

Forecasting Value-at-Risk under Temporal and Portfolio Aggregation

Erik Kole^{1,4}

*Thijs Markwat*²

Anne Opschoor^{3,4}

Dick van Dijk^{1,4,5}

REVISION APRIL 2017

¹ *Econometric Institute, Erasmus University Rotterdam*

² *Robeco Asset Management*

³ *Vrije Universiteit Amsterdam*

⁴ *Tinbergen Institute*

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

Contact: discussionpapers@tinbergen.nl

More TI discussion papers can be downloaded at <http://www.tinbergen.nl>

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 598 4580

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900

Forecasting Value-at-Risk under Temporal and Portfolio Aggregation*

Erik Kole^{†1,4}, Thijs Markwat², Anne Opschoor^{3,4}, and Dick van Dijk^{1,4,5}

¹*Econometric Institute, Erasmus University Rotterdam*

²*Robeco Asset Management*

³*VU University Amsterdam*

⁴*Tinbergen Institute*

⁵*Erasmus Research Institute of Management*

April 19, 2017

Abstract

We examine the impact of temporal and portfolio aggregation on the quality of Value-at-Risk (VaR) forecasts over a horizon of ten trading days for a well-diversified portfolio of stocks, bonds and alternative investments. The VaR forecasts are constructed based on daily, weekly or biweekly returns of all constituent assets separately, gathered into portfolios based on asset class, or into a single portfolio. We compare the impact of aggregation to that of choosing a model for the conditional volatilities and correlations, the distribution for the innovations and the method of forecast construction. We find that the level of temporal aggregation is most important. Daily returns form the best basis for VaR forecasts. Modeling the portfolio at the asset or asset class level works better than complete portfolio aggregation, but differences are smaller. The differences from the model, distribution and forecast choices are also smaller compared to temporal aggregation.

Key words: forecast evaluation, aggregation, Value-at-Risk, model comparison

JEL classification: C22, C32, C52, C53, G17

*The authors thank Fabio Trojani (the editor), the associate editor and two anonymous referees, as well as participants at the 9th Annual SoFiE Conference in Hong Kong 2016, the 3rd Annual IAAE Conference in Milan 2016 and the Financial Econometrics and Empirical Asset Pricing Conference at the University of Lancaster 2016 for many helpful comments and suggestions.

[†]Corresponding Author: Erasmus University Rotterdam, P.O. Box 1738, NL-3000 DR Rotterdam, The Netherlands. E-mail addresses are kole@ese.eur.nl, t.markwat@robeco.nl, a.opschoor@vu.nl, and djvandijk@ese.eur.nl.

1 Introduction

Value-at-Risk (VaR) for a horizon of ten trading days has become the standard downside risk measure for investment portfolios since the Basel Committee on Banking Supervision (1996), and is widely used in the financial sector (see Berkowitz and O'Brien, 2002). Because prices can be observed at a higher frequency and for all assets in the portfolio, a risk manager has to choose the degree of aggregation of the observations when constructing the VaR forecast. The typical approach in practice first gathers the assets into asset classes like stocks, bonds, and alternative assets. Next, a model for the daily asset class returns is used to construct a one-step VaR forecast for the portfolio. This forecast is then scaled to the required horizon by the square-root-of-time rule. This scaling of the daily VaR forecast is explicitly advised in the Basel Committee on Banking Supervision (1996) (see also Diebold et al., 1997; Daníelsson and Zigrand, 2006). However, many alternatives with a lower or higher degree of temporal and portfolio aggregation are available.

We investigate the importance of the aggregation choice for the quality of ten-day VaR forecasts. This VaR forecast is a quantile of the portfolio return distribution over the coming ten days. We focus on the 1% and 5% quantiles, so VaR forecasts with 99% and 95% confidence levels. To construct this distribution, a risk manager can use different degrees of both temporal and portfolio aggregation. We compare their importance to other required choices. In particular, we consider different models for the conditional volatilities and correlations of the asset returns, different distributions for the innovations in these models, and different methods to construct cumulative multi-step forecasts (see also McAleer, 2009).

The degree of aggregation determines how detailed the serial and cross-sectional dependence of the returns is modeled. A model for all assets at the highest observed frequency describes very precisely how shocks to one particular asset affect its own future return distribution at different horizons, as well as that of other assets. However, such an extensive model can be tedious to use in practice. Aggregation of assets into portfolios reduces the model dimension, whereas temporal aggregation reduces the number of forecasts that have to be accumulated. It also reduce the complexity of the model, if serial dependence only has short-term consequences.

Consequently, these choices involve a trade-off between precision and efficiency on the one hand and complexity and the risk of misspecification on the other hand. A low degree of aggregation particularly improves precision when different aspects of the return distributions

vary. For instance, when the effect of shocks on the volatility is more persistent in one (group of) asset(s), temporally and cross-sectionally aggregated models will lead to imprecise forecasts. Less temporal aggregation also increases the efficiency, since more observations are available (keeping the estimation window fixed). The flip side of this choice is the risk of misspecification. More complex models are generally less robust to misspecification. The effects of errors in the specification can be amplified by the forecasting horizon. Small errors in a single-period forecast can build up to a large error in a multi-step forecast.¹

We study the impact of the degree of aggregation on ten-day VaR forecasts for a well-diversified portfolio of eight indexes, related to equities, bonds, and alternative assets. We vary the degree of temporal aggregation between modeling daily, weekly and biweekly returns, so from no to full temporal aggregation. In the cross-sectional dimension we analyze aggregation into a single portfolio, into the three main asset classes (stocks, bonds and alternatives), and no aggregation, so modeling the returns on all eight portfolio constituents. We examine all possible combinations of these different levels of temporal and portfolio aggregation. We do not examine the use of intraday data, because they are generally used in models for daily data (see, for example, Giot and Laurent, 2004; Clements et al., 2008; Brownlees and Gallo, 2010). Instead, we focus on methods that use one specific frequency. We leave a comparison of models that use mixed data frequencies with or without intraday data (see also Ghysels et al., 2009) for future research.

We compare the impact of aggregation to that of three other important required choices: the model for the conditional volatilities and correlations, the distribution for the innovations and the method of forecast construction. Regarding the first choice, we examine six well-established, relatively simple models from the GARCH family. We combine the univariate GARCH models of Bollerslev (1986) and Glosten et al. (1993) (GJR) with the CCC model of Bollerslev (1990), the DCC model of Engle (2002) and the rank-1 HDCC model from Bauwens et al. (2016) for the correlations. Hansen and Lunde (2005) and Laurent et al. (2012) show that these models perform just as well as more extensive alternatives (see, for example, Kuester et al., 2006; Brownlees and Gallo, 2010). The RiskMetrics approach from JP Morgan and Reuters (1996), which is popular in practice, completes the set of models. These seven models allow us to determine the importance of asymmetry in the marginal models and of dynamics in the correlations. The RiskMetrics approach offers a one-size-

¹See the surveys by Bhansali (1999) and Chevillon (2007) for a general discussion of iterated versus direct multi-step forecasts.

fits-all alternative. As distributions for the innovations in these models we use a copula approach based on the normal, the Student’s t and the empirical distribution. While the failure of the normal distribution for risk measurement has been widely documented since Mandelbrot (1963) and Fama (1965), it remains popular in practice. The other two are promising alternatives.

The method of forecast construction is relevant when returns are modeled at a higher frequency than the forecast horizon. Cumulative multi-step forecasts can be constructed by an iterative procedure or by scaling one-step forecasts.² In the iterative procedure, h -step forecasts are constructed by iterating forward the one-period-ahead density for h periods.³ This procedure captures the path dependence that the models imply, but can be time-consuming. A faster though perhaps less accurate alternative is the scaling of the one-period risk forecast. Scaling by the square-root-of-time is the industry standard, even though it can lead to overestimation of the ten-day VaR (see Diebold et al., 1997) as well as underestimation (see Daniélsson and Zigrand, 2006; Wang et al., 2011).

The combination of these five choice aspects leads to 255 methods to forecast the VaR. We use the tests proposed by Christoffersen (1998) and Engle and Manganelli (2004) to assess the accuracy of each method in isolation. We examine whether a particular method outperforms another by constructing Model Confidence Sets as proposed by Hansen et al. (2011) based on the asymmetric tick-loss function as in Giacomini and Komunjer (2005). We also conduct pairwise comparisons based on the test of Diebold and Mariano (1995) in the framework of Giacomini and White (2006). These horse races are based on more than five thousand ten-day VaR forecasts for the period 1994–2014. We use the industry standard of a VaR confidence level of 99%, and use 95% as a robustness check.

Our results show that the degree of temporal aggregation is most important, whereas the degree of portfolio aggregation is less consequential. Working with daily returns leads to 99%- and 95%-VaR forecasts that come closest to having correct coverage, and perform significantly better than forecasts based on weekly or biweekly returns. While we also find evidence supporting more detailed multivariate models, in particular for 95%-VaR forecasts, differences here are smaller and in many cases not significant. Our results point at models based on the three main asset classes for 99%-VaR forecasts, and models on the asset level

²In specific cases, closed-form expressions for the multi-step distributions can be used. See for example Drost and Nijman (1993) for the GARCH(1,1) model, Hafner (2008) for multivariate GARCH models, and Sbrana and Silvestrini (2013) for the RiskMetrics method.

³See Marcellino et al. (2006) for iterated forecasts in an autoregressive framework and Ghysels et al. (2009) for iterated forecasts in a GARCH framework.

for the 95% case. We conclude that the methods with lower degrees of aggregation better capture the propagation of shocks, both over time and between the different asset returns.

When forecasts are constructed by iterating based on daily data, the GJR model of Glosten et al. (1993) gives the best forecasts, whereas the distribution choice is of less consequence. When forecasts are constructed from weekly or biweekly data, the distribution choice is important for 99%-VaR forecasts, but not for 95% forecasts. In the 99% case, the empirical (Student's t) distribution performs best for weekly (biweekly) data. So, at the (bi)weekly frequency the innovations capture the fat tails and asymmetry, whereas the models incorporate these features at the daily frequency. The method of forecast construction is not that important. Our evidence tends to favor iterating over scaling to form forecasts, but differences vary widely based on the other settings. When combined with a normal distribution, scaling leads to worse performance than iterating. The same holds for combinations with a GJR model. Modeling time-variation in correlation hardly has any consequences at all.

We contribute to several strands of the literature. Most closely related are the papers that compare different methods for VaR forecasting. We show the importance of the degree of temporal aggregation in the realistic risk management setting of a ten-day VaR forecast, whereas the literature so far only evaluates one-day VaR forecasts based on either daily data or intraday data.⁴ Our results for portfolio aggregation complement McAleer and Da Veiga (2008) and Santos et al. (2013). McAleer and Da Veiga (2008) conclude that VaR forecasts for a small equity-only portfolio are best when full aggregation is applied. We also find that aggregation of the equity part into a single portfolio works well, but that further aggregation of the different asset classes does not improve performance. In line with our findings, Santos et al. (2013) report that multivariate models perform better than univariate models, but they do not consider intermediate degrees of aggregation. The importance of the distribution choice is in line with Giot and Laurent (2004); Kuuster et al. (2006); Bao et al. (2006); Clements et al. (2008); McAleer and Da Veiga (2008).

Second, we add to the literature that examines multi-period forecasts of variances and covariances. Contrary to the papers that study VaR forecasts, the papers in this area do investigate temporal aggregation, but they generally evaluate the forecasts by a loss function

⁴Giacomini and Komunjer (2005); Bao et al. (2006); Kuuster et al. (2006); McAleer and Da Veiga (2008); Santos et al. (2013) use daily data, whereas Giot and Laurent (2004); Brownlees and Gallo (2010); Clements et al. (2008) also use intraday data.

that uses a volatility proxy.⁵ Though Patton (2011) and Laurent et al. (2013) consider general classes of loss functions, our use of the asymmetric tick-loss function adds a new perspective, in particular on the interplay of the volatility forecast and the conditional return distribution.⁶ Also when tail loss is important, iterated forecasts are preferred, followed by scaled and then direct forecasts as in Ghysels et al. (2009). In line with Hansen and Lunde (2005); Brownlees et al. (2011); Laurent et al. (2012) we find that leverage effects in GARCH models are important. Contrary to Laurent et al. (2012), we do not find a preference for DCC over CCC. When tail loss is evaluated, the distribution constitutes an important choice opposite to Brownlees et al. (2011) who evaluate forecasts based on a symmetric loss function.

Third, our research is related to the papers in the large field of economic forecasting that compare iterated with direct forecasts, and forecasts of aggregates with aggregating forecasts. Theoretically, iterated forecasts are more efficient if the model is correctly specified, whereas direct forecasts are more robust to misspecification (see Chevillon, 2007). Empirically, Marcellino et al. (2006) report that iterated forecasts outperform direct ones in an analysis of 170 macro time series, because the models for the higher observation frequency can include more lags. We do not include more lags in models for daily data, but in our case the more precise identification and propagation of shocks in the daily models can explain their outperformance. The same explanation applies to the preference of multivariate models over univariate models that we find. In a macroeconomic setting, Marcellino et al. (2003) conclude that the aggregation of forecasts works better than forecasting aggregates. In a follow-up of Marcellino et al. (2006), Pesaran et al. (2011) show that multivariate models can also improve the forecasts for single non-aggregated variables.

From a practical perspective, our results show that it is best to work with daily data, and to model on the level of either the assets or the asset classes. A GJR-GARCH model with iterated forecasts further improves performance. For 99%-VaR forecasts in particular, it is best to use the empirical distribution function. However, our results also indicate that the RiskMetrics approach based on scaling forecasts constructed from daily data in combination with the empirical distribution is not so bad after all.

The article proceeds as follows. In Section 2 we discuss the different choice aspects in more detail, and present the design of our research. In Section 3 we present the data. We present the results in Section 4 and conclude in Section 5. Appendices provide additional

⁵See the surveys by Andersen et al. (2006); Patton and Sheppard (2009).

⁶See Elliott and Timmermann (2008) for a general discussion about the loss function in relation to forecasting.

information on the methods and results.

2 Methodology

We take the perspective of a risk manager responsible for monitoring the risk of a portfolio of n assets over the coming h time periods, during which the portfolio composition does not change. She observes the asset prices at discrete points in time $t = 1, 2, \dots$. We can approximate the log portfolio return from t to $t + h$ by

$$r_{t,h}^{\text{P}} = \sum_{\tau=1}^h \mathbf{w}' \mathbf{r}_{t+\tau}, \quad (1)$$

where \mathbf{w} and \mathbf{r}_t are $n \times 1$ vectors that contain the portfolio weights and the one-period log asset returns. We concentrate on a fully invested long-only portfolio, so we assume $w_i \geq 0$, $i = 1, 2, \dots, n$ and $\sum_{i=1}^n w_i = 1$, though these assumptions are not crucial for our analysis. The approximation holds reasonably well when the variance of the portfolio return is not too large, which holds for our case of a ten-day horizon and long-only portfolios.

The risk of the portfolio is measured in terms of Value-at-Risk. The h -period VaR with confidence level $1 - \vartheta$ of the portfolio return $r_{t,h}^{\text{P}}$ is defined as

$$\text{VaR}_{\vartheta}(r_{t,h}^{\text{P}}) = -\sup\{l : \Pr[r_{t,h}^{\text{P}} \leq l | \mathcal{F}_t] \leq \vartheta\}, \quad (2)$$

where \mathcal{F}_t denotes the information set at time t . In line with the Basel accords, we set the confidence level equal to 99% ($\vartheta = 0.01$) and the horizon equal to ten days. Additionally, we consider $\vartheta = 0.05$.

Value-at-Risk is the negative of the ϑ -quantile of the conditional distribution of $r_{t,h}^{\text{P}}$. To forecast this distribution, the risk manager has to make a number of choices. We investigate the consequences of five different choice aspects. The first two choices concern the degree of temporal and portfolio aggregation. Together, they establish the basic return variable $\mathbf{r}_{t,k}^{\text{b}}$ that will be modeled. Third, she has to decide on the time series model to describe the return dynamics. Fourth, she has to choose a distribution for (the innovation in) $\mathbf{r}_{t,k}^{\text{b}}$. Fifth, she has to determine how to construct the distribution of $r_{t,h}^{\text{P}}$ from the model for $\mathbf{r}_{t,k}^{\text{b}}$. We discuss these five aspects in turn below.

2.1 Choosing the degree of aggregation

The first choice determines the degree of aggregation of the nh random variables that constitute $r_{t,h}^p$. Aggregation can be done both temporally and cross-sectionally. Temporal aggregation means that k -period returns form the basis of the model instead of single-period returns. Cross-sectional aggregation means a reduction of the portfolio dimension by grouping assets together in m basic portfolios. So instead of the single-period n -dimensional returns \mathbf{r}_t , the k -period returns $\mathbf{r}_{t,k}^b$ with dimension $m_b \leq n$ form the basis of the model, where b indicates the particular set of basic portfolios.

To construct the basic portfolios we define an $m_b \times n$ selection matrix \mathbf{S}^b , with elements $s_{ij}^b = 1$ if asset j is part of portfolio i and zero otherwise. Row i indicates which assets constitute portfolio i . Since each asset has to be part of exactly one portfolio, all columns of \mathbf{S}^b must sum to one. The $m_b \times 1$ vector \mathbf{w}^b contains the weights of the m_b basic portfolios and satisfies $\mathbf{w}^b = \mathbf{S}^b \mathbf{w}$. The returns $\mathbf{r}_{t,k}^b$ on the basic portfolios b over horizon k are constructed as the product of the weight matrix $\widetilde{\mathbf{W}}^b$ and the asset returns summed over k periods:

$$\mathbf{r}_{t,k}^b = \widetilde{\mathbf{W}}^b \sum_{\tau=1}^k \mathbf{r}_{t+\tau} \quad \text{with} \quad \tilde{w}_{ij}^b = s_{ij}^b w_j / \sum_{\kappa=1}^n s_{i\kappa}^b w_{\kappa}. \quad (3)$$

Row i of $\widetilde{\mathbf{W}}^b$ contains the weights of the assets in basic portfolio i , scaled so they sum to one. Assuming that k is a divisor of h , we replace Equation (1) by

$$r_{t,h}^p = \sum_{\tau=1}^{h/k} \mathbf{w}^{b'} \mathbf{r}_{t+(\tau-1)k,k}^b. \quad (4)$$

When $k = 1$ and $m_b = n$ the risk manager aggregates in neither dimension, and models the asset returns in the most detailed way. When $k = h$ she fully aggregates over time, and constructs a one-step ahead forecast of the return distribution. When $m_b = 1$, the cross-section is fully aggregated, and she models the portfolio return as a univariate random variable. Other choices of k and m_b reduce the scale of the model, but the distribution of $r_{t,h}^p$ remains the result of a sum of multivariate return distributions.

For temporal aggregation we investigate the cases of no ($k = 1$), weekly ($k = 5$) and full (biweekly) aggregation ($k = h = 10$). For cross-sectional aggregation we analyze the case of no ($m_b = n$), full ($m_b = 1$), and aggregation by asset class.

2.2 Choosing the time-series model

The financial econometrics literature has established that the conditional distribution of asset returns is not constant over time. For the time-varying mean of returns (V)ARMA models have been proposed. The family of GARCH models has become the standard for modeling the volatilities and correlations of returns. Because the number of models has grown very large, we restrict our attention to a couple of relatively simple models that have been found to perform well.⁷ We specify all models in a multivariate setting, and only discuss the univariate version if necessary.

We model the mean of the return distribution by a VAR(p) model

$$\mathbf{r}_{t,k}^b = \boldsymbol{\phi}_k^b + \sum_{i=1}^p \boldsymbol{\Gamma}_{i,k}^b \mathbf{r}_{t-ik,k}^b + \boldsymbol{\varepsilon}_{t,k}^b, \quad \mathbb{E}_t[\boldsymbol{\varepsilon}_{t,k}^b] = \mathbf{0}, \quad \mathbb{E}_t[\boldsymbol{\varepsilon}_{t,k}^b \boldsymbol{\varepsilon}_{t,k}^{b'}] = \boldsymbol{\Sigma}_{t,k}^b \quad (5)$$

where $\boldsymbol{\phi}_k^b$ is an $m_b \times 1$ vector, $\boldsymbol{\Gamma}_{i,k}^b$, $i = 1, \dots, p$ are $m_b \times m_b$ matrices, and $\boldsymbol{\varepsilon}_{t,k}^b$ is an $m_b \times 1$ vector with innovations over the period from t to $t+k$.⁸ Conditional on the information available at time t , the innovations have mean zero and variance $\boldsymbol{\Sigma}_{t,k}^b$. When no temporal aggregation takes place and returns are daily, we choose $p = 1$. Assets may be traded in different time zones, and a VAR(1) model can accommodate the resulting lagged cross-sectional dependence. In the cases of weekly and fortnightly aggregation, serial dependence in returns becomes negligible, so we use a VAR(0) model, that is, we assume a constant mean.

Because (co)variances have a crucial impact on risk measures, we investigate several specifications from the class of multivariate GARCH models. The literature offers a wide variety of models that differ in the effects that shock have on the variance matrix. We are particularly interested in the importance of asymmetric effects of shocks, and the importance of time-varying correlations. Moreover, estimation should remain feasible when no cross-sectional aggregation is applied. Therefore, we focus on GARCH and GJR-GARCH models combined with CCC and parsimonious DCC specifications, and the RiskMetrics approach. Based on the evidence in Laurent et al. (2012), we do not consider more extensive multivariate models.

⁷See Bauwens et al. (2006) and Silvennoinen and Teräsvirta (2009) for an overview of (multivariate) GARCH models. Hansen and Lunde (2005); Laurent et al. (2012) show that relatively simple models are not outperformed by more extensive alternatives.

⁸In our notation, the unit of t is a day independent of the choice of temporal aggregation, because the forecast horizon h is also independent of it. As a consequence, we denote the first lagged k -period observation as $\mathbf{r}_{t-k,k}^b$. The same structure applies to the volatility models in this subsection.

Conditional correlation models in the style of Bollerslev (1990) and Engle (2002) build the variance matrix $\boldsymbol{\Sigma}_{t,k}^b$ from m_b univariate models for the variances, $\sigma_{i,t,k}^{2b}$ and a separate model for the correlation matrix $\mathbf{R}_{t,k}^b$,

$$\boldsymbol{\Sigma}_{t,k}^b = \text{diag}(\boldsymbol{\sigma}_{t,k}^b) \mathbf{R}_{t,k}^b \text{diag}(\boldsymbol{\sigma}_{t,k}^b), \quad (6)$$

where the diag operator returns a diagonal matrix with the argument vector on the diagonal, and the $m_b \times 1$ vector $\boldsymbol{\sigma}_{t,k}^b$ contains the volatilities $\sigma_{i,t,k}^b$. We define standardized innovations

$$\eta_{i,t,k}^b = \varepsilon_{i,t,k}^b / \sigma_{i,t,k}^b, \quad i = 1, \dots, m_b, \quad (7)$$

that are still correlated, and uncorrelated innovations $\zeta_{i,t,k}^b$ that satisfy

$$\boldsymbol{\eta}_{t,k}^b = (\mathbf{R}_{t,k}^b)^{1/2} \boldsymbol{\zeta}_{t,k}^b, \quad \text{with } \mathbb{E}_{t-k}[\boldsymbol{\zeta}_{t,k}^b \boldsymbol{\zeta}_{t,k}^{b'}] = \mathbf{I}, \quad (8)$$

where $\boldsymbol{\eta}_{t,k}^b$ and $\boldsymbol{\zeta}_{t,k}^b$ are the m_b -vectors with elements $\eta_{i,t,k}^b$ and $\zeta_{i,t,k}^b$, and $(\mathbf{R}_{t,k}^b)^{1/2}$ denotes the lower triangular matrix that results from the Cholesky decomposition of $\mathbf{R}_{t,k}^b$. We discuss the distributions for the innovations in the next subsection.

The univariate models for $\sigma_{i,t,k}^{2b}$ belong to the GARCH family, in particular the standard GARCH(1,1) model of Bollerslev (1986), and the extension proposed by Glosten et al. (1993), referred to as GJR-GARCH. The latter can be written as

$$\sigma_{i,t,k}^{2b} = \omega_{i,k}^b + (\alpha_{i,k}^b + \gamma_{i,k}^b I(\varepsilon_{i,t-k,k}^b < 0)) (\varepsilon_{i,t-k,k}^b)^2 + \beta_{i,k}^b \sigma_{i,t-k,k}^{2b}, \quad i = 1, \dots, m_b, \quad (9)$$

where $\omega_{i,k}^b, \alpha_{i,k}^b, \beta_{i,k}^b > 0$, $\gamma_{i,k}^b > -\alpha_{i,k}^b$, and I is the indicator function. Assuming that the conditional distribution of $\varepsilon_{i,t,k}^b$ is symmetric, the variance is stationary when $\alpha_{i,k}^b + \frac{1}{2}\gamma_{i,k}^b + \beta_{i,k}^b < 1$. Shocks have an asymmetric effect on the volatility when $\gamma_{i,k}^b \neq 0$. The standard GARCH model results when $\gamma_{i,k}^b = 0$.

We consider three models for the correlation matrix: the Constant Conditional Correlation (CCC) of Bollerslev (1990), the Dynamic Conditional Correlation (DCC) of Engle (2002) with the correction of Aielli (2013), and a Hadamard DCC (HDCC) model of Bauwens et al. (2016). The CCC model assumes a constant correlation matrix $\mathbf{R}_{t,b,k}^k = \mathbf{R}_k^b$. The DCC model contains a dynamic specification for the correlation matrix. First, a matrix $\mathbf{Q}_{t,k}^b$

is constructed from

$$\mathbf{Q}_{t,k}^b = (1 - \alpha_{b,k}^q - \beta_{b,k}^q) \bar{\mathbf{Q}}_k^b + \alpha_{b,k}^q \tilde{\boldsymbol{\eta}}_{t-k,k}^b \tilde{\boldsymbol{\eta}}_{t-k,k}^{b'} + \beta_{b,k}^q \mathbf{Q}_{t-k,k}^b, \quad (10)$$

where $\tilde{\boldsymbol{\eta}}_{t-k,k}^b = \text{Diag}(\mathbf{Q}_{t-k,k}^b)^{1/2} \boldsymbol{\eta}_{t-k,k}^b$, and the Diag operator returns a diagonal matrix whose diagonal elements are equal to those of the square matrix given as argument (as in Bauwens et al., 2016). We use correlation targeting, and set $\bar{\mathbf{Q}}_k^b = \text{E}[\tilde{\boldsymbol{\eta}}_{t-k,k}^b \tilde{\boldsymbol{\eta}}_{t-k,k}^{b'}]$. The parameters $\alpha_{b,k}^q, \beta_{b,k}^q$ should be positive and satisfy $\alpha_{b,k}^q + \beta_{b,k}^q < 1$. While $\mathbf{Q}_{t,k}^b$ is positive definite by construction, its diagonal elements are not necessarily equal to one. The correlation matrix $\mathbf{R}_{t,k}^b$ results from

$$\mathbf{R}_{t,k}^b = \text{Diag}(\mathbf{Q}_{t,k}^b)^{-1/2} \mathbf{Q}_{t,k}^b \text{Diag}(\mathbf{Q}_{t,k}^b)^{-1/2}. \quad (11)$$

The DCC model implies the same dynamics for all correlations, which may be restrictive, in particular when the number of assets grows large. We therefore consider the Hadamard DCC extension of Bauwens et al. (2016), replacing Equation (10) by

$$\mathbf{Q}_{t,k}^b = (\mathbf{v}' - \mathbf{A}_{b,k}^q - \mathbf{B}_{b,k}^q) \odot \bar{\mathbf{Q}}_k^b + \mathbf{A}_{b,k}^q \odot \tilde{\boldsymbol{\eta}}_{t-k,k}^b \tilde{\boldsymbol{\eta}}_{t-k,k}^{b'} + \mathbf{B}_{b,k}^q \odot \mathbf{Q}_{t-k,k}^b, \quad (12)$$

where $\mathbf{A}_{b,k}^q, \mathbf{B}_{b,k}^q$ are $m_b \times m_b$ matrices and \odot denotes the Hadamard product. Because Bauwens et al. (2016) finds that a rank-1 specification for $\mathbf{A}_{b,k}^q$ and $\mathbf{B}_{b,k}^q$ is on the one hand parsimonious and on the other hand not outperformed by more extensive specifications, we impose $\mathbf{A}_{b,k}^q = \boldsymbol{\alpha}_{b,k}^q \boldsymbol{\alpha}_{b,k}^{q'}$ and $\mathbf{B}_{b,k}^q = \beta_{b,k}^q \mathbf{v}'$.⁹

Finally, we include the RisMetrics model, which is widely used in practice. This model directly gives the evolution of the variance matrix,

$$\boldsymbol{\Sigma}_{t,k}^b = \boldsymbol{\Sigma}_{b,k}^0 + (1 - \lambda_{b,k}) \boldsymbol{\epsilon}_{t-k,k}^b \boldsymbol{\epsilon}_{t-k,k}^{b'} + \lambda_{b,k} \boldsymbol{\Sigma}_{t-k,k}^b, \quad (13)$$

where $\boldsymbol{\Sigma}_{b,k}^0$ is a constant matrix and $0 \leq \lambda_{b,k} \leq 1$ is the decay parameter (see JP Morgan and Reuters, 1996 and Mina and Xiao, 2001). Based on Zaffaroni (2008) we include $\boldsymbol{\Sigma}_{b,k}^0$ in the specification. Because the dynamics in this model are determined by a single parameter $\lambda_{b,k}$, the variances of all basic portfolio returns and the covariances between all pairs of returns share the same behavior over time.

⁹To prevent empirical problems, we do not apply the more general rank-1 specification of $\mathbf{A}_{b,k}^q$ to $\mathbf{B}_{b,k}^q$.

2.3 Choosing the distribution

The third choice concerns the distribution of the innovations $\varepsilon_{t,k}^b$. We examine three choices for this distribution: the normal, the Student's t , and the empirical distribution. We briefly discuss these options, and the implications of their combination with the models of the previous subsection.

The multivariate normal distribution enjoys wide popularity, despite its exponentially declining tails. Because it is closed under summation, temporal and cross-sectional aggregation of forecasted distributions is straightforward. The combination of the normal distribution with univariate GARCH models and a constant correlation matrix was proposed by Bollerslev (1990). Replacing the constant correlation matrix by the dynamics in Equations (10) and (11) is the DCC model as in Engle (2002). The normal distribution combined with the RiskMetrics model produces the classical RiskMetrics approach (see Mina and Xiao, 2001).

The Student's t -distribution is a popular alternative, because it has fatter tails.¹⁰ To enhance flexibility, we use a univariate Student's t -distribution for the marginal distribution of each $\varepsilon_{i,t,k}^b$, with a specific degrees of freedom parameter $\nu_{i,k}^b$, and model the dependence by the Student's t -copula. We transform the univariate pdf to have expectation μ and variance σ^2 ,

$$\tilde{\psi}(z; \mu, \sigma^2, \nu) = \frac{\Gamma((\nu + 1)/2)}{\Gamma(\nu/2)\sqrt{(\nu - 2)\pi\sigma}} \left(1 + \frac{(z - \mu)^2}{(\nu - 2)\sigma^2}\right)^{-\frac{\nu+1}{2}}, \quad \nu > 2 \quad (14)$$

where ν is the degrees of freedom, and Γ is the Gamma function. We combine these marginal models with the Student's t -copula,

$$C_\psi(v_1, \dots, v_{m_b}; \boldsymbol{\Omega}, \nu) = \Psi_{m_b}(\Psi^{-1}(v_1; \nu), \dots, \Psi^{-1}(v_{m_b}; \nu); \boldsymbol{\Omega}, \nu), \quad \nu > 2 \quad (15)$$

where $\boldsymbol{\Omega}$ is a correlation matrix, ν the degrees of freedom parameter, Ψ_{m_b} the cdf of the (unscaled) multivariate Student's t -distribution, and Ψ^{-1} the inverse of the cdf of the (unscaled) univariate Student's t -distribution. Together, the marginal models and the copula have $m_b + 1$ degrees of freedom parameters, which can all be different. The combination of this copula approach with the DCC specification in Equations (10) and (11) is closely related to one of the models in Jondeau and Rockinger (2006). When the RiskMetrics approach is

¹⁰Though the literature proposes many alternatives with fatter tails than the normal distribution, we restrict ourselves to the Student's t -distribution because of its good performance reported by Mittnik and Paoletta (2000); Giot and Laurent (2004); Bao et al. (2006); Kuester et al. (2006).

selected, we use a multivariate Student's t -distribution for $\boldsymbol{\varepsilon}_{t,k}^b$,

$$\tilde{\psi}_{m_b}(\mathbf{z}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu) = \frac{\Gamma((\nu + m_b)/2)}{\Gamma(\nu/2)((\nu - 2)\pi)^{m_b/2} \sqrt{|\boldsymbol{\Sigma}|}} \left(1 + \frac{(\mathbf{z} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{z} - \boldsymbol{\mu})}{(\nu - 2)} \right)^{-\frac{\nu + m_b}{2}}, \quad (16)$$

with $\nu > 2$ and where $|\boldsymbol{\Sigma}|$ denotes the determinant of $\boldsymbol{\Sigma}$. This specification is again scaled to ensure $\text{Var}[\mathbf{z}] = \boldsymbol{\Sigma}$.

