Journal of Forecasting, *J. Forecast.* **36**, 181–206 (2017)
Published online 9 May 2016 in Wiley Online Library (wileyonlinelibrary.com) **DOI**: 10.1002/for.2423

On the Modelling and Forecasting of Multivariate Realized Volatility: Generalized Heterogeneous Autoregressive (GHAR) Model

FRANTIŠEK ČECH[1,2] AND JOZEF BARUNÍK[1,2]*

[1] *Institute of Economic Studies, Charles University Prague, Czech Republic*
[2] *Institute of Information Theory and Automation, The Czech Academy of Sciences, Prague, Czech Republic*

ABSTRACT
Recent multivariate extensions of the popular heterogeneous autoregressive model (HAR) for realized volatility leave substantial information unmodelled in residuals. We propose to employ a system of seemingly unrelated regressions to model and forecast a realized covariance matrix to capture this information. We find that the newly proposed generalized heterogeneous autoregressive (GHAR) model outperforms competing approaches in terms of economic gains, providing better mean–variance trade-off, while, in terms of statistical precision, GHAR is not substantially dominated by any other model. Our results provide a comprehensive comparison of the performance when realized covariance, subsampled realized covariance and multivariate realized kernel estimators are used. We study the contribution of the estimators across different sampling frequencies, and show that the multivariate realized kernel and subsampled realized covariance estimators deliver further gains compared to realized covariance estimated on a 5-minute frequency. In order to show economic and statistical gains, a portfolio of various sizes is used. Copyright © 2016 John Wiley & Sons, Ltd.

KEY WORDS Multivariate volatility; realized covariance; portfolio optimisation; economic evaluation

INTRODUCTION

The risk of individual financial instruments is crucial for asset pricing, portfolio and risk management. Besides volatility of individual assets, knowledge of covariance structure between assets in portfolio is of great importance. Accurate forecasts of variance–covariance matrices are particularly important in asset allocation and portfolio management.

The traditional approach of obtaining covariance matrix estimates relies on multivariate generalized autoregressive conditional heteroscedasticity (MGARCH) models such as the constant conditional correlation GARCH of Bollerslev (1990), the dynamic conditional correlation GARCH of Engle (2002) or the BEKK of Engle and Kroner (1995) (for a survey of MGARCH models see Bauwens *et al.*, 2006). These models are popular in the literature although they suffer from the curse of dimensionality problem. Increased availability of high-frequency data in the last decade resulted in development of the new non-parametric approach for treating multivariate volatility. A milestone for covariance matrix modelling is the work of Barndorff-Nielsen and Shephard (2004), where the theory of 'realized covariation' is introduced. Realized covariance matrices are ex post measures of daily covariation and they need to be further modelled. The research dedicated to modelling the entire covariance matrices is still lively. From the already established methods, let us mention Wishart autoregression (WAR) of Gouriéroux *et al.* (2009), with numerous extensions presented in Bonato (2009) and Bonato *et al.* (2013). A different approach of realized volatility modelling can be found in Bauer and Vorkink (2011), who model realized stock market volatility using matrix-logarithm transformation and primarily concentrate on forecasting performance of the factor model. A more common approach for obtaining positive definite forecasts of covariance matrices is the use of Cholesky decomposition. The use of Cholesky factors, further estimated by vector autoregressive fractionally integrated moving average (VARFIMA), heterogeneous autoregression (HAR) or WAR-HAR can be found in the work of Chiriac and Voev (2011). More recently, Amendola and Storti (2015) consider combining predictions from multivariate GARCH models and realized covariance matrices.

In this paper, we contribute to the literature by proposing a new model for dynamic covariance matrix modelling and forecasting. We model Cholesky factors of the realized covariance matrix as a system of seemingly unrelated heterogeneous autoregressions. The main motivation is that we may expect the residuals from simple HAR model to be contemporaneously correlated and, moreover, heteroscedastic due to well-known volatility in the volatility effect (Corsi *et al.*, 2008). Estimating the system of HAR equations using generalized least squares allows us to capture these dependencies. Hence the generalized HAR (GHAR) may provide more precise and more efficient forecasts, which

*Correspondence to: Jozef Barunik, Institute of Information Theory and Automation The Czech Academy of Sciences. Prague, Czech Republic. E-mail: barunik@utia.cas.cz

will translate to economic gains directly. On the portfolios of various sizes, we show that the GHAR model delivers significant economic gains and, statistically, is not substantially outperformed when compared to natural benchmark models based on high-frequency data (HAR, VARFIMA), as well as daily data (DCC-GARCH, RiskMetrics). In addition, we study the economic benefits of estimating the realized covariance with more efficient subsampled realized covariance and multivariate realized kernel estimators.

The rest of the paper is structured as follows. We provide the background for estimation of realized covariation from high-frequency data in the next section. The third section describes frameworks for modelling multivariate volatility and presents our GHAR model. The fourth section provides a description of dataset and research design, including economic as well as statistical evaluation criteria. In the fifth section we discuss out-of-sample forecast evaluation, and the sixth section concludes.

ESTIMATION OF COVARIATION FROM HIGH FREQUENCY DATA

We assume that the q-dimensional efficient price process p_t evolves over time $0 \leq t \leq T$ according to the following dynamics

$$\mathrm{d}p_t = \mu_t\,\mathrm{d}t + \Sigma_t\,\mathrm{d}W_t + \mathrm{d}J_t \tag{1}$$

where μ_t is predictable component, Σ_t is a real-value $q \times q$ volatility process, $W_1, \ldots, W_q$ is a q-dimensional Brownian motion and $\mathrm{d}J_t$ is a jump process. A central object of interest is the integrated covariation, which measures the covariance of asset returns over a particular period. Andersen *et al.* (2003) and Barndorff-Nielsen and Shephard (2004) suggest estimating the quadratic covariation matrix analogously to the realized variance, by taking the outer product of the observed high-frequency return over the period. This estimation, however, assumes synchronized equidistant data.

In practice, trading is non-synchronous, delivering fresh prices at irregularly spaced times that differ across stocks. In order to estimate the covariance, the data need to be synchronized, meaning that the prices of the q assets need to be collected at the same time stamp. Research of non-synchronous trading has been an active field of financial econometrics in past years: see, for example, Hayashi and Yoshida (2005) and Voev and Lunde (2007). This practical issue induces bias in the estimators and may be partly responsible for the Epps effect (Epps, 1979), a phenomenon of decreasing empirical correlation between the returns of two different stocks with increasing data sampling frequency. Ait-Sahalia *et al.* (2010) compare various synchronization schemes and find that the estimates do not differ significantly from estimates using the so-called refresh time scheme when dealing with highly liquid assets. The data used further in our study consist of the most liquid US stocks; hence we can restrict ourselves to the refresh time synchronization scheme in our work.

Let $N_{(q)t}$ be the counting process governing the number of observations in the qth asset up to time t, with times of trades $t_{(q)1}, t_{(q)2}, \ldots$. Following Barndorff-Nielsen *et al.* (2011), we define the first refresh time as

$$\tau_1 = \max(t_{(1)1}, \ldots, t_{(d)1}) \tag{2}$$

for $d = 1, \ldots, q$ assets, and all subsequent refresh times as

$$\tau_{j+1} = \max\left(t_{(1)N_{(1)\tau_j}+1}, \ldots, t_{(d)N_{(d)\tau_j}+1}\right) \tag{3}$$

with the resulting refresh time sample being of length N. τ_1 is thus the first time that all assets record prices, while τ_2 is the first time that all asset prices are refreshed. In the following analysis, we will always set our clock time to τ_j when using the estimators.

Having synchronized the data, let us denote by $\Delta_k p_t = p_{t-1+\tau_k/N} - p_{t-1+\tau_{k-1}/N}$ a discretely sampled vector of kth intraday log-returns in $[t-1, t]$, with N intraday observations available for each asset q. A simple estimator of realized covariance is then constructed as

$$\widehat{\Sigma}_t^{(\mathrm{RC})} = \sum_{k=1}^{N} (\Delta_k p_t)\,(\Delta_k p_t)' \tag{4}$$

As shown by Barndorff-Nielsen and Shephard (2004), realized covariance is a consistent estimator of integrated covariance and is asymptotically mixed normal. However, the estimator is biased and becomes inconsistent in the case that micro-structure noise is present in the data. Sparse sampling is used to mitigate the trade-off between the bias due to noise and variance of the estimator.

To effectively use all available high-frequency data, Zhang *et al.* (2005) propose using subsampling and averaging for realized variance calculation. In their setup the whole sample is divided into M non-overlapping subsamples, and in each subsample realized variance is calculated and averaged across the subsampled estimates to form the final estimate:

$$\widehat{\Sigma}_t^{(\mathrm{RC\ SS})} = \frac{1}{M} \sum_{i=1}^{M} \widehat{\Sigma}_{t,i}^{(\mathrm{RC})} \tag{5}$$

In addition, the covariance matrix estimated by realized covariance might not necessarily be positive semi-definite. To overcome these problems, Barndorff-Nielsen *et al.* (2011) introduced the multivariate realized kernels (MRK) estimator, which guaranties the covariance matrix to be positive semi-definite. Moreover, MRK is more efficient, and it is able to deal with noise. Following Barndorff-Nielsen *et al.* (2011), the MRK estimator is defined as

$$\widehat{\Sigma}_t^{(\mathrm{MRK})} = \sum_{h=-n}^{n} k\left(\frac{h}{H}\right) \Gamma_h \tag{6}$$

where Γ_h stands for the hth realized autocovariance and $k(x)$ is a non-stochastic weight function. In the empirical implementation, we need to choose the kernel function and bandwidth parameter. Following Barndorff-Nielsen *et al.* (2011), we use a Parzen kernel,[1] which satisfies the smoothness conditions, $K'(0) = K'(1) = 0$, and guarantees $\widehat{\Sigma}_t^{(\mathrm{MRK})}$ to be positive semi-definite. We use the optimal bandwidth derived in Barndorff-Nielsen *et al.* (2011).

Recently, many new approaches to covariance matrix estimation using high-frequency data have emerged in the literature. In addition to estimators used in this study, realized co-range (Bannouh *et al.*, 2009) or two-scale realized covariance (Zhang, 2011) are also becoming increasingly popular. Today, the literature also pays attention to disentangling jumps, common jumps and true covariation (see Boudt *et al.*, 2012, or Elst and Veredas, 2015). When the dimension of the problem is high, the estimator of Hautsch *et al.* (2012), which estimates covariance using block-wise multivariate realized kernels, might be of interest.

While the number of recently proposed estimators is growing, we restrict our study to a comparison of the main estimators used in the literature,[2] and focus on the actual estimator of the proposed model.

MODELLING AND FORECASTING MULTIVARIATE VOLATILITY

Modelling and forecasting a conditional covariance matrix of asset returns Σ_t is pivotal to asset allocation, risk management and option pricing. In order to have a valid multivariate forecasting model, one needs to specify a model that produces symmetric and positive semi-definite covariance matrix predictions. Whereas it is still relatively scarce to use high-frequency data in multivariate modelling, the literature dealing with challenging issues is growing quickly. There are three types of approach proposed recently: modelling the Cholesky factorization of the covariance matrix (Chiriac and Voev, 2011), its matrix-log transformation with the use of latent factors (Bauer and Vorkink, 2011), and direct modelling of the covariance dynamics as a Wishart autoregressive model (Bonato, 2009; Jin and Maheu, 2013).

