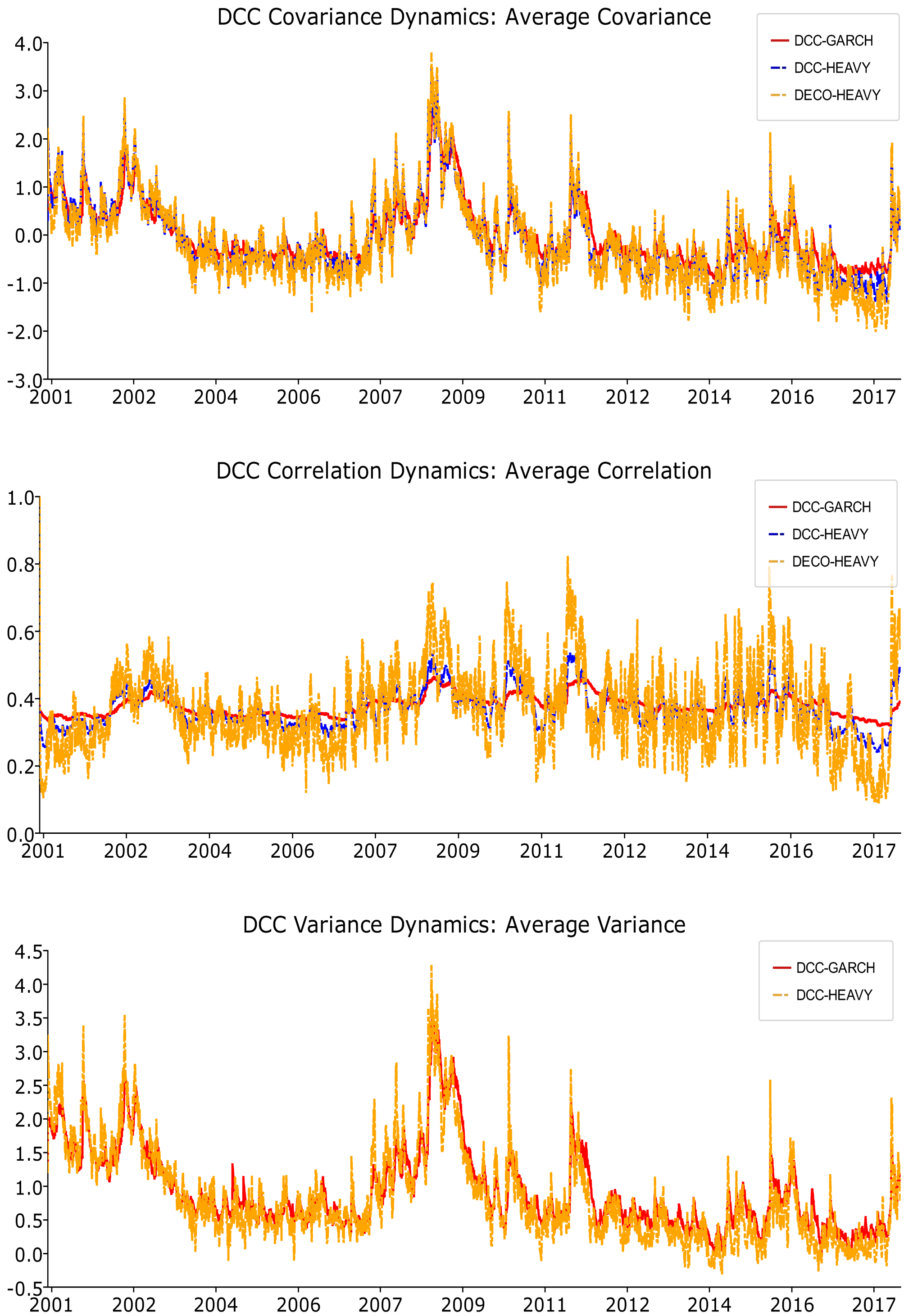

Figure 3: The figure shows the pairwise average of log-covariances and correlations estimated in the second step with DCC-GARCH and DCC/DECO-HEAVY-H, over the sample period 03/01/2001-16/04/2018. The bottom graph shows the average of the log-variances of the 29 stocks, estimated in the first step of DCC-GARCH and DCC-HEAVY.

Supplementary Appendix

A Forecasts of the ADCC-HEAVY Model

In this section, formulas for multiple-step forecasts of the ADCC-HEAVY model are derived. The model is defined in Section 3.3 of the paper. The aim is to forecast the conditional covariance of daily returns, H_t. Its s-step forecasts $H_{t+s|t}$ is computed as defined by (32), so $E_t[h_{t+s}]$ and $E_t[R_{t+s}]$ must be computed.

The joint conditional expectation for r_t^2 and v_t is formed by stacking the conditional and realized variance equations (35) and (38). Defining the vectors of $2k \times 1$ elements $x_t = [(r_t^2)', v_t']'$, $\mu_t = [h_t', m_t']'$, $x_t^w = [(0', (v_t^w)')]'$, and $x_t^m = [0', (v_t^m)']'$, the joint conditional expectation is

$$
\begin{aligned}
E(x_t|\mathcal{F}_{t-1}) &:= \mu_t \\
\mu_t &= \omega + A x_{t-1} + B\mu_{t-1} + \Gamma d_{t-1} \odot x_{t-1} + A_w x_{t-1}^w + A_m x_{t-1}^m
\end{aligned}
$$

where

$$
\begin{aligned}
\omega &= \begin{bmatrix} \omega_h \\ \omega_m \end{bmatrix}, A = \begin{bmatrix} 0 & A_h \\ 0 & A_m \end{bmatrix}, B = \begin{bmatrix} B_h & 0 \\ 0 & B_m \end{bmatrix}, \\
\Gamma &= \begin{bmatrix} 0 & \Gamma_h \\ 0 & \Gamma_m \end{bmatrix}, A_w = \begin{bmatrix} 0 & 0 \\ 0 & A_m^w \end{bmatrix}, A_m = \begin{bmatrix} 0 & 0 \\ 0 & A_m^m \end{bmatrix}.
\end{aligned}
$$

The matrices A, B, Γ, A_w and A_m are all square of order $2k$, and ω is a column vector of $2k$ elements.