As the third option we propose the empirical distribution. It can capture both the fat tails and the asymmetry that are typically present in return distributions.¹¹ It does not suffer from the risk of misspecification, but the resulting loss of efficiency may particularly affect forecasts of tail events. For the CCC and DCC specifications, we use it for the marginal distributions of the standardized innovations $\eta_{i,t,k}^b$ in Equation (7) in combination with the Gaussian copula. For a set of P realizations in the vector \mathbf{x} , the empirical cumulative probability function is given by

$$F_P^{\text{emp}}(z; \mathbf{x}) = \frac{1}{P + 1} \sum_{\tau=1}^P I(x_\tau \leq z). \quad (17)$$

The Gaussian copula is given by

$$C_\phi(v_1, \dots, v_{m_b}; \boldsymbol{\Omega}) = \Phi_{m_b}(\Phi^{-1}(v_1), \dots, \Phi^{-1}(v_{m_b}); \boldsymbol{\Omega}), \quad (18)$$

where $\forall i, v_i \in [0, 1]$, $\boldsymbol{\Omega}$ is a correlation matrix, Φ_{m_b} the cdf of the m_b -variate normal distribution, and Φ^{-1} the inverse of the cdf of the univariate normal distribution. This design corresponds with the SCOMDY models in Chen and Fan (2006). In a univariate setting this approach reduces to Filtered Historical Simulation (see also Christoffersen, 2009). When the RiskMetrics-specification is used, the empirical distribution applies to $\zeta_{i,t,k}^b$ as defined in Equation (8). We do not assume that $\zeta_{i,t,k}^b$ and $\zeta_{j,t,k}^b$ are independent for $i \neq j$, and use the multivariate empirical distribution

$$F_{P,m_b}^{\text{emp}}(\mathbf{z}; \mathbf{X}) = \frac{1}{P + 1} \sum_{\tau=1}^P \prod_{i=1}^{m_b} I(x_{i,\tau} \leq z_i), \quad (19)$$

where \mathbf{z} is an $m_b \times 1$ vector and \mathbf{X} is a $m_b \times P$ matrix that contains the set of P observations.

¹¹The skewed t distributions and copulas are also good candidates to capture these aspects. See for example Hansen (1994); Demarta and McNeil (2005) and Bauwens and Laurent (2005).

2.4 Choosing the method of forecast construction

The final choice relates to the construction of forecasts. The risk manager has to choose how to construct a multi-period forecast, in case she has not applied full temporal aggregation. When $h = k$ she constructs a one-step-ahead forecast, and we call this forecast “direct”. When $k < h$, she can choose between “iterated” and “scaled” forecasts. Iterated forecasts take into account the path dependence that results from the serial dependence in the return models. The forecasted distribution for $\mathbf{r}_{t+jk,k}^b$, $j \geq 1$ is constructed based on the forecasted distribution for the previous horizon, $\mathbf{r}_{t+(j-1)k,k}^b$. Generally, the convolution of these distributions does not have a closed-form expression, which means that simulations have to be used. Scaling forecasts implies the assumption that the distributions of $\mathbf{r}_{t,k}^b$ and $\mathbf{r}_{t,h}^b$ have the same shape, but the location and scale parameters of the aggregated distribution $\mathbf{r}_{t,h}^b$ are adjusted for the horizon h . When VAR-effects in Equation (5) are absent, the mean and variance are scaled by a factor h/k . Formulated for volatility, this is the well-known “square root of time” rule, that is advocated in the Basel Accords. We discuss the technical aspects of the forecast construction in Appendix B.2, and show in Appendix B.3 how to include the serial dependence implied by the VAR in Equation (5) when scaled forecasts are used.

Iterated forecasts can be more precise but also more time-consuming than scaled forecasts. When portfolios are large, iterated forecasts may become computationally infeasible, which explains the popularity of scaling in practice. However, Diebold et al. (1997); Danielsson and Zigrand (2006); Wang et al. (2011) show that scaled forecasts become worse when the horizon increases. For some of the models we consider, for instance the univariate GARCH or RiskMetrics specifications, multi-step ahead forecasts can be constructed in closed form (see Drost and Nijman, 1993 for the first and Sbrana and Silvestrini, 2013 for the second), but generally this is not the case. Therefore, scaling remains a relevant forecasting method. Ghysels et al. (2009) show that MIDAS forecasts dominate other forecasting methods for horizons longer than 10 days (see also Ghysels et al., 2006). Because our forecasting horizon is shorter, we do not investigate the performance of MIDAS techniques.

2.5 Estimation and Empirical Design

We estimate the parameters in all the different models with a multi-step procedure. First, we estimate the VAR-parameters in Equation (5) by ordinary least squares. The VAR residuals are the input for the next estimation step. For the CCC and DCC models we estimate the

parameters of the marginal models in Equation (9) (GARCH or GJR-GARCH) separately for each basic portfolio by (Quasi) Maximum Likelihood (QML). We construct standardized residuals as in Equation (7), and use these as the basis of our Maximum Likelihood estimation of the (unconditional) correlation matrix, and if needed the parameters of the (H)DCC models and the degrees of freedom in Equation (15). For the RiskMetrics approach, we estimate the parameters in Equation (13) by (Q)ML. We provide a more detailed description in Appendix B.1.

The estimation procedure for the CCC and DCC models is the common approach that is advocated in the literature (see Bollerslev, 1990; Engle, 2002; Chen and Fan, 2006; Aielli, 2013). For copulas this multi-step estimation is referred to as the Inference Functions for Margins method (see Joe, 1997, Ch. 10), or Multi-Stage Maximum Likelihood (see Patton, 2013). Originally, the decay factor λ of the RiskMetrics models is determined by minimizing the root mean squared error (see JP Morgan and Reuters, 1996), but Zaffaroni (2008) shows that this approach is unreliable, and advocates QML instead.

Each forecast is based on a moving window of 1,000 days (so approximately four years of data). When we apply temporal aggregation ($k > 1$), we aggregate the daily observations to 200 weekly ($k = 5$) or 100 biweekly ($k = 100$) observations. We reestimate all model parameters, when 50 days have passed, and assume that the parameters do not change at intermediate points in time. When simulations are needed, we use 10,000 draws. We also investigate the effect of doubling the size of the estimation windows.

2.6 Evaluation

At each point in time t , we construct an out-of-sample Value-At-Risk forecast for the portfolio return $r_{t,h}^P$ over the next h days for all combinations of choices. We refer to each choice combination as a method (cf. Giacomini and White, 2006), and denote the forecast by method j as $q_t^j \equiv -\text{VaR}_\vartheta(r_{t,h}^P)$. We evaluate the quality of each forecasting method itself, and compare them with each other.

The unconditional coverage (UC) test of Christoffersen (1998); Kupiec (1995) tests whether the actual fraction of VaR-violations equals the theoretical fraction ϑ . We define the VaR violation (also called hit)

$$x_{t+h}^j \equiv I(r_{t,h}^P \leq q_t^j), \tag{20}$$

where I denotes again the indicator function. If VaR method j works correctly, $E[x_{t+h}^j] = \vartheta$. We test this hypothesis against the alternative $E[x_{t+h}^j] > \vartheta$, as regulators care mostly about too many violations.

The dynamic quantile (DQ) test of Engle and Manganelli (2004) tests whether the VaR-violations are predictable by past information. If VaR method j works correctly, the violation x_{t+h}^j should be unrelated to the forecasted Value-at-Risk q_t^j . Because of the overlap in the forecasting horizon, violations x_{t+h}^j and x_{t+h+1}^j are related by construction. Therefore, we only include q_t^j in the test. Based on Herwartz and Waichman (2010), we implement the test by a linear regression

$$\tilde{x}_{t+h}^j = c_{0j} + c_{1j}q_t^j, \quad (21)$$

where $\tilde{x}_{t+h}^j = x_{t+h}^j - \vartheta$ is the centered hit. We then test whether $c_{1j} = 0$ against the alternative $c_{1j} \neq 0$. As advocated by Herwartz and Waichman (2010), we construct the critical values for the test based on bootstraps. They show that using the asymptotic distribution of the test statistic leads to considerable size distortions. Because of the overlap in the centered hit series, we use a block bootstrap with a block length of 10. We do not use a logit (or probit) model, because the maximum likelihood estimator is inconsistent because of the serial correlation in x_{t+h}^j .¹²

To determine the differences between the different choice aspects, we apply the procedure for constructing the Model Confidence Set as in Hansen et al. (2011, Sec. 3.1.2). We also compare pairs of forecasts by the test of Diebold and Mariano (1995) (DM-test) in the framework of Giacomini and White (2006). We need their framework, because some of the methods that we compare are nested (for example, the GARCH models are nested in the GJR models, and the CCC models in the DCC models). Additionally, it allows for parameter estimation uncertainty. Our use of a rolling window to estimate the model parameters fits in this framework.

We use the asymmetric tick loss function as in Giacomini and Komunjer (2005)

$$L_\vartheta(e_{t+h}^j) = (\vartheta - I(e_{t+h}^j < 0))e_{t+h}^j, \quad (22)$$

where $e_{t+h}^j \equiv r_{t,h}^p - q_t^j$. Because of the asymmetry of the loss function, realizations below q_t^j , that is, VaR violations, lead to a larger loss. Next, we define the loss differential between

¹²We are grateful to a referee for pointing out this issue.

methods j and j' as

$$d_{\vartheta,t}^{j,j'} \equiv L_{\vartheta}(e_t^j) - L_{\vartheta}(e_t^{j'}). \quad (23)$$

We test $E[d_{\vartheta,t}^{j,j'}] = 0$ by the statistic proposed by Diebold and Mariano (1995). When the null-hypothesis is rejected and $E[d_{\vartheta,t}^{j,j'}] < 0$, method j outperforms method j' .

Because we make ten-day VaR forecasts for every day in our sample, all tests in this section suffer from overlapping data. Therefore, we use the procedure of Newey and West (1987) with a Bartlett kernel and 10 leads and lags in the calculation of variances and standard errors. We use a significance level of 5% in our tests, unless stated otherwise.

3 Data

We base our empirical analysis on a well-diversified investment portfolio. This portfolio invests 50% in bonds, 30% in equities, 10% in real estate, and the remaining 10% in commodities. The investments in bonds are split between US government bonds (30%) and US corporate bonds (20%). The equity part can be further divided based on regions: Europe (12%), United States (10.5%), Pacific (3.75%) and Emerging Markets (3.75%). Appendix A provides details about our used indexes. Our sample covers the period January 1, 1990–February 28, 2014. We delete holidays, giving a sample size of 6,030 daily observations.¹³ Because we use an estimation window of 1,000 observations and make a ten-day forecast, we evaluate 5,021 forecasts. The first forecast is for the portfolio return over the ten working days from December 17 to 31, 1993.

We analyze our data at the daily, weekly and biweekly frequency, for each asset and for the aggregated portfolio. We summarize the main aspects here, and present more details in Appendix A. Generally, bonds show considerably lower returns, but also lower volatilities than the other assets. All returns exhibit negative skewness and excess kurtosis and autocorrelation. The correlation matrix shows a block structure, with high correlation between bonds and between equities. Real estate is highly correlated with equities and mildly with commodities. The portfolio returns are less volatile than the individual asset returns, which shows the effect of diversification. However, they are stronger left skewed than almost all

¹³Because we analyse an international portfolio, we define a holiday as day when the values of two or more indexes do not change. Because of the dominance of the US in the portfolio, most holidays are US public or government holidays.

constituent series, and the kurtosis coefficient is close to the maximum among them, which shows the effect that diversification breaks down in time of crises.

Our comparison of daily, weekly and biweekly returns shows that returns are far from i.i.d. Lau and Wingender (1989) show that when k i.i.d. returns are aggregated, the skewness coefficient is scaled by $1/\sqrt{k}$ and the coefficient of excess kurtosis by $1/k$. Instead of decreasing, the left skewness of all series increases when moving from daily to weekly or biweekly returns. The kurtosis coefficients for weekly returns decrease, but much less than for the i.i.d. case. For the portfolio as a whole these effects are even stronger. Taken together, these results indicate that extreme and negative returns cluster over time and in the cross section. So, scaling densities may not work well for measures of tail risk. Moreover, portfolio aggregation does not mitigate the effect.

4 Results

Our analysis boils down to a large-scale horse race. We analyze all combinations of three possible levels of temporal aggregation (daily, weekly and biweekly), three levels of portfolio aggregation (no aggregation, aggregation into the three main asset classes, and aggregation into a single portfolio), five model specifications (CCC-GARCH, DCC-GARCH, CCC-GJR, DCC-GJR and RiskMetrics), three distributions (normal, empirical, and Student's t), and two forecasting options (iterated and scaling). The forecasting options are available when daily or weekly returns are used. For biweekly returns only a one-step-ahead forecast is needed. In total this leads to 255 different methods. We evaluate the quality of the resulting VaR forecasts in different ways, where we focus on the impact of the different choices. We do not aim at finding the specific combination of options that beats all others. We examine the VaR forecasts with a 99% confidence level in detail, and briefly consider the forecasts with the 95% confidence level. We use a 5% significance level unless otherwise noted.

4.1 VaR forecasts with a 99% confidence level

4.1.1 Absolute forecast quality

We first investigate how the five different choices impact the coverage of the VaR forecasts and test equality to the nominal 1% level against the one-sided alternative of more violations. Too many violations can lead to penalties for a financial institution, whereas too few

violations may be undesirable, but are less likely to cause problems. To see how the different choices influence the results, we report the total number of rejections for a particular significance level in Table 1. The detailed results are available in Table C.1 in Appendix C.1. The interpretation of the summarized results deserves some care, because the tests are not independent.

[Table 1 about here.]

Table 1a shows that temporal aggregation leads to more violations. For daily returns combined with iterated forecasts, 15 methods out of 51 leads to violations that significantly exceed 1% with p -values below 0.05. For biweekly aggregation, 39 methods lead to significantly more violations than 1%. The results for weekly returns lie in between. Together, they indicate that less temporal aggregation, so a more detailed observation of the return dynamics, leads to better coverage ratios.

The relation of portfolio aggregation with the coverage generally shows a U-shape, though with some exceptions. A split into three basic assets, so modeling the main asset classes, leads to the lowest violation frequency. For 57 out of 105 forecast methods, the frequency of VaR violations deviates significantly from 1%. Aggregating the assets into a single portfolio generally leads to larger violation frequencies, with 30 out of 45 being significant. Modeling all asset leads to slightly more violations (62 out of 105) than modeling at the level of the main asset classes.

The model choice shows some interesting results. First, the one-size-fits-all approach of RiskMetrics works quite well, both univariately and multivariately, with 8 out of 15 and 15 out of 30 rejections at the 0.05 level. Second, replacing GARCH by GJR, so including an asymmetric effect of shocks, improves the coverage ratios, and leads to 9 rejections in the univariate setting. In the multivariate setting, exchanging the CCC, DCC and HDCC specifications does not influence the coverage ratios much. When combined with the GJR specification their performance is at par with RiskMetrics, with 15 out of 30 ratios significant at the 0.05 level. Though the performance of the RiskMetrics and DCC-GJR models is similar, the RiskMetrics model works particularly well when combined with scaling, whereas the DCC-GJR model works better in combination with iterated forecasts.

The choice for the distribution impacts coverage substantially. The empirical distribution comes first, the Student's t -distribution second, and the normal distribution last. When daily returns are combined with iterated forecasts, the differences between the distributions are

not that large, but when weekly and biweekly form the basis the differences are considerable (see Table C.1). When the empirical distribution is used, only 7 out of 85 ratios differ significantly from 1%, compared to no less than 67 for the Student's t and even 75 for the normal distribution. Moreover, when the violation frequencies exceed 1%, those of the empirical distribution are the lowest.

The choice between scaled or iterated forecasts particularly influences the results when daily data is used. Scaling a forecast from one to ten steps ahead generally leads to worse coverage ratios than constructing iterated forecasts. This effect holds particularly when the dynamics are specified in more detail, and are largest for the DCC-GJR model. When RiskMetrics is used constructing scaled or iterated forecasts does not matter much. Differences are also smaller for weekly data. Because the variance process in the RiskMetrics method is not mean-reverting, this result is not surprising.

Next, we turn to the dynamic quantile (DQ) test based on Equation (21) to investigate whether the forecast quality varies with the magnitude of the VaR forecast. The summary results are in Table 1, and the full results in Table C.2. Mostly, the relation between a violation and the forecast is not significant, with only 40 rejections out of 255 methods. Nevertheless, the estimates and in particular their signs give useful insights into the forecast quality. A positive sign means that risk is overestimated in times of distress, because a larger than average VaR forecast (a value for q_t^j below average) decreases the probability of a violation. Risk is then underestimated for quiet times. For negative signs, the interpretation is reversed.

Temporal aggregation has a clear effect on the forecast quality. The use of daily returns leads to a slight underestimation of risk in times of distress for 46 methods, but to overestimation for only five. These results indicate that models based on daily returns generally revert a bit too fast to the steady-state volatility. A closer inspection of Table C.20a shows that the five positive coefficients are all for forecasts based on RiskMetrics, where mean-reversion is absent. The negative coefficients are mostly insignificant, but a richer GARCH specification may improve the forecasts. For weekly returns, the number of positive and negative coefficients are about equal. For biweekly returns all coefficients are positive, and 30 out of 51 are significant. Here, mean reversion is clearly too slow.

Full portfolio aggregation leads more often to positive than to negative coefficients (27 versus 18), but this pattern reverses for less aggregation. When all assets are considered, 73 methods lead to negative coefficients and 32 to positive ones. When three basic assets

are used, positive and negative coefficients are more balanced. Also in the cross-sectional dimension, methods with a low degree of aggregation lead more often to an underestimation of VaR in times of distress.

The results for the different model specifications show again the effect of mean reversion. In the univariate case, we find more evidence for overestimation of VaR in times of distress, compared to the multivariate cases. A more precise identification of the nature of the shocks may help, but it is not clear which specification is best, as the differences between the six multivariate GARCH models are quite small. The asymmetry captured by the GJR model has small effects on iterated forecasts (daily or weekly) but leads to underestimation of risk in times of distress when forecasts are scaled. The difference between CCC, DCC and HDCC models is generally larger when eight assets are considered than three. In the CCC models correlations are assumed to be constant, so it leads more often to underestimation. The RiskMetrics model, which does not imply mean reversion, leads to overestimation.

Overall, we conclude that the results of the DQ-test are mostly related to the speed of mean reversion implied by the models. When mean reversion is relatively fast, the DQ-test indicates a slight underestimation of risk in times of distress, though mostly insignificant. This effect is clearest when daily data and/or CCC models are chosen. When mean reversion is slower and shocks become more persistent, the DQ-tests signal overestimation in times of distress. Significance is a bit stronger, though the majority of methods show insignificant coefficients. The differences between the distributions are small. All show about as many positive as negative coefficients. The same conclusion applies to iterating versus scaling.

Summarizing, we find that temporal aggregation has a large impact on the forecast quality. Using weekly or biweekly data leads to too frequent violations, though more often during quiet times. Portfolio aggregation has less impact. Modeling on the level of asset classes gives the best results for coverage and time-variation. The model choice is also of less importance. Differences between the combinations of CCC/DCC/HDCC and GARCH/GJR methods are small, but still indicate a preference for the DCC/HDCC-GJR method. Differences with RiskMetrics are somewhat larger. The choice for the distribution is again more important. In particular distributions with fat tails lead to better coverage ratios. The forecasting method is important when daily data are used, with iterated forecasts outperforming scaled ones.

We conclude that more detailed information on shocks as provided by daily data on the asset or asset class level, and a more detailed specification of their propagation in the form of

a richer model and iterated forecasts, improve forecast quality. Increased aggregation means that sometimes shocks are missed, whereas the one-size-fits-all approach of the RiskMetrics methods, and the scaling of forecasts lead to a too strong extrapolation of the effects of shocks. On the other hand, we find that detailed information, models and forecasting increase the risk of forecast errors. So far, the benefits of an increased level of detail outweigh these errors.

4.1.2 Relative forecast quality

We now turn to a statistical comparison of the different methods. For each method, we calculate the average asymmetric tick loss as in Equation (22), reported in Table 2. The lowest loss (5.10×10^{-2}) is obtained by constructing iterated forecasts based on daily data for eight assets, with a DCC-GJR model and the empirical distribution. The largest loss (8.95×10^{-2}) results from a direct forecast based on biweekly data for eight assets with the RiskMetrics approach and the empirical distribution.

[Table 2 about here.]

[Table 2 (continued) about here.]

We construct the Model Confidence Set (MCS) (Hansen et al., 2011) with a significance level of 10% to determine the importance of the different choice aspects. The results in Table 3 show that the observation frequency is the most important choice. None of the methods based on daily observations are removed from the MCS, neither when iterated nor when scaled forecasts are constructed. A lower frequency leads to more removals: 7 for weekly observations combined with iterated forecasts, 13 for the combination with scaled forecasts, and 31 for biweekly observations. So, less temporal aggregation clearly leads to better results.

[Table 3 about here.]

When weekly data are used to construct iterated forecasts, the seven models that are excluded typically do not allow for enough dynamics (Table 3a). Five of the excluded models use RiskMetrics, the other two use a univariate GARCH specification. Along the dimension of portfolio aggregation, five excluded models model at the portfolio level, the other two at the asset class level. The exclusions seem not particularly related to the distributions.

The removals from the subset of scaled forecasts based on weekly data in Table 3b show a slightly clearer pattern. Of the 13 removed methods five use RiskMetrics, and three use CCC-GARCH. The other five are spread over the other model choices. Looking at the portfolio aggregation, seven removals model returns at the portfolio level, five at the asset class level and only one at the asset level. We now see that the Student's t -distribution counts only two removals, whereas the normal counts six removals, and the empirical five.

From the 51 methods based on biweekly data in Table 3c, 31 are removed and 20 remain. In particular, all 17 models based on the empirical distribution are excluded, compared to five for the normal distribution and nine for the Student's t -distribution. The sample size of 100 observations is too small to reliably construct the empirical distribution, and even gives trouble for the Student's t -distribution.

From the subset of methods that use either the normal or Student's t -distribution, all models on the portfolio level are excluded. From the models based on the asset class level, six are removed, compared to only two from the models based on the asset level. With regard to the model specification, the RiskMetrics approach often leads to exclusion (four out of six). We do not observe a clear pattern in the in- or exclusion of the other model specifications.

We conclude that using observations at the asset level, and to a lesser extent at the asset class level, can compensate for the low detail in the temporal dimension. Also, a model specification with richer dynamics can offer some compensation. Working with biweekly data typically leads to larger losses than using data with a higher frequency, whatever the portfolio details or model specification.

We now investigate whether the methods that use daily data can yield some further insight into the consequences of the remaining choices. The MCS results shows that no particular combination stands out, as all models are included in the MCS. Also when we start the MCS procedure with only the 110 daily methods, no methods are removed from this set. Therefore, we conduct pairwise comparisons of those methods that differ in only one choice aspect based on the Diebold-Mariano test. For example, we test whether the DCC-GJR specification leads to the same average loss as the RiskMetrics specification, when iterated forecasts are constructed based on daily data at the asset level using the empirical distribution. Detailed results are available in Appendix C.1. We summarize them in Table 4.

[Table 4 about here.]

Our results show a preference for iterated forecasts. For 38 out of 51 comparisons, the losses from iterated forecasts are lower than from scaled forecasting. Six (14) loss differentials are significant at the 5% (10%) level. Because of this difference, we separately report summary statistics for iterated and scaled forecasts.

The results for portfolio aggregation (panels b and c) confirm our conclusion from the MCS that a low degree of portfolio aggregation leads to better performance. Modeling at the asset level leads to lower losses for 16 (18) out of 21 methods that use iterated forecasts, when compared to modeling at the asset class (portfolio) level. Modeling at the asset class level beats modeling at the portfolio level also 21 times. However, differences are generally not significant. For scaled forecasts, results are similar. These results are consistent with the preference for multivariate models reported by Santos et al. (2013).

With regard to the distribution choice, the normal distribution loses from the empirical and Student's t -distributions in 13 (15) out of 17 comparisons based on iterated forecasts, and for all 17 comparisons based on scaled forecasts. The empirical distribution is preferred to the Student's t -distribution. This result may be due to the skewness that the empirical distribution can capture, and contrasts with the bad results for the empirical distribution when biweekly data are used. The empirical distribution only works well when a large number of observations is available.

The number of comparisons of model specifications is large, because we analyze 7 specifications. We therefore summarize our results such that we can see the consequences of the parsimonious specification of RiskMetrics versus the richer (multivariate) GARCH specifications, the asymmetric effects in the marginal specification (GARCH vs. GJR), and the dynamics in the multivariate part (CCC, DCC and HDCC).

Panels f and g clearly show a preference for richer model specifications than RiskMetrics. All 36 multivariate models produce a lower average loss than RiskMetrics, which is significant for 15 (12) methods with iterated (scaled) forecasts. For the univariate case, the results are more mixed. Given the preference for multivariate models in panels b and c, we conclude that allowing for a separate specification of the dynamics in the marginal volatilities and in the dependence is worthwhile.

Our results also show a clear preference for the asymmetric specification of the volatility dynamics in the GJR-GARCH model. When iterated forecasts are used, the univariate GJR specification beats GARCH for 3 out of 3 methods. For the multivariate specifications, it does so for 53 out of 54 methods, with 20 loss differentials that are significant at the 10%

level. When scaled forecasts are used, the evidence is also in favor of the GJR specification, though less pronounced.

We do not find a clear winner for the dynamics in the dependence. The CCC specification leads to losses that are not that different from the DCC/HDCC specifications. The HDCC specification performs less than DCC, in particular when scaled forecasts are constructed.

Summarizing, the DM tests indicate that methods that have more information available and make better use of it lead to lower losses, though differences are often not significant. For portfolio aggregation, we find a preference for models at the asset or asset class level. For the distribution, the empirical one is preferred. Multivariate models with asymmetric effects in the volatility are preferred, but correlation dynamics do not matter much.

4.2 VaR forecasts with a 95% confidence level

We repeat our analyses for VaR forecasts with a confidence level of 95%. We focus here on the main findings, and provide the full set of results in Appendix C.2. In the discussion below, we always use the test outcomes at the 5% significance level, except for the results of the Model Confidence Set, which uses again the 10% significance level.

In terms of absolute forecasting quality, the different methods show better coverage ratios than for the 99% VaR forecasts. Based on their 95% VaR-forecasts, only 49 methods (out of 255) are rejected, compared to 149 based on their 99% VaR forecasts. This drop in the number of rejections is concentrated in the methods that use the normal distribution, as it goes down from 75 to only 11. For the Student's t -distribution, the number of rejections decreases from 67 to 38, indicating that the normal distribution now leads to better coverage. When we use a significance level of 10%, the empirical distribution outperforms both the normal and Student's t -distribution with 4 compared to 47 and 61 rejections.

The rejections based on the 95% VaR forecasts are more concentrated in the methods that use biweekly observations (19 out of 51), which further strengthens the preference for less time aggregation. The options for the portfolio aggregation and for the model are more or less equally affected, and show similar drops in the number of rejections.

The DQ-tests find more evidence of time variation of the VaR exceedances. Based on the 95% VaR forecasts, 65 coefficients are significantly positive, compared to 35 based on the 99% VaR forecasts. We also find more positive coefficients for the 95% case (198 compared to 117). In particular, the coefficients are positive when weekly and biweekly data are used to construct forecasts. This finding indicates more overestimation of risk in times of distress.

The differences with the 99% case may point at changes in the shape of the distributions, but we do not investigate this issue in more detail.

The average losses can be estimated more precisely based on the 95% VaR forecasts. As a consequence, the MCS becomes smaller: 126 methods are maintained, and 129 are removed. Recall that for the 99% VaR forecasts, only 51 models are removed. The removals are again concentrated in the methods that use a larger degree of temporal aggregation. Many methods that generate forecasts based on weekly data are removed (75 out of 102), and even all that use biweekly data. From the 110 methods that use daily data only 3 are removed.

The pattern that we observe for the removals from the subset of methods that use weekly data for 99% VaR forecasts is stronger for 95% forecasts. All methods that model on the asset class or portfolio level are out, as are all methods that use RiskMetrics, and most methods that use the empirical distribution. The three excluded methods from the subset based on daily data all use RiskMetrics.

The pairwise comparison of the different methods based on daily data also show a stronger preference for no portfolio aggregation. Losses are again lower for models at the asset level, but now the differences are significant for 5 and 13 out of 21 comparisons with models at the asset class and portfolio level.

The method of constructing forecasts and the distribution for innovations are less consequential for 95% VaR forecasts. Though losses tend to be lower for scaled forecasts of 95% VaR (42 times compared to 17 times for the 99% VaR), only one differential is significant. Also, the normal distribution does not lead as often to negative losses.

The pairwise comparisons of the different model specifications yields similar results as for 99% VaR forecasts. RiskMetrics loses from richer multivariate GARCH specifications. Within the latter subset, models with the asymmetric GJR specification for the marginal volatility process outperform models with the symmetric specification, and the choice for CCC, DCC or HDCC does not matter much.

We conclude that the degree of time aggregation is the most important choice, both for 99% and 95% VaR forecasts. Using daily data leads to the best coverage and the lowest losses. For 95% VaR forecasts, the degree of portfolio aggregation becomes more important, though it remains less consequential than the degree of temporal aggregation. Models at the asset level lead to reasonable coverage, and to lower losses than models at the asset class or portfolio level. Daily observations at the asset level are best combined with models that

specify the effects of shocks more precisely. The choice for the distribution is important for 99% VaR forecasts, with the empirical and Student's t -distribution outperforming the normal distribution. For the 95% VaR forecasts, the distribution matters less. The same holds for the forecasting method.

4.3 Longer estimation window

Our main result that methods with daily data outperform those with (bi)weekly data can have two causes. Methods based on daily data can better capture the propagation of shocks, but the better performance can also reflect increased efficiency because more observations are available for estimation. In particular, 100 biweekly observations may be too few to reliably estimate the parameters in multivariate GARCH models. Therefore, we investigate how our results change when we double the estimation window from approximately four to eight years. So, instead of 1,000, 200 and 100 observations, we now use 2,000, 400 and 200 observations for estimation in the daily, weekly and biweekly methods. We discuss the main findings here, and present the detailed results in Appendix C.3. If efficiency is the main driver of our result, we should see an improvement in the forecast quality of methods based on weekly and especially biweekly data, both in absolute sense and in comparison with methods based on daily data.

Regarding absolute forecast quality, we observe only slight improvements for biweekly data. Using the 4-year estimation window, correct unconditional coverage is rejected for 39 out of 51 methods at the 5% significance level (see Table 1). The DQ-test leads to 30 rejections. For the 8-year window, these tests produce 36 and 11 rejections. For weekly data results become a bit worse.

Our analyses of relative forecast quality show smaller differences between all methods when using an 8-year estimation window. The MCS contains all methods based on daily and weekly data, and only six methods based on biweekly data are removed. Pairwise comparisons with DM-tests show that daily methods with iterated (scaled) forecasts significantly (at 5% level) outperform weekly methods 14 (18) times using a 4-year estimation window. These numbers go down to 8 (11) using a 8-year window. The comparison of daily versus biweekly methods shows a reduction of 34 (26) to 19 (19). However, at the 10% level, results are the same for both windows. Moreover, the loss differentials actually increase for the 8-year window, and so do their standard errors, which explains the decreased significance.

Based on these results, we conclude that a low degree of temporal aggregation is prefera-

ble, because it leads to methods that better capture the propagation of shocks. Our results for portfolio aggregation point in the same direction. An increase in the number of modeled assets also leads to models that better capture shocks, without affecting efficiency.

5 Conclusion

In this paper we investigate the importance of the degree of aggregation for forecasting 99% and 95% ten-day ahead Value-at-Risk (VaR). We analyze the effect of aggregating daily returns to weekly and biweekly returns, and aggregation of assets into the main asset classes and into a single portfolio. We compare the importance of aggregation to those of choosing a time-series model (asymmetric or symmetric GARCH, constant or time-varying correlation, or the RiskMetrics approach), choosing a distribution (the Gaussian, Student's t , or the empirical distribution), and choosing a method for forecast construction (iterated or scaled).