To ensure positive semi-definiteness of covariance matrix forecasts, we adopt the approach from Chiriac and Voev (2011): we apply the Cholesky decomposition to the covariance matrix. This approach is attractive as it also helps to reduce the curse of dimensionality, especially in the model structures we are going to use in this study. Following Chiriac and Voev (2011), we model the lower triangular elements of the Cholesky factorization:

$$X_t = \mathrm{vech}\,(P_t) \tag{7}$$

where P_t are Cholesky factors $P_t' P_t = \Sigma_t$ and X_t is $m \times 1$ vector, with $m = \frac{q(q+1)}{2}$. Forecasts of the covariance matrix are then obtained by reverse transformation.

Generalized heterogeneous autoregressive (GHAR) model

A simple approximate long-memory model for realized volatility—heterogeneous autoregression (HAR)—has been introduced by Corsi (2009). Whereas the approach has been introduced for the univariate volatility modelling, its extension to multivariate volatility has been recently used in the literature (see, for example, Chiriac and Voev, 2011;

[1] The Parzen kernel function is given by $k(x) = \begin{cases} 1 - 6x^2 + 6x^3 & 0 \le x \le 1/2 \\ 2(1-x)^3 & 1/2 \le x \le 1 \\ 0 & x > 1 \end{cases}$.

[2] Realized covariance sampled at a 5-minute frequency is the industry standard; subsampled realized covariance enables us to use all data points, resulting in a more efficient estimator, and MRK is able to handle noise and non-synchronous trading

Bauer and Vorkink, 2011). The original HAR model has an autoregressive structure, and combines volatilities measured at different frequencies (daily, weekly, monthly). Chiriac and Voev (2011) propose a multivariate extension of HAR to model a vector of Cholesky factors X_t, as

$$X_{t+1}^{(1)} = c + \beta^{(1)} X_t^{(1)} + \beta^{(5)} X_t^{(5)} + \beta^{(22)} X_t^{(22)} + \epsilon_t, \;\; \epsilon_t \sim \text{i.i.d.} \tag{8}$$

where 1,5 and 22 stand for day, week (5 days) and month (22 days) respectively, c is an $m \times 1$ vector of constants, $\beta^{(.)}$ are scalar parameters and $X_t^{(.)}$ are averages of lagged daily volatility, e.g. $X_t^{(5)} = \frac{1}{5}\sum_{i=0}^{4} X_{t-i}$. To obtain parameter estimates, ordinary least squares (OLS) is used.

One of the disadvantages of this modelling strategy is that we are assuming the same structure for all elements of the Cholesky factors in X_t. Much more importantly, we are leaving a significant amount of information in the error term. One can expect the error term to be heteroscedastic due to volatility of volatility (Corsi *et al.*, 2008) present in the realized measures. More importantly, a common structure of X_t elements may be left unmodelled in residuals. Hence it may be more natural to estimate the model in Eq. 8 as a system of equations with some covariance structure of the error terms.

To deal with this problem, we propose to build a system of seemingly unrelated HAR regressions (Zellner, 1962) for all elements of X_t. The advantage of this approach is that we estimate a multivariate HAR model, which will capture the separate dynamics of the variances and covariances, but also possible common structure. Moreover, it will also yield more efficient estimates. As we know, error terms from HAR are heteroscedastic (Corsi *et al.*, 2008), which makes the coefficient estimates less efficient. Moreover, where there is no information about dependence between equations left in the residuals from regression Eq. 8, the estimator will converge to a simple OLS estimate, as the diagonal weighting matrix in generalized regression will reduce the estimates to OLS. On the other hand, the possible disadvantage is in a larger number of parameters to be estimated, which may render the model unreliable with highly dimensional portfolios.

Let us consider the system of $i = 1, \ldots, m$ equations, where $m = \frac{q(q+1)}{2}$:

$$X_{i,t+1}^{(1)} = \beta_i^{(c)} + \beta_i^{(1)} X_{i,t}^{(1)} + \beta_i^{(5)} X_{i,t}^{(5)} + \beta_i^{(22)} X_{i,t}^{(22)} + \epsilon_{i,t}, \;\; \epsilon_{i,t} \sim \text{i.i.d.} \tag{9}$$

There are m equations representing elements of the Cholesky factors, with T observations. Define the $mT \times 1$ vector of disturbances $\epsilon = \left(\epsilon_1', \ldots, \epsilon_m'\right)'$, and rewrite the model as

$$\begin{pmatrix} X_{1,t+1}^{(1)} \\ \vdots \\ X_{m,t+1}^{(1)} \end{pmatrix} = \begin{pmatrix} X_{1,t} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & X_{m,t} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} + \begin{pmatrix} \epsilon_{1,t} \\ \vdots \\ \epsilon_{m,t} \end{pmatrix} \tag{10}$$

where $X_{i,t} = \left(e \;\; X_{i,t}^{(1)} \;\; X_{i,t}^{(5)} \;\; X_{i,t}^{(22)} \right)$ is the ith element of X_t and e a vector of ones, $\beta_i = \left(\beta_i^{(c)} \;\; \beta_i^{(1)} \;\; \beta_i^{(5)} \;\; \beta_i^{(22)} \right)'$ and $\beta_i^{(c)}$ estimates of the intercept. It is more convenient to work with this system in the following form:

$$y = Z\beta + \epsilon \tag{11}$$

where $y = \left(X_{1,t+1}^{(1)}, \ldots, X_{m,t+1}^{(1)}\right)'$ and ϵ are of dimension $mT \times 1$, $Z = \text{diag}\{X_{1,t}, \ldots, X_{m,t}\}$ is a block diagonal matrix of dimension $mT \times 4m$, and the matrix of parameters $\beta = (\beta_1, \ldots, \beta_m)'$ is of dimension $4m \times 1$.

The disturbances will satisfy strict exogeneity $E[\epsilon|Z] = 0$, but will be correlated across equations, $E[\epsilon_i'\epsilon_j|Z] = \sigma_{ij} I_T$ or

$$\Omega = \begin{pmatrix} \sigma_{11} I_T & \cdots & \sigma_{1m} I_T \\ \vdots & \ddots & \vdots \\ \sigma_{m1} I_T & \cdots & \sigma_{mm} I_T \end{pmatrix} = \Sigma \otimes I_T \tag{12}$$

where $\Sigma = \sigma_{ij}$ for $i, j = 1, \ldots, m$, $\otimes$ is a Kronecker product and I_T is an identity matrix of dimension $T \times T$. The model parameters are estimated in two-step feasible generalized least squares. We run OLS regression in the first step to obtain estimates $\hat{\sigma}_{ij}$ from residuals. In the second step, we run generalized least squares regression using the variance matrix $\widehat{\Omega} = \widehat{\Sigma} \otimes I_T$ as

$$\widehat{\beta} = \left(Z'\widehat{\Omega}^{-1} Z\right)^{-1} Z'\widehat{\Omega}^{-1} y \tag{13}$$

The estimator $\widehat{\beta}$ is unbiased, and a consistent estimator of β with asymptotically normal limiting distribution:

$$\sqrt{T}\left(\widehat{\beta}-\beta\right) \xrightarrow{d} \mathcal{N}\left(0, \left(\frac{1}{T} Z'\widehat{\Omega}^{-1} Z\right)^{-1}\right) \tag{14}$$

While this is a standard estimation technique, we will refrain from discussing any further details about the properties of the generalized least squares estimator.

Competing models

To show the contribution of the GHAR model, we compare the forecasts to several competing alternatives. The first natural choice of benchmark model is a multivariate extension of the original HAR. By comparing these two models, we will see the portion of the contribution brought by allowing for correlated residuals in the estimation. Another natural candidate is vector ARFIMA, as Chiriac and Voev (2011) find it outperforms the HAR model slightly, but conclude that HAR performs reasonably well in comparison to VARFIMA. Hence we may have reason to believe that our approach will provide better results than the VARFIMA model.

These three main models share the same framework of modelling elements of Cholesky factors from the realized covariance matrix. Hence we also contrast them to two benchmark models, namely the popular DCC-GARCH[3] of Engle (2002) and the risk metrics standard widely used in the business industry. These benchmark models operate on daily data, so we will have a direct comparison of gains from high-frequency data.

HAR

A first, natural competing model to our generalized HAR strategy is the multivariate extension of an original HAR, which models a vector of Cholesky factors X_t as

$$X_{t+1}^{(1)} = c + \beta^{(1)} X_t^{(1)} + \beta^{(5)} X_t^{(5)} + \beta^{(22)} X_t^{(22)} + \epsilon_t, \;\; \epsilon_t \sim \text{i.i.d.} \tag{15}$$

where 1, 5 and 22 stand for day, week (5 days) and month (22 days) respectively, c is an $m \times 1$ vector of constants, $\beta^{()}$ is an $m \times 1$ vector of parameters and $X_t^{()}$ are averages of lagged daily volatility, e.g. $X_t^{(5)} = \frac{1}{5}\sum_{i=0}^{4} X_{t-i}$. To obtain parameter estimates, OLS is used.

Vector ARFIMA model

A second competing model to the HAR family is the vector autoregressive fractionally integrated moving average (VARFIMA) model of Chiriac and Voev (2011), who use a restricted VARFIMA$(1, d, 1)$ specification to model and forecast dynamics of X_t directly. The authors find that ARFIMA provides a slightly better forecast in comparison to the HAR model, which makes it a natural candidate for our modelling strategy. We consider the vector ARFIMA model

$$(1-\phi L)\, D(L)\, [X_t - c] = (1-\theta L)\, \epsilon_t, \;\; \epsilon_t \sim N(0, \Sigma) \tag{16}$$

where ϕ and θ are scalars, c is an $m \times 1$ vector of constants and $D(L) = (1-L)^d I_m$ with a common parameter of fractional integration d for all constituents of X_t. In our case we reject the hypothesis about equality of d; thus we estimated each element of X_t using a unique $d_t : D(L) = \text{diag}\left\{(1-L)^{d_1}, \ldots, (1-L)^{d_m}\right\}$. Hence we use model 1 in Chiriac and Voev (2011). In addition, we have experimented with a general VARFIMA(p, d, q), not restricting $p = q = 1$.[4] Comparing the models through information criteria decisively yields VARFIMA$(1, d, 1)$ as the best model; hence we use it as a benchmark for our modelling strategy in the empirical section of the paper.

RiskMetrics

RiskMetrics of JP Morgan Chase, based on an exponentially weighted moving average (EWMA), is a financial industry standard and common benchmark for any volatility model (univariate or multivariate). In our work we use the specification of Longerstaey and Spencer (1996) with decay factor λ set to 0.94. We assume a $q \times 1$ vector of daily returns $r_t = \sum_{k=1}^{n} (\Delta_k p_t)$ for $t = 1, \ldots, T$ such that $r_t \sim N\left(\mu_t, \sigma_t^2\right)$, where μ_t is the conditional mean and σ_t^2 the conditional variance of daily returns. Moreover, if we assume $\mu_t = 0$, the conditional covariance has the form

[3] DCC-GARCH is an industry standard and we decided to implement it in its original form, despite the known problem with consistency of the estimator. For more information about the inconsistency of DCC see Aielli (2013) and Caporin and McAleer (2013).