From the previous expression, the s-step forecast $E_t(\mu_{t+s})$ can be deduced, so that $E_t[h_{t+s}]$ can be extracted from it. The one-step forecast is

$$
\mu_{t+1|t} = \omega + A x_t + B\mu_t + \Gamma d_t \odot x_t + A_w x_t^w + A_m x_t^m.
$$

To forecast more than one step ahead, it is assumed that each return r_{it} has a symmetric distribution, uncorrelated with v_t, so that $E(d_{it} v_{it}) = 0.5E(v_{it})$. Since x_{t+s}, x_{t+s}^w and x_{t+s}^m for $s > 0$ are not fully known, they are replaced by their corresponding conditional expectations. Hence for $s > 1$, the forecast formula is

$$
\begin{aligned}
\mu_{t+s|t} &= \omega + (A + B + 0.5\Gamma)\mu_{t+s-1|t} + A_w x_{t+s-1|t}^w + A_m x_{t+s-1|t}^m \quad \text{for } 2 \leq s < 22, \\
\mu_{t+s|t} &= \omega + (A + B + 0.5\Gamma + A_w + A_m)\mu_{t+s-1|t} \quad \text{for } s \geq 22,
\end{aligned}
$$

where $x_{t+s-1|t}^w = \frac{1}{5}\sum_{j=1}^{5} x_{t+s-j|t}$, $x_{t+s-1|t}^m = \frac{1}{22}\sum_{j=1}^{22} x_{t+s-j|t}$ and $x_{t+s-j|t} = \mu_{t+s-j|t}$ if $s > j$.

The joint conditional expectation of $Y_t = [u_t u_t', RL_t]'$ is obtained by stacking the conditional and realized correlation equations (36)-(37) and (39):

$$\begin{aligned} E(Y_t|\mathcal{F}_{t-1}) &:= \Phi_t \\ \Phi_t &= W + (\alpha \otimes J_k)Y_{t-1} + (\beta \otimes J_k)\Phi_{t-1} + (\gamma \otimes J_k)(D_{t-1} \odot Y_{t-1}) \\ &\quad +(\alpha_w \otimes J_k)Y_{t-1}^w + (\alpha_m \otimes J_k)Y_{t-1}^m \end{aligned}$$

where

$$\begin{aligned} W &= \begin{bmatrix} \tilde{R} \\ \tilde{P} \end{bmatrix}, \ \alpha = \begin{bmatrix} 0 & \alpha_r \\ 0 & \alpha_p \end{bmatrix}, \ \beta = \begin{bmatrix} \beta_r & 0 \\ 0 & \beta_p \end{bmatrix}, \\ \gamma &= \begin{bmatrix} 0 & \gamma_r \\ 0 & \gamma_p \end{bmatrix}, \ \alpha_w = \begin{bmatrix} 0 & 0 \\ 0 & \alpha_p^w \end{bmatrix}, \ \alpha_m = \begin{bmatrix} 0 & 0 \\ 0 & \alpha_p^m \end{bmatrix}. \end{aligned}$$

The s-step ahead forecast $E_t(\Phi_{t+s})$ is then

$$\begin{aligned} \Phi_{t+1|t} &= W + (\alpha \otimes J_k)Y_t + (\beta \otimes J_k)\Phi_t + (\gamma \otimes J_k)(D_t \odot Y_t) \\ &\quad +(\alpha_w \otimes J_k)Y_t^w + (\alpha_m \otimes J_k)Y_t^m \\ \Phi_{t+s|t} &= W + [(\alpha + \beta + 0.25\gamma) \otimes J_k)]\Phi_{t+s-1|t} \\ &\quad +(\alpha_w \otimes J_k)Y_{t+s-1|t}^w + (\alpha_m \otimes J_k))Y_{t+s-1|t}^m \quad \text{for } 2 \le s < 22, \\ \Phi_{t+s|t} &= W + [(\alpha + \beta + 0.25\gamma + \alpha_w + \alpha_m) \otimes J_k)]\Phi_{t+s-1|t} \quad \text{for } s \ge 22, \end{aligned}$$

where $Y_{t+s-1|t}^w = \frac{1}{5}\sum_{j=1}^{5} Y_{t+s-j|t}$, $Y_{t+s-1|t}^m = \frac{1}{22}\sum_{j=1}^{22} Y_{t+s-j|t}$ and $Y_{t+s-j|t} = \Phi_{t+s-j|t}$ if $s > j$. We assume that $E(D_{t-1} \odot Y_{t-1}) = E(D_{t-1}) \odot E(Y_{t-1})$ and $E(D_{t-1}) = I_k + 0.25\text{OFFDIAG}(U_k)$ where U_k is a $k \times k$ matrix of ones. The value 0.25 results from the assumption that $E(d_{it}d_{jt}) = E(d_{it})E(d_{jt})$.

B Additional Tables for Section 5

Table B1: Data descriptive statistics

	r_t^2		v_t		RC_t		RL_t	
Stock	Mean	StDev	Mean	StDev	Mean	StDev	Mean	StDev
AAPL	5.488	15.118	3.151	5.319	0.324	0.147	0.324	0.147
AXP	4.996	18.254	2.645	6.536	0.349	0.152	0.349	0.152
BA	3.381	9.966	1.827	2.922	0.321	0.157	0.321	0.157
CAT	3.930	10.319	2.118	3.525	0.341	0.155	0.341	0.155
CSCO	5.475	19.464	2.579	4.565	0.342	0.148	0.342	0.148
CVX	2.535	9.408	1.464	3.039	0.314	0.170	0.314	0.170
DIS	3.364	11.319	1.899	3.730	0.325	0.163	0.325	0.163
DWDP	4.656	14.401	1.711	2.752	0.346	0.158	0.346	0.158
GE	3.550	12.067	2.076	5.053	0.354	0.152	0.354	0.152
GS	5.120	20.686	2.761	9.618	0.338	0.145	0.338	0.145
HD	3.339	9.951	1.956	3.490	0.333	0.153	0.333	0.153
IBM	2.301	6.889	1.259	2.584	0.353	0.153	0.353	0.153
INTC	4.873	15.478	2.536	3.898	0.345	0.146	0.345	0.146
JNJ	1.278	5.833	0.842	1.614	0.305	0.157	0.305	0.157
JPM	6.128	25.275	3.057	7.765	0.352	0.146	0.352	0.146
KO	1.417	4.944	0.913	1.657	0.295	0.154	0.295	0.154
MCD	1.902	5.741	1.272	2.418	0.286	0.156	0.286	0.156
MMM	1.933	5.312	1.179	2.611	0.370	0.160	0.370	0.160
MRK	2.914	17.349	1.543	4.397	0.288	0.157	0.288	0.157
MSFT	3.254	10.082	1.677	2.606	0.343	0.149	0.343	0.149
NKE	3.104	11.323	1.623	2.521	0.292	0.159	0.292	0.159
PFE	2.296	7.765	1.420	2.126	0.296	0.155	0.296	0.155
PG	1.265	3.724	0.832	1.525	0.296	0.153	0.296	0.153
TRV	3.182	14.660	1.687	4.464	0.298	0.161	0.298	0.161
UNH	3.813	19.125	1.943	3.491	0.249	0.156	0.249	0.156
UTX	2.636	15.604	1.373	2.487	0.353	0.160	0.353	0.160
VZ	2.185	6.472	1.471	2.784	0.288	0.158	0.288	0.158
WMT	1.813	5.507	1.177	2.121	0.302	0.153	0.302	0.153
XOM	2.257	8.378	1.373	3.159	0.329	0.167	0.329	0.167