Our main finding is that the degree of temporal aggregation is the most important choice. Using daily returns leads to better VaR forecasts than (bi)weekly returns. The dynamics in the distribution of returns is best captured at the daily level. A higher degree of aggregation leads to a loss of important details. The chosen model and distribution also matter for the forecasts. When daily returns are chosen, the model is the second-most important choice. Because the dynamics are present in great detail, the model should accurately capture them. Because weekly or biweekly returns obscure the return dynamics, the chosen distribution becomes more important than the model.

We find that multivariate models outperform univariate ones, confirming Santos et al. (2013), though the differences are not that large, and often not significant. The degree of portfolio aggregation affects forecast quality less than the degree of temporal aggregation, the model and distribution choice. Iterating forecasts are better than scaled and direct forecasts. This holds in particular for the 99% VaR. Scaled forecasts generally perform better than direct forecasts, in line with Ghysels et al. (2009).

We advise to construct models on the asset class level based on daily data. The model had best contain asymmetric effects on the marginal level, but correlation dynamics are not consequential. The distribution should allow for fat tails, in particular for 99% VaR forecasts. It is also best to iterate forecasts. However, applying the RiskMetrics approach with scaling does not make forecasts that much worse.

References

- Aielli, G. P. (2013). Dynamic conditional correlation: on properties and estimation. *Journal of Business & Economic Statistics*, 31(3):282–299.
- Andersen, T. G., Bollerslev, T., Christoffersen, P. F., and Diebold, F. X. (2006). Volatility and correlation forecasting. In Elliot, G., Granger, C. W., and Timmermann, A., editors, *Handbook of Economic Forecasting*, volume 1, chapter 15, pages 777–878. Elsevier North Holland, Amsterdam, Netherlands.
- Bao, Y., Lee, T.-H., and Saltoglu, B. (2006). Evaluating predictive performance of value-at-risk models in emerging markets: A reality check. *Journal of Forecasting*, 25(2):101–128.
- Basel Committee on Banking Supervision (1996). Amendment to the capital accord to incorporate market risks. Technical Report Committee Report 24, Basel Committee on Banking Supervision, Basel, Switzerland.
- Bauwens, L., Grigoryeva, L., and Ortega, J.-P. (2016). Estimation and empirical performance of non-scalar dynamic conditional correlation models. *Computational Statistics & Data Analysis*, 100:17–36.
- Bauwens, L. and Laurent, S. (2005). A new class of multivariate skew densities, with application to generalized autoregressive conditional heteroscedasticity models. *Journal of Business & Economic Statistics*, 23(3):346–354.
- Bauwens, L., Laurent, S., and Rombouts, J. (2006). Multivariate GARCH models: a survey. *Journal of Applied Econometrics*, 21(1):79–109.
- Berkowitz, J. and O’Brien, J. (2002). How accurate are value-at-risk models at commercial banks? *Journal of Finance*, 57(3):1093–1111.
- Bhansali, R. (1999). Parameter estimation and model selection for multistep prediction of a time series. In Ghosh, S., editor, *Asymptotics, Nonparametrics, and Time Series*, Statistics: textbooks and monographs, chapter 6, pages 201–225. Marcel Dekker, New York.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327.
- Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: A multivariate generalized arch model. *Review of Economics and Statistics*, 72(3):498–505.
- Brownlees, C. T., Engle, R. F., and Kelly, B. T. (2011). A practical guide to volatility forecasting through calm and storm. *Journal of Risk*, 14(2):1–20.
- Brownlees, C. T. and Gallo, G. M. (2010). Comparison of volatility measures: a risk management perspective. *Journal of Financial Econometrics*, 8(1):29–56.
- Campbell, J. Y. and Viceira, L. M. (2005). Long-horizon mean-variance analysis: User guide. *Financial Analysts Journal*, 61(1):Supplement, 29 pages.
- Chen, X. and Fan, Y. (2006). Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification. *Journal of Econometrics*, 135(1):125–154.
- Chevillon, G. (2007). Direct multi-step estimation and forecasting. *Journal of Economic Surveys*, 21(4):746–785.
- Christoffersen, P. (2009). Value-at-risk models. In Andersen, T. G., Davis, R. A., Kreiss, J.-P., and Mikosch, T., editors, *Handbook of Financial Time Series*, pages 753–766. Springer.

- Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review*, 39(4):841–862.
- Clements, M. P., Galvão, A. B., and Kim, J. H. (2008). Quantile forecasts of daily exchange rate returns from forecasts of realized volatility. *Journal of Empirical Finance*, 15(4):729–750.
- Daniélsson, J. and Zigrand, J. (2006). On time-scaling of risk and the square-root-of-time rule. *Journal of Banking & Finance*, 30(10):2701–2713.
- Demarta, S. and McNeil, A. J. (2005). The t copula and related copulas. *International Statistical Review*, 73(1):111–130.
- Diebold, F., Hickman, A., Inoue, A., and Schuermann, T. (1997). Converting 1-day volatility to h -day volatility: Scaling by \sqrt{h} is worse than you think. *Wharton Working Paper*, 97-34.
- Diebold, F. X. and Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3):134–144.
- Drost, F. C. and Nijman, T. E. (1993). Temporal aggregation of garch processes. *Econometrica: Journal of the Econometric Society*, 61(4):909–927.
- Elliott, G. and Timmermann, A. (2008). Economic forecasting. *Journal of Economic Literature*, 46(1):3–56.
- Engle, R. F. (2002). Dynamic conditional correlation - a simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, 20(3):339–350.
- Engle, R. F. and Manganelli, S. (2004). CAViaR: Conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics*, 22(4):367–381.
- Fama, E. F. (1965). The behavior of stock market prices. *Journal of Business*, 38(1):34–105.
- Ghysels, E., Rubia, A., and Valkanov, R. (2009). Multi-period forecasts of volatility: Direct, iterated, and mixed-data approaches. Working paper, Rady School of Management, University of California San Diego.
- Ghysels, E., Santa-Clara, P., and Valkanov, R. (2006). Predicting volatility: Getting the most out of return data sampled at different frequencies. *Journal of Econometrics*, 131(1):59–95.
- Giacomini, R. and Komunjer, I. (2005). Evaluation and combination of conditional quantile forecasts. *Journal of Business & Economic Statistics*, 23(4):416–431.
- Giacomini, R. and White, H. (2006). Tests of conditional predictive ability. *Econometrica*, 74(6):1545–1578.
- Giot, P. and Laurent, S. (2004). Modelling daily value-at-risk using realized volatility and ARCH type models. *Journal of Empirical Finance*, 11(3):379–398.
- Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*, 48(5):1779–1801.
- Hafner, C. M. (2008). Temporal aggregation of multivariate garch processes. *Journal of Econometrics*, 142(1):467–483.
- Hansen, B. E. (1994). Autoregressive conditional density estimation. *International Economic Review*, 35(3):705–730.

- Hansen, P. R. and Lunde, A. (2005). A forecast comparison of volatility models: does anything beat a GARCH (1, 1)? *Journal of Applied Econometrics*, 20(7):873–889.
- Hansen, P. R., Lunde, A., and Nason, J. M. (2011). The model confidence set. *Econometrica*, 79(2):453–497.
- Herwartz, H. and Waichman, I. (2010). A comparison of bootstrap and monte-carlo testing approaches to Value-at-Risk diagnosis. *Computational Statistics*, 25(4):725–732.
- Joe, H. (1997). *Multivariate Models and Dependence Concepts*, volume 73 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London, UK.
- Jondeau, E. and Rockinger, M. (2006). The copula-GARCH model of conditional dependencies: An international stock market application. *Journal of International Money and Finance*, 25(5):827–853.
- JP Morgan and Reuters (1996). Riskmetrics—technical document. Technical report, Morgan Guaranty Trust Company and Reuters Ltd.
- Kuester, K., Mittnik, S., and Paolella, M. S. (2006). Value-at-risk prediction: A comparison of alternative strategies. *Journal of Financial Econometrics*, 4(1):53–89.
- Kupiec, P. (1995). Techniques for verifying the accuracy of risk measurement models. *The Journal of Derivatives*, 3(2):73–84.
- Lau, H.-S. and Wingender, J. R. (1989). The analytics of the intervaling effect on skewness and kurtosis of stock returns. *The Financial Review*, 24(2):215–233.
- Laurent, S., Rombouts, J. V., and Violante, F. (2012). On the forecasting accuracy of multivariate garch models. *Journal of Applied Econometrics*, 27(6):934–955.
- 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):1–10.
- Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business*, 36(4):394–419.
- Marcellino, M., Stock, J. H., and Watson, M. W. (2003). Macroeconomic Forecasting in the Euro Area. *European Economic Review*, 47(1):1–18.
- Marcellino, M., Stock, J. H., and Watson, M. W. (2006). A comparison of direct and iterated multistep ar methods for forecasting macroeconomic time series. *Journal of Econometrics*, 135(1):499–526.
- McAleer, M. (2009). The ten commandments for optimizing value-at-risk and daily capital charges. *Journal of Economic Surveys*, 23(5):831–849.
- McAleer, M. and Da Veiga, B. (2008). Single-index and portfolio models for forecasting value-at-risk thresholds. *Journal of Forecasting*, 27(3):217.
- Mina, J. and Xiao, J. Y. (2001). Return to RiskMetrics: The evolution of a standard. Technical report, RiskMetrics, New York, NY, USA.
- Mittnik, S. and Paolella, M. S. (2000). Conditional density and value-at-risk prediction of Asian currency exchange rates. *Journal of Forecasting*, 19(4):313–333.
- Newey, W. K. and West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708.

- Patton, A. (2013). Copula methods for forecasting multivariate time series. In Elliott, G. and Timmermann, A., editors, *Handbook of Economic Forecasting*, volume 2B, chapter 16, pages 899–960. Elsevier North Holland, Amsterdam, Netherlands.
- Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160(1):246–256.
- Patton, A. J. and Sheppard, K. (2009). Evaluating volatility and correlation forecasts. In Andersen, T., Davis, R., Kreiss, J.-P., and Mikosch, T., editors, *Handbook of Financial Time Series*, pages 801–838. Springer-Verlag, Berlin, Heidelberg.
- Pesaran, M. H., Pick, A., and Timmermann, A. (2011). Variable selection, estimation and inference for multi-period forecasting problems. *Journal of Econometrics*, 164(1):173–187.
- Santos, A. A. P., Nogales, F. J., and Ruiz, E. (2013). Comparing univariate and multivariate models to forecast portfolio value-at-risk. *Journal of Financial Econometrics*, 11(2):400–441.
- Sbrana, G. and Silvestrini, A. (2013). Aggregation of exponential smoothing processes with an application to portfolio risk evaluation. *Journal of Banking & Finance*, 37(5):1437–1450.
- Silvennoinen, A. and Teräsvirta, T. (2009). Multivariate GARCH models. In Andersen, T. G., Davis, R. A., Kreiss, J.-P., and Mikosch, T., editors, *Handbook of Financial Time Series*, pages 201–229. Springer-Verlag, Berlin, Heidelberg.
- Wang, J.-N., Yeh, J.-H., and Cheng, N. Y.-P. (2011). How accurate is the square-root-of-time rule in scaling tail risk: A global study. *Journal of Banking & Finance*, 35(5):1158–1169.
- Zaffaroni, P. (2008). Large-scale volatility models: theoretical properties of professionals’ practice. *Journal of Time Series Analysis*, 29(3):581–599.

Table 1: Summary results of unconditional coverage and dynamic quantile tests

# methods	UC-test results rejections at			DQ-test results				
	1%	5%	10%	neg.	neg. sign.	pos. sign.	pos.	
(a) Temporal aggregation								
Daily, iterated	51	1	15	22	46	3	5	0
Daily, scaled	51	12	29	37	46	2	5	0
Weekly, iterated	51	7	32	37	23	0	28	2
Weekly, scaled	51	18	34	37	23	0	28	3
Biweekly, direct	51	26	39	39	0	0	51	30
(b) Portfolio aggregation								
Portfolio level	45	15	30	37	18	3	27	10
Asset class level	105	24	57	65	47	0	58	17
Asset level	105	25	62	70	73	2	32	8
(c) Model specification								
GARCH, uv.	15	6	13	14	6	0	9	3
GJR, uv.	15	5	9	12	6	1	9	3
RiskMetrics, uv.	15	4	8	11	6	2	9	7
CCC-GARCH	30	8	19	21	17	0	13	1
DCC-GARCH	30	10	20	20	20	0	10	3
HDCC-GARCH	30	12	20	22	19	0	11	3
CCC-GJR	30	5	15	19	22	2	8	1
DCC-GJR	30	6	15	17	21	0	9	4
HDCC-GJR	30	6	15	18	19	0	11	3
RiskMetrics, mv.	30	2	15	18	2	0	28	7
(d) Distribution								
Normal	85	50	75	78	44	0	41	7
Empirical	85	2	7	21	47	2	38	18
Student's t	85	12	67	73	47	3	38	10

This table summarizes the results of the unconditional coverage (UC) and dynamic quantile (DQ) tests of Christoffersen (1998) and Engle and Manganelli (2004) for the 255 different methods that we analyze. In each panel, we categorize the methods based on the choice aspect given in the panel heading. For each option that can be chosen, we report the total number of methods. Next, we report for the number of rejections from the UC tests for significance levels of 1, 5 and 10%. Under the null (alternative) hypothesis, the actual coverage ratio equals (exceeds) the theoretical coverage ratio of 0.01. To conduct the DQ test, we regress the centered hit series on the negative VaR forecasts. We report the number of negative and positive coefficients, and the number of significant ones at the 10% level. Significance is determined by bootstraps as in Herwartz and Waichman (2010). The full results are available in Tables C.1 and C.2 in Appendix C.1.

Table 2: Average asymmetric tick loss

(a) daily returns, iterated forecasts									
Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-GARCH	5.84	5.47	5.59	5.81	5.39	5.41	5.82	5.47	5.54
DCC-GARCH		5.54	5.51		5.43	5.34		5.50	5.46
HDCC-GARCH		5.60	5.58		5.47	5.38		5.54	5.56
CCC-GJR	5.54	5.30	5.27	5.51	5.28	5.18	5.48	5.27	5.26
DCC-GJR		5.34	5.17		5.33	5.10		5.23	5.18
HDCC-GJR		5.38	5.24		5.39	5.20		5.35	5.22
RiskMetrics	5.59	6.02	6.07	5.66	5.96	5.82	5.64	6.00	6.13

(b) daily returns, scaled forecasts									
Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-GARCH	5.92	5.59	5.65	5.66	5.39	5.35	5.78	5.46	5.47
DCC-GARCH		5.66	5.58		5.50	5.33		5.53	5.42
HDCC-GARCH		5.70	5.66		5.47	5.31		5.56	5.48
CCC-GJR	5.74	5.59	5.54	5.59	5.48	5.29	5.66	5.45	5.35
DCC-GJR		5.64	5.44		5.51	5.21		5.40	5.28
HDCC-GJR		5.68	5.47		5.56	5.25		5.56	5.32
RiskMetrics	5.74	6.06	6.13	5.60	5.96	6.00	5.64	6.00	6.06

(c) weekly returns, iterated forecasts									
Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-GARCH	7.13	6.69	6.73	6.86	6.33	6.34	6.88	6.58	6.68
DCC-GARCH		6.70	6.67		6.32	6.29		6.50	6.62
HDCC-GARCH		6.68	6.74		6.36	6.39		6.58	6.67
CCC-GJR	6.62	6.47	6.40	6.43	6.21	6.10	6.38	6.24	6.29
DCC-GJR		6.50	6.36		6.22	6.11		6.25	6.25
HDCC-GJR		6.48	6.41		6.26	6.11		6.32	6.29
RiskMetrics	7.08	7.07	7.33	6.97	7.00	7.51	6.69	7.07	7.27

This table shows the average value of the asymmetric tick loss function of Giacomini and Komunjer (2005) as in Equation (22) for $\vartheta = 0.01$. All values have been multiplied by 100.

Table 2: Average asymmetric tick loss – *continued*

(d) weekly returns, scaled forecasts									
Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-GARCH	7.28	6.79	6.79	6.75	6.37	6.33	6.84	6.61	6.63
DCC-GARCH		6.79	6.75		6.36	6.30		6.58	6.62
HDCC-GARCH		6.81	6.80		6.42	6.31		6.60	6.66
CCC-GJR	6.75	6.68	6.55	6.27	6.18	6.07	6.37	6.32	6.32
DCC-GJR		6.69	6.50		6.19	6.02		6.34	6.27
HDCC-GJR		6.71	6.57		6.19	6.05		6.37	6.34
RiskMetrics	7.15	7.06	7.36	6.90	6.98	7.30	6.67	7.04	7.22

(e) biweekly returns, direct forecasts									
Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-GARCH	7.42	7.16	7.18	7.55	7.60	7.16	7.22	7.06	7.04
DCC-GARCH		7.12	7.17		7.57	7.16		7.03	7.01
HDCC-GARCH		7.17	7.19		7.61	7.20		7.05	7.00
CCC-GJR	7.17	7.00	6.95	7.15	7.33	6.90	6.88	6.90	6.82
DCC-GJR		6.96	6.93		7.28	6.89		6.82	6.81
HDCC-GJR		7.01	6.96		7.29	6.92		6.88	6.80
RiskMetrics	7.59	7.59	8.45	7.80	8.59	8.95	7.48	7.66	8.40

See table note on previous page.

Table 3: Methods removed from the Model Confidence Set

(a) weekly returns, iterated forecasts									
Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-GARCH	0.087			0.087					
DCC-GARCH									
HDCC-GARCH									
CCC-GJR									
DCC-GJR									
HDCC-GJR									
RiskMetrics	0.098			0.089	0.082		0.070	0.073	

(b) weekly returns, scaled forecasts									
Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-GARCH	0.086			0.079			0.096		
DCC-GARCH									
HDCC-GARCH				0.095					
CCC-GJR	0.089	0.096							
DCC-GJR			0.091						
HDCC-GJR			0.088						
RiskMetrics	0.089			0.082	0.073	0.065	0.075		

(c) biweekly returns, direct forecasts									
Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-GARCH	0.082			0.060	0.054	0.078	0.084		
DCC-GARCH							0.091		
HDCC-GARCH	0.097						0.089		
CCC-GJR	0.087			0.069	0.057	0.081	0.087		
DCC-GJR							0.090		
HDCC-GJR							0.087		
RiskMetrics	0.083		0.091	0.066	0.074	0.075	0.068		0.091

(d) Number of removed and maintained methods			
Temporal aggr.	Out	In	Total
Daily, iterated	0	51	51
Daily, scaled	0	51	51
Weekly, iterated	7	44	51
Weekly, scaled	13	38	51
Biweekly, direct	31	20	51

Portfolio aggr.	Out	In	Total
Portfolio level	21	24	45
Asset class level	19	86	105
Asset level	11	94	105

Disitribution	Out	In	Total
Normal	13	72	85
Empirical	26	59	85
Student's t	12	73	85

Model	Out	In	Total
GARCH, uv.	8	7	15
GJR, uv.	4	11	15
RiskMetrics, uv.	9	6	15
CCC-GARCH	2	28	30
DCC-GARCH	3	27	30
HDCC-GARCH	5	25	30
CCC-GJR	4	26	30
DCC-GJR	4	26	30
HDCC-GJR	4	26	30
RiskMetrics, mv.	8	22	30

This table presents the MCS p -value with which a model has been removed from the Model Confidence Set. We follow the procedure of Hansen et al. (2011, Sec. 3.1.2), with the tick loss function in (22), $\vartheta = 0.01$, and a significance level of 10%. The procedure starts with the complete set of 255 methods. Panels for daily observations combined with iterated or scaled forecasts are absent, because no methods with these combinations have been removed. Panel (d) presents the number of methods with a particular choice that have been removed from (“Out”) and are maintained in (“In”) the Model Confidence Set.

Table 4: Results of the Diebold-Mariano tests

Method A	Method B	# of tests	neg.	sign. 5%	neg. 10%	pos.	sign. 5%	pos. 10%
(a) Forecasting method								
Iterated	Scaled	51	38	6	14	13	0	0
(b) Portfolio aggregation, iterated forecasts								
Asset level	Asset class level	21	15	1	3	6	0	0
Asset level	Portfolio level	21	18	0	6	3	1	2
Asset class level	Portfolio level	21	18	0	0	3	1	2
(c) Portfolio aggregation, scaled forecasts								
Asset level	Asset class level	21	16	1	3	5	0	0
Asset level	Portfolio level	21	18	0	6	3	1	2
Asset class level	Portfolio level	21	18	0	0	3	1	2
(d) Distribution, iterated forecasts								
Normal	Empirical	17	4	0	0	13	1	2
Normal	Student's t	17	2	0	1	15	0	3
Empirical	Student's t	17	12	0	2	5	1	1
(e) Distribution, scaled forecasts								
Normal	Empirical	17	0	0	0	17	1	2
Normal	Student's t	17	0	0	1	17	0	3
Empirical	Student's t	17	15	0	2	2	1	1
(f) Model, iterated forecasts								
RiskMetrics, uv.	other, uv.	6	3	0	0	3	0	0
RiskMetrics, mv.	other, mv.	36	0	0	0	36	15	32
GARCH, uv.	GJR, uv.	3	0	0	0	3	0	1
GARCH, mv.	GJR, mv.	54	1	0	0	53	0	20
CCC	DCC	12	5	0	0	7	0	0
CCC	HDCC	12	8	0	1	4	0	0
DCC	HDCC	12	12	0	3	0	0	0
(g) Model, scaled forecasts								
RiskMetrics, uv.	other, uv.	6	5	0	0	1	0	0
RiskMetrics, mv.	other, mv.	36	0	0	0	36	12	28
GARCH, uv.	GJR, uv.	3	0	0	0	3	0	0
GARCH, mv.	GJR, mv.	54	14	0	0	40	2	14
CCC	DCC	12	5	0	0	7	0	1
CCC	HDCC	12	8	0	0	4	0	0
DCC	HDCC	12	10	1	5	2	0	0

This table shows summary results of the tests that the expected value of the loss function of methods A and B are equal, $L^A = L^B$. We calculate the loss differential as $L^A - L^B$. A negative loss differential means that method A is preferable to method B. We use the asymmetric tick loss function of Giacomini and Komunjer (2005) as in Equation (22) with $\vartheta = 0.01$. We report the number of negative and positive average loss differentials, and the number of times these differentials are significant for the 5% and 10% significance levels, based on the statistic proposed by Diebold and Mariano (1995), evaluated in the setting of Giacomini and White (2006). Standard errors of the average loss differential are based on Newey and West (1987) with 10 leads and lags. We compare pairs of methods that differ in only one choice aspect (indicated by the panel heading), so the other four choices are the same. All methods use daily data.

A Data Details

A.1 Portfolio construction

This section contains detailed information on the construction of the typical pension fund portfolio. For all series, returns in USD are based on the closing prices of each day. All return series are obtained from Thomson Reuters Datastream. For the returns on the US bond portfolios we use the Barclays US Aggregate Bond Index (Mnemonic: LHGOVBD) and the Barclays Capital US Corporate Bond Index (Mn: LHCCORP). The Aggregate Bond Index is a predominate index benchmark for US bond investors, and acts as a benchmark index for many US index funds. It comprises four major subindexes: US Government Index, US Credit Index, US Mortgage Backed Securities Index and US Asset Backed Securities Index. The Corporate Bond Index covers investment-grade bonds that are denominated in US Dollar.

We use the MSCI Europe Index to cover the European stock market (Mn: MSEROP\$). This index captures large and mid cap representation across the following 15 Developed Markets (DM) countries in Europe: Austria, Belgium, Denmark, Finland, France, Germany, Ireland, Italy, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland and the UK. The US equity part is based on the MSCI USA Index (Mn: MSUSAML), a free float adjusted market capitalization index that is designed to measure large and mid cap US equity market performance. The equity part for the Pacific area represents the large and mid cap across Australia, Hong Kong, Japan, New Zealand and Singapore (Mn: MSPACF\$). Likewise, the MSCI EM Index summarizes the equity returns from 23 Emerging Markets (Mn: MSEMKF\$). All retur

The return on the real estate investment is constructed from the FTSE EPRA/NAREIT Developed that incorporates Real Estate Investment Trusts (REITs) and Real Estate Holding & Development companies (Mn: FEGLOB\$). The index represent general trends in eligible real estate equities worldwide.

The investment in commodities is approximated by the the S&P GSCI Index (formerly the Goldman Sachs Commodity Index) (Mn: GSCITOT). This Index serves as a benchmark for investment in the commodity markets and as a measure of commodity performance over time. Currently, it comprises 24 commodities from all commodity sectors - energy products, industrial metals, agricultural products, livestock products and precious metals with a high exposure to energy products.

A.2 Initial data analysis

The summary statistics for daily returns in Table A.1 indicate substantial differences between bonds and the other assets. Bonds show a considerably lower return than the other asset classes. The lowest average return comes from Pacific equities (-0.50% per year), the highest from US equities. Bond volatilities range from 4.48% to 5.40% per year, whereas the volatilities of the other classes range from 16.57% to 21.71%. Almost all series exhibit negative skewness (between -0.18 and -0.61), except the Pacific equity series. Excess kurtosis is present in each return series (ranging from 2.48 to 10.18). Consequently, the Jarque-Bera test rejects the hypothesis of normality for all series. The significance of Ljung-Box statistics indicates significant autocorrelation for all series. The autocorrelation is largely driven by the first lag and is positive for all series, except for US equity and for commodity returns.

[Table A.1 about here.]

Because we are interested in the effects of temporal and portfolio aggregation on predictive densities in general and risk in particular, we calculate summary statistics for the portfolio as a whole, and for weekly and biweekly returns. Of course, the average portfolio return corresponds directly with a weighted average of the returns of each asset class. The volatilities of the portfolio returns shows the effects of diversification, as it is much lower than the non-bond classes. However, diversification breaks down when extreme returns occur. The portfolio returns are stronger left skewed than almost all constituent series, and the kurtosis coefficient is close to the maximum among them. Clustering over time is also stronger for the portfolio, with a high first order autocorrelation coefficient and a high Ljung-Box statistic.

Using weekly or biweekly returns instead of daily returns has a small effect on the volatility of the series. If returns were i.i.d., the scaled volatilities would be independent of the interval over which the returns are observed. When returns exhibit positive (negative) autocorrelation, the volatility of aggregated returns will be higher (lower). For Emerging Market equities, real estate, and the portfolio as a whole, we see an increase in volatility, which implies positive autocorrelation. For the other classes, volatilities go down, implying negative autocorrelation.

The coefficients for skewness and kurtosis show large deviations from the values that would result if daily returns were i.i.d. Lau and Wingender (1989) derive that when k i.i.d. returns are aggregated, the skewness coefficient is scaled by $1/\sqrt{k}$ and the coefficient of excess kurtosis by $1/k$. Instead of a decrease for skewness, the left skewness of all series increases for biweekly and weekly returns compared to daily returns. The kurtosis coefficients for weekly

returns decrease, but much less than for the i.i.d. case. These results indicate that extreme returns, and in particular negative extreme returns tend to cluster. Also, the hypothesis of a normal distribution is rejected for (bi)weekly data.

The statistics for the portfolio as a whole show that extreme and negative returns tend to cluster over time and in the cross section. The volatility estimates do not vary much with the frequency of observation, and are considerably lower than the average volatilities of the asset classes. When returns are aggregated, the degree of left skewness goes up instead of down, and the coefficient of excess kurtosis does not decrease either. Hence, while scaling daily or weekly volatilities to biweekly volatilities may work, scaling the density seems to be incorrect, in particular when risk measures are needed. The results for the portfolio returns show that cross-sectional aggregation does not mitigate this effect.

We show the correlations between the different asset classes at different frequencies in Table A.2. We see a clear block structure in the correlation matrix. Corporate bonds and government bonds are highly correlated with each other, but correlations with the other assets is slightly negative. The equity returns are also highly correlated with each other and with real estate. Commodities is a category on its own. Aggregating the assets into the larger categories of bonds, equities and alternatives corresponds with this block structure.

[Table A.2 about here.]

Correlations that are calculated with weekly or a biweekly returns generally show the same picture as the correlations for daily data, though some equity correlations show a substantial increase. This applies in particular to all correlations with Pacific Equities, reflecting that daily returns are not synchronized.

We conclude from this initial analysis that asset returns show several forms of temporal dependence. Linear dependence over time, as measured by autocorrelation, is small but significant. Scaling volatility is therefore not correct, but deviations are generally not very large. The results for skewness and kurtosis on the other hand are more troublesome. Instead of going down, both measures increase, both for temporal and for portfolio aggregation. These increases indicate that extreme, and in particular extreme negative returns tend to cluster. As a consequence, risks are larger for a (bi)weekly than for a daily horizon. Simply scaling daily risk measures might therefore not be appropriate.

Table A.1: Summary Statistics of Asset and Portfolio Returns

	Govt. bonds	Corp. bonds	EU equities	US equities	Pacific equities	EM equities	real estate	commo- dities	portfolio
Weight (in %)	30	20	12	10.5	3.75	3.75	10	10	100

(a) daily returns									
Mean (% p.a.)	-0.02	0.45	5.24	7.06	-0.50	6.24	3.16	4.00	2.38
Volatility (% p.a.)	4.48	5.40	20.18	18.35	21.24	19.15	16.57	21.71	6.87
Skewness	-0.18	-0.27	-0.14	-0.30	-0.01	-0.54	-0.50	-0.61	-0.59
Excess kurtosis	2.19	2.48	9.24	8.43	5.27	10.18	10.13	8.73	9.45
Jarque-Bera	1240	1610	21485	17963	6974	26308	26026	19527	22796
ACF(1)	0.022	0.004	0.007	-0.065	0.015	0.212	0.137	-0.012	0.154
ACF(2)	-0.031	-0.012	-0.039	-0.023	-0.031	0.044	0.035	0.004	0.008
ACF(3)	-0.018	-0.008	-0.022	0.010	-0.016	0.018	0.026	-0.002	0.016
Ljung-Box	46.80	35.30	77.76	87.67	49.94	371.1	176.8	39.64	205.0

(b) weekly returns									
Volatility (% p.a.)	4.42	5.33	19.73	17.11	20.98	22.73	18.74	21.55	7.74
Skewness	-0.25	-0.39	-0.73	-0.66	-0.20	-0.80	-1.05	-0.60	-0.98
- implied by daily	-0.08	-0.12	-0.06	-0.13	0.00	-0.24	-0.22	-0.27	-0.26
Excess kurtosis	0.85	2.04	5.01	6.10	2.42	5.59	8.79	3.30	7.56
- implied by daily	0.44	0.50	1.85	1.69	1.05	2.04	2.03	1.75	1.89
Jarque-Bera	49.59	262.2	1729	2103	329.4	1927	4235	661.3	3645
ACF(1)	-0.064	-0.027	-0.079	-0.103	-0.052	0.029	0.010	-0.032	0.008
ACF(2)	0.058	0.090	0.025	0.045	0.030	0.092	0.058	0.043	0.067
Ljung-Box	30.31	40.13	50.31	42.63	31.42	54.99	53.71	37.13	48.96

(c) biweekly returns									
Volatility (% p.a.)	4.27	5.26	18.93	16.20	20.42	23.05	18.82	21.20	7.77
Skewness	-0.28	-0.64	-1.06	-0.97	-0.21	-0.97	-1.35	-0.68	-1.45
- implied by weekly	-0.18	-0.27	-0.51	-0.47	-0.14	-0.56	-0.74	-0.42	-0.69
- implied by daily	-0.06	-0.08	-0.04	-0.09	0.00	-0.17	-0.16	-0.19	-0.19
Excess kurtosis	0.66	2.73	5.41	6.47	2.34	4.22	9.28	2.92	10.63
- implied by weekly	0.42	1.02	2.50	3.05	1.21	2.80	4.40	1.65	3.78
- implied by daily	0.22	0.25	0.92	0.84	0.53	1.02	1.01	0.87	0.95
Jarque-Bera	22.87	235.8	1004	1483	188.2	587.5	2635	285.4	3538
ACF(1)	0.05	0.11	0.00	-0.01	0.02	0.13	0.09	0.07	0.09
Ljung-Box	25.74	27.55	29.57	23.66	21.73	42.97	45.06	28.32	36.60

Means and volatilities for daily returns in panel(a) are annualized assuming 250 trading days. $ACF(q)$ denotes the autocorrelation for lag q . Ljung-Box statistics are calculated for 20 lags. Critical values for 95% and 99% confidence levels are 31.4 and 37.6 based on a χ^2_{20} distribution. The statistics in panel (b) are the averages of the statistics for each day of the week. The statistics in panel (c) are the averages of the statistics for each of the ten possible starting days. Implied coefficients of skewness and of excess kurtosis for the aggregation of k returns are calculated by scaling the non-aggregated coefficients by $1/\sqrt{k}$ and $1/k$, based on Lau and Wingender (1989).