[4] These results are available upon request from the authors.

$$\sigma_{i,j} = (1-\lambda)\sum_{t=1}^{T}\lambda^{t-1}r_i r_j \tag{17}$$

The previous equation can be rewritten in recursive form:

$$\sigma_{i,j,t} = \lambda\sigma_{i,j,t-1} + (1-\lambda)r_{i,t-1}r_{j,t-1} \tag{18}$$

where the expression $\sigma_{i,j,t}$ stands for covariance between assets i and j in time t.

DCC-GARCH

The dynamic conditional correlation generalized autoregressive conditional heteroscedasticity (DCC-GARCH) of Engle (2002) is a widely used multivariate GARCH model in practice. It is a generalization of Bollerslev (1990) constant conditional correlation GARCH, with time-varying correlation matrix R. The model is defined as

$$H_t = D_t R_t D_t \tag{19}$$

where D_t is a diagonal matrix of conditional time-varying standard deviations, $D_t = \text{diag}\left(\sqrt{h_{i,t}}\right)$ and $h_{i,t}$ are univariate GARCH processes, $h_{i,t} = \omega_i + \sum_{p=1}^{P_i}\alpha_{i,p}r_{i,t-p}^2 + \sum_{q=1}^{Q_i}\beta_{i,q}h_{i,t-q}$. The dynamics of the correlation matrix are given by transformation:

$$R_t = Q_t^{*-1}Q_t Q_t^{*-1} \tag{20}$$

where $Q_t = \left(1 - \sum_{m=1}^{M}\alpha_m - \sum_{n=1}^{N}\beta_n\right)\bar{Q} + \sum_{m=1}^{M}A_m\left(\epsilon_{t-m}\epsilon_{t-m}^T\right) + \sum_{n=1}^{N}B_n Q_{t-n}$, $\bar{Q}$ is the unconditional covariance matrix of the standardized residuals from the univariate GARCH processes and $Q_t^* = \text{diag}\left(\sqrt{q_{ii,t}}\right)$. In our work we use the two-stage estimator presented in Engle (2002) and Engle and Sheppard (2001).

DATA AND RESEARCH DESIGN

The dataset consists of tick prices of 15 S&P 500 index constituents with highest liquidity and market capitalization. The final portfolio thus consists[5] of Apple Inc. (AAPL), Exxon Mobile Corp. (XOM), Google Inc. (GOOG), Wal-Mart Stores (WMT), Microsoft Corp. (MSFT), General Electric Co. (GE), International Business Machines Corp. (IBM), Johnson & Johnson (JNJ), Chevron Corp. (CVX), Procter & Gamble (PG), Pfizer Inc. (PFE), AT&T Inc. (T), Wells Fargo & Co. (WFC), JP Morgan Chase & Co. (JPM) and Coca-Cola Co. (KO). We obtain 390, 78, 39, 26 and 19 time-synchronized intraday observations using refresh-time, resulting in 1-, 5-, 10-, 15- and 20-minute intraday returns. Besides 1- to 20-minute returns we also construct open-to-close returns that are used for RiskMetrics and DCC-GARCH models. Moreover, we create sub-portfolios consisting of 5, 10 and 15 assets (assets chosen according to market capitalization). Hence, in total, we study 18 different datasets.

The sample covers the period from 1 July 2005 to 3 January 2012 (1623 trading days), and we consider trades between 9:30 and 16:00 EST time. To ensure sufficient liquidity on the market we explicitly exclude weekends and holidays (New Year's Day, Independence Day, Thanksgiving Day, Christmas). For estimation and forecasting purposes we divide our sample into in-sample, spanning from 1 July 2005 to 9 July 2008, and out-of-sample from 10 July 2008 to 3 January 2012. For the forecasting, we use rolling window estimation with a fixed length of 750 days. Summary statistics of all returns are presented in Appendix D.

Accuracy of the forecasts is evaluated primarily according to economic criteria. The rationale behind this is the importance of well-conditioned and invertible forecasts rather than focusing on unbiasedness, as an unbiased forecast does not necessarily translate into an unbiased inverse (Bauwens *et al.*, 2012). As a robustness check we also provide ranking of the models based on statistical loss functions.

Economic forecast evaluation

For economic evaluation of volatility forecasts, we use the approach of Markowitz (1952). There are two possibilities for constructing an optimal portfolio. In the first one we specify the expected portfolio return and try to find asset weights minimizing the risk. In the second one the expected return of the portfolio is maximized according to a certain risk. Asset weights, $w = (w_1, \ldots, w_q)'$, maximizing the utility of a risk-averse investor, can be found by solving the following problem:

[5] Assets are ordered according to market capitalization.

$$\begin{aligned} \min_{w_{t+1}} \quad & w'_{t+1}\widehat{\Sigma}_{t+1|t}w_{t+1} \\ \text{s.t.} \quad & l'w_{t+1} = 1 \\ & w'_{t+1}\widehat{\mu}_{t+1|t} = \mu_P \end{aligned} \tag{21}$$

where w_{t+1} is a $q \times 1$ vector of asset weights, $\widehat{\Sigma}_{t+1|t}$ represents a covariance matrix forecast, l denotes a $q \times 1$ vector of ones, $\widehat{\mu}_{t+1|t}$ is a vector of mean forecasts and μ_P stands for portfolio return. Once the optimization problem is solved for different risk levels, we are able to construct an efficient frontier. The Markowitz-type portfolio relies heavily on mean forecasts. As these forecasts might be noisy, portfolio weights and variance can become notably sensitive to changes in assets mean. To overcome these difficulties we also consider the problem of finding the global minimum variance portfolio (GMVP). The specification of the optimization problem is similar to the Markowitz setup:

$$\begin{aligned} \min_{w_{t+1}} \quad & w'_{t+1}\widehat{\Sigma}_{t+1|t}w_{t+1} \\ \text{s.t.} \quad & l'w_{t+1} = 1 \end{aligned} \tag{22}$$

which can be solved analytically:[6]

$$w_{t+1}^{\mathrm{GMV}} = \frac{\widehat{\Sigma}_{t+1|t}^{-1} l}{l'\widehat{\Sigma}_{t+1|t}^{-1} l} \tag{23}$$

with expected return variance being

$$\sigma^{2\,\mathrm{GMV}}_{t+1} = w_{t+1}^{\mathrm{GMV}\prime}\widehat{\Sigma}_{t+1|t}w_{t+1}^{\mathrm{GMV}} = \frac{1}{l'\widehat{\Sigma}_{t+1|t}^{-1} l} \tag{24}$$

Statistical forecast evaluation

For statistical evaluation of covariance forecasts, we employ root mean squared error (RMSE) loss functions based on the Frobenius norm.[7] As a volatility proxy we use realized covariance, subsampled realized covariance (RCOV SS) and MRK estimates at given frequencies; i.e. to calculate loss function for forecasts based on 5-minute realized covariance we use realized covariance estimates based on 5-minute data as a benchmark. In the case of DCC-GARCH and RiskMetrics forecasts we calculate loss functions using all RCOV, RCOV SS and MRK estimates at all frequencies. The measures are calculated for the $t = 1, \ldots, T$ forecasts as

$$e_{t,t+h} = \Sigma_{t+h} - \widehat{\Sigma}_{t+h|t} \tag{25}$$

$$\mathcal{L}^{\mathrm{RMSE}} = \sqrt{\frac{1}{T-1}\sum_{t=1}^{T}\sum_{i,j}\left|e_{t_{i,j}}\right|^2} \tag{26}$$

where $\widehat{\Sigma}_{t+h|t}$ is a covariance matrix forecast and Σ_{t+h} is the volatility proxy.

To test the significant differences of competing models, we use the model confidence set (MCS) methodology of Hansen *et al.* (2011). Given a set of forecasting models, $\mathcal{M}_0$, we identify the model confidence set $\widehat{\mathcal{M}}^*_{1-\alpha} \subset \mathcal{M}_0$, which is the set of models that contain the 'best' forecasting model given a level of confidence α. For a given model $i \in \mathcal{M}_0$, the p-value is the threshold confidence level. Model i belongs to the MCS only if $\widehat{p}_i \geq \alpha$. MCS methodology repeatedly tests the null hypothesis of equal forecasting accuracy

$$H_{0,\mathcal{M}} : E[\mathcal{L}_{i,t} - \mathcal{L}_{j,t}] = 0, \quad \text{for all } i, j \in \mathcal{M}$$

with $L_{i,t}$ being an appropriate loss function of the ith model. Starting with the full set of models, $\mathcal{M} = \mathcal{M}_0$, this procedure sequentially eliminates the worst-performing model from $\mathcal{M}$ when the null is rejected. The surviving set of models then belong to the model confidence set $\widehat{\mathcal{M}}^*_{1-\alpha}$. Following Hansen *et al.* (2011), we implement the MCS using a stationary bootstrap with an average block length of 10 days.[8]

[6] Kempf and Memmel (2006).

[7] The Frobenius norm of the $m \times n$ matrix A is defined as $\|A\|_F{}^2 = \sum_{i,j} |a_{i,j}|^2$.

[8] We have used different block lengths, including those dependent on the forecasting horizons, to assess the robustness of the results, without any change in the final results. These results are available from the authors upon request.

Table I. Cumulative version of GMVP: portfolio of five stocks.

	MRK	RCOV		Subsampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	**30.50**	**30.50**	**30.50**	**30.50**	**30.50**	30.50	30.50
RiskMetrics	40.64	40.64	40.64	40.64	40.64	40.64	40.64
VARFIMA	30.76	34.47	32.44	32.84	31.04	29.86	29.31
GHAR	30.60	34.14	32.22	32.53	30.83	**29.65**	**29.08**
HAR	31.42	34.84	33.05	33.35	31.61	30.50	29.99

Note: The model with the overall best performance is highlighted; for the given frequency the model with the lowest risk is presented in bold; values represent percentage level of risk.

Table II. Annualized version of GMVP: portfolio of five stocks.

	MRK	RCOV		Subsampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	17.38	**17.38**	**17.38**	**17.38**	17.38	17.38	17.38
RiskMetrics	23.13	23.13	23.13	23.13	23.13	23.13	23.13
VARFIMA	17.62	19.39	18.44	18.61	17.68	17.04	16.77
GHAR	**17.32**	19.08	18.08	18.27	**17.36**	**16.69**	**16.38**
HAR	18.01	19.61	18.79	18.91	18.01	17.40	17.14

Note: The model with the overall best performance is highlighted; for the given frequency the model with the lowest risk is presented in bold; values represent percentage level of risk.

RESULTS

For clarity of presentation, we begin with a discussion of the results of one-step-ahead forecasts for the portfolio of five stocks (AAPL, XOM, GOOG, WMT, MSFT), whereas we leave portfolios of 10 and 15 stocks and also five- and 10-step-ahead forecasts as a robustness check showing that the methodology also works well at larger dimensions and different forecasting horizons. Focusing on the economic evaluation, we first discuss the results from GMVP,[9] followed by the Markowitz approach and statistical evaluation.