This table provides descriptive statistics for the dataset of 29 Dow Jones Industrial Average Stocks. Column 1 show the stock tickers. Columns 2-3 report the corresponding time-series averages (Mean) and standard deviations (StDev) of the squared returns (r_t^2). Columns 4-5 report the same statistics for the realized variances (v_t). Columns 6-7 show the meanss (Mean) and standard deviations (StDev) of the time series averages of the realized covariances of each stock with the other 28 stocks. Columns 8-9 show the statistics for the corresponding realized correlations.

Table B2: Parameter estimates (and robust t-statistics) of the variance equations

	GARCH (26)-(28)		GARCHX (26)-(29)			HEAVY-h (10)		HEAVY-m (20)	
	A_h	B_h	A_h	A_{hm}	B_h	A_h	B_h	A_m	B_m
AAPL	0.039	0.956	0.006	0.404	0.724	0.417	0.721	0.369	0.627
	(2.41)	(50.82)	(0.48)	(4.14)	(11.75)	(4.48)	(11.76)	(8.07)	(14.24)
AXP	0.086	0.907	0.000	1.006	0.446	1.006	0.446	0.386	0.598
	(5.21)	(54.85)	(.)	(3.44)	(2.88)	(3.44)	(2.88)	(7.47)	(11.27)
BA	0.087	0.888	0.032	0.685	0.507	0.787	0.481	0.312	0.658
	(4.66)	(39.91)	(1.88)	(3.85)	(4.38)	(4.03)	(3.81)	(5.59)	(11.41)
CAT	0.043	0.945	0.000	0.353	0.770	0.353	0.770	0.348	0.624
	(2.33)	(34.81)	(.)	(2.36)	(7.52)	(2.36)	(7.51)	(11.10)	(18.20)
CSCO	0.024	0.971	0.000	1.748	0.199	1.748	0.199	0.391	0.599
	(2.51)	(79.04)	(.)	(7.39)	(2.56)	(7.39)	(2.56)	(8.75)	(14.07)
CVX	0.080	0.900	0.024	0.434	0.671	0.496	0.656	0.396	0.584
	(7.49)	(70.79)	(1.77)	(4.41)	(10.57)	(5.66)	(11.01)	(15.94)	(22.58)
DIS	0.086	0.898	0.000	0.756	0.549	0.756	0.549	0.393	0.583
	(3.31)	(34.21)	(.)	(2.97)	(3.72)	(2.97)	(3.72)	(9.28)	(13.80)
DWDP	0.072	0.920	0.068	0.103	0.885	0.767	0.682	0.367	0.606
	(4.96)	(55.82)	(3.74)	(1.69)	(21.41)	(3.57)	(7.48)	(14.39)	(23.00)
GE	0.077	0.916	0.045	0.943	0.394	0.942	0.441	0.386	0.601
	(2.32)	(25.27)	(1.31)	(4.55)	(3.16)	(4.22)	(3.66)	(10.90)	(17.50)
GS	0.052	0.940	0.005	0.383	0.769	0.382	0.775	0.337	0.631
	(3.97)	(60.57)	(0.28)	(3.94)	(12.51)	(3.87)	(13.50)	(4.68)	(8.46)
HD	0.075	0.916	0.038	0.825	0.457	0.865	0.476	0.359	0.619
	(3.48)	(36.02)	(1.43)	(8.90)	(7.76)	(8.76)	(8.94)	(10.60)	(17.95)
IBM	0.117	0.830	0.028	1.271	0.218	1.323	0.220	0.363	0.605
	(3.22)	(18.07)	(1.18)	(4.58)	(1.42)	(4.81)	(1.40)	(10.57)	(17.02)
INTC	0.047	0.945	0.002	1.112	0.403	1.111	0.406	0.417	0.560
	(4.71)	(85.53)	(0.08)	(5.37)	(3.70)	(5.29)	(3.67)	(7.64)	(10.31)
JNJ	0.108	0.866	0.052	0.398	0.629	0.490	0.612	0.355	0.627
	(6.23)	(43.30)	(2.65)	(5.36)	(11.55)	(6.38)	(10.96)	(8.32)	(15.86)
JMP	0.080	0.914	0.019	0.621	0.646	0.681	0.635	0.466	0.514
	(4.13)	(45.08)	(1.36)	(2.97)	(5.79)	(3.22)	(5.62)	(13.46)	(14.99)
KO	0.082	0.901	0.050	0.692	0.427	0.781	0.408	0.347	0.631
	(1.66)	(14.92)	(2.00)	(5.72)	(5.04)	(6.51)	(5.14)	(11.25)	(19.70)
MCD	0.046	0.944	0.017	0.698	0.469	0.691	0.494	0.297	0.695
	(4.51)	(75.42)	(0.56)	(2.76)	(2.36)	(2.49)	(2.53)	(7.88)	(18.49)
MMM	0.060	0.917	0.000	0.803	0.466	0.803	0.466	0.425	0.570
	(2.85)	(30.59)	(.)	(2.81)	(2.53)	(2.81)	(2.53)	(6.47)	(10.96)
MRK	0.061	0.866	0.028	0.878	0.338	0.923	0.337	0.470	0.519
	(1.94)	(10.62)	(0.83)	(4.56)	(3.67)	(4.96)	(2.72)	(5.95)	(6.66)
MSFT	0.040	0.947	0.000	1.394	0.234	1.394	0.234	0.380	0.587
	(1.84)	(31.32)	(.)	(6.42)	(2.18)	(6.42)	(2.18)	(10.85)	(15.55)
NKE	0.060	0.910	0.000	1.127	0.344	1.125	0.345	0.422	0.543
	(2.17)	(23.16)	(.)	(4.68)	(2.74)	(4.68)	(2.76)	(8.26)	(10.35)
PFE	0.093	0.893	0.062	0.582	0.549	0.778	0.478	0.337	0.636
	(1.78)	(14.92)	(1.11)	(3.10)	(3.68)	(1.93)	(1.75)	(8.15)	(13.76)
PG	0.057	0.921	0.037	0.738	0.374	0.794	0.372	0.386	0.588
	(1.86)	(20.18)	(1.70)	(4.70)	(3.05)	(4.75)	(2.81)	(9.83)	(14.61)
TRV	0.050	0.942	0.012	0.525	0.684	0.536	0.691	0.371	0.618
	(1.84)	(26.99)	(0.75)	(5.50)	(12.18)	(5.74)	(13.47)	(11.15)	(18.47)
UNH	0.062	0.926	0.010	0.856	0.481	0.894	0.470	0.327	0.649
	(2.16)	(26.32)	(0.43)	(3.18)	(3.62)	(3.97)	(3.87)	(9.48)	(18.37)
UTX	0.091	0.903	0.000	0.796	0.539	0.796	0.539	0.343	0.640
	(3.51)	(46.15)	(.)	(4.48)	(6.63)	(4.48)	(6.63)	(11.64)	(22.20)
VZ	0.080	0.899	0.035	0.388	0.660	0.437	0.661	0.317	0.662
	(3.93)	(34.21)	(2.03)	(4.29)	(9.08)	(4.22)	(8.41)	(9.91)	(20.01)
WMT	0.040	0.952	0.008	0.526	0.562	0.508	0.587	0.300	0.677
	(3.25)	(61.63)	(0.30)	(1.00)	(1.21)	(1.24)	(1.74)	(6.46)	(13.54)
XOM	0.082	0.903	0.032	0.513	0.611	0.590	0.591	0.391	0.597
	(6.46)	(59.79)	(1.92)	(4.45)	(7.95)	(4.96)	(7.46)	(12.93)	(21.07)