Table A.2: Correlations of Asset Returns

	Govt. bonds	Corp. bonds	EU equities	US equities	Pacific equities	EM equities	Real estate	Commo- dities
(a) daily returns								
Govt. bonds	1	0.91	-0.15	-0.16	-0.02	-0.14	-0.13	-0.14
Corp. bonds		1	-0.08	-0.06	0.06	-0.02	-0.04	-0.12
EU equities			1	0.51	0.12	0.43	0.56	0.14
US equities				1	0.38	0.63	0.61	0.25
Pac equities					1	0.53	0.52	0.10
EM equities						1	0.62	0.22
Real estate							1	0.19
Commodities								1
(b) weekly returns								
Govt. bonds	1	0.85	-0.11	-0.10	-0.06	-0.14	-0.05	-0.15
Corp. bonds		1	0.05	0.06	0.07	0.05	0.12	-0.07
EU equities			1	0.74	0.59	0.73	0.74	0.25
US equities				1	0.46	0.64	0.67	0.18
Pacific equities					1	0.61	0.67	0.19
EM equities						1	0.71	0.25
Real estate							1	0.23
Commodities								1
(c) biweekly returns								
Govt. bonds	1	0.80	-0.10	-0.10	-0.07	-0.13	-0.02	-0.16
Corp. bonds		1	0.15	0.16	0.13	0.15	0.24	-0.03
EU equities			1	0.77	0.62	0.73	0.75	0.26
US equities				1	0.53	0.67	0.69	0.18
Pacific equities					1	0.63	0.68	0.22
EM equities						1	0.71	0.27
Real estate							1	0.25
Commodities								1

The correlations in panel (b) are the averages of the correlations based on weekly returns for each day of the week. The correlations in panel (c) are the averages of the correlations based on biweekly returns for each of the ten possible starting days.

B Methodological details

B.1 Esimation

In this subsection, we discuss the parameter estimation for all combinations of univariate and multivariate models on the one hand and distributions on the other hand.

Univariate case

We first estimate the parameters of the univariate version of Equation (5) by ordinary least squares (OLS), and construct a series of residuals $\hat{\varepsilon}_{t,k}^b$. Based on this series, we estimate the parameters of the GJR or GARCH model in Equation (9) or of the RiskMetrics model in Equation (13) as the second pass. When these models are combined with the normal or Student's t -distribution for the innovations we use Maximum Likelihood (ML) estimation. When the Student's t -distribution is used, we base the likelihood function on the scaled pdf in Equation (14). When these models are combined with the empirical distribution, we use Quasi Maximum Likelihood (QML) estimation. For all models, we use variance targeting to determine $\omega_{i,k}^b$.

Multivariate RiskMetrics

We estimate the parameters of the VAR-model of Equation (5) by OLS, and construct a series of residuals $\hat{\varepsilon}_{t,k}^b$. In the second pass, we estimate the parameters in Equation (13) based on this series. When RiskMetrics is combined with the normal or Student's t -distribution, we apply ML estimation. In the case of the Student's t -distribution, the likelihood function is scaled, as in Equation (16). When the empirical distribution is chosen, we use QML estimation.

CCC model

We estimate the parameters of the VAR-model of Equation (5) by OLS, and construct a series of residuals $\hat{\varepsilon}_{t,k}^b$. In the second pass, we estimate the parameters of the marginal models (GJR or GARCH) for each series $\hat{\varepsilon}_{i,t,k}^b$, $i = 1, 2, \dots, m_b$ in the same way as in the univariate case, and construct series of standardized residuals $\hat{\eta}_{i,t,k}^b$. The third pass depends on the distribution. When the normal distribution is used, we estimate the correlation matrix by the sample correlation of $\hat{\eta}_{t,k}^b$ (cf. Bollerslev, 1990). When the Student's t or empirical distribution is used, we first apply the probability integral transform (PIT) to the residuals of each marginal model based on Equation (14) or Equation (17). Next, we use ML based

on the transformed series to estimate the parameters of the Student's t in Equation (15) (cf. Patton, 2013, Sec. 3.1) or the parameters of the Gaussian copula in Equation (18) (cf. Patton, 2013, Sec. 3.2; Chen and Fan, 2006).

DCC model

In this case, we follow the same first two passes as for the CCC model. The final pass depends again on the distribution. When the normal distribution is used, we estimate the DCC parameters in (10) by ML estimation based on the normal distribution function (cf. Engle, 2002). For the other two distributions, we apply the PIT, and then use ML based on the transformed series to estimate the DCC parameters (and the degrees of freedom in the case of the Student's t -copula). In all methods we use correlation targeting as proposed by Aielli (2013).

B.2 Forecasting methods

We build the forecasts of the portfolio return distribution from forecasts for the expected returns, and for the unexpected part, $u_{t,h}^p \equiv r_{t,h}^p - E_t[r_{t,h}^p]$. The Value-at-Risk for the portfolio return follows as

$$\text{VaR}_\alpha(r_{t,h}^p) = -E_t[r_{t,h}^p] + \text{VaR}_\alpha(u_{t,h}^p). \quad (\text{B.1})$$

The expected portfolio return obeys

$$E_t[r_{t,h}^p] = \mathbf{w}' \boldsymbol{\mu}_{t,h/k,k}, \quad (\text{B.2})$$

where $\boldsymbol{\mu}_{t,\kappa,k}$ is the conditional κ -steps ahead cumulative expected return with $\kappa = h/k$ that follows from Equation (5).¹⁴ We show in Appendix B.3 that

$$\begin{aligned} \boldsymbol{\mu}_{t,\kappa,k} &\equiv E_t \left[\sum_{j=0}^{\kappa-1} \mathbf{r}_{t+jk,k} \right] \\ &= (\mathbf{I}_m - \boldsymbol{\Gamma})^{-2} (\kappa \mathbf{I}_m - (\kappa + 1) \boldsymbol{\Gamma} + \boldsymbol{\Gamma}^{\kappa+1}) \boldsymbol{\phi} + \\ &\quad (\mathbf{I}_m - \boldsymbol{\Gamma})^{-1} (\boldsymbol{\Gamma} - \boldsymbol{\Gamma}^{\kappa+1}) \mathbf{r}_{t-k,k}. \end{aligned} \quad (\text{B.3})$$

The forecasts for the unexpected part are derived in closed form or based on simulations. We discuss the different cases in the next subsections.

¹⁴We suppress the superscript b that indicates the basic portfolios to ease the notation in this subsection.

B.2.1 One-step-ahead forecasts

One-step ahead forecasts result when full temporal aggregation is applied to the modeled returns. They are also used when forecasts are scaled. For some cases, VaR can be calculated analytically, for others simulations are needed. We provide an overview below, and discuss the construction for a one-step-ahead forecast k

	Normal	Student's t		Empirical	
		$m_b = 1$	$m_b > 1$	$m_b = 1$	$m_b > 1$
CCC or DCC	Closed form	Closed form	Simulation C	Closed form	Simulation C
RiskMetrics	Closed form	Closed form	Closed form	Closed form	Simulation E

In the closed-form cases, Value-at-Risk is calculated as

$$\text{VaR}_\alpha(u_{t,k}^p) = -\sqrt{\mathbf{w}' \boldsymbol{\Sigma}_{t,k} \mathbf{w}} F^{-1}(\alpha; \boldsymbol{\theta}). \quad (\text{B.4})$$

Here $\boldsymbol{\Sigma}_{t,k}$ denotes the forecasted variance, which depends on the specification, CCC, DCC or RiskMetrics. F^{-1} denotes the inverse of a standardized cumulative distribution function, and $\boldsymbol{\theta}$ is a vector containing all parameters needed to calculate this inverse. When a normal distribution is assumed, this vector is empty. When the Student's t -distribution is used, the cdf should be based on Equation (14), and $\boldsymbol{\theta}$ contains the degree of freedom parameter. When the empirical distribution is used, $\boldsymbol{\theta}$ contains a set of realizations (cf. Equation (17)). We use here that the normal and Student's t -distribution are closed under summation.

When the CCC or DCC approaches are combined with the Student's t - or empirical distribution, copula-based simulations are needed to calculate $\text{VaR}_\alpha(u_{t,k})$. Each simulation s consist of four steps.

- C1. Generate a one-step-ahead forecast for the correlation matrix $\mathbf{R}_{t,k}$ and the volatilities $\sigma_{i,t,k}$.
- C2. Draw a random $m_b \times 1$ vector $\tilde{\mathbf{v}}_s$ from the copula (Gaussian or Student's t -), using the forecasted correlation matrix $\mathbf{R}_{t,k}$.
- C3. Transform each draw $\tilde{v}_{s,i}$ by applying the standardized inverse cdf of the respective marginal distribution, $\tilde{\eta}_{s,i} = F_i^{-1}(\tilde{v}_{s,i}; \boldsymbol{\theta}_i)$.
- C4. Multiply $\tilde{\eta}_{s,i}$ by the forecasted volatility, $\tilde{\varepsilon}_{s,i} = \sigma_{i,t,k} \tilde{\eta}_{s,i}$.
- C5. Pre-multiply the vector $\tilde{\boldsymbol{\varepsilon}}_s$ by the weights to produce a random draw from the distribution of $u_{t,h}^p$, $\tilde{u}_s^p = \mathbf{w}' \tilde{\boldsymbol{\varepsilon}}_s$.

Repeating these steps S times approximates the distribution of $u_{t,k}^p$. The α -quantile of the simulated distribution is a forecast for $\text{VaR}_\alpha(u_{t,k}^p)$.

When the empirical distribution is combined with RiskMetrics, a different simulation procedure is needed.

- E1. Generate a one-step-ahead forecast for the variance matrix $\Sigma_{t,k}$.
- E2. Take a random draw $\tilde{\zeta}_s$ from the empirical distribution that corresponds with $\zeta_{t,k}$.
- E3. Multiply $\tilde{\zeta}_s$ with the Cholesky decomposition of the forecasted variance matrix, $\tilde{\varepsilon}_s = \Sigma_{t,k}^{1/2} \tilde{\zeta}_s$.
- E4. As in step C5 of the copula-based simulation.

This procedure is again repeated S times to approximate the distribution of $u_{t,k}^p$, from which the α -quantile is taken.

B.2.2 Iterated multi-step-ahead forecasts

Iterated forecasts take the complete serial dependence that is implied by the different models into account. Because of this serial dependence, closed-form expression for VaR are not available and simulations are needed. For the CCC or DCC models, a simulation s consists of the following steps

- I1. Set $j = 1$. Generate a one-step-ahead forecast for the correlation matrix $\mathbf{R}_{t,k,s} = \mathbf{R}_{t,k}$ and the volatilities $\sigma_{i,t,k,s} = \sigma_{i,t,k}$.
- I2. Draw a random $m_b \times 1$ vector $\tilde{\mathbf{v}}_{s,j}$ from the copula (Gaussian or Student's t -), using the forecasted correlation matrix $\mathbf{R}_{t+(j-1)k,k,s}$.
- I3. Transform each draw $\tilde{v}_{s,j,i}$ by applying the standardized inverse cdf of the respective marginal distribution, $\tilde{\eta}_{s,j,i} = F_i^{-1}(\tilde{v}_{s,j,i}; \boldsymbol{\theta}_i)$.
- I4. Multiply $\tilde{\eta}_{s,j,i}$ by the forecasted volatility, $\tilde{\varepsilon}_{s,j,i} = \sigma_{i,t+(j-1)k,k,s} \tilde{\eta}_{s,j,i}$.
- I5. Construct one-step-ahead forecasts for the correlation matrix $\mathbf{R}_{t+jk,k,s}$ based on $\mathbf{R}_{t+(j-1)k,k,s}$ and $\tilde{\boldsymbol{\eta}}_{s,j}$, and the volatilities $\sigma_{i,t+jk,k,s}$ based on $\sigma_{i,t+(j-1)k,k,s}$ and $\tilde{\varepsilon}_{s,j,i}$. If $j < \kappa$, increase j by one, and go to step I2. Otherwise continue.
- I6. Construct the draw from the distribution $u_{t,h}^p$ as

$$\tilde{u}_s^p = \mathbf{w}' \sum_{j=1}^{\kappa} \tilde{\boldsymbol{\varepsilon}}_{s,j}.$$

Repeating these steps S times approximates the distribution of $u_{t,h}^p$. The α -quantile of the simulated distribution is a forecast for $\text{VaR}_\alpha(u_{t,h}^p)$.

For the RiskMetrics approach, the procedure for simulation s goes as follows.

- R1. Set $j = 1$. Generate a one-step-ahead forecast for the variance matrix $\boldsymbol{\Sigma}_{t,k,s} = \boldsymbol{\Sigma}_{t,k}$.
- R2. Draw a random $m_b \times 1$ vector $\tilde{\boldsymbol{\zeta}}_{s,j}$ from the standardized distribution for $\boldsymbol{\zeta}_{t,k}$.
- R3. Multiply $\tilde{\boldsymbol{\zeta}}_{s,j}$ with the Cholesky decomposition of the forecasted variance matrix, $\tilde{\boldsymbol{\epsilon}}_{s,j} = \boldsymbol{\Sigma}_{t+(j-1)k,k,s}^{1/2} \tilde{\boldsymbol{\zeta}}_{s,j}$.
- R4. Construct a one-step-ahead forecast for the variance matrix $\boldsymbol{\Sigma}_{t+jk,k,s}$ based on $\boldsymbol{\Sigma}_{t+(j-1)k,k,s}$ and $\tilde{\boldsymbol{\epsilon}}_{s,j}$. If $j < \kappa$, increase j by one, and go to step R2. Otherwise continue with step I6.

Repeating this procedure S times approximates the distribution $u_{t,h}^p$, from which the α -quantile is taken.

B.2.3 Scaled multi-step-ahead forecasts

Contrary to iterated forecasts, scaled forecasts ignore the serial dependence between the innovations in the return process. Scaled forecasts are based on the assumption that the distribution of $u_{t,h}^p$ is the same as the distribution of $u_{t,k}^p$ with a scale parameter that is adjusted for the horizon h . A scaled forecast is hence a one-step-ahead forecasts combined with an adjustment of the scale. When the one-step-ahead forecast can be constructed in closed form as in Equation (B.4), the scaled forecast takes the form

$$\text{VaR}_\alpha(u_{t,h}^p) = -\sqrt{(\mathbf{w} \otimes \mathbf{w})' \mathbf{S}_{h/k}(\boldsymbol{\Gamma}) \text{vec}(\boldsymbol{\Sigma}_{t,k})} F^{-1}(\alpha; \boldsymbol{\theta}), \quad (\text{B.5})$$

where $\mathbf{S}_{h/k}(\boldsymbol{\Gamma})$ is the $m_b^2 \times m_b^2$ scaling factor, and $\boldsymbol{\Sigma}_{t,k}$ is the one-step-ahead forecast of the variance matrix. We derive this scaling factor in Appendix B.3.

When simulations are used to construct a one-step-ahead forecast, scaling takes place in the initialization steps C1 and E1. In both cases, a scaled forecast of the variance matrix is constructed as

$$\boldsymbol{\Sigma}_{t,h} = \text{unvec}(\mathbf{S}_{h/k}(\boldsymbol{\Gamma}) \text{vec}(\boldsymbol{\Sigma}_{t,k})). \quad (\text{B.6})$$

This scaled variance matrix can be directly used in the RiskMetrics simulation procedure. For the copula simulation procedure, the (scaled) variance matrix is split in a (scaled) correlation matrix and scaled volatility forecasts.

When VAR-effects are absent, so $\mathbf{\Gamma} = \mathbf{O}$, the scaling factor simplifies to $\mathbf{S}_{h/k}(\mathbf{O}) = h/k\mathbf{I}_{m^2}$, which implies a multiplication of the variance matrix by h/k . In that case, the VaR-forecast itself can be scaled,

$$\text{VaR}_\alpha(u_{t,h}^p) = \sqrt{h/k} \text{VaR}_\alpha(u_{t,k}^p),$$

which is the familiar “square root of time”-rule.

B.3 VAR-forecasts and variance scaling

For the construction of the Value-at-Risk forecasts, analytical results for forecasts by VAR models are relevant. When the variance of the innovations in a VAR-model are constant, multi-step-ahead forecasts for the (cumulative) process can be constructed in closed form. These forecasts indicate how the variance should be scaled, moving from one- to n -step-ahead forecasts. We derive the scaling factor in this appendix (see also Campbell and Viceira, 2005).

We use slightly different notation in this section for clarity. We consider an $m \times 1$ vector \mathbf{y}_t that is observed at each period t , and follows a VAR of order 1,

$$\mathbf{y}_{t+1} = \boldsymbol{\phi} + \mathbf{\Gamma}\mathbf{y}_t + \boldsymbol{\varepsilon}_{t+1}, \quad \text{E}_t[\boldsymbol{\varepsilon}_{t+1}] = \mathbf{0}_m \quad \text{E}_t[\boldsymbol{\varepsilon}_{t+1}\boldsymbol{\varepsilon}'_{t+1}] = \boldsymbol{\Omega}, \quad (\text{B.7})$$

where $\boldsymbol{\phi}$ is an m -vector, $\mathbf{\Gamma}$ is a $m \times m$ matrix, and $\boldsymbol{\varepsilon}_{t+1}$ is an m -vector with innovations that are independent over time. We assume that the eigenvalues of $\mathbf{\Gamma}$ are inside the unit circle, and that $\boldsymbol{\Omega}$ is invertible.

The forward solution \mathbf{y}_{t+n} follows as

$$\begin{aligned} \mathbf{y}_{t+n} &= \sum_{i=0}^{n-1} \mathbf{\Gamma}^i (\boldsymbol{\phi} + \boldsymbol{\varepsilon}_{t+n-i}) + \mathbf{\Gamma}^n \mathbf{y}_t \\ &= (\mathbf{I}_m - \mathbf{\Gamma})^{-1} (\mathbf{I}_m - \mathbf{\Gamma}^n) \boldsymbol{\phi} + \mathbf{\Gamma}^n \mathbf{y}_t + \sum_{i=0}^{n-1} \mathbf{\Gamma}^i \boldsymbol{\varepsilon}_{t+n-i}, \end{aligned} \quad (\text{B.8})$$

where \mathbf{I}_m is an identity matrix of size m , and the inverse of $(\mathbf{I}_m - \mathbf{\Gamma})$ exists because the $\mathbf{\Gamma}$ is convergent. The conditional expectation of \mathbf{y}_{t+n} is thus given by

$$\text{E}[\mathbf{y}_{t+n} | \mathbf{y}_t] = (\mathbf{I}_m - \mathbf{\Gamma})^{-1} (\mathbf{I} - \mathbf{\Gamma}^n) \boldsymbol{\phi} + \mathbf{\Gamma}^n \mathbf{y}_t. \quad (\text{B.9})$$

Next, we define the cumulative process

$$\mathbf{z}_{t,n} \equiv \sum_{i=1}^n \mathbf{y}_{t+i}. \quad (\text{B.10})$$

Its expectation conditional on a value \mathbf{y}_t follows as

$$\begin{aligned} \mathbb{E}[\mathbf{z}_{t,n} | \mathbf{y}_t] &= \sum_{i=1}^n \mathbb{E}[\mathbf{y}_{t+i} | \mathbf{y}_t] = \sum_{i=1}^n ((\mathbf{I}_m - \mathbf{\Gamma})^{-1}(\mathbf{I} - \mathbf{\Gamma}^i)\boldsymbol{\phi} + \mathbf{\Gamma}^i \mathbf{y}_t) \\ &= (\mathbf{I}_m - \mathbf{\Gamma})^{-1} \left(n\mathbf{I}_m - \mathbf{\Gamma} \sum_{i=0}^{n-1} \mathbf{\Gamma}^i \right) \boldsymbol{\phi} + \mathbf{\Gamma} \sum_{i=0}^{n-1} \mathbf{\Gamma}^i \mathbf{y}_t \\ &= (\mathbf{I}_m - \mathbf{\Gamma})^{-2} (n(\mathbf{I}_m - \mathbf{\Gamma}) - \mathbf{\Gamma}(\mathbf{I}_m - \mathbf{\Gamma}^n)) \boldsymbol{\phi} + \mathbf{\Gamma}(\mathbf{I}_m - \mathbf{\Gamma})^{-1}(\mathbf{I}_m - \mathbf{\Gamma}^n) \mathbf{y}_t \\ &= (\mathbf{I}_m - \mathbf{\Gamma})^{-2} (n\mathbf{I}_m - (n+1)\mathbf{\Gamma} + \mathbf{\Gamma}^{n+1}) \boldsymbol{\phi} + \\ &\quad (\mathbf{I}_m - \mathbf{\Gamma})^{-1}(\mathbf{\Gamma} - \mathbf{\Gamma}^{n+1}) \mathbf{y}_t. \end{aligned} \quad (\text{B.11})$$

The deviation of $\mathbf{z}_{t,n}$ from its conditional average can be written as

$$\begin{aligned} \mathbf{z}_{t,n} - \mathbb{E}[\mathbf{z}_{t,n} | \mathbf{y}_t] &= \sum_{i=1}^n \sum_{j=0}^{i-1} \mathbf{\Gamma}^j \boldsymbol{\varepsilon}_{t+i-j} = \sum_{i=1}^n \sum_{j=0}^{n-i} \mathbf{\Gamma}^j \boldsymbol{\varepsilon}_{t+i} \\ &= \sum_{i=1}^n (\mathbf{I}_m - \mathbf{\Gamma})^{-1} (\mathbf{I}_m - \mathbf{\Gamma}^{n-i+1}) \boldsymbol{\varepsilon}_{t+i}. \end{aligned} \quad (\text{B.12})$$

Because the innovation vectors are i.i.d., the variance follows as

$$\text{Var}[\mathbf{z}_{t,n}] = \sum_{i=1}^n (\mathbf{I}_m - \mathbf{\Gamma})^{-1} (\mathbf{I}_m - \mathbf{\Gamma}^i) \boldsymbol{\Omega} ((\mathbf{I}_m - \mathbf{\Gamma})^{-1} (\mathbf{I}_m - \mathbf{\Gamma}^i))',$$

For comfortable matrices \mathbf{A} , \mathbf{B} , and \mathbf{C} , $\text{vec}(\mathbf{ABC}) = (\mathbf{C} \otimes \mathbf{A}) \text{vec}(\mathbf{B})$, which applied to the previous equation yields

$$\begin{aligned} \text{vec}(\text{Var}[\mathbf{z}_{t,n}]) &= \sum_{i=1}^n (((\mathbf{I}_m - \mathbf{\Gamma})^{-1} (\mathbf{I}_m - \mathbf{\Gamma}^i)) \otimes ((\mathbf{I}_m - \mathbf{\Gamma})^{-1} (\mathbf{I}_m - \mathbf{\Gamma}^i))) \text{vec}(\boldsymbol{\Omega}) \\ &= ((\mathbf{I}_m - \mathbf{\Gamma})^{-1} \otimes (\mathbf{I}_m - \mathbf{\Gamma})^{-1}) \cdot \\ &\quad \sum_{i=1}^n ((\mathbf{I}_m - \mathbf{\Gamma}^i) \otimes (\mathbf{I}_m - \mathbf{\Gamma}^i)) \text{vec}(\boldsymbol{\Omega}) \end{aligned} \quad (\text{B.13})$$

The summation in this expression can be reduced as follows¹⁵

$$\begin{aligned}
& \sum_{i=1}^n ((\mathbf{I}_m - \mathbf{\Gamma}^i) \otimes (\mathbf{I}_m - \mathbf{\Gamma}^i)) \\
&= n\mathbf{I}_{m^2} - \mathbf{I}_m \otimes \sum_{i=1}^n \mathbf{\Gamma}^i - \sum_{i=1}^n \mathbf{\Gamma}^i \otimes \mathbf{I}_m + \sum_{i=1}^n \mathbf{\Gamma}^i \otimes \mathbf{\Gamma}^i \\
&= n\mathbf{I}_{k^2} - \mathbf{I}_m \otimes (\mathbf{I}_m - \mathbf{\Gamma})^{-1}(\mathbf{\Gamma} - \mathbf{\Gamma}^{n+1}) - (\mathbf{I}_m - \mathbf{\Gamma})^{-1}(\mathbf{\Gamma} - \mathbf{\Gamma}^{n+1}) \otimes \mathbf{I}_m + \\
& \quad (\mathbf{I}_{k^2} - \mathbf{\Gamma} \otimes \mathbf{\Gamma})^{-1}((\mathbf{\Gamma} \otimes \mathbf{\Gamma}) - (\mathbf{\Gamma} \otimes \mathbf{\Gamma})^{n+1}). \tag{B.14}
\end{aligned}$$

The conditional variance of the weighted sum $\mathbf{w}'\mathbf{z}_{t,n}$ can be calculated as

$$\text{Var}[\mathbf{w}'\mathbf{z}_{t,n}|\mathbf{y}_t] = \mathbf{w}' \text{Var}[\mathbf{z}_{t,n}]\mathbf{w} = (\mathbf{w} \otimes \mathbf{w})' \text{vec}(\text{Var}[\mathbf{z}_{t,n}]) = (\mathbf{w} \otimes \mathbf{w})' \mathbf{S}_n(\mathbf{\Gamma}) \text{vec}(\mathbf{\Omega}), \tag{B.15}$$

where \mathbf{S}_n is the $m^2 \times m^2$ scaling matrix for horizon n depending on $\mathbf{\Gamma}$,

$$\mathbf{S}_n(\mathbf{\Gamma}) = ((\mathbf{I}_m - \mathbf{\Gamma}) \otimes (\mathbf{I}_m - \mathbf{\Gamma}))^{-1} \sum_{i=1}^n ((\mathbf{I}_m - \mathbf{\Gamma}^i) \otimes (\mathbf{I}_m - \mathbf{\Gamma}^i)), \tag{B.16}$$

as follows from Equation (B.13). Equation (B.14) can be substituted for the summation. Because the one-period variance $\mathbf{\Omega}$ and the weights \mathbf{w} are on opposite sides of the scaling matrix $\mathbf{S}_n(\mathbf{\Phi})$, we cannot scale the variance of the weighted sum $\mathbf{w}'\mathbf{\Omega}\mathbf{w}$.

This scaling factor is an extension of the traditional square-root-of-time rule for scaling the volatility. When VAR-effects are absent, i.e., $\mathbf{\Gamma} = \mathbf{O}$, $\mathbf{S}_n(\mathbf{O}) = n\mathbf{I}_{m^2}$, and scaling is simply by n , $\text{Var}[\mathbf{w}'\mathbf{z}_{t,n}|\mathbf{y}_t] = n\mathbf{w}'\mathbf{\Omega}\mathbf{w} = n \text{Var}[\mathbf{w}'\mathbf{z}_{t,1}|\mathbf{y}_t]$. When $m = 1$, the VAR model reduces to an AR(1)-model with autoregressive coefficient ρ . In that case, the scaling factor reduces to (cf. Wang et al., 2011)

$$\mathbf{S}_n(\rho) = \frac{1}{(1 - \rho)^2} \left(n - 2\rho \frac{1 - \rho^n}{1 - \rho} + \rho^2 \frac{1 - \rho^{2n}}{1 - \rho^2} \right). \tag{B.17}$$

C Additional results

C.1 Detailed results for the 99% confidence level

[Table C.1 about here.]

[Table C.1 (continued) about here.]

¹⁵See <http://www.ee.ic.ac.uk/hp/staff/dmb/matrix/intro.html> for an overview of matrix operations.

[Table C.2 about here.]

[Table C.2 (continued) about here.]

[Table C.3 about here.]

[Table C.4 about here.]

[Table C.5 about here.]

[Table C.5 (continued) about here.]

[Table C.6 about here.]

[Table C.7 about here.]

[Table C.7 (continued) about here.]

C.2 Results 95% VaR forecasts

[Table C.8 about here.]

[Table C.9 about here.]

[Table C.9 (continued) about here.]

[Table C.10 about here.]

[Table C.10 (continued) about here.]

[Table C.11 about here.]

[Table C.11 (continued) about here.]

[Table C.12 about here.]

[Table C.12 (continued) about here.]

[Table C.13 about here.]

[Table C.14 about here.]

[Table C.15 about here.]

[Table C.16 about here.]

[Table C.16 (continued) about here.]

[Table C.17 about here.]

C.3 Results based on 8-year estimation window

[Table C.18 about here.]

[Table C.19 about here.]

[Table C.19 (continued) about here.]

[Table C.20 about here.]

[Table C.20 (continued) about here.]

[Table C.21 about here.]

[Table C.21 (continued) about here.]

[Table C.22 about here.]

[Table C.23 about here.]

[Table C.24 about here.]

[Table C.24 (continued) about here.]

Table C.1: Results for the unconditional coverage test

(a) daily returns, iterated forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC- GARCH	1.71** (0.39)	1.73** (0.38)	1.95*** (0.40)	1.53* (0.37)	1.27 (0.32)	1.53* (0.34)	1.69** (0.38)	1.63** (0.36)	1.85** (0.38)
DCC- GARCH		1.73** (0.38)	1.83** (0.39)		1.23 (0.32)	1.35 (0.32)		1.67** (0.37)	1.67** (0.37)
HDCC- GARCH		1.79** (0.39)	1.91** (0.40)		1.33 (0.32)	1.55* (0.34)		1.65** (0.37)	1.79** (0.39)
CCC- GJR	1.18 (0.31)	1.27 (0.31)	1.47* (0.33)	1.04 (0.30)	1.00 (0.27)	1.16 (0.30)	1.08 (0.30)	1.18 (0.30)	1.41 (0.33)
DCC- GJR		1.21 (0.31)	1.37 (0.32)		0.94 (0.27)	1.14 (0.28)		1.10 (0.29)	1.25 (0.31)
HDCC- GJR		1.19 (0.31)	1.45* (0.34)		0.96 (0.26)	1.21 (0.31)		1.21 (0.30)	1.39 (0.33)
RiskMetrics	1.37 (0.34)	1.43 (0.37)	1.75** (0.40)	1.45* (0.35)	1.33 (0.35)	1.39 (0.34)	0.78 (0.24)	1.37 (0.37)	1.59* (0.40)

(b) daily returns, scaled forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC- GARCH	1.97*** (0.42)	1.87** (0.39)	2.07*** (0.41)	1.31 (0.33)	1.33 (0.32)	1.41 (0.33)	1.69** (0.38)	1.57* (0.36)	1.73** (0.37)
DCC- GARCH		1.95*** (0.40)	1.97*** (0.40)		1.35 (0.34)	1.35 (0.33)		1.75** (0.38)	1.63** (0.37)
HDCC- GARCH		1.99*** (0.42)	2.11*** (0.42)		1.39 (0.34)	1.43* (0.34)		1.73** (0.38)	1.65** (0.37)
CCC- GJR	1.93*** (0.40)	1.83** (0.38)	2.01*** (0.39)	1.51* (0.35)	1.47* (0.34)	1.49* (0.34)	1.63** (0.37)	1.63** (0.35)	1.73** (0.36)
DCC- GJR		1.91*** (0.39)	1.81** (0.38)		1.51* (0.34)	1.39 (0.32)		1.67** (0.35)	1.67** (0.36)
HDCC- GJR		1.93*** (0.39)	1.95*** (0.39)		1.55* (0.35)	1.41 (0.33)		1.71** (0.37)	1.63** (0.35)
RiskMetrics	1.75** (0.38)	1.59* (0.39)	2.01*** (0.43)	1.18 (0.32)	1.31 (0.36)	1.18 (0.34)	0.84 (0.26)	1.33 (0.36)	1.35 (0.36)

This table shows the empirical coverage (in %) and results of the unconditional coverage tests of Christoffersen (1998). Models are estimated with a moving window of 1,000 daily returns, 200 weekly returns or 100 biweekly returns. For every day in the sample, we construct a violation indicator that equals one when the ten-day realized portfolio loss exceeds the forecasted VaR_ϑ with $\vartheta = 0.01$, and zero otherwise, as in Equation (20). We report the coverage (in %), and test whether it is equal to ϑ against the alternative of strictly more violations. We report Newey and West (1987) standard errors calculated with 10 leads and lags in parentheses. Superscripts *, **, *** denote rejection of the null-hypothesis with a significance level of 10%, 5% and 1%. The results are based on 5,021 forecasts.