We present GMVP comparison through cumulative and annualized risk. In the cumulative approach we use covariance forecasts for daily rebalancing of our portfolio: at each step we calculate optimal asset weights and using these weights we calculate corresponding daily portfolio risk. The results presented in Table I are sums of portfolio risk $\sigma_{\text{cum.}}$ for the whole out-of sample period. Table I is divided into seven parts according to realized measures and frequencies used for the calculation. For RiskMetrics and DCC-GARCH the corresponding $\sigma_{\text{cum.}}$ are constant for all frequencies because they are calculated using open–close returns. We present the results of DCC-GARCH and RiskMetrics in all columns of Table I so that we can compare the performance of covariance-based models estimated on different frequencies with daily data-based models.

From Table I we can see that the model with the best performance and thus lowest level of risk is GHAR. We can also observe that for various frequencies on which realized measures are calculated DCC-GARCH outperformed covariance-based models. However, these results do not indicate superiority of DCC-GARCH compared to covariance-based models, but highlight the importance of selecting realized measures properly.

A disadvantage of model comparison according to cumulative risk is daily rebalancing, implying high transaction costs. A more comprehensive method of model comparison is to use annualized portfolio risks. In Table II we present the results for the annualized version of GMVP.

Similar to cumulative GMVP, the model with the overall lowest achievable risk is GHAR. Remaining results from Tables I and II partly match the results presented in Chiriac and Voev (2011). The model that scored second is VARFIMA, followed by HAR for subsampled RCOV estimated at 15- and 20-minute frequencies. For the remaining frequencies and realized measures, DCC-GARCH outperforms covariance-based models. Overall, we can say that covariance-based models with appropriate choice of realized measure outperform return-based models.

To assess the performance of models not only from the risk-minimizing point of view but also return maximization, we present efficient frontiers. In contrast to GMVP, we do not allow short selling here.[10] For the calculation of efficient frontiers we use annualized forecasts of covariance matrices and annualized returns.

[9] With short selling allowed.
[10] In the case that short selling is allowed, the ranking of the models is unchanged, and only the magnitude differs.

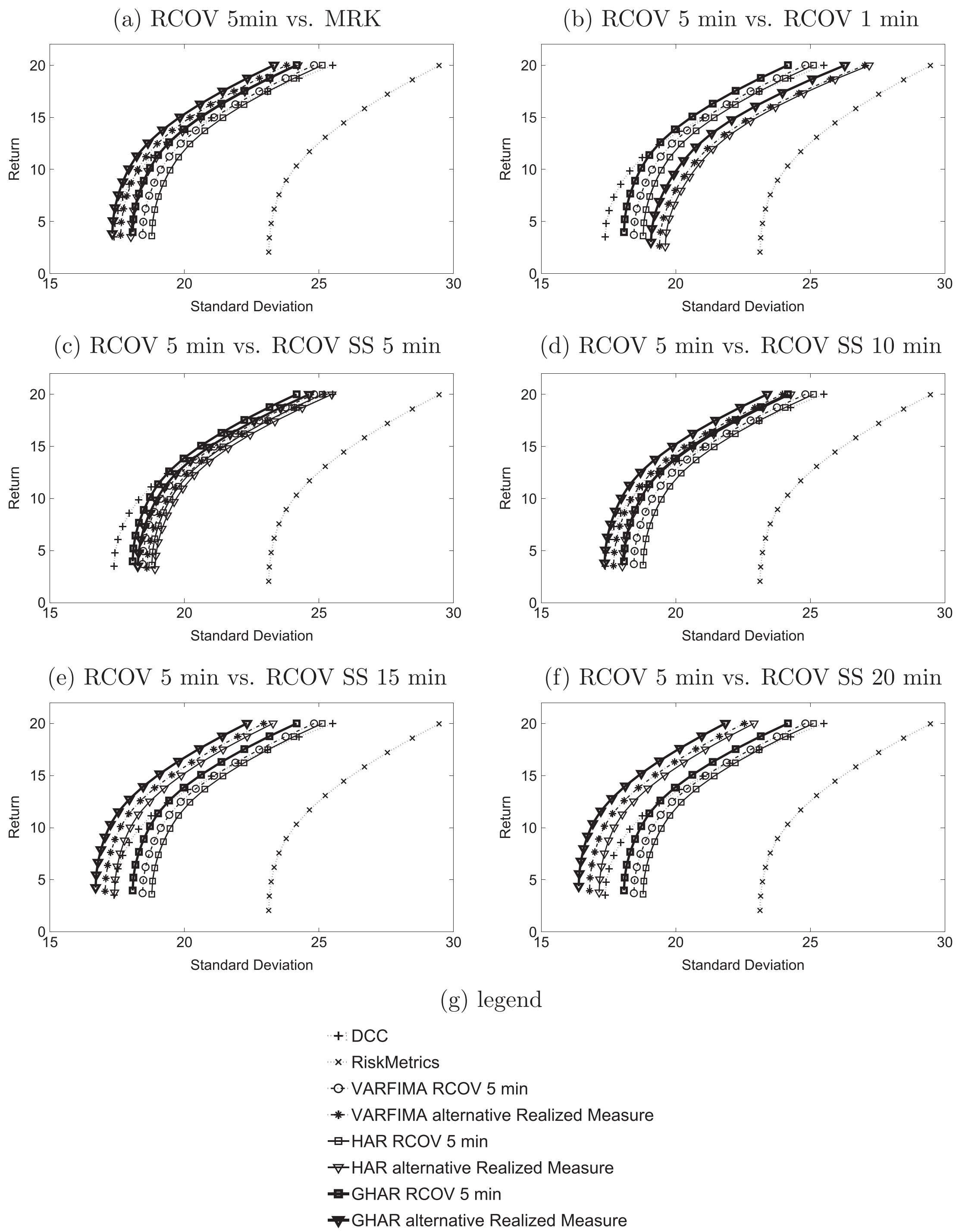

Figure 1. Efficient frontiers: portfolio of five stocks: (a) RCOV 5-minute vs. MRK; (b) RCOV 5-minute vs. RCOV 1-minute; (c) RCOV 5-minute vs. RCOV SS 5-minute; (d) RCOV 5-minute vs. RCOV SS 10-minute; (e) RCOV 5-minute vs. RCOV SS 15-minute; (f) RCOV 5-minute vs. RCOV SS 20-minute

Similar to the results from the GMVP evaluation model with the best risk–return trade-off is the model proposed in this paper: GHAR. The second-best-performing model is VARFIMA, followed by HAR. From Figure 1 we can see that for estimates at 1 minute RCOV and 5 minutes RCOV the score of DCC-GARCH is better than all covariance-based models, which is not in line with results presented in Chiriac and Voev (2011), where DCC-GARCH ended in the penultimate position. We can attribute this difference to a different dataset and period that includes a financial crisis during which periods of high intraday volatility are observable.

As a robustness check to the economic evaluation, we also provide results from a statistical comparison of forecasting performance of the competing models. In Table III a comparison based on the RMSE loss function is presented.

Table III. RMSE: portfolio of five stocks.

	MRK	RCOV		Subsampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	1.593	1.730	1.914	1.707	1.547	1.481	1.474
RiskMetrics	1.668	1.728	1.866	1.709	1.646	1.636	1.633
VARFIMA	1.406	1.537	1.682	1.473	1.363	1.331	1.328
GHAR	1.490	1.401	1.740	1.509	1.438	1.430	1.445
HAR	1.190	1.100	1.380	1.162	1.125	1.144	1.158

Note: Values are scaled by 10^{-3}; highlighted cells belong to 5% MCS.

From the RMSE perspective the lowest error is shown by the HAR model, followed by VARFIMA and GHAR. These models always belong to 5% MCS irrespective of the realized measure used for comparison. The worst performance was shown by RiskMetrics, which does not belong to 5% MCS in two cases and has the highest RMSE in five out of seven cases.

Robustness check

Having discussed the results of one-step-ahead forecasts for a portfolio consisting of five stocks, we now turn to evaluation of one-step-ahead forecasts for a portfolio consisting of 10 (AAPL, XOM, GOOG, WMT, MSFT, GE, IBM, JNJ, CVX, PG), and 15 (AAPL, XOM, GOOG, WMT, MSFT, GE, IBM, JNJ, CVX, PG, PFE, T, WFC, JPM, KO) stocks and five- and 10-step-ahead forecasts for portfolios consisting of five, 10 and 15 stocks. We will concentrate on the main differences compared to the smaller portfolio, as we use these results as a robustness check. We also relegate the tables and figures to Appendices A, B and C.

Portfolio of 10 and 15 stocks

According to GMVP criteria for a portfolio consisting of 10 stocks, results do not differ from results obtained using a portfolio of five stocks. The model with the lowest cumulative and annualized risk is GHAR, estimated on 20-minute subsampled RCOV. In the case of the portfolio consisting of 15 stocks, the only difference is that GHAR estimated on MRK covariance matrices outperformed DCC-GARCH.

From the risk–return trade-off point of view there is a notable difference for portfolio consisting of 10 stocks when the data for higher frequencies (1, 5 and 10 minutes) are used. For these frequencies, the model with the best risk–return trade-off is DCC-GARCH. The order of the remaining models is identical to the portfolio of five stocks: GHAR followed by VARFIMA and HAR. If the 15-minute data are used for optimization, GHAR shares first place with DCC-GARCH. These two models are closely followed by VARFIMA and HAR. For the 20-minute data ordering of the models is similar to the portfolio consisting of five stocks.

Concentrating on statistical evaluation, results of RMSE model comparison for the portfolio consisting of 10 stocks are almost identical to results for the portfolio of five stocks, the only difference being that RiskMetrics does not belong to 5% MCS in any of the cases. On the other hand, a notable difference occurs in a comparison of the portfolio consisting of 15 stocks, where GHAR belongs to 5% MCS only in one case (estimated at 5-minute RCOV), and DCC-GARCH and RiskMetrics do not belong to 5% MCS at all. We address unambiguous results of statistical evaluation to the problem of selecting the 'correct' proxy. These results are also consistent with findings in Kyj *et al.* (2010), who show that for large portfolios, the pure high frequency based covariance forecasts need to be conditioned in order to achieve the benefits of the high frequency data.

This points us to the result that unmodelled dependence from HAR and VARFIMA models is increasing with increasing dimension of the portfolio. Hence the GHAR model delivers significant economic gains with increasing dimension of portfolio.

Five- and 10-step-ahead forecasts

Extension of forecasting horizon from 1 to 5 to 10 days does not substantially change the results of our analysis. The only notable difference is absence of GHAR in 5% MCS in the case of 10-step-ahead forecasts of a portfolio consisting of 15 stocks. Remaining results support our previous findings that application of seemingly unrelated regression for HAR estimation delivers significant economic gains regardless of the size of the portfolio and/or forecasting horizon.[11]

CONCLUSION

In this paper we propose to employ the seemingly unrelated regression of Zellner (1962) to estimate multivariate extension of the heterogeneous autoregression model in order to improve the variance matrix forecasts. The resulting

[11] To make the results comparable, we scale them according to the forecasting horizon.

model—generalized HAR (GHAR)—inherits all the favourable properties of HAR, and provides us with a more efficient estimator that accounts for otherwise hidden dependencies among variables.