Table B3: MCS for several loss functions of in-sample covariance forecasting

	FN	MCS Rank	QLIK	MCS Rank	GMV	MCS Rank	MV	MCS Rank
$s = 1$								
DCC-GARCH	28.467	6	24.518	6	0.417	6	0.679	8
DCC-HEAVY	**26.124**	1	**22.329**	1	**0.351**	1	**0.604**	1
DECO-GARCH	31.762	9	26.337	8	0.477	10	0.809	10
DECO-HEAVY	27.856	5	23.820	3	0.377	4	0.624	3
BEKK-GARCH	35.182	10	27.082	9	0.416	8	0.706	7
BEKK-HEAVY	31.553	8	24.940	7	0.365	3	0.633	5
ADCC-GARCH	27.659	3	24.462	5	0.414	7	0.676	6
ADCC-HEAVY	26.156	2	22.348	2	0.354	2	0.607	2
ADECO-GARCH	29.368	7	25.640	10	0.410	9	0.688	9
ADECO-HEAVY	27.903	4	23.819	4	0.377	5	0.624	4
$s = 5$								
DCC-GARCH	29.620	6	24.582	4	0.384	4	0.638	4
DCC-HEAVY	**28.620**	2	**23.340**	2	**0.340**	2	**0.590**	3
DECO-GARCH	29.599	5	24.587	6	0.384	5	0.638	6
DECO-HEAVY	31.417	8	25.613	5	0.558	9	0.855	10
BEKK-GARCH	35.880	9	26.386	8	0.384	7	0.658	5
BEKK-HEAVY	37.470	10	29.028	10	0.519	10	0.860	9
ADCC-GARCH	29.087	3	24.630	7	0.382	6	0.639	7
ADCC-HEAVY	28.007	1	23.239	1	0.338	1	0.586	1
ADECO-GARCH	30.550	7	25.776	9	0.380	8	0.649	8
ADECO-HEAVY	28.892	4	23.880	3	0.344	3	0.589	2
$s = 22$								
DCC-GARCH	**31.703**	4	**27.354**	3	0.432	5	**0.723**	3
DCC-HEAVY	**32.596**	7	28.404	7	**0.420**	4	0.729	6
DECO-GARCH	**31.625**	3	**27.371**	4	0.432	7	**0.723**	4
DECO-HEAVY	32.954	8	28.399	8	0.427	6	0.747	9
BEKK-GARCH	37.994	9	27.907	5	**0.430**	8	0.741	7
BEKK-HEAVY	42.490	10	33.417	10	0.671	10	1.047	10
ADCC-GARCH	31.656	5	27.759	6	0.430	9	0.736	8
ADCC-HEAVY	30.347	2	26.791	1	0.408	2	0.707	2
ADECO-GARCH	32.587	6	28.823	9	0.422	3	0.731	5
ADECO-HEAVY	28.975	1	27.078	2	0.401	1	0.698	1

Values of loss functions in bold identify the models in the 99% level MCS when the comparison is limited to the first six models. Underlined values identify the models in the 99% level MCS when the comparison is done for all models. A value in bold but not underlined is thus in the MCS of the first six models, but is excluded when considering all models. The MCS rankings are for the global comparison.

Table B4: MCS for loss functions of in-sample variance and correlation forecasting