Table C.1: Results for the unconditional coverage test – *continued*

(c) weekly returns, iterated forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC- GARCH	2.25*** (0.46)	1.95** (0.41)	2.03*** (0.43)	1.79** (0.42)	1.14 (0.33)	1.12 (0.33)	2.03** (0.45)	1.85** (0.41)	1.91** (0.42)
DCC- GARCH		1.99*** (0.42)	2.09*** (0.43)		1.10 (0.33)	1.06 (0.33)		1.77** (0.40)	1.91** (0.41)
HDCC- GARCH		2.03*** (0.43)	2.09*** (0.43)		1.08 (0.32)	1.06 (0.33)		1.97*** (0.41)	1.89** (0.41)
CCC- GJR	1.85** (0.39)	1.71** (0.38)	1.75** (0.39)	1.47* (0.36)	1.02 (0.31)	0.98 (0.30)	1.73** (0.39)	1.53* (0.37)	1.61** (0.37)
DCC- GJR		1.69** (0.38)	1.73** (0.39)		1.06 (0.31)	1.00 (0.30)		1.51* (0.36)	1.65** (0.37)
HDCC- GJR		1.67** (0.38)	1.75** (0.40)		1.12 (0.31)	1.06 (0.31)		1.53* (0.37)	1.65** (0.38)
RiskMetrics	2.01** (0.45)	1.81** (0.42)	2.03** (0.46)	1.67* (0.41)	1.33 (0.39)	1.29 (0.38)	1.79** (0.40)	1.71** (0.42)	1.75** (0.43)

(d) weekly returns, scaled forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC- GARCH	2.41*** (0.47)	2.15*** (0.43)	2.21*** (0.44)	1.75** (0.41)	0.94 (0.31)	1.00 (0.33)	2.13*** (0.46)	1.91** (0.42)	1.91** (0.41)
DCC- GARCH		2.19*** (0.43)	2.21*** (0.44)		0.98 (0.32)	1.04 (0.33)		1.85** (0.42)	1.93** (0.41)
HDCC- GARCH		2.31*** (0.45)	2.23*** (0.44)		0.98 (0.31)	1.00 (0.32)		1.99*** (0.42)	1.99*** (0.42)
CCC- GJR	2.31*** (0.43)	2.23*** (0.44)	2.13*** (0.43)	1.59* (0.37)	1.10 (0.32)	1.10 (0.32)	1.91** (0.40)	1.83** (0.39)	1.81** (0.39)
DCC- GJR		2.29*** (0.44)	2.07*** (0.42)		1.10 (0.31)	1.08 (0.31)		1.81** (0.39)	1.87** (0.39)
HDCC- GJR		2.27*** (0.44)	2.15*** (0.43)		1.18 (0.33)	1.04 (0.31)		1.89** (0.40)	1.87** (0.40)
RiskMetrics	2.15*** (0.46)	1.91** (0.43)	2.03** (0.47)	1.67* (0.41)	1.39 (0.37)	1.27 (0.36)	1.77** (0.40)	1.69** (0.41)	1.65* (0.42)

(e) biweekly returns, direct forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC- GARCH	2.55*** (0.48)	2.21*** (0.44)	2.33*** (0.46)	1.83** (0.42)	0.92 (0.30)	1.06 (0.32)	2.25*** (0.47)	2.15*** (0.44)	1.99** (0.43)
DCC- GARCH		2.27*** (0.45)	2.33*** (0.46)		0.88 (0.30)	1.10 (0.32)		2.05*** (0.42)	2.05*** (0.44)
HDCC- GARCH		2.33*** (0.46)	2.35*** (0.46)		0.94 (0.30)	1.06 (0.32)		2.21*** (0.45)	1.99** (0.43)
CCC- GJR	2.75*** (0.48)	2.31*** (0.44)	2.17*** (0.44)	2.07*** (0.43)	0.90 (0.29)	1.16 (0.32)	2.33*** (0.45)	1.97** (0.42)	1.95** (0.42)
DCC- GJR		2.29*** (0.45)	2.13*** (0.44)		0.88 (0.28)	1.19 (0.32)		2.03*** (0.42)	1.91** (0.41)
HDCC- GJR		2.29*** (0.44)	2.19*** (0.44)		0.98 (0.29)	1.16 (0.33)		1.95** (0.41)	1.91** (0.42)
RiskMetrics	2.29*** (0.48)	1.93** (0.45)	2.29*** (0.49)	2.09*** (0.44)	1.79** (0.43)	2.09** (0.48)	2.07*** (0.45)	1.81** (0.44)	2.11** (0.50)

See table note on previous page.

Table C.2: Results of the dynamic quantile tests

(a) daily returns, iterated forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-	-0.23	-0.11	-0.27	-0.28	-0.19	-0.22	-0.26	-0.15	-0.13
GARCH	(0.14)	(0.27)	(0.14)	(0.11)	(0.13)	(0.14)	(0.12)	(0.20)	(0.24)
DCC-		-0.11	-0.08		-0.15	-0.17		-0.10	-0.07
GARCH		(0.26)	(0.31)		(0.16)	(0.15)		(0.26)	(0.30)
HDCC-		-0.11	-0.19		-0.14	-0.24		-0.15	-0.12
GARCH		(0.26)	(0.20)		(0.19)	(0.12)		(0.19)	(0.26)
CCC-	-0.11	-0.07	-0.19	-0.22*	-0.08	-0.21*	-0.15	-0.08	-0.18
GJR	(0.20)	(0.26)	(0.14)	(0.08)	(0.21)	(0.10)	(0.13)	(0.22)	(0.15)
DCC-		-0.07	-0.10		-0.13	-0.15		-0.11	-0.11
GJR		(0.25)	(0.23)		(0.12)	(0.13)		(0.17)	(0.20)
HDCC-		-0.08	-0.16		-0.07	-0.19		-0.06	-0.12
GJR		(0.22)	(0.18)		(0.22)	(0.13)		(0.26)	(0.21)
RiskMetrics	-0.08	0.03	0.05	-0.13	0.01	0.02	-0.17*	0.01	0.00
	(0.25)	(0.49)	(0.52)	(0.21)	(0.45)	(0.50)	(0.08)	(0.43)	(0.44)

(b) daily returns, scaled forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-	-0.25	-0.17	-0.24	-0.14	-0.08	-0.13	-0.23	-0.11	-0.15
GARCH	(0.15)	(0.20)	(0.16)	(0.18)	(0.27)	(0.21)	(0.14)	(0.24)	(0.22)
DCC-		-0.11	-0.11		-0.08	-0.08		-0.13	-0.08
GARCH		(0.26)	(0.27)		(0.27)	(0.25)		(0.23)	(0.30)
HDCC-		-0.19	-0.22		-0.12	-0.08		-0.10	-0.07
GARCH		(0.19)	(0.19)		(0.22)	(0.26)		(0.27)	(0.31)
CCC-	-0.15	-0.17	-0.30	-0.19	-0.13	-0.19	-0.14	-0.11	-0.29*
GJR	(0.21)	(0.18)	(0.11)	(0.15)	(0.18)	(0.14)	(0.20)	(0.23)	(0.10)
DCC-		-0.14	-0.16		-0.11	-0.09		-0.10	-0.11
GJR		(0.21)	(0.20)		(0.22)	(0.24)		(0.25)	(0.25)
HDCC-		-0.13	-0.26		-0.12	-0.16		-0.09	-0.21
GJR		(0.23)	(0.13)		(0.20)	(0.17)		(0.27)	(0.14)
RiskMetrics	-0.17	0.06	0.07	-0.10	0.03	0.03	-0.18*	0.01	-0.02
	(0.18)	(0.55)	(0.56)	(0.21)	(0.50)	(0.54)	(0.09)	(0.45)	(0.37)

This table reports the results of the dynamic quantile test of Engle and Manganelli (2004). We conduct a linear regression of the centered VaR-violations on a constant and the forecasted VaR_ϑ (in %) for $\vartheta = 0.01$, as in Equation (21). We report the estimated coefficient on the forecasted VaR_ϑ (expressed as a decimal rate), and quantile of its t -statistic in the distribution of the statistic under the null hypothesis of no predictability. Following Herwartz and Waichman (2010), we construct this distribution by 10,000 block bootstrapped samples with a block length of 10. Superscripts *, **, *** denote quantiles in the 10%, 5% and 1% of the left or right tail.

Table C.2: Results of the dynamic quantile test – *continued*

(c) weekly returns, iterated forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-	0.09	0.02	0.00	0.13	0.01	-0.04	0.08	0.01	-0.05
GARCH	(0.60)	(0.47)	(0.46)	(0.72)	(0.48)	(0.34)	(0.58)	(0.48)	(0.38)
DCC-		0.05	-0.03		0.01	-0.08		0.02	-0.07
GARCH		(0.54)	(0.40)		(0.46)	(0.27)		(0.49)	(0.35)
HDCC-		0.11	-0.05		-0.01	-0.03		0.08	-0.03
GARCH		(0.64)	(0.38)		(0.40)	(0.37)		(0.59)	(0.40)
CCC-	0.09	0.01	-0.12	0.06	-0.08	-0.05	0.04	-0.03	-0.13
GJR	(0.64)	(0.47)	(0.26)	(0.57)	(0.21)	(0.29)	(0.52)	(0.36)	(0.23)
DCC-		0.00	-0.12		-0.06	-0.02		0.00	-0.06
GJR		(0.44)	(0.26)		(0.24)	(0.37)		(0.43)	(0.33)
HDCC-		0.05	-0.11		-0.01	-0.03		0.02	-0.05
GJR		(0.56)	(0.27)		(0.40)	(0.36)		(0.47)	(0.35)
RiskMetrics	0.24	0.20	0.21	0.18	0.13	0.15**	0.21*	0.20	0.20
	(0.89)	(0.81)	(0.78)	(0.87)	(0.83)	(0.97)	(0.94)	(0.80)	(0.84)

(d) weekly returns, scaled forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-	0.12	0.05	-0.08	0.16	0.00	-0.02	0.05	0.01	-0.08
GARCH	(0.63)	(0.53)	(0.36)	(0.83)	(0.43)	(0.37)	(0.52)	(0.46)	(0.34)
DCC-		0.08	-0.05		0.00	-0.02		0.00	-0.06
GARCH		(0.56)	(0.37)		(0.41)	(0.39)		(0.45)	(0.36)
HDCC-		0.11	-0.06		0.03	-0.02		0.09	-0.02
GARCH		(0.63)	(0.38)		(0.53)	(0.37)		(0.60)	(0.41)
CCC-	0.05	-0.04	-0.19	0.08	0.00	-0.03	0.06	0.00	-0.19
GJR	(0.53)	(0.39)	(0.22)	(0.64)	(0.44)	(0.34)	(0.55)	(0.42)	(0.20)
DCC-		-0.02	-0.23		0.04	-0.04		0.03	-0.09
GJR		(0.42)	(0.18)		(0.57)	(0.31)		(0.50)	(0.30)
HDCC-		0.08	-0.19		0.03	-0.02		0.05	-0.16
GJR		(0.57)	(0.23)		(0.53)	(0.36)		(0.54)	(0.21)
RiskMetrics	0.22	0.20	0.21	0.18*	0.17*	0.14**	0.16	0.23	0.17
	(0.82)	(0.79)	(0.77)	(0.91)	(0.94)	(0.97)	(0.83)	(0.86)	(0.81)

(e) biweekly returns, direct forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-	0.35*	0.30	0.26	0.26***	0.09*	0.10	0.39**	0.24	0.30
GARCH	(0.91)	(0.86)	(0.79)	(1.00)	(0.95)	(0.85)	(0.97)	(0.83)	(0.89)
DCC-		0.31	0.26		0.09**	0.12*		0.30*	0.23
GARCH		(0.88)	(0.78)		(0.95)	(0.91)		(0.91)	(0.81)
HDCC-		0.35*	0.26		0.10**	0.13*		0.29	0.25
GARCH		(0.91)	(0.80)		(0.97)	(0.94)		(0.89)	(0.85)
CCC-	0.26	0.28	0.15	0.27***	0.08*	0.10	0.27*	0.27*	0.14
GJR	(0.89)	(0.89)	(0.68)	(1.00)	(0.94)	(0.83)	(0.94)	(0.93)	(0.69)
DCC-		0.31*	0.16		0.09**	0.12*		0.25*	0.16
GJR		(0.92)	(0.69)		(0.97)	(0.90)		(0.91)	(0.74)
HDCC-		0.33*	0.17		0.10**	0.10		0.26*	0.19
GJR		(0.94)	(0.71)		(0.97)	(0.83)		(0.92)	(0.77)
RiskMetrics	0.35**	0.32**	0.60***	0.30***	0.21***	0.24***	0.38***	0.33**	0.50**
	(0.96)	(0.96)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(0.97)	(0.97)

See table note on previous page.

Table C.3: Results of the DM-tests for iterated vs. scaled forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-	-0.08	-0.06	-0.13**	0.16	0.06	-0.01	0.04	0.07	0.01
GARCH	(0.07)	(0.07)	(0.07)	(0.14)	(0.11)	(0.04)	(0.06)	(0.08)	(0.05)
DCC-		-0.07	-0.12**		0.02	-0.06*		0.04	-0.03
GARCH		(0.05)	(0.07)		(0.10)	(0.05)		(0.06)	(0.03)
HDCC-		-0.07*	-0.10*		0.07	0.00		0.09	-0.02
GARCH		(0.06)	(0.07)		(0.10)	(0.04)		(0.08)	(0.03)
CCC-	-0.20	-0.26**	-0.29*	-0.08	-0.11	-0.19	-0.18	-0.09	-0.19
GJR	(0.17)	(0.14)	(0.21)	(0.10)	(0.11)	(0.21)	(0.15)	(0.08)	(0.16)
DCC-		-0.27**	-0.30*		-0.11	-0.18		-0.10	-0.17
GJR		(0.16)	(0.22)		(0.11)	(0.22)		(0.09)	(0.16)
HDCC-		-0.23**	-0.29*		-0.05	-0.17		-0.10	-0.21
GJR		(0.13)	(0.22)		(0.10)	(0.21)		(0.09)	(0.17)
RiskMetrics	-0.15*	-0.06*	-0.04	0.07	-0.17***	0.00	0.01	0.06	-0.01
	(0.09)	(0.03)	(0.06)	(0.08)	(0.07)	(0.10)	(0.06)	(0.07)	(0.02)

This table shows summary results of the tests that the expected value of the loss function of methods A and B are equal, $L^A = L^B$. We report the loss differential $L^A - L^B$ (multiplied by 100). A negative loss differential means that method A is preferable to method B. We use the asymmetric tick loss function of Giacomini and Komunjer (2005) as in Equation (22) with $\vartheta = 0.01$. The two methods differ in their forecast construction, with A using iterated and B using scaled forecasts. Both methods use daily data. We report standard errors of the average loss differential in parentheses, based on Newey and West (1987) with 10 leads and lags. We test that the average loss differential is zero based on the statistic proposed by Diebold and Mariano (1995), evaluated in the setting of Giacomini and White (2006). Superscripts *, **, *** denote rejection of the null-hypothesis with a significance level of 10%, 5% and 1%.

Table C.4: Results of the DM-tests for portfolio aggregation

(a) daily returns, iterated forecasts									
Test	Normal			Empirical			Student's t		
	1 vs. 3	1 vs. 8	3 vs. 8	1 vs. 3	1 vs. 8	3 vs. 8	1 vs. 3	1 vs. 8	3 vs. 8
CCC-	0.38	0.25	-0.13	0.43	0.40*	-0.03	0.36	0.29	-0.07
GARCH	(0.36)	(0.25)	(0.20)	(0.41)	(0.30)	(0.17)	(0.36)	(0.27)	(0.17)
DCC-	0.31	0.33	0.03	0.38	0.47*	0.09	0.33	0.36	0.04
GARCH	(0.27)	(0.31)	(0.11)	(0.34)	(0.36)	(0.10)	(0.31)	(0.33)	(0.11)
HDCC-	0.24	0.26	0.02	0.34	0.43	0.09	0.28	0.26	-0.02
GARCH	(0.24)	(0.25)	(0.12)	(0.32)	(0.34)	(0.11)	(0.29)	(0.27)	(0.14)
CCC-	0.24	0.27*	0.03	0.23	0.33*	0.10	0.21	0.22	0.01
GJR	(0.28)	(0.19)	(0.19)	(0.33)	(0.25)	(0.16)	(0.29)	(0.23)	(0.18)
DCC-	0.20	0.37*	0.16*	0.18	0.41*	0.23**	0.25	0.30	0.04
GJR	(0.23)	(0.26)	(0.11)	(0.29)	(0.31)	(0.10)	(0.29)	(0.26)	(0.15)
HDCC-	0.16	0.30	0.14	0.12	0.31	0.19*	0.13	0.26	0.13
GJR	(0.21)	(0.24)	(0.13)	(0.25)	(0.24)	(0.13)	(0.22)	(0.24)	(0.13)
RiskMetrics	-0.43**	-0.48**	-0.05	-0.30*	-0.16*	0.13	-0.35	-0.49	-0.13
	(0.26)	(0.23)	(0.18)	(0.20)	(0.11)	(0.24)	(0.40)	(0.50)	(0.14)

(b) daily returns, scaled forecasts									
Test	Normal			Empirical			Student's t		
	1 vs. 3	1 vs. 8	3 vs. 8	1 vs. 3	1 vs. 8	3 vs. 8	1 vs. 3	1 vs. 8	3 vs. 8
CCC-	0.33	0.27	-0.05	0.27	0.30	0.04	0.32	0.31	-0.01
GARCH	(0.33)	(0.26)	(0.15)	(0.30)	(0.29)	(0.11)	(0.36)	(0.30)	(0.15)
DCC-	0.26	0.34	0.08	0.16	0.33	0.17	0.25	0.36	0.10
GARCH	(0.25)	(0.30)	(0.11)	(0.20)	(0.34)	(0.16)	(0.28)	(0.34)	(0.13)
HDCC-	0.22	0.26	0.04	0.19	0.35	0.16	0.22	0.31	0.09
GARCH	(0.23)	(0.25)	(0.11)	(0.21)	(0.32)	(0.14)	(0.27)	(0.29)	(0.12)
CCC-	0.15	0.21	0.05	0.11	0.30	0.19*	0.21	0.32	0.11
GJR	(0.24)	(0.22)	(0.15)	(0.21)	(0.26)	(0.12)	(0.27)	(0.28)	(0.13)
DCC-	0.10	0.30	0.20*	0.08	0.38	0.29*	0.26	0.38	0.12
GJR	(0.17)	(0.25)	(0.14)	(0.14)	(0.31)	(0.19)	(0.27)	(0.30)	(0.13)
HDCC-	0.07	0.27	0.21*	0.03	0.34	0.31**	0.10	0.35	0.25*
GJR	(0.15)	(0.25)	(0.14)	(0.13)	(0.27)	(0.17)	(0.19)	(0.30)	(0.16)
RiskMetrics	-0.33*	-0.39**	-0.07	-0.36**	-0.40***	-0.04	-0.37	-0.43	-0.06
	(0.22)	(0.18)	(0.16)	(0.18)	(0.14)	(0.21)	(0.33)	(0.37)	(0.12)

This table shows summary results of the tests that the expected value of the loss function of methods A and B are equal, $L^A = L^B$. We report the loss differential $L^A - L^B$ (multiplied by 100). A negative loss differential means that method A is preferable to method B. We use the asymmetric tick loss function of Giacomini and Komunjer (2005) as in Equation (22) with $\vartheta = 0.01$. The two methods differ in their degree of portfolio aggregation. We consider aggregation into a portfolio (labeled “1”), three asset classes (labeled “3”), and no aggregation (labeled “8”). All methods use daily data. We report standard errors of the average loss differential in parentheses, based on Newey and West (1987) with 10 leads and lags. We test that the average loss differential is zero based on the statistic proposed by Diebold and Mariano (1995), evaluated in the setting of Giacomini and White (2006). Superscripts *, **, *** denote rejection of the null-hypothesis with a significance level of 10%, 5% and 1%.

Table C.5: Results of the DM-tests for the model specification

(a) daily returns, iterated forecasts									
Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-GARCH vs.		-0.07	0.08		-0.05	0.07		-0.03	0.08
DCC-GARCH		(0.10)	(0.10)		(0.08)	(0.09)		(0.06)	(0.09)
CCC-GARCH vs.		-0.14	0.01		-0.09	0.03		-0.08	-0.03
HDCC-GARCH		(0.13)	(0.03)		(0.10)	(0.07)		(0.08)	(0.04)
CCC-GARCH vs.	0.30	0.16	0.32*	0.30	0.10	0.23	0.34*	0.20	0.28*
CCC-GJR	(0.25)	(0.20)	(0.20)	(0.24)	(0.18)	(0.19)	(0.26)	(0.21)	(0.18)
CCC-GARCH vs.		0.13	0.42*		0.06	0.31		0.24	0.35*
DCC-GJR		(0.18)	(0.28)		(0.16)	(0.25)		(0.23)	(0.24)
CCC-GARCH vs.		0.08	0.35*		0.00	0.21		0.11	0.32*
HDCC-GJR		(0.17)	(0.24)		(0.15)	(0.17)		(0.18)	(0.21)
CCC-GARCH vs.	0.25	-0.56*	-0.48***	0.15	-0.57	-0.41***	0.18	-0.53*	-0.59*
RiskMetrics	(0.21)	(0.41)	(0.18)	(0.15)	(0.46)	(0.14)	(0.34)	(0.39)	(0.37)
DCC-GARCH vs.		-0.07*	-0.07		-0.04	-0.04		-0.04	-0.11
HDCC-GARCH		(0.04)	(0.08)		(0.04)	(0.04)		(0.04)	(0.09)
DCC-GARCH vs.		0.24	0.24*		0.15	0.16		0.23	0.20*
CCC-GJR		(0.26)	(0.15)		(0.23)	(0.14)		(0.25)	(0.14)
DCC-GARCH vs.		0.20	0.34*		0.10	0.24*		0.27	0.27*
DCC-GJR		(0.22)	(0.21)		(0.20)	(0.18)		(0.26)	(0.17)
DCC-GARCH vs.		0.15	0.27*		0.04	0.14		0.14	0.24*
HDCC-GJR		(0.21)	(0.17)		(0.18)	(0.12)		(0.20)	(0.15)
DCC-GARCH vs.		-0.49*	-0.56**		-0.52*	-0.48***		-0.50*	-0.67*
RiskMetrics		(0.32)	(0.25)		(0.40)	(0.20)		(0.35)	(0.44)
HDCC-GARCH vs.		0.30	0.31*		0.19	0.20		0.27	0.31*
CCC-GJR		(0.29)	(0.20)		(0.25)	(0.17)		(0.26)	(0.20)
HDCC-GARCH vs.		0.27	0.41*		0.14	0.28*		0.32	0.38*
DCC-GJR		(0.25)	(0.27)		(0.22)	(0.21)		(0.27)	(0.25)
HDCC-GARCH vs.		0.22	0.34*		0.08	0.18		0.19	0.35*
HDCC-GJR		(0.24)	(0.23)		(0.20)	(0.14)		(0.21)	(0.22)
HDCC-GARCH vs.		-0.42*	-0.49***		-0.48*	-0.44***		-0.46*	-0.57*
RiskMetrics		(0.29)	(0.19)		(0.37)	(0.18)		(0.33)	(0.37)
CCC-GJR vs.		-0.03	0.10		-0.05	0.08		0.04	0.07
DCC-GJR		(0.05)	(0.10)		(0.05)	(0.08)		(0.04)	(0.07)
CCC-GJR vs.		-0.08	0.03		-0.11*	-0.02		-0.08	0.04
HDCC-GJR		(0.08)	(0.07)		(0.08)	(0.04)		(0.08)	(0.05)
CCC-GJR vs.	-0.05	-0.72*	-0.80**	-0.15	-0.67	-0.64**	-0.16	-0.73*	-0.87**
RiskMetrics	(0.15)	(0.55)	(0.35)	(0.19)	(0.59)	(0.30)	(0.14)	(0.56)	(0.52)
DCC-GJR vs.		-0.05	-0.07*		-0.06	-0.10		-0.13*	-0.03
HDCC-GJR		(0.04)	(0.05)		(0.05)	(0.08)		(0.08)	(0.03)
DCC-GJR vs.		-0.69*	-0.90**		-0.63	-0.72**		-0.77*	-0.94**
RiskMetrics		(0.50)	(0.43)		(0.56)	(0.35)		(0.57)	(0.57)
HDCC-GJR vs.		-0.64*	-0.83**		-0.57	-0.62**		-0.64*	-0.91**
RiskMetrics		(0.48)	(0.39)		(0.52)	(0.28)		(0.49)	(0.55)

This table shows summary results of the tests that the expected value of the loss function of methods A and B are equal, $L^A = L^B$. We report the loss differential $L^A - L^B$ (multiplied by 100). A negative loss differential means that method A is preferable to method B. We use the asymmetric tick loss function of Giacomini and Komunjer (2005) as in Equation (22) with $\vartheta = 0.01$. The two methods differ in their model specification. All methods use daily data. We report standard errors of the average loss differential in parentheses, based on Newey and West (1987) with 10 leads and lags. We test that the average loss differential is zero based on the statistic proposed by Diebold and Mariano (1995), evaluated in the setting of Giacomini and White (2006). Superscripts *, **, *** denote rejection of the null-hypothesis with a significance level of 10%, 5% and 1%.

Table C.5: Results of the DM-tests for the model specification – *continued*

(b) daily returns, scaled forecasts									
Asset split	Normal			Empirical			Student's <i>t</i>		
	1	3	8	1	3	8	1	3	8
CCC-GARCH vs.		-0.06	0.07		-0.10	0.03		-0.07	0.05
DCC-GARCH		(0.08)	(0.07)		(0.11)	(0.08)		(0.09)	(0.07)
CCC-GARCH vs.		-0.11	-0.01		-0.08	0.05		-0.10	0.00
HDCC-GARCH		(0.11)	(0.04)		(0.10)	(0.05)		(0.10)	(0.04)
CCC-GARCH vs.	0.18	0.01	0.11	0.07	-0.09	0.06	0.12	0.00	0.12
CCC-GJR	(0.15)	(0.12)	(0.10)	(0.12)	(0.13)	(0.06)	(0.14)	(0.12)	(0.10)
CCC-GARCH vs.		-0.04	0.21*		-0.12	0.14*		0.06	0.19*
DCC-GJR		(0.13)	(0.13)		(0.17)	(0.09)		(0.11)	(0.12)
CCC-GARCH vs.		-0.08	0.18*		-0.17	0.10**		-0.11	0.16**
HDCC-GJR		(0.14)	(0.12)		(0.18)	(0.06)		(0.15)	(0.11)
CCC-GARCH vs.	0.19	-0.47	-0.48**	0.06	-0.57*	-0.64***	0.14	-0.55*	-0.59*
RiskMetrics	(0.15)	(0.41)	(0.23)	(0.10)	(0.39)	(0.19)	(0.25)	(0.42)	(0.39)
DCC-GARCH vs.		-0.04*	-0.08		0.03	0.02		-0.04	-0.05
HDCC-GARCH		(0.03)	(0.07)		(0.03)	(0.03)		(0.03)	(0.07)
DCC-GARCH vs.		0.07	0.04		0.02	0.03		0.07	0.07
CCC-GJR		(0.16)	(0.11)		(0.13)	(0.10)		(0.15)	(0.10)
DCC-GARCH vs.		0.02	0.14*		-0.01	0.11**		0.13	0.14*
DCC-GJR		(0.12)	(0.09)		(0.12)	(0.07)		(0.14)	(0.08)
DCC-GARCH vs.		-0.02	0.11		-0.06	0.07		-0.04	0.11*
HDCC-GJR		(0.12)	(0.09)		(0.13)	(0.08)		(0.12)	(0.08)
DCC-GARCH vs.		-0.40	-0.55**		-0.46*	-0.67***		-0.48*	-0.64*
RiskMetrics		(0.34)	(0.28)		(0.30)	(0.24)		(0.35)	(0.44)
HDCC-GARCH vs.		0.12	0.12		-0.01	0.02		0.11	0.13
CCC-GJR		(0.18)	(0.12)		(0.14)	(0.09)		(0.15)	(0.12)
HDCC-GARCH vs.		0.07	0.22*		-0.04	0.09		0.16	0.19*
DCC-GJR		(0.13)	(0.14)		(0.13)	(0.07)		(0.15)	(0.13)
HDCC-GARCH vs.		0.03	0.19*		-0.09	0.05		0.00	0.16*
HDCC-GJR		(0.13)	(0.14)		(0.13)	(0.07)		(0.12)	(0.12)
HDCC-GARCH vs.		-0.36	-0.47**		-0.49*	-0.69***		-0.44*	-0.59*
RiskMetrics		(0.32)	(0.22)		(0.30)	(0.22)		(0.34)	(0.39)
CCC-GJR vs.		-0.05	0.10		-0.03	0.08		0.05*	0.07
DCC-GJR		(0.08)	(0.08)		(0.08)	(0.08)		(0.04)	(0.06)
CCC-GJR vs.		-0.09	0.07		-0.08	0.04		-0.11	0.03
HDCC-GJR		(0.10)	(0.07)		(0.09)	(0.03)		(0.10)	(0.05)
CCC-GJR vs.	0.01	-0.48	-0.59**	-0.01	-0.48*	-0.71***	0.03	-0.55	-0.72*
RiskMetrics	(0.16)	(0.45)	(0.29)	(0.14)	(0.36)	(0.20)	(0.21)	(0.44)	(0.45)
DCC-GJR vs.		-0.04*	-0.03*		-0.05*	-0.04		-0.16**	-0.03
HDCC-GJR		(0.03)	(0.02)		(0.03)	(0.06)		(0.09)	(0.03)
DCC-GJR vs.		-0.43	-0.69**		-0.45*	-0.78***		-0.60*	-0.78*
RiskMetrics		(0.38)	(0.34)		(0.30)	(0.25)		(0.44)	(0.48)
HDCC-GJR vs.		-0.39	-0.66**		-0.40*	-0.74***		-0.44	-0.75*
RiskMetrics		(0.36)	(0.33)		(0.28)	(0.21)		(0.36)	(0.47)

See table note on the previous page.

Table C.6: Results of the DM-tests for the distribution

(a) daily returns, iterated forecasts									
Asset split	Normal vs. t			Normal vs. Empirical			Empirical vs. t		
	1	3	8	1	3	8	1	3	8
CCC-	0.02	0.00	0.05	0.03	0.08	0.18*	-0.01	-0.08	-0.12*
GARCH	(0.02)	(0.02)	(0.05)	(0.06)	(0.11)	(0.13)	(0.06)	(0.11)	(0.09)
DCC-		0.04	0.05		0.10	0.17*		-0.06	-0.12
GARCH		(0.05)	(0.04)		(0.13)	(0.13)		(0.10)	(0.10)
HDCC-		0.06	0.02		0.13	0.21		-0.07	-0.18*
GARCH		(0.07)	(0.04)		(0.13)	(0.17)		(0.09)	(0.14)
CCC-	0.06**	0.03	0.01	0.03	0.02	0.09	0.03	0.01	-0.08
GJR	(0.03)	(0.04)	(0.05)	(0.05)	(0.10)	(0.12)	(0.04)	(0.07)	(0.10)
DCC-		0.11*	-0.01		0.01	0.07		0.10**	-0.08
GJR		(0.09)	(0.02)		(0.10)	(0.10)		(0.05)	(0.10)
HDCC-		0.03	0.02		-0.01	0.04		0.04	-0.02
GJR		(0.03)	(0.03)		(0.09)	(0.08)		(0.07)	(0.06)
RiskMetrics	-0.05	0.03	-0.06	-0.07*	0.07	0.25*	0.02	-0.04	-0.31
	(0.22)	(0.03)	(0.20)	(0.05)	(0.06)	(0.18)	(0.24)	(0.05)	(0.33)

(b) daily returns, scaled forecasts									
Asset split	Normal vs. t			Normal vs. Empirical			Empirical vs. t		
	1	3	8	1	3	8	1	3	8
CCC-	0.14**	0.14*	0.18**	0.27*	0.20	0.29*	-0.12	-0.07	-0.12
GARCH	(0.08)	(0.10)	(0.10)	(0.19)	(0.17)	(0.20)	(0.12)	(0.09)	(0.11)
DCC-		0.13*	0.15*		0.16	0.25		-0.03	-0.10
GARCH		(0.09)	(0.10)		(0.15)	(0.21)		(0.08)	(0.11)
HDCC-		0.14*	0.18**		0.23*	0.35*		-0.09	-0.17
GARCH		(0.10)	(0.10)		(0.17)	(0.25)		(0.09)	(0.15)
CCC-	0.08*	0.13*	0.19**	0.15*	0.11	0.25*	-0.08	0.02	-0.06
GJR	(0.06)	(0.08)	(0.11)	(0.11)	(0.11)	(0.16)	(0.07)	(0.05)	(0.06)
DCC-		0.24**	0.15**		0.13	0.23*		0.11	-0.07
GJR		(0.14)	(0.09)		(0.11)	(0.16)		(0.10)	(0.08)
HDCC-		0.11*	0.15**		0.12	0.22**		-0.01	-0.06
GJR		(0.08)	(0.09)		(0.11)	(0.13)		(0.06)	(0.06)
RiskMetrics	0.10	0.06	0.07	0.14	0.11	0.13	-0.04	-0.05	-0.07
	(0.25)	(0.09)	(0.18)	(0.14)	(0.17)	(0.26)	(0.12)	(0.10)	(0.33)

This table shows summary results of the tests that the expected value of the loss function of methods A and B are equal, $L^A = L^B$. We report the loss differential $L^A - L^B$ (multiplied by 100). A negative loss differential means that method A is preferable to method B. We use the asymmetric tick loss function of Giacomini and Komunjer (2005) as in Equation (22) with $\vartheta = 0.01$. The two methods differ in the distribution of the innovations. We consider the normal, empirical and Student's t -distributions. All methods use daily data. We report standard errors of the average loss differential in parentheses, based on Newey and West (1987) with 10 leads and lags. We test that the average loss differential is zero based on the statistic proposed by Diebold and Mariano (1995), evaluated in the setting of Giacomini and White (2006). Superscripts *, **, *** denote rejection of the null-hypothesis with a significance level of 10%, 5% and 1%.