In our setup we closely follow Chiriac and Voev (2011) and model elements of Cholesky decomposed covariance matrices to test the economic and statistical value of the proposed modelling strategy. Moreover, we perform our analysis on portfolios consisting of five, 10 and 15 assets, we include three covariance matrix estimators (realized covariation, subsampled realized covariation and multivariate realized kernels), and we obtain covariance matrix estimates using high-frequency data of five different frequencies (1, 5, 10, 15 and 20 minutes). Overall, we test the performance of the GHAR estimator on 15 different high-frequency datasets. The resulting forecasts of GHAR prove to perform significantly better than benchmark models according to global minimum variance portfolio and mean–variance evaluation criteria irrespective of frequency or size of the portfolio. Whereas our study focuses on more important economic evaluation of the forecasts, statistical evaluation is used as a robustness check of the results. According to statistical criteria for comparison of models, we find that GHAR is not systematically dominated by any benchmark model, which is a supportive result for economic evaluation.

ACKNOWLEDGEMENTS

We are indebted to the Editor and two anonymous referees for many useful comments, suggestions and discussions. Support from the Czech Science Foundation under the 13-32263S project and support from the Grant Agency of Charles University under the 1198214 project is gratefully acknowledged.

REFERENCES

Aielli GP. 2013. Dynamic conditional correlation: on properties and estimation. *Journal of Business and Economic Statistics* **31**(3): 282–299.

Ait-Sahalia Y, Fan J, Xiu D. 2010. High frequency covariance estimates with noisy and asynchronous financial data. *Journal of the American Statistical Association* **105**: 1504–1517.

Amendola A, Storti G. 2015. Model uncertainty and forecast combination in high- dimensional multivariate volatility prediction. *Journal of Forecasting* **34**: 83–91.

Andersen TG, Bollerslev T, Diebold FX, Labys P. 2003. Modeling and forecasting realized volatility. *Econometrica* **71**(2): 579–625.

Bannouh K, van Dijk D, Martens M. 2009. Range-based covariance estimation using high-frequency data: the realized co-range. *Journal of Financial Econometrics* **7**(4): 341–372.

Barndorff-Nielsen OE, Shephard N. 2004. Econometric analysis of realized covariation: high frequency based covariance, regression, and correlation in financial economics. *Econometrica* **72**(3): 885–925.

Barndorff-Nielsen OE, Hansen PR, Lunde A, Shephard N. 2011. Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. *Journal of Econometrics* **162**(2): 149–169.

Bauer GH, Vorkink K. 2011. Forecasting multivariate realized stock market volatility. *Journal of Econometrics* **160**(1): 93–101.

Bauwens L, Laurent S, Rombouts JV. 2006. Multivariate GARCH models: a survey. *Journal of Applied Econometrics* **21**(1): 79–109.

Bauwens L, Hafner CM, Laurent S. 2012. *Handbook of Volatility Models and their Applications*, Vol. 3. John Wiley & Sons: Chichester.

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

Bonato M. 2009. *Estimating the degrees of freedom of the realized volatility Wishart autoregressive model*. Available at SSRN 1357044.

Bonato M, Caporin M, Ranaldo A. 2013. Risk spillovers in international equity portfolios. *Journal of Empirical Finance* **24**: 121–137.

Boudt K, Cornelissen J, Croux C. 2012. Jump robust daily covariance estimation by disentangling variance and correlation components. *Computational Statistics and Data Analysis* **56**(11): 2993–3005.

Caporin M, McAleer M. 2013. Ten things you should know about the dynamic conditional correlation representation. *Econometrics* **1**(1): 115–126.

Chiriac R, Voev V. 2011. Modelling and forecasting multivariate realized volatility. *Journal of Applied Econometrics* **26**(6): 922–947.

Corsi F. 2009. A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics* **7**(2): 174–196.

Corsi F, Mittnik S, Pigorsch C, Pigorsch U. 2008. The volatility of realized volatility. *Econometric Reviews* **27**(1–3): 46–78.

Elst VH, Veredas D. 2015. *Smoothing it out: empirical and simulation results for disentangled realized covariances*. SSRN Scholarly Paper ID 2490335, Social Science Research Network: Rochester, NY.

Engle R. 2002. Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business and Economic Statistics* **20**(3): 339–350.

Engle RF, Kroner KF. 1995. Multivariate simultaneous generalized ARCH. *Econometric Theory* **11**(01): 122–150.

Engle RF, Sheppard K. 2001. Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH. *Technical report* National Bureau of Economic Research.

Epps TW. 1979. Comovements in stock prices in the very short run. *Journal of the American Statistical Association* **74**: 291–296.

Gouriéroux C, Jasiak J, Sufana R. 2009. The Wishart autoregressive process of multivariate stochastic volatility. *Journal of Econometrics* **150**(2): 167–181.

Hansen PR, Lunde A, Nason JM. 2011. The model confidence set. *Econometrica* **79**(2): 453–497.

Hautsch N, Kyj L M, Oomen RCA. 2012. A blocking and regularization approach to high-dimensional realized covariance estimation. *Journal of Applied Econometrics* **27**(4): 625–645.

Hayashi T, Yoshida N. 2005. On covariance estimation of non-synchronously observed diffusion processes. *Bernoulli* **11**: 359–379.

Jin X, Maheu JM. 2013. Modeling realized covariances and returns. *Journal of Financial Econometrics* **11**(2): 335–369.

Kempf A, Memmel C. 2006. Estimating the global minimum variance portfolio. Schmalenbach Business Review. *Available at SSRN* **58**(4): 332–348.

Kyj LM, Ostdiek B, Ensor K. Covariance estimation in dynamic portfolio optimization: a realized single factor model, In *American Finance Association 2010 Atlanta Meetings*, 2010.

Longerstaey J, Spencer M. 1996. *RiskMetrics™: Technical document*. Morgan Guaranty Trust Company of New York.

Markowitz H. 1952. Portfolio selection. *Journal of Finance* **7**(1): 77–91.

Voev V, Lunde A. 2007. Integrated covariance estimation using high-frequency data in the presence of noise. *Journal of Financial Econometrics* **5**: 68–104.

Zellner A. 1962. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American statistical Association* **57**(298): 348–368.

Zhang L. 2011. Estimating covariation: Epps effect, microstructure noise. *Journal of Econometrics* **160**(1): 33–47.

Zhang L, Mykland P A, Ait-Sahalia Y. 2005. A tale of two time scales: determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association* **100**: 1394–1411.

APPENDIX A: ONE-STEP-AHEAD FORECASTS

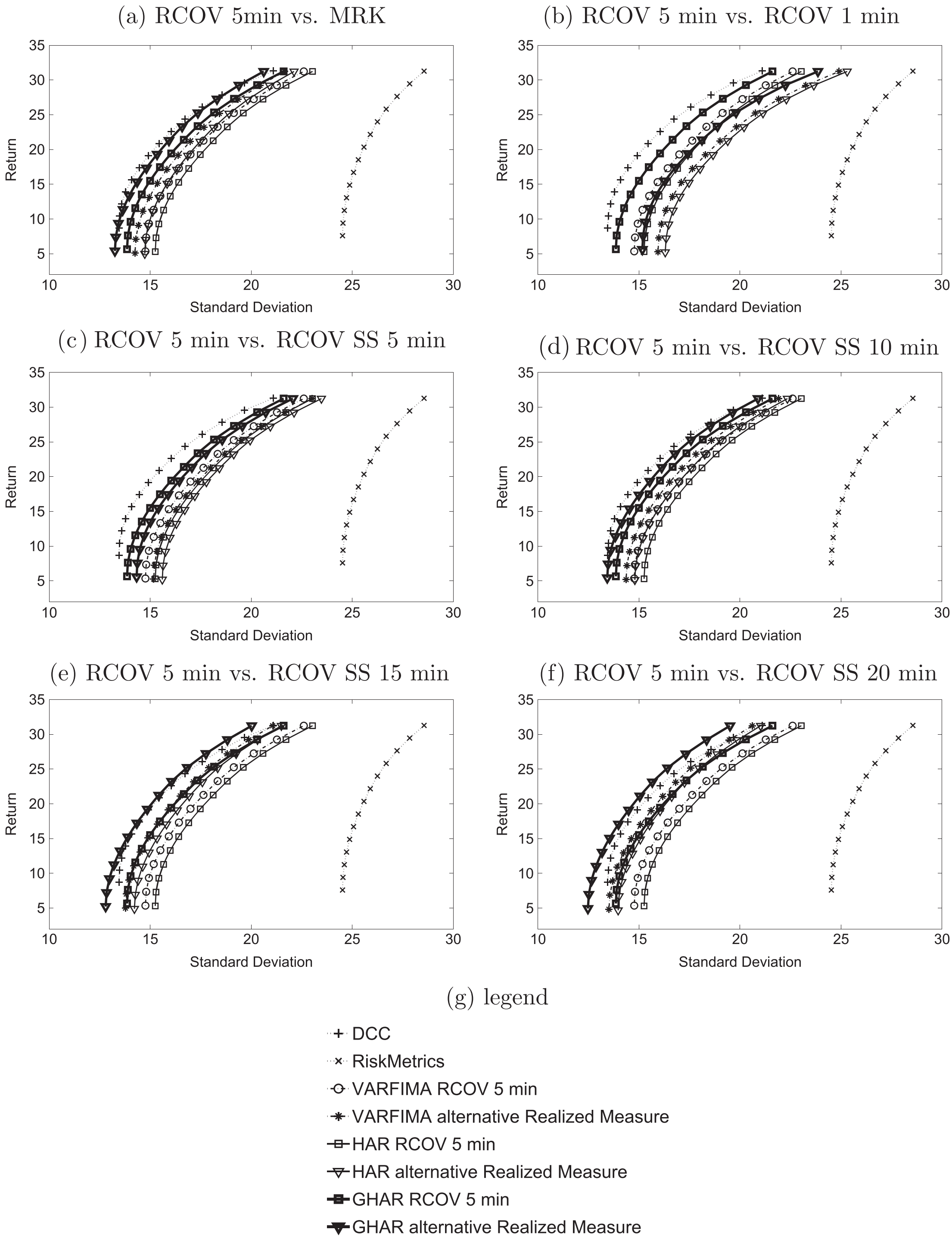

Figure A.1. Efficient frontiers: portfolio of 10 stocks: (a) RCOV 5-minute vs. MRK; (b) RCOV 5-minute vs. RCOV 1-minute; (c) RCOV 5-minute vs. RCOV SS 5-minute; (d) RCOV 5-minute vs. RCOV SS 10-minute; (e) RCOV 5-minute vs. RCOV SS 15-minute (f) RCOV 5-minute vs. RCOV SS 20-minute