	Correlation				Variance			
	FN	MCS Rank	QLIK	MCS Rank	FN	MCS Rank	QLIK	MCS Rank
$s = 1$								
DCC-GARCH	5.361	7	19.844	5	4.418	4	0.417	4
DCC-HEAVY	$\underline{\mathbf{4.960}}$	2	$\underline{\mathbf{19.570}}$	2	**4.252**	2	**0.344**	2
DECO-GARCH	5.880	8	29.244	8				
DECO-HEAVY	5.018	4	19.920	6				
ADCC-GARCH	5.307	6	19.800	3	4.365	3	0.406	3
ADCC-HEAVY	$\underline{4.960}$	1	$\underline{19.554}$	1	$\underline{4.136}$	1	$\underline{0.343}$	1
ADECO-GARCH	5.220	5	20.046	7				
ADECO-HEAVY	$\underline{5.013}$	3	19.915	4				
$s = 5$								
DCC-GARCH	5.388	7	20.039	7	4.652	4	0.458	4
DCC-HEAVY	$\underline{\mathbf{5.187}}$	1	$\underline{\mathbf{19.638}}$	1	**4.408**	2	**0.421**	2
DECO-GARCH	5.391	8	20.030	6				
DECO-HEAVY	**5.402**	6	26.023	8				
ADCC-GARCH	5.344	5	19.998	4	4.422	3	0.452	3
ADCC-HEAVY	$\underline{5.195}$	2	19.960	3	$\underline{4.358}$	1	$\underline{0.416}$	1
ADECO-GARCH	5.292	4	19.980	2				
ADECO-HEAVY	5.263	3	19.994	5				
$s = 5$								
DCC-GARCH	5.525	8	19.856	5	**4.509**	2	**0.548**	2
DCC-HEAVY	$\underline{\mathbf{5.419}}$	1	$\underline{\mathbf{19.809}}$	3	4.539	4	0.582	4
DECO-GARCH	**5.463**	4	**19.831**	4				
DECO-HEAVY	5.482	7	20.831	8				
ADCC-GARCH	5.469	6	20.181	7	4.500	3	0.555	3
ADCC-HEAVY	$\underline{5.426}$	2	$\underline{19.577}$	1	$\underline{4.361}$	1	$\underline{0.521}$	1
ADECO-GARCH	5.471	5	20.189	6				
ADECO-HEAVY	$\underline{5.451}$	3	$\underline{19.789}$	2				

Values of loss functions in bold identify the models in the 99% level MCS when the comparison is limited to the first six models. Underlined values identify the models in the 99% level MCS when the comparison is done for all models. A value in bold but not underlined is thus in the MCS of the first six models, but is excluded when considering all models. The MCS rankings are for the global comparison.

Table B5: Parameter estimates (and robust t-statistics) of the asymmetric variance equations

	A-HEAVY-h (35)			HAR-A-HEAVY-m (38)				
	A_h	B_h	Γ_h	A_m	B_m	Γ_m	A_m^w	A_m^m
AAPL	0.349	0.737	0.084	0.295	0.096	0.163	0.239	0.267
	(3.01)	(11.82)	(1.03)	(6.17)	(0.80)	(4.65)	(1.84)	(6.04)
AXP	0.558	0.628	0.244	0.335	0.361	0.152	0.023	0.177
	(0.49)	(1.08)	(1.09)	(10.83)	(3.99)	(5.19)	(0.36)	(5.33)
BA	0.625	0.515	0.191	0.269	0.152	0.118	0.247	0.203
	(3.13)	(3.98)	(1.86)	(7.02)	(1.41)	(3.81)	(3.06)	(3.32)
CAT	0.314	0.785	0.032	0.287	0.138	0.130	0.283	0.166
	(1.12)	(5.23)	(0.24)	(9.81)	(2.01)	(5.97)	(4.11)	(5.91)
CSCO	1.748	0.199	0.000	0.355	0.276	0.116	0.093	0.203
	(7.39)	(2.56)	(.)	(12.41)	(5.42)	(4.49)	(2.25)	(7.27)
CVX	0.359	0.682	0.177	0.316	0.143	0.122	0.316	0.121
	(3.47)	(10.73)	(2.68)	(12.32)	(2.24)	(6.43)	(4.96)	(5.44)
DIS	0.756	0.549	0.000	0.338	0.199	0.104	0.209	0.158
	(2.97)	(3.72)	(.)	(7.28)	(2.27)	(3.39)	(2.40)	(4.75)
DWDP	0.570	0.689	0.348	0.277	0.359	0.139	0.130	0.122
	(3.20)	(8.82)	(2.54)	(10.10)	(3.10)	(6.63)	(1.36)	(4.18)
GE	0.904	0.448	0.054	0.312	0.417	0.173	0.032	0.133
	(3.02)	(3.30)	(0.38)	(12.48)	(5.18)	(7.29)	(0.62)	(5.03)
GS	0.228	0.797	0.218	0.323	0.065	0.174	0.216	0.229
	(2.85)	(17.59)	(3.22)	(4.98)	(0.43)	(2.78)	(1.77)	(4.40)
HD	0.742	0.503	0.153	0.310	0.401	0.120	0.066	0.127
	(5.70)	(8.86)	(1.50)	(12.10)	(4.79)	(5.55)	(1.18)	(4.69)
IBM	1.180	0.251	0.172	0.275	0.305	0.142	0.191	0.101
	(3.72)	(1.55)	(0.95)	(8.88)	(2.95)	(5.40)	(1.94)	(4.00)
INTC	1.078	0.407	0.057	0.370	0.229	0.114	0.109	0.200
	(4.66)	(3.67)	(0.38)	(10.98)	(3.36)	(4.19)	(2.00)	(5.76)
JNJ	0.332	0.653	0.201	0.304	0.457	0.118	0.077	0.077
	(3.40)	(10.94)	(2.71)	(5.89)	(3.84)	(3.93)	(0.84)	(2.25)
JMP	0.481	0.694	0.170	0.398	0.249	0.119	0.105	0.154
	(1.88)	(5.46)	(2.04)	(13.57)	(4.46)	(5.08)	(2.05)	(6.70)
KO	0.630	0.447	0.193	0.298	0.402	0.116	0.080	0.132
	(4.17)	(4.87)	(1.74)	(12.23)	(5.32)	(5.84)	(1.45)	(4.75)
MCD	0.654	0.447	0.233	0.281	0.215	0.106	0.202	0.230
	(4.05)	(3.48)	(1.59)	(7.93)	(2.29)	(3.83)	(3.05)	(5.10)
MMM	0.803	0.466	0.000	0.325	0.358	0.195	0.103	0.105
	(2.81)	(2.53)	(.)	(10.20)	(7.17)	(3.00)	(2.60)	(4.30)

Continued in Table B6.