Table C.7: Results of the DM-tests for temporal aggregation

(a) daily vs. weekly returns, iterated forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-	-1.29**	-1.14*	-1.22*	-1.05**	-0.93*	-0.95*	-1.05**	-1.15*	-1.11*
GARCH	(0.62)	(0.77)	(0.85)	(0.49)	(0.66)	(0.65)	(0.54)	(0.79)	(0.79)
DCC-		-1.16*	-1.16*		-0.94*	-0.89*		-1.16*	-1.00*
GARCH		(0.81)	(0.76)		(0.67)	(0.57)		(0.82)	(0.71)
HDCC-		-1.16*	-1.08*		-1.01*	-0.88*		-1.11*	-1.04*
GARCH		(0.79)	(0.72)		(0.72)	(0.56)		(0.78)	(0.71)
CCC-	-1.08**	-1.13*	-1.17*	-0.91**	-0.92*	-0.93*	-0.90**	-1.03*	-0.97*
GJR	(0.64)	(0.74)	(0.79)	(0.51)	(0.59)	(0.60)	(0.51)	(0.73)	(0.70)
DCC-		-1.19*	-1.16*		-1.00*	-0.89*		-1.07*	-1.02*
GJR		(0.79)	(0.75)		(0.65)	(0.55)		(0.74)	(0.71)
HDCC-		-1.17*	-1.10*		-0.91*	-0.87**		-1.07*	-0.97*
GJR		(0.79)	(0.70)		(0.58)	(0.51)		(0.74)	(0.65)
RiskMetrics	-1.49**	-1.26*	-1.05**	-1.31**	-1.68**	-1.05**	-1.05**	-1.14*	-1.07**
	(0.86)	(0.93)	(0.63)	(0.73)	(0.80)	(0.49)	(0.58)	(0.73)	(0.63)

(b) daily vs. weekly returns, scaled forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-	-1.36**	-1.14*	-1.19*	-1.09**	-0.98*	-0.98*	-1.06**	-1.16*	-1.15*
GARCH	(0.66)	(0.82)	(0.86)	(0.49)	(0.70)	(0.61)	(0.55)	(0.84)	(0.85)
DCC-		-1.17*	-1.13*		-0.97*	-0.87**		-1.20*	-1.05*
GARCH		(0.84)	(0.78)		(0.75)	(0.52)		(0.87)	(0.72)
HDCC-		-1.14*	-1.11*		-1.01*	-0.95**		-1.18*	-1.04*
GARCH		(0.82)	(0.75)		(0.74)	(0.53)		(0.84)	(0.73)
CCC-	-1.01**	-1.02*	-1.09*	-0.68**	-0.78*	-0.70**	-0.71**	-0.97*	-0.87*
GJR	(0.56)	(0.72)	(0.72)	(0.36)	(0.57)	(0.42)	(0.40)	(0.71)	(0.64)
DCC-		-1.06*	-1.05*		-0.81*	-0.68**		-0.99*	-0.94*
GJR		(0.73)	(0.65)		(0.58)	(0.34)		(0.70)	(0.64)
HDCC-		-1.10*	-1.03**		-0.80*	-0.63**		-1.03*	-0.81*
GJR		(0.77)	(0.62)		(0.56)	(0.33)		(0.75)	(0.55)
RiskMetrics	-1.41**	-1.23*	-0.99*	-1.30**	-1.30**	-1.02**	-1.04**	-1.15*	-1.04**
	(0.81)	(0.92)	(0.62)	(0.69)	(0.71)	(0.49)	(0.52)	(0.75)	(0.63)

(c) daily vs. biweekly returns, iterated forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-	-1.57**	-1.59**	-1.70*	-1.74***	-1.75**	-2.22***	-1.40**	-1.50*	-1.60*
GARCH	(0.69)	(0.95)	(1.07)	(0.57)	(0.81)	(0.86)	(0.63)	(0.97)	(1.00)
DCC-		-1.66*	-1.59**		-1.82**	-2.14***		-1.56*	-1.53**
GARCH		(1.02)	(0.95)		(0.88)	(0.75)		(1.01)	(0.93)
HDCC-		-1.61**	-1.56**		-1.82**	-2.13***		-1.44*	-1.51**
GARCH		(0.96)	(0.93)		(0.88)	(0.74)		(0.95)	(0.89)
CCC-	-1.63**	-1.68**	-1.69*	-1.64***	-1.72**	-2.05**	-1.40**	-1.57*	-1.63*
GJR	(0.83)	(1.01)	(1.15)	(0.69)	(0.88)	(0.93)	(0.71)	(0.99)	(1.05)
DCC-		-1.76*	-1.63*		-1.78**	-1.95**		-1.62*	-1.59*
GJR		(1.09)	(1.07)		(0.93)	(0.84)		(1.03)	(1.00)
HDCC-		-1.72*	-1.63*		-1.72**	-1.90***		-1.58*	-1.53**
GJR		(1.07)	(1.06)		(0.87)	(0.79)		(1.01)	(0.92)
RiskMetrics	-2.00**	-2.38**	-1.57*	-2.14***	-3.13**	-2.63***	-1.84**	-2.27**	-1.67**
	(0.99)	(1.40)	(1.02)	(0.82)	(1.37)	(0.99)	(0.91)	(1.21)	(0.99)

This table shows summary results of the tests that the expected value of the loss function of methods A and B are equal, $L^A = L^B$. We report the loss differential $L^A - L^B$ (multiplied by 100). A negative loss differential means that method A is preferable to method B. We use the asymmetric tick loss function of Giacomini and Komunjer (2005) as in Equation (22) with $\vartheta = 0.01$. The two methods differ in their degree of temporal aggregation as stated in the headings of the different panels. We report standard errors of the average loss differential in parentheses, based on Newey and West (1987) with 10 leads and lags. We test that the average loss differential is zero based on the statistic proposed by Diebold and Mariano (1995), evaluated in the setting of Giacomini and White (2006). Superscripts *, **, *** denote rejection of the null-hypothesis with a significance level of 10%, 5% and 1%.

Table C.7: Results of the DM-tests for temporal aggregation – *continued*

(d) daily vs. biweekly returns, scaled forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC- GARCH	-1.50** (0.70)	-1.53* (0.99)	-1.57* (1.05)	-1.89*** (0.69)	-1.81** (0.92)	-2.21*** (0.87)	-1.44** (0.67)	-1.57* (1.04)	-1.60* (1.04)
DCC- GARCH		-1.60* (1.04)	-1.47* (0.95)		-1.84** (0.97)	-2.08*** (0.73)		-1.59* (1.06)	-1.50* (0.93)
HDCC- GARCH		-1.53* (0.99)	-1.46* (0.93)		-1.89** (0.97)	-2.14*** (0.75)		-1.53* (1.02)	-1.48* (0.91)
CCC- GJR	-1.43** (0.72)	-1.42* (0.95)	-1.41* (1.02)	-1.56*** (0.66)	-1.61** (0.84)	-1.85*** (0.78)	-1.22** (0.61)	-1.47* (0.98)	-1.45* (0.95)
DCC- GJR		-1.49* (0.99)	-1.33* (0.91)		-1.67** (0.89)	-1.77*** (0.66)		-1.52* (0.99)	-1.42* (0.89)
HDCC- GJR		-1.49* (1.01)	-1.33* (0.91)		-1.67** (0.86)	-1.73*** (0.64)		-1.48* (0.99)	-1.32* (0.80)
RiskMetrics	-1.86** (0.92)	-2.32** (1.39)	-1.53* (1.00)	-2.20*** (0.88)	-2.95** (1.41)	-2.63*** (1.07)	-1.85** (0.86)	-2.34** (1.26)	-1.66** (1.01)

(e) weekly vs. biweekly returns, iterated forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC- GARCH	-0.28** (0.17)	-0.45** (0.22)	-0.47** (0.27)	-0.69*** (0.16)	-0.82*** (0.17)	-1.27*** (0.24)	-0.35** (0.17)	-0.36** (0.20)	-0.49** (0.26)
DCC- GARCH		-0.51** (0.24)	-0.42** (0.25)		-0.88*** (0.23)	-1.25*** (0.22)		-0.40** (0.21)	-0.53** (0.27)
HDCC- GARCH		-0.45** (0.21)	-0.49** (0.26)		-0.81*** (0.18)	-1.25*** (0.22)		-0.33** (0.19)	-0.47** (0.23)
CCC- GJR	-0.55** (0.25)	-0.56** (0.31)	-0.52 (0.41)	-0.72*** (0.23)	-0.80*** (0.30)	-1.12*** (0.35)	-0.50** (0.24)	-0.53** (0.29)	-0.66** (0.38)
DCC- GJR		-0.58** (0.33)	-0.47 (0.38)		-0.78*** (0.30)	-1.06*** (0.31)		-0.55** (0.31)	-0.57** (0.31)
HDCC- GJR		-0.55** (0.31)	-0.53* (0.41)		-0.81*** (0.31)	-1.03*** (0.30)		-0.51** (0.29)	-0.56** (0.30)
RiskMetrics	-0.52** (0.24)	-1.12** (0.54)	-0.52 (0.44)	-0.83*** (0.21)	-1.44*** (0.60)	-1.59*** (0.55)	-0.79** (0.38)	-1.14** (0.53)	-0.59* (0.42)

(f) weekly vs. biweekly returns, scaled forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC- GARCH	-0.14 (0.14)	-0.39** (0.21)	-0.38* (0.25)	-0.80*** (0.24)	-0.83*** (0.23)	-1.23*** (0.29)	-0.39** (0.20)	-0.41** (0.22)	-0.45** (0.26)
DCC- GARCH		-0.43** (0.23)	-0.34* (0.23)		-0.86*** (0.25)	-1.21*** (0.25)		-0.40** (0.21)	-0.45** (0.26)
HDCC- GARCH		-0.39** (0.20)	-0.36* (0.24)		-0.89*** (0.25)	-1.18*** (0.26)		-0.34** (0.21)	-0.44** (0.23)
CCC- GJR	-0.42** (0.22)	-0.40* (0.26)	-0.32 (0.38)	-0.87*** (0.34)	-0.83*** (0.29)	-1.15*** (0.39)	-0.51** (0.27)	-0.51** (0.30)	-0.58** (0.34)
DCC- GJR		-0.43* (0.29)	-0.28 (0.35)		-0.86*** (0.33)	-1.09*** (0.35)		-0.53** (0.31)	-0.48** (0.28)
HDCC- GJR		-0.39* (0.26)	-0.30 (0.36)		-0.87*** (0.32)	-1.09*** (0.36)		-0.46** (0.26)	-0.51** (0.28)
RiskMetrics	-0.44** (0.22)	-1.09** (0.53)	-0.54 (0.43)	-0.90*** (0.27)	-1.65** (0.75)	-1.61*** (0.64)	-0.81** (0.39)	-1.18** (0.55)	-0.62* (0.43)

See table note on previous page.

Table C.8: Summary results of UC and DQ tests, $\vartheta = 5\%$

# methods	UC-test results rejections at			DQ-test results				
	1%	5%	10%	neg.	neg. sign.	pos. sign.	pos.	
(a) Temporal aggregation								
Daily, iterated	51	0	6	10	35	0	16	1
Daily, scaled	51	0	12	22	22	0	29	2
Weekly, iterated	51	0	3	19	0	0	51	7
Weekly, scaled	51	0	9	26	0	0	51	6
Biweekly, direct	51	0	19	35	0	0	51	49
(b) Portfolio aggregation								
Portfolio level	45	0	7	17	10	0	35	14
Asset class level	105	0	15	39	15	0	90	24
Asset level	105	0	27	56	32	0	73	27
(c) Model specification								
GARCH, uv.	15	0	4	7	5	0	10	3
GJR, uv.	15	0	2	7	3	0	12	3
RiskMetrics, uv.	15	0	1	3	2	0	13	8
CCC-GARCH	30	0	8	16	9	0	21	6
DCC-GARCH	30	0	9	19	10	0	20	5
HDCC-GARCH	30	0	9	19	11	0	19	5
CCC-GJR	30	0	4	11	7	0	23	6
DCC-GJR	30	0	5	12	3	0	27	6
HDCC-GJR	30	0	6	13	7	0	23	6
RiskMetrics, mv.	30	0	1	5	0	0	30	17
(d) Distribution								
Normal	85	0	11	47	18	0	67	23
Empirical	85	0	0	4	20	0	65	24
Student's t	85	0	38	61	19	0	66	18

This table summarizes the results of the unconditional coverage (UC) and dynamic quantile (DQ) tests of Christoffersen (1998) and Engle and Manganelli (2004) for the 255 different methods that we analyze. In each panel, we categorize the methods based on the choice aspect given in the panel heading. For each option that can be chosen, we report the total number of methods. Next, we report for the number of rejections from the UC tests for significance levels of 1, 5 and 10%. Under the null (alternative) hypothesis, the actual coverage ratio equals (exceeds) the theoretical coverage ratio of 0.05. To conduct the DQ test, we regress the centered hit series on the negative VaR forecasts. We report the number of negative and positive coefficients, and the number of significant ones at the 10% level. Significance is determined by bootstraps as in Herwartz and Waichman (2010). The full results are available in Tables C.9 and C.10.

Table C.9: Results for the unconditional coverage test, $\vartheta = 5\%$

(a) daily returns, iterated forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC- GARCH	5.66 (0.74)	5.90 (0.73)	6.35** (0.78)	5.50 (0.72)	5.02 (0.68)	5.68 (0.73)	5.56 (0.73)	6.01* (0.75)	6.53** (0.80)
DCC- GARCH		5.86 (0.74)	6.27** (0.77)		5.20 (0.71)	5.66 (0.73)		6.11* (0.76)	6.27** (0.77)
HDCC- GARCH		5.86 (0.75)	6.31** (0.78)		5.28 (0.72)	5.78 (0.73)		6.07* (0.76)	6.41** (0.78)
CCC- GJR	4.88 (0.64)	5.04 (0.65)	5.74 (0.71)	4.92 (0.65)	4.46 (0.61)	5.22 (0.66)	4.92 (0.65)	5.16 (0.65)	5.76 (0.71)
DCC- GJR		5.10 (0.66)	5.58 (0.69)		4.66 (0.63)	5.06 (0.65)		5.14 (0.66)	5.76 (0.70)
HDCC- GJR		5.02 (0.65)	5.70 (0.70)		4.54 (0.63)	5.26 (0.67)		5.02 (0.66)	5.76 (0.70)
RiskMetrics	5.26 (0.71)	5.32 (0.73)	6.11* (0.80)	5.56 (0.72)	5.34 (0.72)	5.76 (0.76)	3.21 (0.55)	5.30 (0.73)	5.95 (0.78)

(b) daily returns, scaled forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC- GARCH	5.62 (0.73)	5.90 (0.73)	6.35** (0.78)	5.64 (0.73)	5.34 (0.69)	5.94 (0.74)	5.76 (0.74)	6.23* (0.76)	6.67** (0.80)
DCC- GARCH		5.99* (0.75)	6.33** (0.77)		5.44 (0.71)	5.88 (0.73)		6.33** (0.77)	6.61** (0.79)
HDCC- GARCH		5.88 (0.74)	6.45** (0.79)		5.52 (0.71)	6.03* (0.76)		6.19* (0.76)	6.59** (0.80)
CCC- GJR	5.60 (0.68)	5.60 (0.68)	6.11* (0.72)	5.54 (0.68)	5.30 (0.66)	5.84 (0.71)	5.82 (0.70)	5.90* (0.70)	6.41** (0.74)
DCC- GJR		5.74 (0.69)	6.19** (0.73)		5.28 (0.66)	5.70 (0.69)		6.05* (0.71)	6.41** (0.74)
HDCC- GJR		5.58 (0.68)	6.09* (0.72)		5.42 (0.67)	5.80 (0.70)		5.95* (0.70)	6.31** (0.73)
RiskMetrics	5.36 (0.71)	5.44 (0.74)	6.19* (0.81)	5.62 (0.73)	5.48 (0.74)	5.97 (0.78)	3.60 (0.59)	5.58 (0.75)	6.33** (0.81)

This table shows the empirical coverage (in %) and results of the unconditional coverage tests of Christoffersen (1998). Models are estimated with a moving window of 1,000 daily returns, 200 weekly returns or 100 biweekly returns. For every day in the sample, we construct a violation indicator that equals one when the ten-day realized portfolio loss exceeds the forecasted VaR_{ϑ} with $\vartheta = 0.05$, and zero otherwise, as in Equation (20). We report the coverage (in %), and test whether it is equal to ϑ against the alternative of strictly more violations. We report Newey and West (1987) standard errors calculated with 10 leads and lags in parentheses. Superscripts a, b, c denote rejection of the null-hypothesis with a significance level of 1%, 5% and 10%. The results are based on 5,021 forecasts.

Table C.9: Results for the unconditional coverage test, $\vartheta = 5\%$ – *continued*

(c) weekly returns, iterated forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC- GARCH	6.09* (0.77)	5.97* (0.75)	5.94 (0.76)	5.66 (0.74)	4.72 (0.65)	4.64 (0.65)	6.35** (0.79)	6.29* (0.79)	6.15* (0.78)
DCC- GARCH		6.05* (0.75)	5.97* (0.75)		4.72 (0.65)	4.78 (0.66)		6.35** (0.79)	6.17* (0.78)
HDCC- GARCH		6.05* (0.75)	6.01* (0.76)		4.62 (0.63)	4.68 (0.66)		6.55** (0.81)	6.25* (0.78)
CCC- GJR	5.95* (0.74)	5.54 (0.70)	5.86 (0.73)	5.58 (0.71)	4.52 (0.63)	4.58 (0.63)	6.21* (0.76)	5.74 (0.71)	6.01* (0.74)
DCC- GJR		5.66 (0.71)	5.84 (0.72)		4.52 (0.63)	4.64 (0.63)		5.88 (0.72)	6.09* (0.74)
HDCC- GJR		5.80 (0.71)	5.88 (0.73)		4.54 (0.62)	4.70 (0.64)		5.95* (0.73)	5.95* (0.74)
RiskMetrics	5.50 (0.74)	5.32 (0.73)	5.48 (0.73)	5.50 (0.73)	4.88 (0.68)	4.68 (0.67)	5.74 (0.75)	5.34 (0.73)	5.60 (0.74)

(d) weekly returns, scaled forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC- GARCH	6.13* (0.78)	5.95 (0.75)	5.97* (0.76)	5.54 (0.73)	4.72 (0.65)	4.64 (0.65)	6.43** (0.80)	6.33** (0.79)	6.29** (0.78)
DCC- GARCH		6.07* (0.75)	5.99* (0.76)		4.70 (0.64)	4.62 (0.64)		6.35** (0.78)	6.21* (0.77)
HDCC- GARCH		6.11* (0.75)	5.99* (0.76)		4.58 (0.64)	4.70 (0.65)		6.45** (0.79)	6.31** (0.79)
CCC- GJR	6.15* (0.75)	5.92 (0.73)	6.03* (0.74)	5.60 (0.71)	4.56 (0.62)	4.72 (0.64)	6.19* (0.75)	6.11* (0.74)	6.13* (0.75)
DCC- GJR		6.01* (0.73)	5.94* (0.72)		4.58 (0.62)	4.86 (0.64)		6.27** (0.75)	6.17* (0.74)
HDCC- GJR		6.11* (0.73)	6.05* (0.74)		4.68 (0.62)	4.60 (0.62)		6.25** (0.75)	6.25** (0.76)
RiskMetrics	5.60 (0.74)	5.32 (0.73)	5.52 (0.73)	5.50 (0.74)	4.94 (0.68)	4.80 (0.69)	5.86 (0.76)	5.38 (0.73)	5.74 (0.75)

(e) biweekly returns, direct forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC- GARCH	6.43** (0.79)	6.03* (0.76)	6.19* (0.78)	5.97* (0.74)	4.30 (0.64)	4.22 (0.63)	6.81** (0.83)	6.41** (0.79)	6.51** (0.80)
DCC- GARCH		6.15* (0.77)	6.27* (0.78)		4.34 (0.63)	4.22 (0.62)		6.43** (0.79)	6.45** (0.79)
HDCC- GARCH		6.13* (0.76)	6.21* (0.78)		4.40 (0.63)	4.30 (0.63)		6.61** (0.80)	6.39** (0.79)
CCC- GJR	6.57** (0.78)	6.19* (0.75)	6.25** (0.76)	6.21* (0.74)	4.46 (0.64)	4.42 (0.63)	6.75** (0.79)	6.45** (0.77)	6.51** (0.77)
DCC- GJR		6.15* (0.75)	6.23* (0.76)		4.56 (0.64)	4.46 (0.63)		6.51** (0.77)	6.45** (0.77)
HDCC- GJR		6.35** (0.77)	6.15* (0.75)		4.68 (0.65)	4.56 (0.64)		6.55** (0.77)	6.35** (0.76)
RiskMetrics	6.03* (0.78)	5.40 (0.73)	6.03* (0.80)	6.05* (0.77)	5.54 (0.72)	5.94 (0.76)	6.49** (0.79)	5.70 (0.74)	6.33* (0.82)

See table note on previous page.

Table C.10: Results of the dynamic quantile tests, $\vartheta = 5\%$

(a) daily returns, iterated forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-	-0.16	-0.13	-0.49	-0.25	-0.34	-0.74	-0.23	-0.18	-0.48
GARCH	(0.36)	(0.38)	(0.24)	(0.33)	(0.25)	(0.14)	(0.32)	(0.35)	(0.24)
DCC-		-0.16	-0.44		-0.25	-0.61		-0.06	-0.27
GARCH		(0.36)	(0.24)		(0.31)	(0.17)		(0.43)	(0.32)
HDCC-		-0.24	-0.47		-0.31	-0.85		-0.07	-0.41
GARCH		(0.33)	(0.24)		(0.28)	(0.11)		(0.42)	(0.26)
CCC-	-0.05	0.16	-0.16	-0.20	-0.20	-0.47	-0.09	0.12	-0.10
GJR	(0.41)	(0.60)	(0.36)	(0.32)	(0.31)	(0.20)	(0.38)	(0.56)	(0.40)
DCC-		0.18	0.00		-0.12	-0.37		0.13	0.02
GJR		(0.63)	(0.47)		(0.36)	(0.23)		(0.57)	(0.48)
HDCC-		0.12	-0.15		-0.18	-0.55		0.15	-0.05
GJR		(0.56)	(0.37)		(0.32)	(0.17)		(0.59)	(0.42)
RiskMetrics	0.06	0.39	0.80*	0.08	0.28	0.60	-0.30	0.35	0.48
	(0.50)	(0.74)	(0.92)	(0.51)	(0.67)	(0.86)	(0.22)	(0.72)	(0.76)

(b) daily returns, scaled forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-	-0.03	0.04	-0.35	0.08	0.09	-0.42	-0.02	0.04	-0.37
GARCH	(0.43)	(0.49)	(0.28)	(0.51)	(0.53)	(0.24)	(0.44)	(0.49)	(0.28)
DCC-		-0.12	-0.26		0.00	-0.07		0.09	-0.24
GARCH		(0.39)	(0.31)		(0.47)	(0.42)		(0.53)	(0.33)
HDCC-		-0.15	-0.36		0.00	-0.49		-0.01	-0.27
GARCH		(0.36)	(0.27)		(0.46)	(0.21)		(0.46)	(0.32)
CCC-	0.12	0.28	-0.05	0.19	0.20	-0.12	0.11	0.33	-0.21
GJR	(0.55)	(0.70)	(0.44)	(0.61)	(0.65)	(0.40)	(0.54)	(0.72)	(0.34)
DCC-		0.31	0.00		0.27	-0.04		0.39	0.11
GJR		(0.72)	(0.47)		(0.70)	(0.44)		(0.76)	(0.55)
HDCC-		0.25	-0.17		0.28	0.00		0.39	-0.03
GJR		(0.67)	(0.35)		(0.70)	(0.46)		(0.77)	(0.45)
RiskMetrics	0.16	0.54	0.78*	0.15	0.71*	0.57	-0.29	0.44	0.64
	(0.58)	(0.84)	(0.91)	(0.57)	(0.94)	(0.85)	(0.24)	(0.77)	(0.83)

This table reports the results of the dynamic quantile test of Engle and Manganelli (2004). We conduct a linear regression of the centered VaR-violations on a constant and the forecasted VaR_ϑ (in %) for $\vartheta = 0.01$, as in Equation (21). We report the estimated coefficient on the forecasted VaR_ϑ (expressed as a decimal rate), and its quantile in the distribution of this coefficient under the null hypothesis of no predictability. Following Herwartz and Waichman (2010), we construct this distribution by 10,000 block bootstrapped samples with a block length of 10. Superscripts a, b, c denote quantiles in the 1%, 5 % and 10% of the left or right tail.

Table C.10: Results of the dynamic quantile test, $\vartheta = 5\%$ – *continued*

(c) weekly returns, iterated forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-	0.94	0.50	0.34	0.75	0.39	0.24	0.89	0.59	0.41
GARCH	(0.83)	(0.53)	(0.36)	(0.75)	(0.53)	(0.33)	(0.79)	(0.58)	(0.40)
DCC-		0.64	0.42		0.54	0.21		0.71	0.48
GARCH		(0.66)	(0.44)		(0.70)	(0.28)		(0.68)	(0.48)
HDCC-		0.63	0.35		0.56	0.24		0.69	0.50
GARCH		(0.66)	(0.37)		(0.74)	(0.32)		(0.67)	(0.48)
CCC-	0.82	0.56	0.28	0.36	0.27	0.14	0.77	0.49	0.40
GJR	(0.85)	(0.67)	(0.32)	(0.46)	(0.42)	(0.20)	(0.81)	(0.58)	(0.43)
DCC-		0.62	0.43		0.35	0.15		0.49	0.56
GJR		(0.72)	(0.48)		(0.53)	(0.21)		(0.58)	(0.59)
HDCC-		0.66	0.27		0.38	0.31		0.66	0.42
GJR		(0.77)	(0.31)		(0.57)	(0.42)		(0.74)	(0.46)
RiskMetrics	0.93*	0.88	1.02*	0.95*	0.84*	0.85**	0.95*	0.93	0.93*
	(0.91)	(0.87)	(0.92)	(0.92)	(0.92)	(0.96)	(0.91)	(0.88)	(0.92)

(d) weekly returns, scaled forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-	0.91	0.63	0.30	0.74	0.57	0.28	0.88	0.72	0.37
GARCH	(0.83)	(0.65)	(0.31)	(0.83)	(0.75)	(0.39)	(0.79)	(0.68)	(0.36)
DCC-		0.75	0.36		0.61	0.53		0.70	0.52
GARCH		(0.75)	(0.38)		(0.80)	(0.69)		(0.69)	(0.51)
HDCC-		0.71	0.30		0.57	0.31		0.96	0.46
GARCH		(0.73)	(0.31)		(0.79)	(0.43)		(0.84)	(0.45)
CCC-	0.77	0.62	0.28	0.71	0.37	0.28	0.71	0.59	0.29
GJR	(0.81)	(0.70)	(0.31)	(0.82)	(0.61)	(0.40)	(0.77)	(0.65)	(0.31)
DCC-		0.72	0.34		0.39	0.28		0.61	0.47
GJR		(0.78)	(0.40)		(0.61)	(0.42)		(0.68)	(0.51)
HDCC-		0.71	0.40		0.50	0.28		0.62	0.41
GJR		(0.79)	(0.45)		(0.74)	(0.43)		(0.69)	(0.46)
RiskMetrics	0.94*	0.90	1.06*	0.90*	0.94**	0.95**	0.88	0.85	0.95*
	(0.92)	(0.89)	(0.93)	(0.94)	(0.96)	(0.97)	(0.88)	(0.84)	(0.92)

(e) biweekly returns, direct forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-	1.04**	1.23**	0.97*	1.09**	0.71**	0.88**	1.00*	1.29**	0.98*
GARCH	(0.97)	(0.98)	(0.91)	(0.98)	(0.96)	(0.99)	(0.93)	(0.97)	(0.90)
DCC-		1.28**	1.10*		0.72**	0.93***		1.26**	0.99
GARCH		(0.98)	(0.95)		(0.96)	(0.99)		(0.97)	(0.90)
HDCC-		1.38***	1.05*		0.81**	0.98***		1.32**	0.92
GARCH		(0.99)	(0.94)		(0.98)	(0.99)		(0.98)	(0.90)
CCC-	1.24***	1.38***	1.18**	1.03***	0.78**	0.72**	1.08**	1.36***	1.12**
GJR	(1.00)	(1.00)	(0.98)	(1.00)	(0.99)	(0.96)	(0.99)	(0.99)	(0.95)
DCC-		1.40***	1.14**		0.76**	0.71**		1.32***	1.16**
GJR		(1.00)	(0.97)		(0.99)	(0.96)		(1.00)	(0.97)
HDCC-		1.46***	1.14**		0.82***	0.77**		1.43***	1.21**
GJR		(1.00)	(0.97)		(0.99)	(0.98)		(1.00)	(0.97)
RiskMetrics	1.22***	1.20***	1.75***	1.03**	1.17***	1.58***	1.55***	1.20***	2.50***
	(1.00)	(1.00)	(1.00)	(0.98)	(1.00)	(1.00)	(1.00)	(0.99)	(1.00)

See table note on previous page.

Table C.11: Average asymmetric tick loss, $\vartheta = 5\%$

(a) daily returns, iterated forecasts									
Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-GARCH	18.25	17.77	17.87	18.29	17.80	17.84	18.22	17.79	17.84
DCC-GARCH		17.89	17.74		17.94	17.71		17.94	17.72
HDCC-GARCH		17.96	17.84		18.00	17.78		17.94	17.84
CCC-GJR	17.99	17.67	17.47	18.02	17.79	17.55	17.94	17.62	17.50
DCC-GJR		17.79	17.43		17.89	17.46		17.68	17.43
HDCC-GJR		17.79	17.44		17.91	17.50		17.78	17.48
RiskMetrics	18.04	18.31	18.61	18.14	18.26	18.42	18.71	18.24	18.46

(b) daily returns, scaled forecasts									
Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-GARCH	18.15	17.74	17.76	18.17	17.76	17.70	18.14	17.79	17.81
DCC-GARCH		17.84	17.65		17.86	17.60		17.88	17.69
HDCC-GARCH		17.90	17.75		17.90	17.65		17.95	17.79
CCC-GJR	17.95	17.68	17.50	18.00	17.72	17.50	17.95	17.69	17.53
DCC-GJR		17.76	17.43		17.80	17.43		17.70	17.47
HDCC-GJR		17.79	17.44		17.86	17.46		17.81	17.51
RiskMetrics	18.04	18.28	18.61	18.08	18.36	18.32	18.37	18.22	18.51

(c) weekly returns, iterated forecasts									
Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-GARCH	19.48	19.04	18.96	19.42	19.00	18.87	19.30	19.03	18.98
DCC-GARCH		19.06	18.92		19.00	18.84		19.02	18.91
HDCC-GARCH		19.12	18.97		19.09	18.87		19.10	19.00
CCC-GJR	18.96	18.87	18.67	18.90	18.90	18.65	18.81	18.71	18.66
DCC-GJR		18.90	18.65		18.93	18.65		18.70	18.64
HDCC-GJR		18.95	18.70		18.96	18.69		18.81	18.69
RiskMetrics	19.53	19.25	19.55	19.46	19.37	19.89	19.41	19.25	19.66

This table shows the average value of the asymmetric tick loss function of Giacomini and Komunjer (2005) as in Equation (22) for $\vartheta = 0.05$. All values have been multiplied by 100. See Table 2.