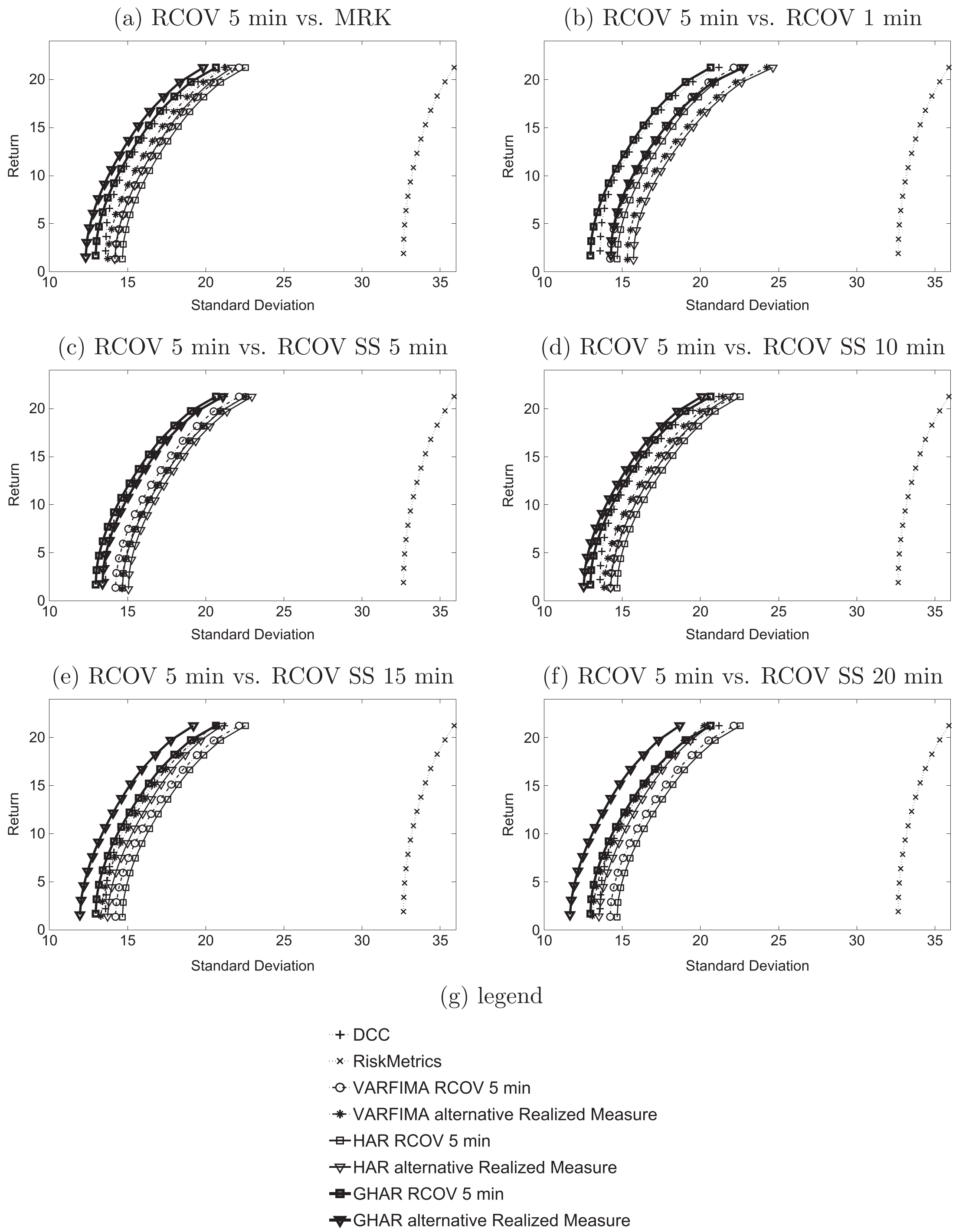

Figure A2. Efficient frontiers: portfolio of 15 stocks: (a) RCOV 5-minute vs. MRK; (b) RCOV 5-minute vs. RCOV 1-minute; (c) RCOV 5-minute vs. RCOV SS 5-minute; (d) RCOV 5-minute vs. RCOV SS 10-minute; (e) RCOV 5-minute vs. RCOV SS 15-minute; (f) RCOV 5-minute vs. RCOV SS 20-minute

Table A.1. GMVP: portfolio of 10 stocks

	MRK	RCOV		Subsampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
Cumulative							
DCC	**22.14**	**22.14**	**22.14**	**22.14**	**22.14**	22.14	22.14
RiskMetrics	42.15	42.15	42.15	42.15	42.15	42.15	42.15
VARFIMA	23.34	27.70	24.75	25.64	23.82	22.52	21.85
GHAR	22.50	26.71	23.90	24.79	22.98	**21.66**	**20.98**
HAR	24.28	28.30	25.66	26.40	24.63	23.39	22.79
Annualized							
DCC	13.12	**13.12**	**13.12**	**13.12**	13.12	13.12	13.12
RiskMetrics	24.32	24.32	24.32	24.32	24.32	24.32	24.32
VARFIMA	13.74	15.76	14.40	14.84	13.90	13.21	12.88
GHAR	**12.82**	15.00	13.53	14.04	**13.03**	**12.30**	**11.91**
HAR	14.31	16.14	14.96	15.31	14.40	13.74	13.43

Note: The model with the overall best performance is highlighted; for the given frequency the model with the lowest risk is presented in bold; values represent percentage level of risk.

Table A.2. RMSE: portfolio of 10 stocks

	MRK	RCOV		Subsampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	3.242	3.624	3.896	3.600	3.162	3.044	3.085
RiskMetrics	3.808	4.006	4.167	3.949	3.803	3.822	3.846
VARFIMA	2.592	3.028	3.228	2.903	2.551	2.494	2.539
GHAR	3.101	3.109	3.639	3.237	2.988	2.965	3.057
HAR	2.295	2.271	2.837	2.405	2.181	2.213	2.307

Note: Values are scaled by 10^{-3}; highlighted cells belong to 5% MCS.

Table A.3. GMVP: portfolio of 15 stocks

	MRK	RCOV		Subsampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
Cumulative							
DCC	20.72	**20.72**	**20.72**	**20.72**	**20.72**	20.72	20.72
RiskMetrics	56.67	56.67	56.67	56.67	56.67	56.67	56.67
VARFIMA	21.34	25.63	22.71	23.71	21.91	20.62	19.93
GHAR	**20.37**	24.46	21.75	22.59	20.90	**19.66**	**18.97**
HAR	22.25	26.21	23.52	24.42	22.69	21.47	20.83
Annualized							
DCC	12.64	**12.64**	**12.64**	**12.64**	12.64	12.64	12.64
RiskMetrics	32.19	32.19	32.19	32.19	32.19	32.19	32.19
VARFIMA	12.88	14.80	13.52	13.99	13.06	12.40	12.07
GHAR	**11.64**	13.82	12.39	12.86	**11.91**	**11.22**	**10.83**
HAR	13.43	15.21	14.06	14.45	13.56	12.92	12.62

Note: The model with the overall best performance is highlighted; for the given frequency the model with the lowest risk is presented in bold; values represent percentage level of risk.

Table A.4. RMSE: portfolio of 15 stocks

	MRK	RCOV		Subsampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	5.323	5.601	6.064	5.793	5.158	5.023	5.058
RiskMetrics	11.905	11.881	12.030	11.902	11.952	12.044	12.030
VARFIMA	4.555	4.809	5.207	4.900	4.374	4.276	4.323
GHAR	5.881	5.352	6.342	5.918	5.565	5.521	5.677
HAR	4.285	3.599	4.832	4.226	3.948	4.005	4.150

Note: Values are scaled by 10^{-3}; highlighted cells belong to 5% MCS.

APPENDIX B: FIVE-STEP-AHEAD FORECASTS

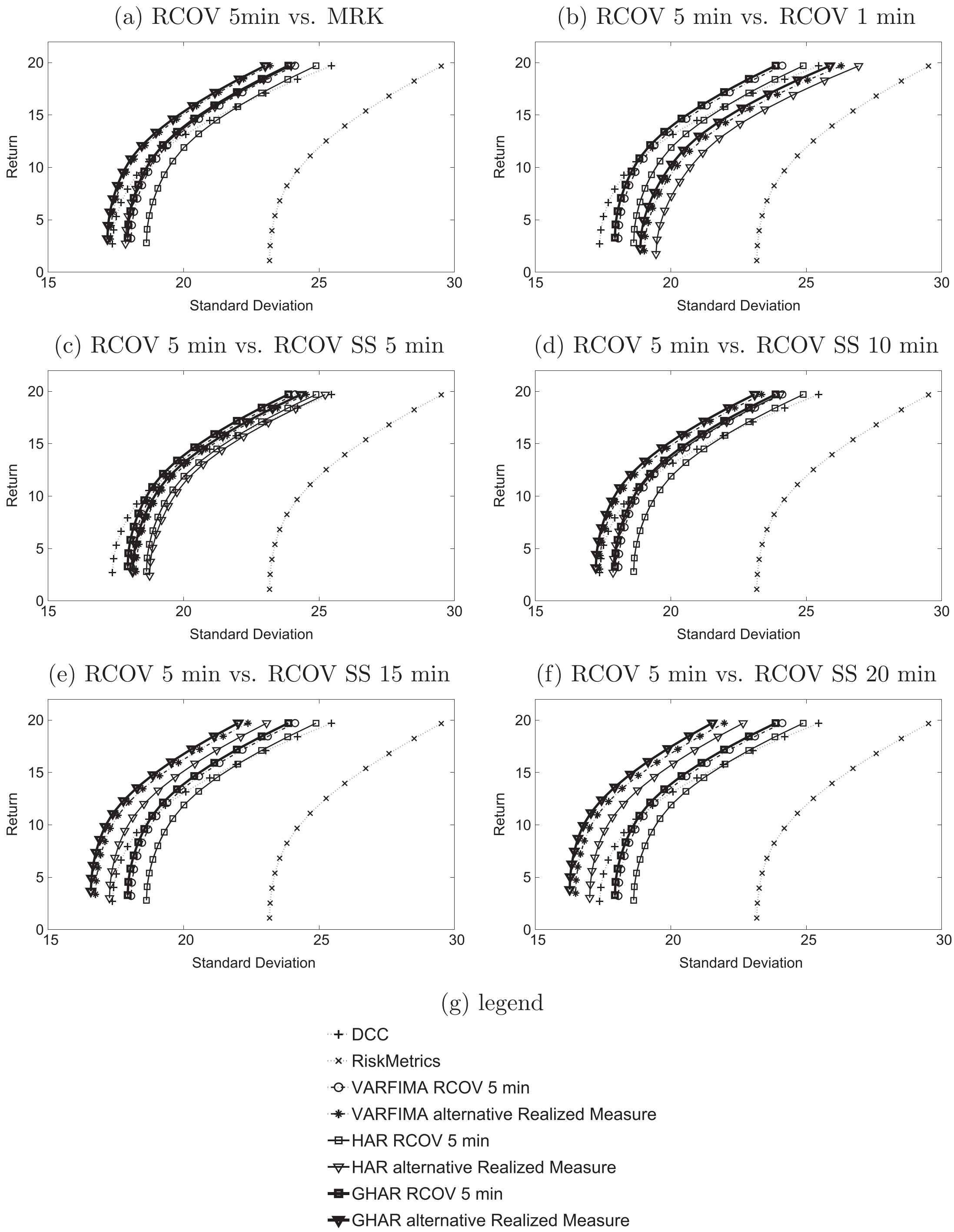

Figure B.1. Efficient frontiers: portfolio of five stocks: (a) RCOV 5-minute vs. MRK; (b) RCOV 5-minute vs. RCOV 1-minute; (c) RCOV 5-minute vs. RCOV SS 5-minute; (d) RCOV 5-minute vs. RCOV SS 10-minute; (e) RCOV 5-minute vs. RCOV SS 15-minute; (f) RCOV 5-minute vs. RCOV SS 20-minute