Table B6: Parameter estimates (and robust t-statistics) of the asymmetric variance equations

	A-HEAVY-h (35)			HAR-A-HEAVY-m (38)				
	A_h	B_h	Γ_h	A_m	B_m	Γ_m	A_m^w	A_m^m
MMRK	0.923	0.337	0.000	0.341	0.139	0.161	0.232	0.190
	(4.96)	(2.72)	(.)	(5.06)	(0.64)	(2.98)	(1.10)	(2.73)
MSFT	1.394	0.234	0.000	0.318	0.333	0.127	0.102	0.131
	(6.42)	(2.18)	(.)	(11.25)	(6.27)	(5.45)	(2.23)	(6.09)
NKE	1.079	0.355	0.055	0.326	0.380	0.177	0.015	0.150
	(4.30)	(2.95)	(0.31)	(10.55)	(5.58)	(6.40)	(0.35)	(6.35)
PFE	0.778	0.478	0.000	0.321	0.348	0.063	0.125	0.133
	(1.93)	(1.74)	(.)	(8.44)	(2.73)	(2.18)	(1.19)	(4.08)
PG	0.599	0.410	0.309	0.325	0.200	0.119	0.268	0.105
	(3.08)	(2.57)	(3.13)	(10.38)	(2.50)	(4.06)	(3.08)	(3.35)
TRV	0.430	0.712	0.130	0.351	0.280	0.105	0.141	0.153
	(3.48)	(12.70)	(1.45)	(12.99)	(4.11)	(4.64)	(2.56)	(5.52)
UNH	0.636	0.558	0.205	0.293	0.578	0.080	-0.022	0.083
	(2.14)	(3.82)	(1.40)	(8.81)	(2.92)	(3.05)	(-0.17)	(1.37)
UTX	0.419	0.640	0.444	0.247	0.488	0.173	0.065	0.089
	(2.58)	(7.77)	(3.27)	(10.11)	(6.48)	(7.08)	(1.13)	(3.98)
VZ	0.437	0.661	0.000	0.292	0.427	0.081	0.088	0.119
	(4.22)	(8.41)	(.)	(10.59)	(3.36)	(3.42)	(1.08)	(2.92)
WMT	0.481	0.569	0.101	0.265	0.061	0.138	0.360	0.187
	(0.99)	(1.25)	(0.56)	(7.16)	(0.37)	(2.88)	(2.98)	(2.90)
XOM	0.471	0.611	0.166	0.287	0.173	0.154	0.320	0.112
	(3.88)	(7.68)	(2.29)	(10.15)	(2.75)	(5.75)	(4.76)	(4.02)
Minimum	0.228	0.199	0.000	0.247	0.061	0.063	-0.022	0.077
Medium	0.625	0.549	0.153	0.312	0.280	0.122	0.125	0.133
Maximum	1.748	0.797	0.444	0.398	0.578	0.195	0.360	0.267

Minimum, Medium and Maximum are for all stocks of Tables B5 and B6.

C Empirical Results for NSS Data

NSS Data

The data as Noureldin, Shephard and Sheppard (2012), hereby NSS, are available from JAE data archive. NSS report they selected the ten most liquid stocks of the Dow Jones Industrial Average (DJIA) index. These are: Alcoa (AA), American Express (AXP), Bank of America (BAC), Coca Cola (KO), Du Pont (DD), General Electric (GE), International Business Machines (IBM), JP Morgan (JPM), Microsoft (MSFT), and Exxon Mobil (XOM). The sample period is 1 February 2001 to 31 December 2009 with a total of 2242 trading days.

Estimation Results

Tables C1 (variance equations) and C2 (correlation equations) show the full sample estimation results. The results are broadly similar to those reported in Section 5.2. As shown by the reported maximum and minimum values in Table C1 (compared to the values in Table 1), there is less heterogeneity between the parameter estimates of the variance equations for the 10 stocks of the NSS data than for the 29 stocks, which is not surprising given the smaller number of stocks and their selection criterion. Nevertheless, the median values of both sets of estimates are very similar.

For the correlation equations, a comparison of the estimates reported in Tables C2 and 2) does not reveal important differences. The higher t-statistics for the 29 stocks are due to the larger sample size.

Table C3 (comparable to Table 6 for the 29 stocks) shows the estimates of the asymmetric correlation equations, with HAR terms in the realized correlation part (HEAVY-M). The estimates of the coefficients of the asymmetric terms ((γ_r and γ_p) are similar, except for the ADECO-HEAVY-R, where the estimate of γ_r is positive and significant for the NSS dataset. For the HAR terms, the estimates for the ten stocks are larger in absolute value; notice that α_p^w and α_p^m are both estimated to be significantly negative in the HAR-ADCC-HEAVY-M.

Table C1: Parameter estimates (and robust t-statistics) of the variance equations

	GARCH (26)-(28)		GARCHX (26)-(29)			HEAVY-h (10)		HEAVY-m (20)	
	A_h	B_h	A_h	A_{hm}	B_h	A_h	B_h	A_m	B_m
BAC	0.067	0.927	0.003	0.644	0.604	0.641	0.614	0.474	0.526
	(2.53)	(32.70)	(0.27)	(7.11)	(3.20)	(7.59)	(3.73)	(9.86)	(10.96)
JPM	0.086	0.914	0.007	0.767	0.347	0.763	0.363	0.501	0.499
	(4.89)	(52.30)	(0.47)	(10.68)	(2.82)	(10.83)	(3.19)	(11.30)	(11.27)
IBM	0.106	0.879	0.034	0.513	0.699	0.499	0.782	0.352	0.634
	(1.09)	(8.02)	(1.40)	(3.37)	(2.67)	(3.11)	(2.95)	(9.20)	(16.10)
DD	0.061	0.931	0.005	0.655	0.453	0.652	0.463	0.366	0.618
	(2.98)	(41.60)	(0.22)	(7.99)	(3.34)	(8.65)	(4.22)	(10.71)	(17.39)
XOM	0.082	0.895	0.008	0.695	0.339	0.691	0.354	0.378	0.603
	(5.53)	(47.60)	(0.45)	(10.96)	(4.05)	(11.31)	(4.87)	(11.35)	(17.50)
AA	0.045	0.948	0	0.868	0.197	0.868	0.197	0.314	0.669
	(4.22)	(72.92)	(.)	(4.91)	(0.76)	(4.91)	(0.76)	(9.47)	(19.41)
AXP	0.095	0.905	0.001	0.727	0.442	0.726	0.444	0.397	0.603
	(6.82)	(64.76)	(0.13)	(13.30)	(4.57)	(13.66)	(4.81)	(10.65)	(16.20)
DD	0.061	0.931	0.005	0.655	0.453	0.652	0.463	0.366	0.618
	(2.98)	(41.60)	(0.22)	(7.99)	(3.34)	(8.65)	(4.22)	(10.71)	(17.39)
GE	0.054	0.945	0.011	0.478	0.837	0.509	0.799	0.369	0.631
	(3.29)	(57.56)	(0.12)	(1.16)	(1.31)	(23.81)	(.)	(11.38)	(19.45)
KO	0.105	0.888	0.057	0.556	0.515	0.491	0.659	0.356	0.631
	(1.70)	(14.20)	(1.69)	(4.78)	(2.83)	(4.47)	(4.36)	(10.61)	(18.58)
Minimum	0.045	0.879	0.001	0.478	0.197	0.491	0.197	0.314	0.499
Median	0.075	0.921	0.007	0.655	0.453	0.652	0.463	0.368	0.618
Maximum	0.106	0.948	0.057	0.868	0.837	0.868	0.799	0.501	0.669