Table C.11: Average asymmetric tick loss, $\vartheta = 5\%$ – *continued*

(d) weekly returns, scaled forecasts									
Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-GARCH	19.51	19.03	18.93	19.40	19.02	18.82	19.32	19.04	18.99
DCC-GARCH		19.03	18.87		19.05	18.80		19.00	18.94
HDCC-GARCH		19.08	18.95		19.12	18.84		19.07	18.99
CCC-GJR	18.89	18.85	18.64	18.84	18.88	18.59	18.76	18.70	18.63
DCC-GJR		18.87	18.62		18.86	18.56		18.70	18.62
HDCC-GJR		18.93	18.67		18.89	18.59		18.78	18.66
RiskMetrics	19.53	19.24	19.56	19.57	19.42	19.83	19.43	19.26	19.63

(e) biweekly returns, direct forecasts									
Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-GARCH	19.65	19.44	19.37	19.56	19.50	19.53	19.50	19.47	19.32
DCC-GARCH		19.41	19.33		19.45	19.51		19.45	19.36
HDCC-GARCH		19.48	19.38		19.49	19.52		19.47	19.36
CCC-GJR	19.40	19.16	19.07	19.21	19.19	19.15	19.15	19.25	19.11
DCC-GJR		19.16	19.06		19.17	19.14		19.20	19.08
HDCC-GJR		19.20	19.08		19.24	19.17		19.26	19.10
RiskMetrics	19.93	19.80	20.72	19.87	20.03	20.91	19.84	19.81	20.80

See table note on previous page.

Table C.12: Models removed from the Model Confidence Set, $\vartheta = 5\%$

(a) daily returns, iterated forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-GARCH									
DCC-GARCH									
HDCC-GARCH									
CCC-GJR									
DCC-GJR									
HDCC-GJR									
RiskMetrics			0.072						0.084

(b) daily returns, iterated forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-GARCH									
DCC-GARCH									
HDCC-GARCH									
CCC-GJR									
DCC-GJR									
HDCC-GJR									
RiskMetrics			0.074						

(c) weekly returns, iterated forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-GARCH	0.055	0.063		0.051	0.055	0.074	0.060	0.066	
DCC-GARCH		0.062			0.052	0.071		0.062	
HDCC-GARCH		0.055			0.050	0.081		0.060	
CCC-GJR	0.062	0.063		0.060	0.056		0.064	0.071	
DCC-GJR		0.062			0.055	0.071		0.071	
HDCC-GJR		0.055			0.053	0.077		0.060	
RiskMetrics	0.054	0.074	0.078	0.055	0.062	0.052	0.045	0.097	0.071

This table presents the MCS p -value with which a model has been removed from the Model Confidence Set. We follow the procedure of Hansen et al. (2011, Sec. 3.1.2), with the tick loss function in (22), $\vartheta = 0.05$, and a significance level of 10%. The procedure starts with the complete set of 255 methods. Panel (f) presents the number of methods with a particular choice that have been removed from (“Out”) and are maintained in (“In”) the Model Confidence Set.

Table C.12: Models removed from the Model Confidence Set, $\vartheta = 5\%$ – *continued*

(d) weekly returns, scaled forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-GARCH	0.053	0.062		0.048	0.052	0.079	0.057	0.063	
DCC-GARCH		0.061			0.052	0.071		0.061	
HDCC-GARCH		0.055			0.044	0.074		0.058	
CCC-GJR	0.063	0.068		0.060	0.055		0.069	0.071	
DCC-GJR		0.062			0.057			0.071	
HDCC-GJR		0.056			0.055	0.097		0.062	
RiskMetrics	0.055	0.077	0.079	0.053	0.059	0.052	0.045	0.098	0.071

(e) biweekly returns, direct forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-GARCH	0.051	0.060	0.069	0.045	0.042	0.046	0.052	0.060	0.071
DCC-GARCH		0.060	0.067		0.044	0.048		0.056	0.071
HDCC-GARCH		0.055	0.066		0.041	0.048		0.055	0.071
CCC-GJR	0.051	0.071	0.097	0.052	0.050	0.055	0.055	0.059	0.074
DCC-GJR		0.064	0.082		0.048	0.055		0.057	0.071
HDCC-GJR		0.063	0.083		0.044	0.054		0.054	0.074
RiskMetrics	0.048	0.071	0.055	0.045	0.060	0.051	0.044	0.071	0.057

(f) Number of removed and maintained methods

Temporal aggr.	Out	In	Total
Daily, iterated	2	49	51
Daily, scaled	1	50	51
Weekly, iterated	38	13	51
Weekly, scaled	37	14	51
Biweekly, direct	51	0	51
Portfolio aggr.			
Portfolio level	28	17	45
Asset class level	63	42	105
Asset level	38	67	105
Disitribution			
Normal	41	44	85
Empirical	48	37	85
Student's t	40	45	85

Model	Out	In	Total
GARCH, uv.	9	6	15
GJR, uv.	9	6	15
RiskMetrics, uv.	10	5	15
CCC-GARCH	14	16	30
DCC-GARCH	14	16	30
HDCC-GARCH	14	16	30
CCC-GJR	12	18	30
DCC-GJR	13	17	30
HDCC-GJR	14	16	30
RiskMetrics, mv.	20	10	30

See table note on the previous page.

Table C.13: Results of the Diebold-Mariano tests, $\vartheta = 5\%$

Method A	Method B	# of tests	neg.	sign. 5%	neg. 10%	pos.	sign. 5%	pos. 10%
(a) Forecasting method								
Iterated	Scaled	51	13	0	0	38	1	1
(b) Portfolio aggregation, iterated forecasts								
Asset level	Asset class level	21	15	5	9	6	0	1
Asset level	Portfolio level	21	19	13	18	2	1	1
Asset class level	Portfolio level	21	19	0	5	2	0	0
(c) Portfolio aggregation, scaled forecasts								
Asset level	Asset class level	21	17	5	9	4	0	1
Asset level	Portfolio level	21	18	13	18	3	1	1
Asset class level	Portfolio level	21	19	0	5	2	0	0
(d) Distribution, iterated forecasts								
Normal	Empirical	17	5	1	1	12	3	5
Normal	Student's t	17	12	0	2	5	1	1
Empirical	Student's t	17	4	1	1	13	2	3
(e) Distribution, scaled forecasts								
Normal	Empirical	17	12	1	1	5	3	5
Normal	Student's t	17	11	0	2	6	1	1
Empirical	Student's t	17	11	1	1	6	2	3
(f) Model, iterated forecasts								
RiskMetrics, uv.	other, uv.	6	2	0	0	4	1	1
RiskMetrics, mv.	other, mv.	36	0	0	0	36	15	19
GARCH, uv.	GJR, uv.	3	0	0	0	3	0	0
GARCH, mv.	GJR, mv.	54	4	0	0	50	21	27
CCC	DCC	12	6	0	2	6	0	0
CCC	HDCC	12	6	0	3	6	0	0
DCC	HDCC	12	11	0	1	1	0	0
(g) Model, scaled forecasts								
RiskMetrics, uv.	other, uv.	6	2	0	0	4	0	1
RiskMetrics, mv.	other, mv.	36	0	0	0	36	15	18
GARCH, uv.	GJR, uv.	3	0	0	0	3	0	0
GARCH, mv.	GJR, mv.	54	5	0	0	49	11	22
CCC	DCC	12	6	0	0	6	0	0
CCC	HDCC	12	6	0	0	6	0	0
DCC	HDCC	12	12	0	3	0	0	0

This table shows summary results of the tests that the expected value of the loss function of methods A and B are equal, $L^A = L^B$. We calculate the loss differential as $L^A - L^B$. A negative loss differential means that method A is preferable to method B. We use the asymmetric tick loss function of Giacomini and Komunjer (2005) as in Equation (22) with $\vartheta = 0.05$. We report the number of negative and positive average loss differentials, and the number of times these differentials are significant for the 5% and 10% significance levels, based on the statistic proposed by Diebold and Mariano (1995), evaluated in the setting of Giacomini and White (2006). Standard errors of the average loss differential are based on Newey and West (1987) with 10 leads and lags. We compare pairs of methods that differ in only one choice aspect (indicated by the panel heading), so the other four choices are the same. All methods use daily data.

Table C.14: Results of the DM-tests for iterated vs. scaled forecasts, $\vartheta = 5\%$

(a) daily returns									
Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-	0.09	0.11	0.02	0.13	0.14	0.03	0.08	0.03	0.00
GARCH	(0.09)	(0.10)	(0.07)	(0.12)	(0.12)	(0.11)	(0.09)	(0.07)	(0.07)
DCC-		0.10	0.05		0.11	0.08		0.03	0.05
GARCH		(0.10)	(0.08)		(0.12)	(0.11)		(0.08)	(0.08)
HDCC-		0.09	0.06		0.13	0.09		0.05	-0.01
GARCH		(0.10)	(0.09)		(0.12)	(0.11)		(0.08)	(0.07)
CCC-	0.04	-0.03	-0.01	0.02	0.05	0.07	0.00	-0.03	-0.07
GJR	(0.09)	(0.07)	(0.08)	(0.13)	(0.11)	(0.11)	(0.09)	(0.08)	(0.10)
DCC-		0.01	0.03		0.03	0.09		-0.03	-0.03
GJR		(0.07)	(0.08)		(0.10)	(0.11)		(0.07)	(0.09)
HDCC-		0.01	0.00		0.04	0.06		-0.03	-0.03
GJR		(0.07)	(0.08)		(0.10)	(0.11)		(0.07)	(0.09)
RiskMetrics	0.00	0.00	0.02	0.06	0.11	-0.10	0.34***	-0.05	0.02
	(0.03)	(0.01)	(0.02)	(0.10)	(0.13)	(0.08)	(0.09)	(0.06)	(0.04)

This table shows summary results of the tests that the expected value of the loss function of methods A and B are equal, $L^A = L^B$. We report the loss differential $L^A - L^B$ (multiplied by 100). A negative loss differential means that method A is preferable to method B. We use the asymmetric tick loss function of Giacomini and Komunjer (2005) as in Equation (22) with $\vartheta = 0.05$. The two methods differ in their forecast construction, with A using iterated and B using scaled forecasts. Both methods use daily data. We report standard errors of the average loss differential in parentheses, based on Newey and West (1987) with 10 leads and lags. We test that the average loss differential is zero based on the statistic proposed by Diebold and Mariano (1995), evaluated in the setting of Giacomini and White (2006). Superscripts *, **, *** denote rejection of the null-hypothesis with a significance level of 10%, 5% and 1%.

Table C.15: Results of the DM-tests for portfolio aggregation, $\vartheta = 5\%$

(a) daily returns, iterated forecasts									
Test	Normal			Empirical			Student's t		
	1 vs. 3	1 vs. 8	3 vs. 8	1 vs. 3	1 vs. 8	3 vs. 8	1 vs. 3	1 vs. 8	3 vs. 8
CCC-	0.48*	0.38*	-0.10	0.50*	0.45*	-0.05	0.43*	0.38*	-0.05
GARCH	(0.34)	(0.26)	(0.21)	(0.36)	(0.29)	(0.20)	(0.33)	(0.28)	(0.20)
DCC-	0.35*	0.50**	0.15	0.36	0.58**	0.23*	0.28	0.50*	0.21
GARCH	(0.26)	(0.30)	(0.16)	(0.28)	(0.35)	(0.17)	(0.27)	(0.32)	(0.18)
HDCC-	0.29	0.40**	0.11	0.30	0.52**	0.22*	0.28	0.38*	0.10
GARCH	(0.23)	(0.24)	(0.16)	(0.25)	(0.30)	(0.16)	(0.25)	(0.26)	(0.18)
CCC-	0.32	0.52**	0.20	0.23	0.48**	0.25*	0.33*	0.45**	0.12
GJR	(0.26)	(0.23)	(0.18)	(0.26)	(0.25)	(0.18)	(0.25)	(0.24)	(0.18)
DCC-	0.20	0.55**	0.35**	0.13	0.56**	0.43**	0.27	0.51**	0.24*
GJR	(0.19)	(0.26)	(0.17)	(0.19)	(0.28)	(0.19)	(0.23)	(0.27)	(0.18)
HDCC-	0.20	0.54**	0.34**	0.11	0.52**	0.41**	0.16	0.46**	0.30**
GJR	(0.18)	(0.25)	(0.18)	(0.18)	(0.24)	(0.18)	(0.18)	(0.24)	(0.17)
RiskMetrics	-0.26	-0.57**	-0.31*	-0.12	-0.29	-0.16	0.47	0.25	-0.22
	(0.31)	(0.33)	(0.21)	(0.33)	(0.32)	(0.22)	(0.44)	(0.58)	(0.21)

(b) daily returns, scaled forecasts									
Test	Normal			Empirical			Student's t		
	1 vs. 3	1 vs. 8	3 vs. 8	1 vs. 3	1 vs. 8	3 vs. 8	1 vs. 3	1 vs. 8	3 vs. 8
CCC-	0.41	0.40*	-0.02	0.41	0.47*	0.06	0.35	0.34	-0.02
GARCH	(0.33)	(0.27)	(0.20)	(0.37)	(0.30)	(0.21)	(0.32)	(0.28)	(0.21)
DCC-	0.31	0.51*	0.19	0.31	0.57*	0.26*	0.26	0.45*	0.19
GARCH	(0.25)	(0.31)	(0.17)	(0.28)	(0.35)	(0.18)	(0.26)	(0.31)	(0.18)
HDCC-	0.26	0.41*	0.15	0.27	0.52**	0.25*	0.19	0.35*	0.16
GARCH	(0.23)	(0.25)	(0.16)	(0.26)	(0.30)	(0.17)	(0.23)	(0.26)	(0.18)
CCC-	0.27	0.45**	0.18	0.28	0.50**	0.22	0.26	0.42*	0.16
GJR	(0.25)	(0.24)	(0.18)	(0.27)	(0.25)	(0.19)	(0.26)	(0.26)	(0.17)
DCC-	0.19	0.52**	0.33**	0.21	0.58**	0.37**	0.24	0.48**	0.24*
GJR	(0.18)	(0.26)	(0.17)	(0.19)	(0.26)	(0.18)	(0.22)	(0.27)	(0.18)
HDCC-	0.16	0.51**	0.35**	0.15	0.55***	0.40***	0.13	0.43**	0.30**
GJR	(0.17)	(0.25)	(0.18)	(0.18)	(0.23)	(0.17)	(0.17)	(0.24)	(0.18)
RiskMetrics	-0.24	-0.56**	-0.32*	-0.28	-0.24	0.04	0.15	-0.14	-0.29*
	(0.31)	(0.32)	(0.21)	(0.33)	(0.29)	(0.29)	(0.42)	(0.57)	(0.22)

This table shows summary results of the tests that the expected value of the loss function of methods A and B are equal, $L^A = L^B$. We report the loss differential $L^A - L^B$ (multiplied by 100). A negative loss differential means that method A is preferable to method B. We use the asymmetric tick loss function of Giacomini and Komunjer (2005) as in Equation (22) with $\vartheta = 0.01$. The two methods differ in their degree of portfolio aggregation. We consider aggregation into a portfolio (labeled “1”), three asset classes (labeled “3”), and no aggregation (labeled “8”). All methods use daily data. We report standard errors of the average loss differential in parentheses, based on Newey and West (1987) with 10 leads and lags. We test that the average loss differential is zero based on the statistic proposed by Diebold and Mariano (1995), evaluated in the setting of Giacomini and White (2006). Superscripts *, **, *** denote rejection of the null-hypothesis with a significance level of 10%, 5% and 1%.

Table C.16: Results of the DM-tests for the model specification, $\vartheta = 5\%$

(a) daily returns, iterated forecasts									
Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-GARCH vs.		-0.13	0.13		-0.14*	0.13		-0.15*	0.12
DCC-GARCH		(0.11)	(0.12)		(0.11)	(0.12)		(0.11)	(0.11)
CCC-GARCH vs.		-0.19*	0.03		-0.20*	0.07		-0.15	0.00
HDCC-GARCH		(0.15)	(0.06)		(0.14)	(0.07)		(0.12)	(0.07)
CCC-GARCH vs.	0.26	0.09	0.40**	0.27	0.00	0.30**	0.28	0.17	0.34**
CCC-GJR	(0.24)	(0.18)	(0.20)	(0.22)	(0.17)	(0.17)	(0.24)	(0.18)	(0.18)
CCC-GARCH vs.		-0.02	0.44**		-0.09	0.39**		0.11	0.41**
DCC-GJR		(0.16)	(0.26)		(0.16)	(0.22)		(0.18)	(0.24)
CCC-GARCH vs.		-0.02	0.43**		-0.12	0.34**		0.01	0.36**
HDCC-GJR		(0.17)	(0.24)		(0.17)	(0.18)		(0.17)	(0.20)
CCC-GARCH vs.	0.20	-0.54*	-0.74***	0.16	-0.46	-0.58**	-0.49	-0.45	-0.62*
RiskMetrics	(0.23)	(0.40)	(0.28)	(0.19)	(0.45)	(0.31)	(0.41)	(0.40)	(0.41)
DCC-GARCH vs.		-0.06	-0.10		-0.06	-0.07		0.00	-0.11
HDCC-GARCH		(0.05)	(0.10)		(0.05)	(0.08)		(0.04)	(0.10)
DCC-GARCH vs.		0.22	0.27**		0.14	0.16		0.32*	0.23*
CCC-GJR		(0.25)	(0.16)		(0.22)	(0.15)		(0.25)	(0.16)
DCC-GARCH vs.		0.11	0.31**		0.05	0.25**		0.26	0.29**
DCC-GJR		(0.18)	(0.17)		(0.16)	(0.15)		(0.22)	(0.17)
DCC-GARCH vs.		0.11	0.30**		0.03	0.21*		0.15	0.24*
HDCC-GJR		(0.19)	(0.17)		(0.17)	(0.13)		(0.19)	(0.15)
DCC-GARCH vs.		-0.41	-0.87***		-0.32	-0.71**		-0.30	-0.74*
RiskMetrics		(0.34)	(0.30)		(0.39)	(0.33)		(0.36)	(0.45)
HDCC-GARCH vs.		0.28	0.37**		0.20	0.23*		0.32	0.34**
CCC-GJR		(0.28)	(0.21)		(0.25)	(0.16)		(0.26)	(0.20)
HDCC-GARCH vs.		0.17	0.41*		0.11	0.32**		0.26	0.40**
DCC-GJR		(0.21)	(0.25)		(0.19)	(0.19)		(0.24)	(0.24)
HDCC-GARCH vs.		0.17	0.40**		0.08	0.28**		0.16	0.36**
HDCC-GJR		(0.21)	(0.24)		(0.18)	(0.15)		(0.20)	(0.21)
HDCC-GARCH vs.		-0.35	-0.77***		-0.26	-0.64**		-0.30	-0.62*
RiskMetrics		(0.30)	(0.27)		(0.36)	(0.32)		(0.35)	(0.39)
CCC-GJR vs.		-0.12	0.04		-0.09	0.09		-0.06	0.06
DCC-GJR		(0.10)	(0.09)		(0.10)	(0.10)		(0.08)	(0.10)
CCC-GJR vs.		-0.11	0.03		-0.12	0.05		-0.17*	0.02
HDCC-GJR		(0.11)	(0.08)		(0.12)	(0.07)		(0.11)	(0.06)
CCC-GJR vs.	-0.05	-0.63	-1.14***	-0.11	-0.46	-0.87**	-0.76***	-0.62	-0.96**
RiskMetrics	(0.18)	(0.51)	(0.40)	(0.19)	(0.55)	(0.42)	(0.29)	(0.50)	(0.54)
DCC-GJR vs.		0.00	-0.01		-0.02	-0.04		-0.11*	-0.05
HDCC-GJR		(0.04)	(0.04)		(0.04)	(0.07)		(0.07)	(0.06)
DCC-GJR vs.		-0.52	-1.18***		-0.37	-0.96**		-0.56	-1.03**
RiskMetrics		(0.45)	(0.43)		(0.48)	(0.43)		(0.49)	(0.57)
HDCC-GJR vs.		-0.52	-1.17***		-0.35	-0.92***		-0.45	-0.98**
RiskMetrics		(0.44)	(0.42)		(0.47)	(0.39)		(0.45)	(0.55)

This table shows summary results of the tests that the expected value of the loss function of methods A and B are equal, $L^A = L^B$. We report the loss differential $L^A - L^B$ (multiplied by 100). A negative loss differential means that method A is preferable to method B. We use the asymmetric tick loss function of Giacomini and Komunjer (2005) as in Equation (22) with $\vartheta = 0.05$. The two methods differ in their model specification. All methods use daily data. We report standard errors of the average loss differential in parentheses, based on Newey and West (1987) with 10 leads and lags. We test that the average loss differential is zero based on the statistic proposed by Diebold and Mariano (1995), evaluated in the setting of Giacomini and White (2006). Superscripts *, **, *** denote rejection of the null-hypothesis with a significance level of 10%, 5% and 1%.

Table C.16: Results of the DM-tests for the model specification, $\vartheta = 5\%$ – *continued*

(b) daily returns, scaled forecasts									
Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-GARCH vs. DCC-GARCH		-0.10 (0.11)	0.11 (0.12)		-0.10 (0.11)	0.11 (0.12)		-0.09 (0.10)	0.11 (0.13)
CCC-GARCH vs. HDCC-GARCH		-0.15 (0.13)	0.01 (0.06)		-0.14 (0.13)	0.05 (0.07)		-0.16 (0.13)	0.02 (0.08)
CCC-GARCH vs. CCC-GJR	0.21 (0.19)	0.06 (0.16)	0.26** (0.15)	0.17 (0.21)	0.04 (0.16)	0.20* (0.13)	0.19 (0.19)	0.10 (0.17)	0.28** (0.16)
CCC-GARCH vs. DCC-GJR		-0.02 (0.17)	0.33** (0.20)		-0.03 (0.17)	0.28** (0.17)		0.09 (0.17)	0.34** (0.20)
CCC-GARCH vs. HDCC-GJR		-0.05 (0.18)	0.32** (0.19)		-0.09 (0.18)	0.24** (0.14)		-0.02 (0.18)	0.29** (0.17)
CCC-GARCH vs. RiskMetrics	0.11 (0.16)	-0.54 (0.45)	-0.85*** (0.32)	0.09 (0.14)	-0.59 (0.51)	-0.61** (0.28)	-0.23 (0.31)	-0.43 (0.45)	-0.70* (0.47)
DCC-GARCH vs. HDCC-GARCH		-0.05 (0.04)	-0.10 (0.11)		-0.05 (0.05)	-0.05 (0.08)		-0.07* (0.05)	-0.09 (0.10)
DCC-GARCH vs. CCC-GJR		0.16 (0.21)	0.14 (0.15)		0.13 (0.21)	0.10 (0.16)		0.19 (0.22)	0.17 (0.15)
DCC-GARCH vs. DCC-GJR		0.08 (0.17)	0.22** (0.13)		0.06 (0.16)	0.17* (0.12)		0.18 (0.18)	0.23** (0.13)
DCC-GARCH vs. HDCC-GJR		0.05 (0.17)	0.21* (0.14)		0.00 (0.17)	0.14 (0.13)		0.07 (0.17)	0.18* (0.12)
DCC-GARCH vs. RiskMetrics		-0.44 (0.39)	-0.96*** (0.36)		-0.50 (0.43)	-0.72*** (0.30)		-0.34 (0.40)	-0.81* (0.53)
HDCC-GARCH vs. CCC-GJR		0.21 (0.23)	0.25* (0.17)		0.18 (0.22)	0.15 (0.14)		0.26 (0.24)	0.26* (0.18)
HDCC-GARCH vs. DCC-GJR		0.14 (0.18)	0.32* (0.19)		0.11 (0.17)	0.22* (0.14)		0.25 (0.21)	0.32* (0.20)
HDCC-GARCH vs. HDCC-GJR		0.11 (0.17)	0.31* (0.19)		0.05 (0.17)	0.19* (0.12)		0.14 (0.18)	0.27** (0.16)
HDCC-GARCH vs. RiskMetrics		-0.39 (0.36)	-0.86*** (0.31)		-0.45 (0.42)	-0.67*** (0.28)		-0.27 (0.38)	-0.72** (0.46)
CCC-GJR vs. DCC-GJR		-0.08 (0.10)	0.07 (0.09)		-0.07 (0.10)	0.07 (0.09)		-0.01 (0.08)	0.06 (0.10)
CCC-GJR vs. HDCC-GJR		-0.10 (0.12)	0.07 (0.09)		-0.13 (0.12)	0.04 (0.07)		-0.12 (0.12)	0.01 (0.07)
CCC-GJR vs. RiskMetrics	-0.09 (0.19)	-0.60 (0.52)	-1.10*** (0.41)	-0.08 (0.19)	-0.63 (0.58)	-0.81** (0.36)	-0.42* (0.29)	-0.53 (0.53)	-0.98** (0.58)
DCC-GJR vs. HDCC-GJR		-0.03 (0.04)	-0.01 (0.04)		-0.06* (0.05)	-0.03 (0.06)		-0.11* (0.07)	-0.05 (0.06)
DCC-GJR vs. RiskMetrics		-0.52 (0.47)	-1.18*** (0.44)		-0.56 (0.52)	-0.89*** (0.35)		-0.52 (0.51)	-1.04** (0.61)
HDCC-GJR vs. RiskMetrics		-0.50 (0.45)	-1.17*** (0.44)		-0.50 (0.50)	-0.86*** (0.33)		-0.40 (0.46)	-0.99** (0.58)

See table note on previous page.

Table C.17: Results of the DM-tests for the distribution, $\vartheta = 5\%$

(a) daily returns, iterated forecasts									
Asset split	Normal vs. t			Normal vs. Empirical			Empirical vs. t		
	1	3	8	1	3	8	1	3	8
CCC-	0.03*	-0.02	0.03	-0.05	-0.03	0.02	0.08	0.01	0.00
GARCH	(0.02)	(0.04)	(0.05)	(0.08)	(0.09)	(0.11)	(0.08)	(0.11)	(0.10)
DCC-		-0.04	0.02		-0.04	0.03		0.00	-0.01
GARCH		(0.05)	(0.04)		(0.09)	(0.11)		(0.11)	(0.12)
HDCC-		0.02	0.01		-0.04	0.06		0.06	-0.06
GARCH		(0.07)	(0.05)		(0.10)	(0.13)		(0.11)	(0.14)
CCC-	0.05**	0.06**	-0.03	-0.03	-0.12*	-0.08	0.08	0.18**	0.05
GJR	(0.02)	(0.03)	(0.04)	(0.11)	(0.09)	(0.12)	(0.10)	(0.10)	(0.12)
DCC-		0.11*	0.00		-0.10	-0.02		0.21**	0.03
GJR		(0.07)	(0.04)		(0.09)	(0.10)		(0.12)	(0.12)
HDCC-		0.01	-0.04		-0.12*	-0.06		0.13*	0.02
GJR		(0.04)	(0.04)		(0.08)	(0.09)		(0.09)	(0.11)
RiskMetrics	-0.66***	0.07**	0.15	-0.09	0.05	0.19**	-0.57**	0.02	-0.04
	(0.27)	(0.04)	(0.22)	(0.10)	(0.06)	(0.11)	(0.32)	(0.08)	(0.26)

(b) daily returns, scaled forecasts									
Asset split	Normal vs. t			Normal vs. Empirical			Empirical vs. t		
	1	3	8	1	3	8	1	3	8
CCC-	0.01	-0.05	-0.05*	-0.02	-0.02	0.05	0.03	-0.03	-0.10
GARCH	(0.03)	(0.05)	(0.04)	(0.04)	(0.08)	(0.06)	(0.05)	(0.11)	(0.09)
DCC-		-0.04	-0.05		-0.02	0.05		-0.02	-0.10
GARCH		(0.05)	(0.04)		(0.07)	(0.07)		(0.11)	(0.09)
HDCC-		-0.06	-0.04		-0.01	0.10		-0.05	-0.14*
GARCH		(0.06)	(0.05)		(0.08)	(0.09)		(0.12)	(0.10)
CCC-	0.00	-0.01	-0.02	-0.06*	-0.04	0.00	0.06*	0.04	-0.03
GJR	(0.02)	(0.04)	(0.04)	(0.04)	(0.06)	(0.04)	(0.04)	(0.08)	(0.07)
DCC-		0.06	-0.04		-0.04	0.00		0.09	-0.04
GJR		(0.06)	(0.04)		(0.06)	(0.04)		(0.09)	(0.07)
HDCC-		-0.03	-0.08*		-0.07	-0.02		0.04	-0.06
GJR		(0.04)	(0.05)		(0.06)	(0.07)		(0.08)	(0.07)
RiskMetrics	-0.32*	0.07*	0.10	-0.04	-0.07	0.29**	-0.29	0.14*	-0.19
	(0.21)	(0.04)	(0.24)	(0.07)	(0.07)	(0.14)	(0.25)	(0.09)	(0.35)

This table shows summary results of the tests that the expected value of the loss function of methods A and B are equal, $L^A = L^B$. We report the loss differential $L^A - L^B$ (multiplied by 100). A negative loss differential means that method A is preferable to method B. We use the asymmetric tick loss function of Giacomini and Komunjer (2005) as in Equation (22) with $\vartheta = 0.01$. The two methods differ in the distribution of the innovations. We consider the normal, empirical and Student's t -distributions. All methods use daily data. We report standard errors of the average loss differential in parentheses, based on Newey and West (1987) with 10 leads and lags. We test that the average loss differential is zero based on the statistic proposed by Diebold and Mariano (1995), evaluated in the setting of Giacomini and White (2006). Superscripts *, **, *** denote rejection of the null-hypothesis with a significance level of 10%, 5% and 1%.

Table C.18: Summary results of UC and DQ tests, 8-year estimation window

# methods	UC-test results rejections at			DQ-test results				
	1%	5%	10%	neg.	neg. sign.	pos. sign.	pos.	
(a) Temporal aggregation								
Daily, iterated	51	3	17	30	48	4	3	0
Daily, scaled	51	4	29	42	49	1	2	0
Weekly, iterated	51	10	34	39	34	0	17	2
Weekly, scaled	51	20	36	37	28	0	23	1
Biweekly, direct	51	32	36	37	3	0	48	11
(b) Portfolio aggregation								
Portfolio level	45	16	27	36	18	4	27	8
Asset class level	105	22	55	68	68	0	37	4
Asset level	105	31	70	81	76	1	29	2
(c) Model specification								
GARCH, uv.	15	6	11	14	6	0	9	1
GJR, uv.	15	5	9	12	6	3	9	4
RiskMetrics, uv.	15	5	7	10	6	1	9	7
CCC-GARCH	30	12	21	24	23	0	7	0
DCC-GARCH	30	8	21	25	23	0	7	1
HDCC-GARCH	30	11	20	25	19	0	11	0
CCC-GJR	30	7	19	21	24	1	6	-3
DCC-GJR	30	4	15	18	24	0	6	0
HDCC-GJR	30	6	16	20	24	0	6	0
RiskMetrics, mv.	30	5	13	16	7	0	23	4
(d) Distribution								
Normal	85	49	75	79	52	0	33	4
Empirical	85	3	14	32	53	2	32	6
Student's t	85	17	63	74	57	3	28	4

This table summarizes the results of the unconditional coverage (UC) and dynamic quantile (DQ) tests of Christoffersen (1998) and Engle and Manganelli (2004) for the 255 different methods that we analyze. In each panel, we categorize the methods based on the choice aspect given in the panel heading. For each option that can be chosen, we report the total number of methods. Next, we report for the number of rejections from the UC tests for significance levels of 1, 5 and 10%. Under the null (alternative) hypothesis, the actual coverage ratio equals (exceeds) the theoretical coverage ratio of 0.01. To conduct the DQ test, we regress the centered hit series on the negative VaR forecasts. We report the number of negative and positive coefficients, and the number of significant ones at the 10% level. Significance is determined by bootstraps as in Herwartz and Waichman (2010). The full results are available in Tables C.19 and C.20.