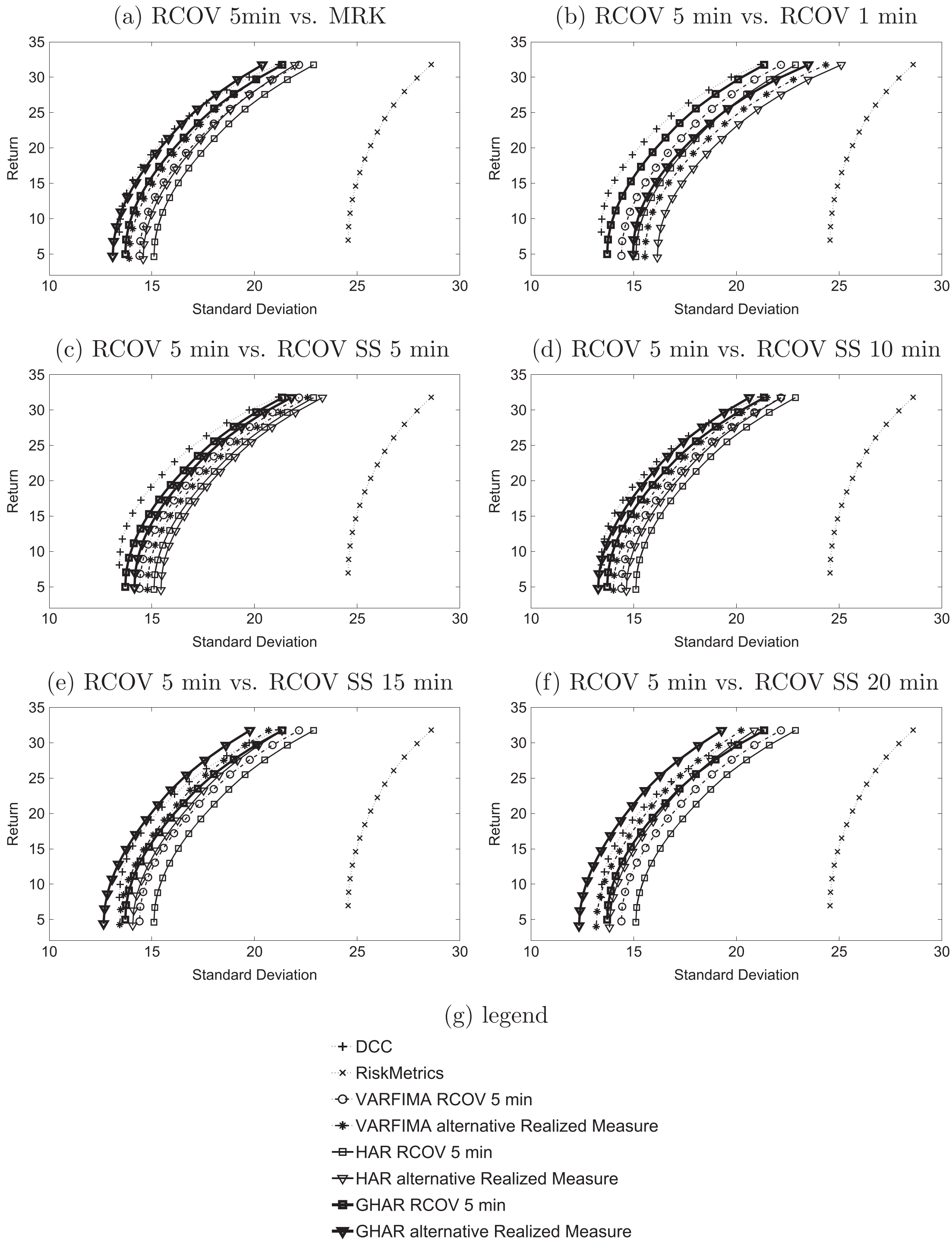

Figure B.2. Efficient frontiers: portfolio of 10 stocks: (a) RCOV 5-minute vs. MRK; (b) RCOV 5-minute vs. RCOV 1-minute; (c) RCOV 5-minute vs. RCOV SS 5-minute; (d) RCOV 5-minute vs. RCOV SS 10-minute; (e) RCOV 5-minute vs. RCOV SS 15-minute; (f) RCOV 5-minute vs. RCOV SS 20-minute

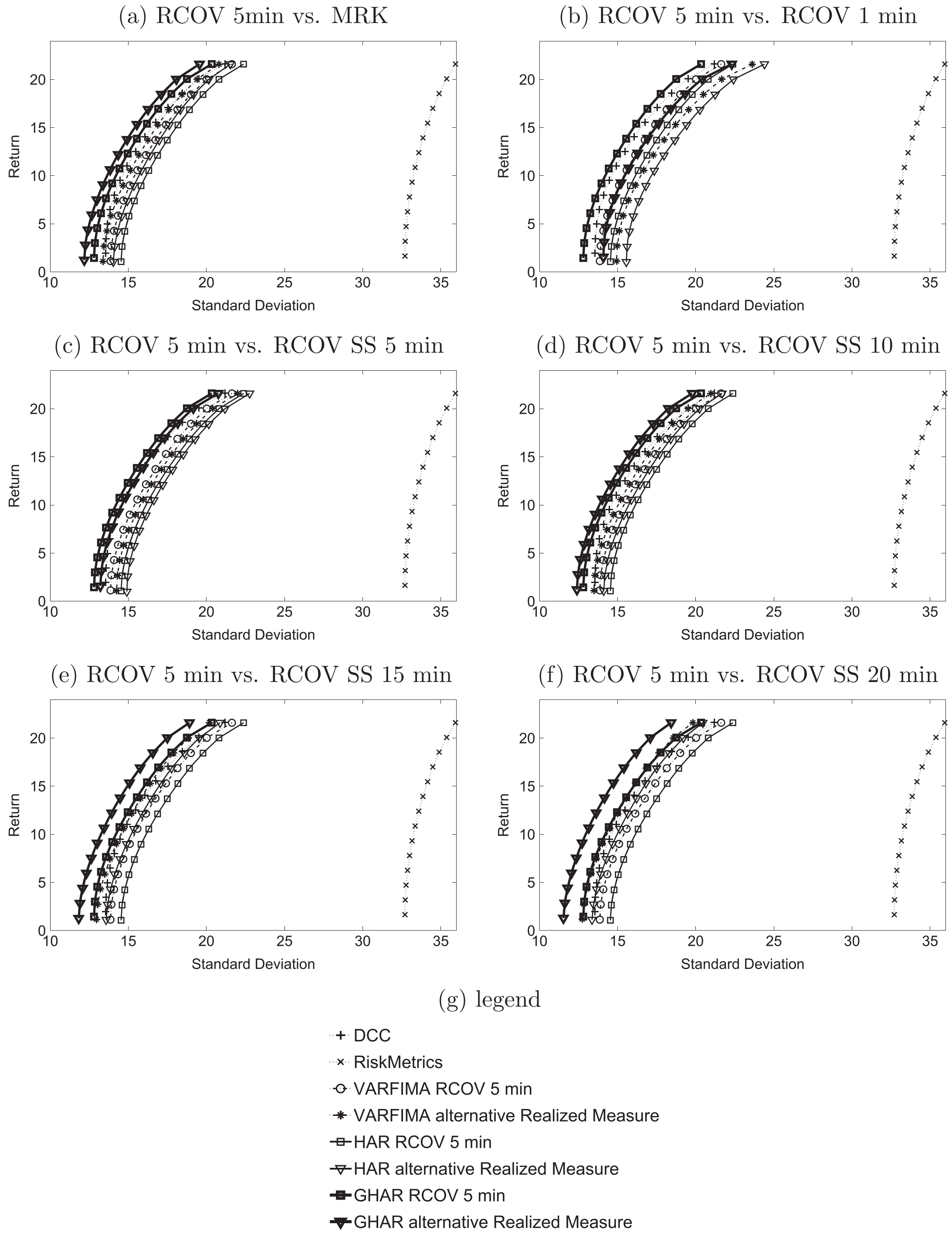

Figure B.3. Efficient frontiers: portfolio of 15 stocks: (a) RCOV 5-minute vs. MRK; (b) RCOV 5-minute vs. RCOV 1-minute; (c) RCOV 5-minute vs. RCOV SS 5-minute; (d) RCOV 5-minute vs. RCOV SS 10-minute; (e) RCOV 5-minute vs. RCOV SS 15-minute; (f) RCOV 5-minute vs. RCOV SS 20-minute

Table B.1. GMVP: portfolio of five stocks

	MRK	RCOV		Subsampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
Cumulative							
DCC	30.50	**30.50**	**30.50**	**30.50**	**30.50**	30.50	30.50
RiskMetrics	40.61	40.61	40.61	40.61	40.61	40.61	40.61
VARFIMA	30.53	34.06	32.09	32.49	30.78	29.64	29.10
GHAR	**30.49**	33.88	32.07	32.36	30.72	**29.54**	**28.96**
HAR	31.30	34.62	32.86	33.19	31.47	30.38	29.87
Annualized							
DCC	17.38	**17.38**	**17.38**	**17.38**	17.38	17.38	17.38
RiskMetrics	23.17	23.17	23.17	23.17	23.17	23.17	23.17
VARFIMA	17.28	19.02	18.06	18.24	17.35	16.73	16.73
GHAR	**17.18**	18.87	17.93	18.12	**17.23**	**16.57**	**16.57**
HAR	17.85	19.45	18.63	18.75	17.86	17.25	17.25

Note: The model with the overall best performance is highlighted; for the given frequency the model with the lowest risk is presented in bold; values represent percentage level of risk; values are scaled by forecasting horizon.

Table B.2. RMSE: portfolio of five stocks

	MRK	RCOV		Subsampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	1.193	1.293	1.376	1.288	1.152	1.081	1.079
RiskMetrics	1.296	1.317	1.330	1.314	1.290	1.288	1.285
VARFIMA	1.043	1.023	1.153	1.055	0.993	0.968	0.978
GHAR	1.261	1.195	1.382	1.273	1.206	1.174	1.189
HAR	1.024	0.980	1.100	1.028	0.968	0.951	0.966

Note: Values are scaled by 10^{-3} and by forecasting horizon; highlighted cells belong to 5% MCS.

Table B.3. GMVP: portfolio of 10 stocks

	MRK	RCOV		Subsampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
Cumulative							
DCC	**22.10**	**22.10**	**22.10**	**22.10**	**22.10**	22.10	22.10
RiskMetrics	42.12	42.12	42.12	42.12	42.12	42.12	42.12
VARFIMA	23.11	27.25	24.44	25.27	23.55	22.30	21.65
GHAR	22.33	26.45	23.72	24.59	22.80	**21.50**	**20.82**
HAR	24.25	28.14	25.56	26.30	24.57	23.35	22.75
Annualized							
DCC	13.07	**13.07**	**13.07**	**13.07**	13.07	13.07	13.07
RiskMetrics	24.36	24.36	24.36	24.36	24.36	24.36	24.36
VARFIMA	13.38	15.36	14.03	14.44	13.54	12.88	12.54
GHAR	**12.67**	14.80	13.38	13.87	**12.88**	**12.16**	**11.78**
HAR	14.15	15.99	14.81	15.15	14.24	13.59	13.28

Note: The model with the overall best performance is highlighted; for the given frequency the model with the lowest risk is presented in bold; values represent percentage level of risk; values are scaled by forecasting horizon.