Table C2: Parameter estimates (and robust t-statistics) of the correlation equations

DCC-GARCH (17)		DCCX-GARCH (30)-(31)			DCC-HEAVY-R (11)-(12)		DCC-HEAVY-P (21)	
α_q	β_q	α_q	β_q	α_r	α_r	β_r	α_m	β_m
0.010	0.974	0.012	0.795	0.106	0.131	0.761	0.077	0.908
(4.58)	(137.1)	(1.33)	(20.3)	(5.04)	(6.24)	(17.16)	(22.45)	(195.8)

Log-likelihood decomposition (DCC-HEAVY vs DCC-GARCH(

	DCC-GARCH	DCC-HEAVY	DCC-HEAVY Gains
Variance	-4316	-4249	67
Correlation	-2714	-2705	9
Total	-7030	-6955	76

DECO-GARCH		DECOX-GARCH			DECO-HEAVY-R		DECO-HEAVY-P	
α_q	β_q	α_q	β_q	α_r	α_r	β_r	α_m	β_m
0.042	0.936	0.022	0.747	0.193	0.247	0.707	0.279	0.684
(2.87)	(41.48)	(1.14)	(10.5)	(2.52)	(4.10)	(10.1)	(11.11)	(21.74)

Log-likelihood decomposition (DECO-HEAVY vs DECO-GARCH)

	DECO-GARCH	DECO-HEAVY	HEAVY Gains
Variance	-4316	-4249	67
Correlation	-2775	-2772	3
Total	-7091	-7021	70

Table C3: Parameter estimates (and robust t-statistics) of the asymmetric correlation equations

ADCC-HEAVY-R (36)-(37)			HAR-ADCC-HEAVY-M (39)				
α_r	β_r	γ_r	α_p	β_p	γ_p	α_p^w	α_p^m
0.118	0.759	0.028	0.102	0.955	0.012	-0.055	-0.014
(5.68)	(17.51)	(1.73)	(24.90)	(138.2)	(6.27)	(10.47)	(3.25)
ADECO-HEAVY-R			**HAR-ADECO-HEAVY-P**				
α_r	β_r	γ_r	α_p	β_p	γ_p	α_p^w	α_p^m
0.127	0.795	0.089	0.194	0.439	0.111	0.122	0.134
(1.92)	(15.71)	(2.09)	(7.69)	(5.91)	(8.69)	(2.08)	(4.06)

Forecasting Comparisons

For out-of-sample forecast results, the models are estimated on rolling windows of 1486 observations, resulting in a total of 756 forecasts for $s = 1$, 752 for $s = 5$, and 735 for $s = 22$. The first rolling window is chosen such that forecasts start on 3 January 2007. The last forecast is for 31 December 2009.

Tables C4 to C8 present the results in the same way as the Tables 3-5 and 7-8 for the 29 stocks. A few comments follow.

1) MCS99 for covariance forecasts (Table C4 compared to Table 3): the main difference is the inclusion in the MCS99 of the BEKK-HEAVY model at the three horizons and for all loss functions. A possible explanation for this is the difference in the number of stocks between the two datasets: the flexibility of DCC with respect to BEKK is more important for 29 stocks than for 10 (the most liquid ones).

Another difference is that the MCS99 include more models than for the 29 stocks. Nevertheless, the results are clearly in favour of HEAVY models: GARCH models are included in MCS99 only at horizon 22 and for the QLIK and MV loss functions. This is reflected in the Covariance panels of Tables C6-C8 where only three ratios (for QLIK loss) are marginally larger than one, whereas the other ratios are all below one; the improvements are important especially for FN loss.

2) MCS99 for correlation forecasts (Table C5 compared to Table 4, Correlation panels): for FN loss, few differences; for QLIK loss, some GARCH models (depending on the horizon) belong to the MCS99, in addition to several HEAVY models. The ratios in the Correlation panels of Tables C6-C8 are in a majority of cases (27 out of 36) smaller than one, but not much (the smallest being 0.914); the 10 ratios that exceed one are very close to one (the largest being 1.031). In brief, even if the HEAVY models do improve correlation forecasts, this is not strongly and some DCC and DECO models do as well.

3) MCS99 for variance forecasts (Table C5 compared to Table 4, Variance panels): When comparing all models, DCC-HEAVY and ADCC-HEAVY are in the MCS99 for both loss functions and all horizons (with one exception for DCC-HEAVY). Their values of their loss functions are close. Loss ratios show that the GARCH models (DCC and ADCC) are dominated by the HEAVY models, even if DCC is in the three MCS99 for MSE loss. The advantage of HEAVY over GARCH models decreases as s increases.

Table C4: MCS for loss functions of out-of-sample covariance forecasting

	FN	MCS Rank	QLIK	MCS Rank	GMV	MCS Rank	MV	MCS Rank
$s = 1$								
DCC-GARCH	48.082	7	18.130	8	1.384	7	2.444	7
DCC-HEAVY	**34.224**	5	**16.940**	1	**1.277**	2	**2.259**	2
DECO-GARCH	46.206	6	18.488	10	1.422	8	2.425	9
DECO-HEAVY	**34.326**	3	17.252	4	1.326	6	2.340	5
BEKK-GARCH	53.422	10	19.605	7	1.634	10	2.712	10
BEKK-HEAVY	**26.467**	1	**17.109**	3	**1.241**	1	**2.138**	1
ADCC-GARCH	48.169	9	17.953	6	1.327	4	2.368	4
ADCC-HEAVY	34.237	4	16.996	2	1.284	3	2.267	3
ADECO-GARCH	47.122	8	18.268	9	1.353	9	2.356	6
ADECO-HEAVY	34.575	2	17.296	5	1.355	5	2.366	8
$s = 5$								
DCC-GARCH	50.778	9	18.736	9	1.438	8	2.557	8
DCC-HEAVY	**37.842**	4	17.808	2	**1.320**	4	2.395	5
DECO-GARCH	48.775	6	18.978	10	1.454	9	2.501	9
DECO-HEAVY	**38.700**	5	**17.714**	1	1.329	5	**2.361**	4
BEKK-GARCH	55.234	8	20.164	8	1.667	10	2.807	10
BEKK-HEAVY	**30.250**	1	**18.130**	5	**1.310**	3	**2.297**	1
ADCC-GARCH	51.362	10	18.559	6	1.387	6	2.493	7
ADCC-HEAVY	35.037	3	17.929	4	1.307	2	2.330	3
ADECO-GARCH	50.057	7	18.769	7	1.397	7	2.443	6
ADECO-HEAVY	34.935	2	17.940	3	1.292	1	2.300	2
$s = 22$								
DCC-GARCH	55.342	7	**20.417**	6	1.508	9	2.593	9
DCC-HEAVY	**42.148**	4	**20.731**	8	**1.442**	4	**2.439**	5
DECO-GARCH	53.727	6	**20.476**	5	1.507	8	2.527	8
DECO-HEAVY	**42.352**	5	**20.696**	7	1.442	5	**2.445**	6
BEKK-GARCH	59.570	8	21.617	10	1.695	10	2.799	10
BEKK-HEAVY	**36.508**	1	**20.850**	9	**1.449**	3	**2.398**	3
ADCC-GARCH	57.837	10	20.190	1	1.465	6	2.518	7
ADCC-HEAVY	38.630	2	20.280	3	1.409	2	2.361	1
ADECO-GARCH	56.670	9	20.316	4	1.464	7	2.451	4
ADECO-HEAVY	39.764	3	20.242	2	1.408	1	2.378	2