Table C.19: Results for the UC tests, 8-year estimation window

(a) daily returns, iterated forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-	1.74**	1.77**	2.24***	1.62*	1.59*	1.91**	1.72*	1.72**	2.14***
GARCH	(0.44)	(0.43)	(0.49)	(0.43)	(0.41)	(0.46)	(0.44)	(0.43)	(0.47)
DCC-		1.79**	1.99**		1.62*	1.79**		1.69**	1.87**
GARCH		(0.44)	(0.46)		(0.42)	(0.44)		(0.42)	(0.45)
HDCC-		1.67*	2.19***		1.57*	1.84**		1.64*	2.06**
GARCH		(0.42)	(0.48)		(0.41)	(0.44)		(0.42)	(0.47)
CCC-	1.22	1.27	1.77**	1.29	1.22	1.67*	1.14	1.19	1.67*
GJR	(0.37)	(0.35)	(0.43)	(0.38)	(0.34)	(0.41)	(0.37)	(0.34)	(0.42)
DCC-		1.07	1.52*		1.04	1.32		1.14	1.47
GJR		(0.34)	(0.40)		(0.33)	(0.36)		(0.35)	(0.39)
HDCC-		1.22	1.69**		1.12	1.57*		1.14	1.57*
GJR		(0.36)	(0.42)		(0.34)	(0.41)		(0.34)	(0.41)
RiskMetrics	1.39	1.37	2.01**	1.44	1.44	1.44	0.82	1.34	1.79*
	(0.39)	(0.40)	(0.50)	(0.40)	(0.41)	(0.43)	(0.31)	(0.40)	(0.49)

(b) daily returns, scaled forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-	2.01**	1.99**	2.34***	1.29	1.64*	1.89**	1.72*	1.69*	2.04**
GARCH	(0.47)	(0.46)	(0.50)	(0.37)	(0.41)	(0.45)	(0.44)	(0.43)	(0.47)
DCC-		1.99**	2.11**		1.64*	1.64*		1.77**	1.89**
GARCH		(0.46)	(0.48)		(0.43)	(0.42)		(0.43)	(0.46)
HDCC-		2.04**	2.19***		1.59*	1.82**		1.72**	2.01**
GARCH		(0.46)	(0.49)		(0.42)	(0.43)		(0.43)	(0.47)
CCC-	1.96**	1.87**	2.26***	1.54*	1.82**	1.87**	1.62*	1.79**	1.96**
GJR	(0.45)	(0.43)	(0.49)	(0.41)	(0.43)	(0.44)	(0.42)	(0.42)	(0.45)
DCC-		1.96**	2.06**		1.49	1.67*		1.59*	1.82**
GJR		(0.44)	(0.47)		(0.40)	(0.41)		(0.41)	(0.43)
HDCC-		1.99**	2.31***		1.52*	1.79**		1.52	1.99**
GJR		(0.45)	(0.50)		(0.40)	(0.43)		(0.40)	(0.46)
RiskMetrics	1.64*	1.57*	2.11**	1.29	1.24	1.34	0.90	1.32	1.49
	(0.43)	(0.42)	(0.52)	(0.39)	(0.38)	(0.43)	(0.32)	(0.40)	(0.45)

This table shows the empirical coverage and results of the unconditional coverage tests of Christoffersen (1998). Models are estimated with a moving window of 2,000 daily returns, 400 weekly returns or 200 biweekly returns. For every day in the sample, we construct a violation indicator that equals one when the ten-day realized portfolio loss exceeds the forecasted VaR_ϑ with $\vartheta = 0.01$, and zero otherwise, as in Equation (20). We report the coverage (in %), and test whether it is equal to ϑ against the alternative of strictly more violations. We report Newey and West (1987) standard errors calculated with 10 leads and lags in parentheses. Superscripts *, **, *** denote rejection of the null-hypothesis with a significance level of 10%, 5% and 1%. The results are based on 4,021 forecasts.

Table C.19: Results for the UC tests, 8-year estimation window – *continued*

(c) weekly returns, iterated forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC- GARCH	2.39*** (0.53)	2.29*** (0.51)	2.49*** (0.53)	1.96** (0.48)	1.42 (0.41)	1.57 (0.44)	2.16** (0.51)	2.06** (0.49)	2.26*** (0.51)
DCC- GARCH		2.19*** (0.50)	2.39*** (0.52)		1.44 (0.42)	1.64* (0.44)		2.01** (0.50)	2.11** (0.50)
HDCC- GARCH		2.19*** (0.49)	2.46*** (0.53)		1.49 (0.44)	1.59* (0.45)		1.96** (0.49)	2.11** (0.49)
CCC- GJR	1.99** (0.47)	1.82** (0.43)	2.06** (0.47)	1.69* (0.44)	1.32 (0.38)	1.44 (0.40)	1.94** (0.46)	1.72** (0.42)	1.96** (0.46)
DCC- GJR		1.79** (0.43)	1.99** (0.46)		1.34 (0.39)	1.24 (0.39)		1.69** (0.42)	1.82** (0.44)
HDCC- GJR		1.77** (0.43)	2.01** (0.46)		1.09 (0.36)	1.32 (0.40)		1.67* (0.42)	1.91** (0.46)
RiskMetrics	2.24*** (0.53)	2.09** (0.52)	2.46*** (0.57)	1.74* (0.48)	1.49 (0.45)	1.27 (0.41)	2.16** (0.54)	1.94** (0.50)	1.96** (0.51)

(d) weekly returns, scaled forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC- GARCH	2.66*** (0.56)	2.51*** (0.54)	2.69*** (0.56)	1.91** (0.47)	1.24 (0.39)	1.44 (0.43)	2.21*** (0.51)	2.11** (0.50)	2.09** (0.49)
DCC- GARCH		2.66*** (0.54)	2.51*** (0.53)		1.37 (0.40)	1.42 (0.42)		2.01** (0.49)	2.04** (0.50)
HDCC- GARCH		2.56*** (0.53)	2.66*** (0.55)		1.27 (0.40)	1.52 (0.46)		2.01** (0.50)	2.19*** (0.50)
CCC- GJR	2.61*** (0.53)	2.39*** (0.50)	2.54*** (0.52)	1.82** (0.45)	1.42 (0.39)	1.42 (0.40)	2.26*** (0.50)	1.89** (0.45)	1.96** (0.47)
DCC- GJR		2.36*** (0.49)	2.41*** (0.50)		1.42 (0.39)	1.32 (0.38)		2.01** (0.45)	2.04** (0.46)
HDCC- GJR		2.39*** (0.49)	2.49*** (0.52)		1.34 (0.39)	1.37 (0.40)		1.91** (0.44)	2.06** (0.47)
RiskMetrics	2.49*** (0.55)	2.26*** (0.53)	2.51*** (0.57)	1.62* (0.46)	1.37 (0.43)	1.09 (0.38)	2.11** (0.53)	1.87** (0.49)	1.79** (0.48)

(e) biweekly returns, direct forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC- GARCH	3.06*** (0.61)	2.79*** (0.58)	2.93*** (0.59)	2.34*** (0.55)	0.92 (0.36)	1.29 (0.42)	2.46*** (0.55)	2.26*** (0.52)	2.34*** (0.52)
DCC- GARCH		2.81*** (0.58)	2.93*** (0.59)		0.97 (0.37)	1.12 (0.38)		2.26*** (0.52)	2.24*** (0.51)
HDCC- GARCH		2.91*** (0.58)	3.03*** (0.60)		0.95 (0.36)	1.22 (0.41)		2.26*** (0.52)	2.26*** (0.52)
CCC- GJR	3.08*** (0.59)	2.54*** (0.52)	2.64*** (0.54)	2.44*** (0.53)	0.95 (0.36)	1.19 (0.39)	2.69*** (0.55)	2.24*** (0.50)	2.14*** (0.49)
DCC- GJR		2.61*** (0.53)	2.61*** (0.53)		0.82 (0.32)	1.17 (0.39)		2.11** (0.49)	2.01** (0.46)
HDCC- GJR		2.61*** (0.52)	2.61*** (0.53)		0.97 (0.33)	1.14 (0.39)		2.19*** (0.48)	2.04** (0.47)
RiskMetrics	2.81*** (0.61)	2.29*** (0.55)	2.56*** (0.58)	2.36*** (0.58)	1.57 (0.46)	1.44 (0.43)	2.36*** (0.55)	1.89** (0.51)	1.79* (0.50)

See table note on previous page.

Table C.20: Results of the DQ test, 8-year estimation window

(a) daily returns, iterated forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-	-0.30	-0.18	-0.33	-0.32	-0.11	-0.34	-0.31	-0.16	-0.29
GARCH	(0.11)	(0.19)	(0.14)	(0.10)	(0.25)	(0.12)	(0.11)	(0.20)	(0.16)
DCC-		-0.15	-0.09		-0.10	-0.07		-0.13	-0.03
GARCH		(0.21)	(0.31)		(0.26)	(0.32)		(0.23)	(0.38)
HDCC-		-0.17	-0.29		-0.14	-0.25		-0.15	-0.16
GARCH		(0.18)	(0.15)		(0.21)	(0.16)		(0.20)	(0.25)
CCC-	-0.21	-0.06	-0.31	-0.29*	-0.06	-0.25	-0.24*	-0.07	-0.30*
GJR	(0.12)	(0.28)	(0.10)	(0.08)	(0.28)	(0.13)	(0.09)	(0.24)	(0.10)
DCC-		-0.09	-0.11		-0.11	-0.12		-0.08	-0.10
GJR		(0.20)	(0.24)		(0.17)	(0.21)		(0.23)	(0.25)
HDCC-		-0.10	-0.19		-0.09	-0.17		-0.08	-0.16
GJR		(0.20)	(0.18)		(0.21)	(0.18)		(0.22)	(0.19)
RiskMetrics	-0.07	-0.01	0.03	-0.08	0.02	-0.02	-0.20*	-0.02	0.03
	(0.28)	(0.39)	(0.46)	(0.26)	(0.45)	(0.37)	(0.09)	(0.38)	(0.47)

(b) daily returns, scaled forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-	-0.29	-0.14	-0.29	-0.21	-0.08	-0.18	-0.31	-0.16	-0.19
GARCH	(0.13)	(0.24)	(0.17)	(0.13)	(0.29)	(0.21)	(0.10)	(0.18)	(0.22)
DCC-		-0.14	-0.10		-0.13	-0.05		-0.14	-0.07
GARCH		(0.22)	(0.31)		(0.22)	(0.35)		(0.22)	(0.32)
HDCC-		-0.12	-0.23		-0.17	-0.11		-0.16	-0.15
GARCH		(0.26)	(0.20)		(0.18)	(0.27)		(0.18)	(0.24)
CCC-	-0.21	-0.13	-0.38	-0.28*	-0.09	-0.25	-0.25	-0.10	-0.31
GJR	(0.16)	(0.23)	(0.11)	(0.10)	(0.29)	(0.15)	(0.11)	(0.26)	(0.12)
DCC-		-0.12	-0.16		-0.15	-0.11		-0.15	-0.08
GJR		(0.23)	(0.23)		(0.18)	(0.24)		(0.17)	(0.29)
HDCC-		-0.18	-0.32		-0.17	-0.20		-0.22	-0.19
GJR		(0.18)	(0.13)		(0.16)	(0.16)		(0.12)	(0.19)
RiskMetrics	-0.07	0.02	0.05	-0.08	-0.02	-0.01	-0.18	-0.02	-0.01
	(0.29)	(0.46)	(0.49)	(0.25)	(0.37)	(0.37)	(0.10)	(0.37)	(0.40)

This table reports the results of the dynamic quantile test of Engle and Manganelli (2004). We conduct a linear regression of the centered VaR-violations on a constant and the forecasted VaR_ϑ (in %) for $\vartheta = 0.01$, as in Equation (21). We report the estimated coefficient on the forecasted VaR_ϑ (expressed as a decimal rate), and quantile of its t -statistic in the distribution of the statistic under the null hypothesis of no predictability. Following Herwartz and Waichman (2010), we construct this distribution by 10,000 block bootstrapped samples with a block length of 10. Superscripts *, **, *** denote quantiles in the 10%, 5% and 1% of the left or right tail.

Table C.20: Results of the DQ test, 8-year estimation window – *continued*

(c) weekly returns, iterated forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC- GARCH	0.06 (0.52)	-0.14 (0.30)	-0.04 (0.40)	0.05 (0.52)	-0.08 (0.30)	-0.21 (0.18)	0.09 (0.57)	-0.05 (0.38)	-0.04 (0.40)
DCC- GARCH		-0.13 (0.30)	-0.05 (0.38)		-0.03 (0.37)	-0.02 (0.41)		-0.02 (0.41)	-0.03 (0.39)
HDCC- GARCH		0.00 (0.44)	-0.01 (0.43)		-0.06 (0.33)	-0.04 (0.35)		0.03 (0.48)	0.04 (0.47)
CCC- GJR	0.08 (0.57)	-0.06 (0.33)	-0.25 (0.18)	0.03 (0.49)	-0.07 (0.26)	-0.15 (0.20)	0.08 (0.58)	-0.14 (0.22)	-0.33 (0.12)
DCC- GJR		-0.07 (0.30)	-0.07 (0.33)		-0.07 (0.28)	-0.10 (0.23)		-0.09 (0.28)	-0.04 (0.38)
HDCC- GJR		-0.07 (0.31)	-0.10 (0.30)		-0.04 (0.32)	-0.11 (0.22)		-0.09 (0.26)	-0.21 (0.19)
RiskMetrics	0.28* (0.91)	0.21 (0.76)	0.31 (0.82)	0.17 (0.80)	0.13 (0.73)	0.14 (0.86)	0.38** (0.99)	0.21 (0.76)	0.19 (0.72)

(d) weekly returns, scaled forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC- GARCH	0.00 (0.45)	-0.09 (0.36)	-0.06 (0.40)	0.14 (0.71)	0.02 (0.46)	-0.17 (0.19)	0.06 (0.52)	0.00 (0.44)	-0.07 (0.36)
DCC- GARCH		-0.07 (0.38)	-0.06 (0.39)		0.03 (0.50)	-0.02 (0.40)		-0.01 (0.42)	-0.01 (0.44)
HDCC- GARCH		-0.01 (0.43)	-0.02 (0.42)		0.06 (0.56)	0.01 (0.44)		0.08 (0.55)	-0.04 (0.40)
CCC- GJR	0.10 (0.57)	0.02 (0.47)	-0.42 (0.13)	0.11 (0.65)	-0.06 (0.30)	-0.15 (0.19)	0.12 (0.63)	-0.09 (0.30)	-0.29 (0.16)
DCC- GJR		0.01 (0.44)	-0.17 (0.25)		-0.04 (0.33)	-0.07 (0.27)		-0.08 (0.31)	-0.08 (0.33)
HDCC- GJR		0.04 (0.50)	-0.30 (0.18)		-0.05 (0.31)	-0.10 (0.25)		-0.09 (0.29)	-0.20 (0.21)
RiskMetrics	0.20 (0.74)	0.28 (0.83)	0.31 (0.81)	0.17 (0.86)	0.14 (0.82)	0.12 (0.86)	0.36** (0.98)	0.25 (0.83)	0.14 (0.67)

(e) biweekly returns, direct forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC- GARCH	0.38 (0.83)	0.36 (0.79)	0.15 (0.59)	0.41*** (1.00)	0.08 (0.83)	0.08 (0.64)	0.30 (0.84)	0.15 (0.63)	0.13 (0.58)
DCC- GARCH		0.30 (0.75)	0.30 (0.73)		0.10* (0.91)	0.14 (0.83)		0.25 (0.76)	0.22 (0.73)
HDCC- GARCH		0.27 (0.73)	0.25 (0.67)		0.09 (0.87)	0.12 (0.76)		0.24 (0.76)	0.09 (0.55)
CCC- GJR	0.26 (0.80)	0.22 (0.75)	0.11 (0.57)	0.32*** (1.00)	0.06 (0.72)	0.07 (0.60)	0.30 (0.89)	0.08 (0.56)	-0.16 (0.26)
DCC- GJR		0.23 (0.77)	0.19 (0.68)		0.06 (0.77)	0.09 (0.68)		0.06 (0.54)	-0.01 (0.44)
HDCC- GJR		0.25 (0.80)	0.09 (0.55)		0.08 (0.86)	0.08 (0.66)		0.09 (0.59)	-0.12 (0.29)
RiskMetrics	0.44** (0.95)	0.43** (0.96)	0.43* (0.91)	0.40*** (1.00)	0.22** (0.99)	0.17** (0.99)	0.35** (0.98)	0.35* (0.94)	0.23 (0.77)

See table note on previous page.

Table C.21: Average asymmetric tick loss, 8-year estimation window

(a) daily returns, iterated forecasts									
Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-GARCH	6.39	5.98	6.31	6.38	5.94	6.18	6.40	5.94	6.26
DCC-GARCH		6.02	6.19		5.94	6.06		6.01	6.16
HDCC-GARCH		6.06	6.29		5.96	6.12		6.04	6.24
CCC-GJR	6.08	5.75	5.94	6.08	5.77	5.91	6.03	5.69	5.81
DCC-GJR		5.73	5.80		5.74	5.72		5.69	5.77
HDCC-GJR		5.75	5.84		5.74	5.78		5.73	5.79
RiskMetrics	6.11	6.20	6.75	6.15	6.22	6.55	6.12	6.25	6.82

(b) daily returns, scaled forecasts									
Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-GARCH	6.42	6.07	6.39	6.11	5.90	6.04	6.29	5.89	6.15
DCC-GARCH		6.11	6.27		5.95	5.94		6.00	6.09
HDCC-GARCH		6.16	6.39		5.97	6.04		5.99	6.17
CCC-GJR	6.24	6.00	6.26	6.08	5.91	5.96	6.15	5.85	6.02
DCC-GJR		6.01	6.09		5.91	5.81		5.87	5.93
HDCC-GJR		6.06	6.22		5.95	5.93		5.90	5.99
RiskMetrics	6.21	6.25	6.79	6.10	6.22	6.59	6.13	6.22	6.71

(c) weekly returns, iterated forecasts									
Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-GARCH	7.88	7.77	7.80	7.56	7.21	7.23	7.67	7.62	7.65
DCC-GARCH		7.71	7.69		7.20	7.18		7.61	7.57
HDCC-GARCH		7.69	7.80		7.18	7.23		7.61	7.65
CCC-GJR	7.26	7.05	7.23	7.08	6.74	6.88	7.25	7.00	7.14
DCC-GJR		7.05	7.09		6.72	6.78		6.95	7.08
HDCC-GJR		6.99	7.20		6.73	6.85		6.93	7.12
RiskMetrics	7.56	7.91	8.09	7.30	7.68	7.86	7.94	7.84	7.79

This table shows the average value of the asymmetric tick loss function of Giacomini and Komunjer (2005) as in Equation (22) for $\vartheta = 0.01$. All values have been multiplied by 100. Models are estimated with a moving window of 2,000 daily returns, 400 weekly returns or 200 biweekly returns. The results are based on 4,021 forecasts.

Table C.21: Average asymmetric tick loss, 8-year estimation window – *continued*

(d) weekly returns, scaled forecasts									
Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-GARCH	7.98	7.89	7.90	7.29	7.09	7.17	7.59	7.65	7.58
DCC-GARCH		7.82	7.77		7.13	7.09		7.62	7.50
HDCC-GARCH		7.79	7.89		7.12	7.18		7.69	7.57
CCC-GJR	7.52	7.36	7.50	7.09	6.81	6.79	7.32	7.16	7.19
DCC-GJR		7.33	7.35		6.78	6.78		7.07	7.14
HDCC-GJR		7.25	7.46		6.74	6.85		7.06	7.23
RiskMetrics	7.62	7.96	8.13	7.23	7.62	7.65	7.88	7.82	7.76

(e) biweekly returns, direct forecasts									
Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-GARCH	8.82	8.59	8.31	8.18	8.29	7.79	8.32	8.23	7.93
DCC-GARCH		8.47	8.16		8.31	7.70		8.09	7.82
HDCC-GARCH		8.39	8.24		8.24	7.76		8.02	7.82
CCC-GJR	8.46	7.97	7.77	7.92	7.70	7.18	8.00	7.78	7.54
DCC-GJR		7.83	7.64		7.73	7.24		7.67	7.39
HDCC-GJR		7.86	7.72		7.67	7.22		7.72	7.45
RiskMetrics	8.98	8.41	8.23	8.69	8.24	8.09	8.22	8.44	7.98

See table note on previous page.

Table C.22: Methods removed from the Model Confidence Set, 8-year estimation window

(a) biweekly returns, direct forecasts									
Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-GARCH						0.094			
DCC-GARCH						0.092			
HDCC-GARCH						0.090			
CCC-GJR									
DCC-GJR						0.097			
HDCC-GJR						0.098			
RiskMetrics									0.099

This table presents the MCS p -value with which a model has been removed from the Model Confidence Set. We follow the procedure of Hansen et al. (2011, Sec. 3.1.2), with the tick loss function in (22) and a significance level of 10%. The procedure starts with the complete set of 255 methods. Panels for daily and weekly observations combined with iterated or scaled forecasts are absent, because no methods with these combinations have been removed. Models are estimated with a moving window of 2,000 daily returns, 400 weekly returns or 200 biweekly returns. The results are based on 4,021 forecasts.

Table C.23: Results of the Diebold-Mariano tests for temporal aggregation, 4 and 8-year estimation window

Method A	Method B	# of tests	neg.	sign. neg.			pos.
				1%	5%	10%	
(a) 4-year estimation window							
Daily, iterated	Weekly, iterated	51	51	0	14	51	0
Daily, scaled	Weekly, scaled	51	51	0	18	51	0
Daily, iterated	Biweekly, direct	51	51	8	34	51	0
Daily, scaled	Biweekly, direct	51	51	10	26	51	0
Weekly, iterated	Biweekly, direct	51	51	17	46	48	0
Weekly, scaled	Biweekly, direct	51	51	16	39	46	0
(b) 8-year estimation window							
Daily, iterated	Weekly, iterated	51	51	3	8	50	0
Daily, scaled	Weekly, scaled	51	51	4	11	50	0
Daily, iterated	Biweekly, direct	51	51	2	19	51	0
Daily, scaled	Biweekly, direct	51	51	3	19	51	0
Weekly, iterated	Biweekly, direct	51	51	9	13	40	0
Weekly, scaled	Biweekly, direct	51	51	10	14	34	0

This table shows summary results of the tests that the expected value of the loss function of methods A and B are equal, $L^A = L^B$. We calculate the loss differential as $L^A - L^B$. A negative loss differential means that method A is preferable to method B. We use the asymmetric tick loss function of Giacomini and Komunjer (2005) as in Equation (22) with $\vartheta = 0.05$. We report the number of negative and positive average loss differentials, and the number of times these differentials are significant for the 1% 5% and 10% significance levels, based on the statistic proposed by Diebold and Mariano (1995), evaluated in the setting of Giacomini and White (2006). Standard errors of the average loss differential are based on Newey and West (1987) with 10 leads and lags. We compare pairs of methods that differ in only in their degree of temporal aggregation, so the other four choices are the same. The full results for the 4-year and 8-year estimation windows are available in Table C.7 and Table C.24.

Table C.24: Results of the DM-tests for temporal aggregation, 8-year estimation window

(a) daily vs. weekly returns, iterated forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-	-1.64***	-1.54*	-1.80*	-1.28***	-1.06	-1.28*	-1.40**	-1.43*	-1.69*
GARCH	(0.70)	(1.01)	(1.19)	(0.54)	(0.83)	(0.97)	(0.66)	(0.99)	(1.19)
DCC-		-1.54*	-1.72*		-1.13*	-1.26*		-1.44*	-1.62*
GARCH		(1.01)	(1.11)		(0.83)	(0.89)		(1.00)	(1.07)
HDCC-		-1.58*	-1.67*		-1.13*	-1.22*		-1.44*	-1.59*
GARCH		(1.04)	(1.07)		(0.86)	(0.85)		(1.02)	(1.07)
CCC-	-1.30**	-1.32*	-1.32*	-1.09***	-0.98*	-0.97*	-1.32**	-1.37*	-1.31*
GJR	(0.58)	(0.85)	(0.91)	(0.43)	(0.74)	(0.70)	(0.57)	(0.87)	(0.94)
DCC-		-1.32*	-1.36*		-1.06*	-0.99*		-1.34*	-1.29*
GJR		(0.88)	(0.88)		(0.74)	(0.69)		(0.90)	(0.87)
HDCC-		-1.39*	-1.28*		-1.08*	-0.99*		-1.36*	-1.22*
GJR		(0.92)	(0.86)		(0.77)	(0.68)		(0.94)	(0.84)
RiskMetrics	-1.57*	-1.43**	-1.73*	-1.26*	-1.34**	-1.47*	-1.96*	-0.99*	-1.61*
	(0.98)	(0.84)	(1.14)	(0.76)	(0.60)	(0.91)	(1.40)	(0.67)	(1.09)

(b) daily vs. weekly returns, scaled forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-	-1.71**	-1.56*	-1.85*	-1.26***	-1.13*	-1.19*	-1.43**	-1.46*	-1.77*
GARCH	(0.75)	(1.05)	(1.21)	(0.53)	(0.85)	(0.88)	(0.70)	(1.04)	(1.22)
DCC-		-1.55*	-1.75*		-1.16*	-1.18*		-1.44*	-1.64*
GARCH		(1.05)	(1.12)		(0.85)	(0.81)		(1.02)	(1.10)
HDCC-		-1.57*	-1.68*		-1.15*	-1.14*		-1.43*	-1.71*
GARCH		(1.07)	(1.09)		(0.88)	(0.76)		(1.06)	(1.12)
CCC-	-1.40***	-1.28*	-1.39*	-1.10***	-0.84	-0.90*	-1.29***	-1.19*	-1.32*
GJR	(0.57)	(0.84)	(0.89)	(0.39)	(0.68)	(0.67)	(0.53)	(0.84)	(0.90)
DCC-		-1.31*	-1.37**		-0.97*	-0.88*		-1.24*	-1.21*
GJR		(0.85)	(0.81)		(0.69)	(0.57)		(0.85)	(0.77)
HDCC-		-1.28*	-1.25*		-0.92*	-0.80*		-1.27*	-1.19*
GJR		(0.86)	(0.79)		(0.71)	(0.53)		(0.88)	(0.76)
RiskMetrics	-1.53*	-1.44**	-1.75*	-1.23**	-1.09**	-1.41**	-1.89*	-1.07*	-1.62*
	(0.97)	(0.84)	(1.12)	(0.71)	(0.53)	(0.76)	(1.31)	(0.71)	(1.10)

(c) daily vs. biweekly returns, iterated forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-	-2.60**	-2.09*	-2.67*	-1.99**	-1.62**	-2.35**	-2.08**	-1.73*	-2.33*
GARCH	(1.21)	(1.30)	(1.66)	(0.96)	(0.98)	(1.12)	(1.07)	(1.20)	(1.53)
DCC-		-2.06*	-2.53*		-1.66**	-2.37**		-1.71*	-2.13*
GARCH		(1.29)	(1.57)		(0.96)	(1.05)		(1.19)	(1.40)
HDCC-		-2.04*	-2.43**		-1.66*	-2.27***		-1.63*	-2.04*
GARCH		(1.31)	(1.46)		(1.01)	(0.91)		(1.20)	(1.32)
CCC-	-2.54**	-1.90*	-2.29*	-2.00**	-1.29*	-1.93**	-2.11**	-1.77*	-2.13*
GJR	(1.26)	(1.21)	(1.46)	(1.01)	(0.89)	(1.02)	(1.08)	(1.18)	(1.44)
DCC-		-1.91*	-2.18*		-1.53*	-1.99**		-1.67*	-2.03*
GJR		(1.25)	(1.39)		(0.93)	(0.94)		(1.18)	(1.39)
HDCC-		-1.94*	-2.21*		-1.45*	-1.93**		-1.69*	-2.05*
GJR		(1.27)	(1.37)		(0.96)	(0.86)		(1.19)	(1.34)
RiskMetrics	-3.06**	-1.66**	-2.27*	-2.77**	-1.68***	-2.05*	-2.28*	-1.23*	-2.23*
	(1.84)	(0.96)	(1.55)	(1.59)	(0.59)	(1.29)	(1.41)	(0.82)	(1.49)

This table shows summary results of the tests that the expected value of the loss function of methods A and B are equal, $L^A = L^B$. We report the loss differential $L^A - L^B$ (multiplied by 100). A negative loss differential means that method A is preferable to method B. We use the asymmetric tick loss function of Giacomini and Komunjer (2005) as in Equation (22) with $\vartheta = 0.01$. The two methods differ in their degree of temporal aggregation as stated in the headings of the different panels. We report standard errors of the average loss differential in parentheses, based on Newey and West (1987) with 10 leads and lags. We test that the average loss differential is zero based on the statistic proposed by Diebold and Mariano (1995), evaluated in the setting of Giacomini and White (2006). Superscripts *, **, *** denote rejection of the null-hypothesis with a significance level of 10%, 5% and 1%.

Table C.24: Results of the DM-tests for temporal aggregation, 8-year estimation window – *continued*

(d) daily vs. biweekly returns, scaled forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-	-2.57**	-2.01*	-2.59*	-2.26**	-1.77**	-2.38**	-2.20**	-1.84*	-2.37*
GARCH	(1.25)	(1.33)	(1.65)	(1.12)	(1.05)	(1.11)	(1.16)	(1.29)	(1.57)
DCC-		-1.98*	-2.44*		-1.78**	-2.36**		-1.79*	-2.14*
GARCH		(1.31)	(1.56)		(1.06)	(1.02)		(1.26)	(1.44)
HDCC-		-1.94*	-2.32*		-1.73*	-2.26***		-1.69*	-2.08*
GARCH		(1.32)	(1.46)		(1.07)	(0.88)		(1.27)	(1.36)
CCC-	-2.37**	-1.58*	-2.04*	-2.00**	-1.24*	-1.79**	-2.00**	-1.56*	-1.97*
GJR	(1.17)	(1.14)	(1.33)	(1.01)	(0.87)	(0.93)	(1.02)	(1.12)	(1.32)
DCC-		-1.62*	-1.90*		-1.44*	-1.82**		-1.51*	-1.85*
GJR		(1.14)	(1.22)		(0.89)	(0.80)		(1.12)	(1.24)
HDCC-		-1.56*	-1.90*		-1.30*	-1.72***		-1.49*	-1.87*
GJR		(1.14)	(1.19)		(0.88)	(0.68)		(1.11)	(1.20)
RiskMetrics	-2.95**	-1.61**	-2.21*	-2.82**	-1.65***	-2.05*	-2.27**	-1.34*	-2.26*
	(1.77)	(0.95)	(1.51)	(1.63)	(0.65)	(1.27)	(1.35)	(0.90)	(1.52)

(e) weekly vs. biweekly returns, iterated forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-	-0.97*	-0.55*	-0.87*	-0.71*	-0.57***	-1.07***	-0.68*	-0.30	-0.63*
GARCH	(0.59)	(0.34)	(0.54)	(0.54)	(0.20)	(0.28)	(0.50)	(0.25)	(0.39)
DCC-		-0.52*	-0.81*		-0.53***	-1.11***		-0.27	-0.50*
GARCH		(0.32)	(0.53)		(0.17)	(0.28)		(0.22)	(0.38)
HDCC-		-0.47*	-0.76*		-0.53***	-1.05***		-0.18	-0.45*
GARCH		(0.32)	(0.46)		(0.19)	(0.25)		(0.21)	(0.32)
CCC-	-1.24*	-0.58*	-0.97*	-0.91	-0.31*	-0.96***	-0.79	-0.41	-0.81*
GJR	(0.81)	(0.41)	(0.61)	(0.73)	(0.20)	(0.38)	(0.63)	(0.34)	(0.56)
DCC-		-0.58*	-0.83*		-0.46**	-1.00***		-0.33	-0.74*
GJR		(0.42)	(0.57)		(0.23)	(0.33)		(0.32)	(0.57)
HDCC-		-0.55*	-0.93*		-0.38**	-0.94***		-0.34	-0.83*
GJR		(0.40)	(0.57)		(0.23)	(0.29)		(0.27)	(0.56)
RiskMetrics	-1.49*	-0.22	-0.53	-1.51**	-0.34**	-0.58*	-0.32*	-0.24	-0.62*
	(0.93)	(0.18)	(0.44)	(0.90)	(0.19)	(0.40)	(0.20)	(0.19)	(0.44)

(f) weekly vs. biweekly returns, scaled forecasts

Asset split	Normal			Empirical			Student's t		
	1	3	8	1	3	8	1	3	8
CCC-	-0.86*	-0.44*	-0.74*	-1.00*	-0.64***	-1.19***	-0.76*	-0.38*	-0.61*
GARCH	(0.61)	(0.32)	(0.52)	(0.70)	(0.23)	(0.33)	(0.56)	(0.28)	(0.41)
DCC-		-0.43*	-0.69*		-0.62***	-1.19***		-0.34*	-0.50
GARCH		(0.31)	(0.52)		(0.23)	(0.32)		(0.27)	(0.40)
HDCC-		-0.38*	-0.64*		-0.58***	-1.12***		-0.26	-0.37
GARCH		(0.29)	(0.44)		(0.22)	(0.28)		(0.25)	(0.31)
CCC-	-0.97	-0.29	-0.65*	-0.91	-0.39**	-0.88***	-0.70	-0.37	-0.65*
GJR	(0.78)	(0.35)	(0.50)	(0.81)	(0.24)	(0.33)	(0.64)	(0.31)	(0.48)
DCC-		-0.31	-0.54		-0.47**	-0.94***		-0.27	-0.63
GJR		(0.35)	(0.48)		(0.23)	(0.31)		(0.30)	(0.52)
HDCC-		-0.28	-0.65*		-0.38**	-0.93***		-0.23	-0.68*
GJR		(0.34)	(0.47)		(0.21)	(0.26)		(0.26)	(0.50)
RiskMetrics	-1.43*	-0.18	-0.47	-1.59*	-0.55***	-0.65	-0.38**	-0.27*	-0.64*
	(0.89)	(0.16)	(0.42)	(0.99)	(0.22)	(0.54)	(0.20)	(0.20)	(0.45)

See table note on previous page.