Values of loss functions in bold identify the models in the 99% level MCS when the comparison is limited to the first six models. Underlined values identify the models in the 99% level MCS when the comparison is done for all models. A value in bold but not underlined is thus in the MCS of the first six models, but is excluded when considering all models. The MCS rankings are for the global comparison.

Table C5: MCS for loss functions of out-of-sample correlation and variance forecasting

	Correlation				Variance			
	FN	MCS Rank	QLIK	MCS Rank	MSE	MCS Rank	UQLIK	MCS Rank
$s = 1$								
DCC-GARCH	1.637	4	**6.709**	2	16.623	3	0.325	4
DCC-HEAVY	**1.511**	2	**6.742**	3	**12.299**	1	**0.214**	1
DECO-GARCH	1.733	7	6.927	7				
DECO-HEAVY	1.687	5	**6.905**	5				
ADCC-GARCH	1.627	3	6.703	1	17.371	4	0.303	3
ADCC-HEAVY	1.505	1	6.762	4	12.324	2	0.214	2
ADECO-GARCH	1.753	8	6.954	8				
ADECO-HEAVY	1.686	6	6.916	6				
$s = 5$								
DCC-GARCH	1.657	6	**6.709**	6	**17.586**	3	0.380	4
DCC-HEAVY	**1.582**	1	**6.706**	5	**13.991**	2	**0.288**	1
DECO-GARCH	1.730	7	**6.818**	7				
DECO-HEAVY	**1.589**	2	**6.712**	3				
ADCC-GARCH	1.649	5	6.701	4	18.698	4	0.359	3
ADCC-HEAVY	1.609	4	6.631	2	13.406	1	0.289	2
ADECO-GARCH	1.745	8	6.824	8				
ADECO-HEAVY	1.595	3	6.623	1				
$s = 22$								
DCC-GARCH	**1.681**	5	**6.641**	5	**19.128**	3	**0.536**	4
DCC-HEAVY	**1.660**	1	**6.633**	3	**15.643**	2	**0.529**	3
DECO-GARCH	1.726	8	**6.636**	4				
DECO-HEAVY	**1.664**	2	6.680	7				
ADCC-GARCH	1.676	4	6.631	2	21.172	4	0.512	2
ADCC-HEAVY	1.729	6	6.690	8	14.646	1	0.502	1
ADECO-GARCH	1.746	7	6.704	6				
ADECO-HEAVY	1.683	3	6.628	1				

Values of loss functions in bold identify the models in the 99% level MCS when the comparison is limited to the first six models. Underlined values identify the models in the 99% level MCS when the comparison is done for all models. A value in bold but not underlined is thus in the MCS of the first six models, but is excluded when considering all models. The MCS rankings are for the global comparison.

Table C6: Out-of-sample loss ratios between DCC-HEAVY and DCC-GARCH

		DCC-HEAVY vs DCC-GARCH		DECO-HEAVY vs DECO-GARCH	
		FN	QLIK	FN	QLIK
Covariance	s=1	0.712	0.934	0.743	0.933
	s=5	0.745	0.950	0.793	0.933
	s=22	0.762	1.015	0.788	1.011
Correlation	s=1	0.923	1.005	0.974	0.997
	s=5	0.955	1.000	0.919	0.984
	s=22	0.988	0.999	0.964	1.007
		MSE	UQLIK		
Variance	s=1	0.740	0.811		
	s=5	0.796	0.871		
	s=22	0.818	0.993		

For the Variance panel, results of DCC and DECO are identical.

Table C7: Out-of-sample loss ratios between ADCC-HEAVY and DCC-GARCH

		ADCC-HEAVY vs DCC-GARCH		ADECO-HEAVY vs DECO-GARCH	
		FN	QLIK	FN	QLIK
Covariance	s=1	0.712	0.937	0.748	0.936
	s=5	0.690	0.957	0.716	0.945
	s=22	0.698	0.993	0.740	0.989
Correlation	s=1	0.919	1.008	0.973	0.998
	s=5	0.971	0.988	0.922	0.971
	s=22	1.029	1.007	0.975	0.999
		MSE	UQLIK		
Variance	s=1	0.741	0.811		
	s=5	0.762	0.873		
	s=22	0.766	0.968		

For the Variance panel, results of DCC and DECO are identical.

Table C8: Out-of-sample loss ratios between ADCC-HEAVY and ADCC-GARCH

		ADCC-HEAVY vs ADCC-GARCH		ADECO-HEAVY vs ADECO-GARCH	
		FN	QLIK	FN	QLIK
Covariance	s=1	0.711	0.947	0.734	0.947
	s=5	0.682	0.966	0.698	0.956
	s=22	0.668	1.004	0.702	0.996
Correlation	s=1	0.925	1.009	0.962	0.995
	s=5	0.976	0.990	0.914	0.970
	s=22	1.031	1.009	0.964	0.989
		MSE	UQLIK		
Variance	s=1	0.709	0.841		
	s=5	0.717	0.897		
	s=22	0.692	0.990		

For the Variance panel, results of DCC and DECO are identical.