Table B.4. RMSE: portfolio of 10 stocks

	MRK	RCOV		Subsampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	2.487	2.683	2.773	2.690	2.402	2.290	2.309
RiskMetrics	3.232	3.250	3.217	3.222	3.242	3.278	3.267
VARFIMA	1.952	1.966	2.166	2.024	1.867	1.833	1.872
GHAR	2.598	2.480	2.759	2.611	2.481	2.445	2.501
HAR	1.950	1.881	2.103	1.984	1.845	1.826	1.877

Note: Values are scaled by 10^{-3} and by forecasting horizon; highlighted cells belong to 5% MCS.

Table B.5. GMVP: portfolio of 15 stocks

	MRK	RCOV		Subsampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
Cumulative							
DCC	20.70	**20.70**	**20.70**	**20.70**	**20.70**	20.70	20.70
RiskMetrics	56.64	56.64	56.64	56.64	56.64	56.64	56.64
VARFIMA	21.23	25.28	22.52	23.44	21.75	20.52	19.86
GHAR	**20.31**	24.30	21.65	22.45	20.83	**19.62**	**18.92**
HAR	22.31	26.13	23.51	24.40	22.72	21.51	20.89
Annualized							
DCC	12.60	**12.60**	12.60	**12.60**	12.60	12.60	12.60
RiskMetrics	32.25	32.25	32.25	32.25	32.25	32.25	32.25
VARFIMA	12.53	14.43	13.17	13.60	12.72	12.07	11.74
GHAR	**11.53**	13.66	**12.26**	12.70	**11.79**	**11.12**	**10.73**
HAR	13.29	15.07	13.94	14.31	13.42	12.78	12.48

Note: The model with the overall best performance is highlighted; for the given frequency the model with the lowest risk is presented in bold; values represent percentage level of risk; values are scaled by forecasting horizon.

Table B.6. RMSE: portfolio of 15 stocks

	MRK	RCOV		Subsampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	4.110	4.251	4.384	4.329	3.992	3.919	3.949
RiskMetrics	11.404	11.318	11.262	11.260	11.487	11.599	11.573
VARFIMA	3.453	3.223	3.596	3.422	3.239	3.201	3.283
GHAR	4.913	4.490	4.961	4.821	4.644	4.590	4.706
HAR	3.575	3.216	3.644	3.489	3.331	3.314	3.421

Note: Values are scaled by 10^{-3} and by forecasting horizon; highlighted cells belong to 5% MCS.

APPENDIX C: TEN-STEP-AHEAD FORECASTS

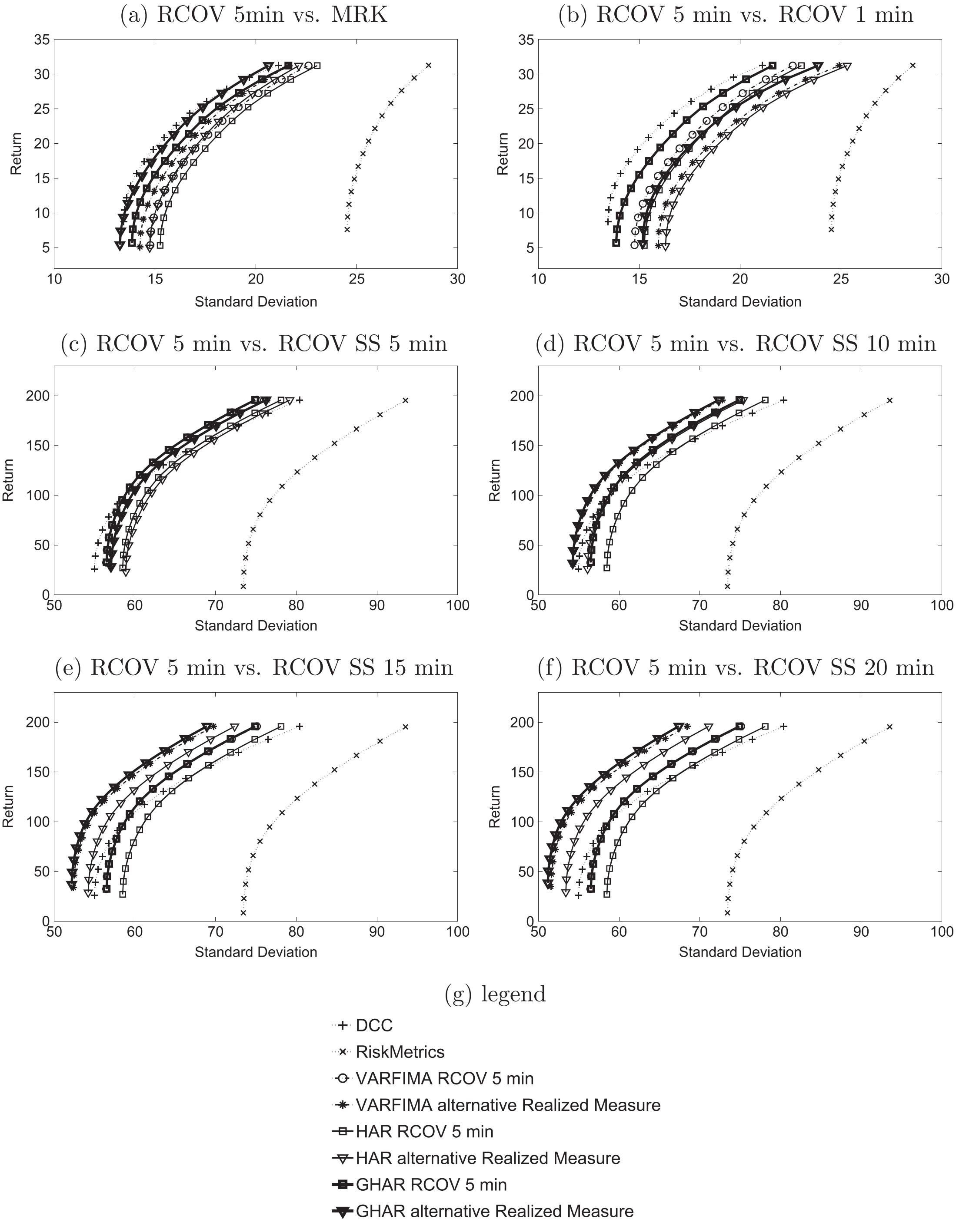

Figure C.1. Efficient frontiers: portfolio of five stocks: (a) RCOV 5-minute vs. MRK; (b) RCOV 5-minute vs. RCOV 1-minute; (c) RCOV 5-minute vs. RCOV SS 5-minute; (d) RCOV 5-minute vs. RCOV SS 10-minute; (e) RCOV 5-minute vs. RCOV SS 15-minute; (f) RCOV 5-minute vs. RCOV SS 20-minute

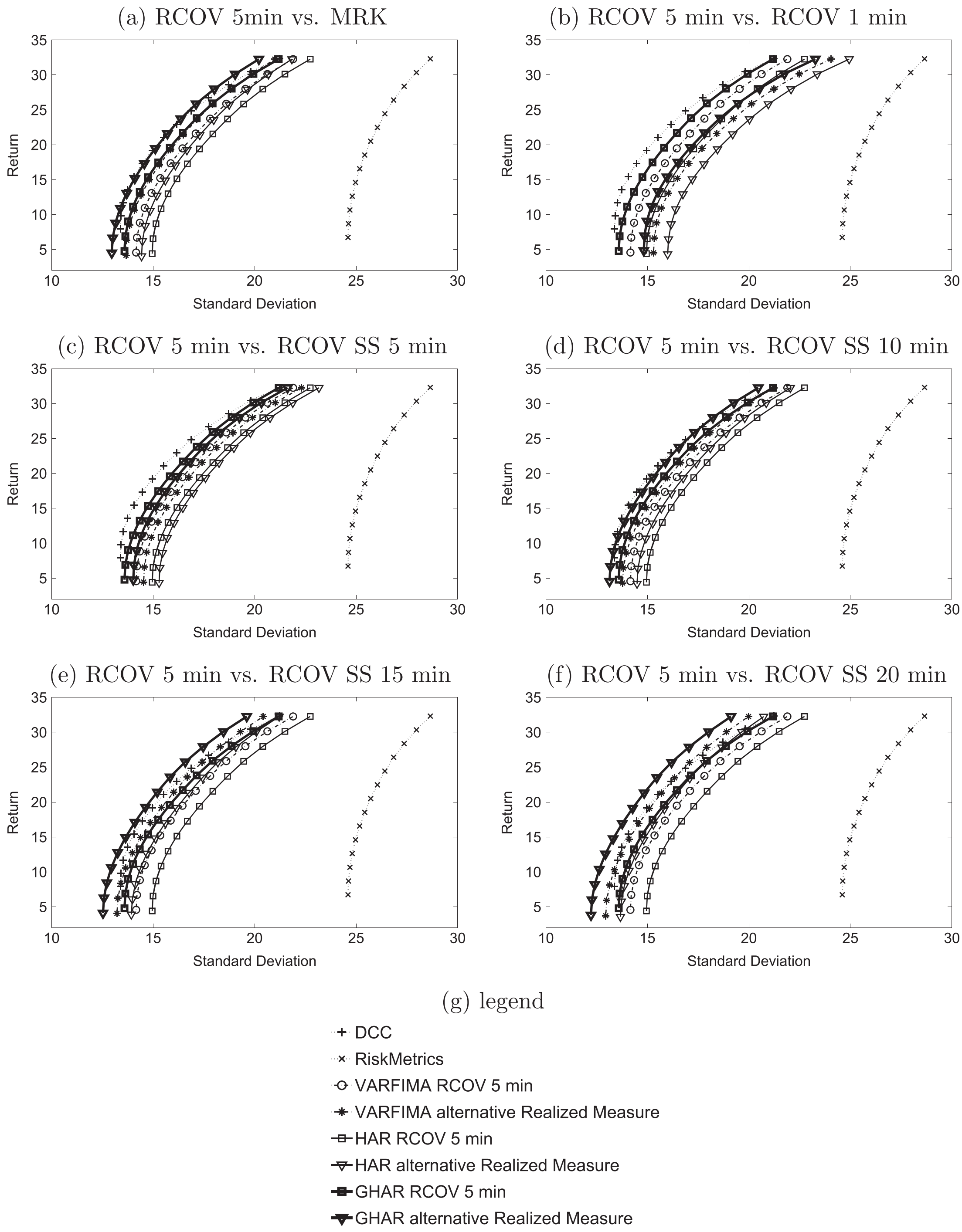

Figure C.2. Efficient frontiers: portfolio of 10 stocks: (a) RCOV 5-minute vs. MRK; (b) RCOV 5-minute vs. RCOV 1-minute; (c) RCOV 5-minute vs. RCOV SS 5-minute; (d) RCOV 5-minute vs. RCOV SS 10-minute; (e) RCOV 5-minute vs. RCOV SS 15-minute; (f) RCOV 5-minute vs. RCOV SS 20-minute

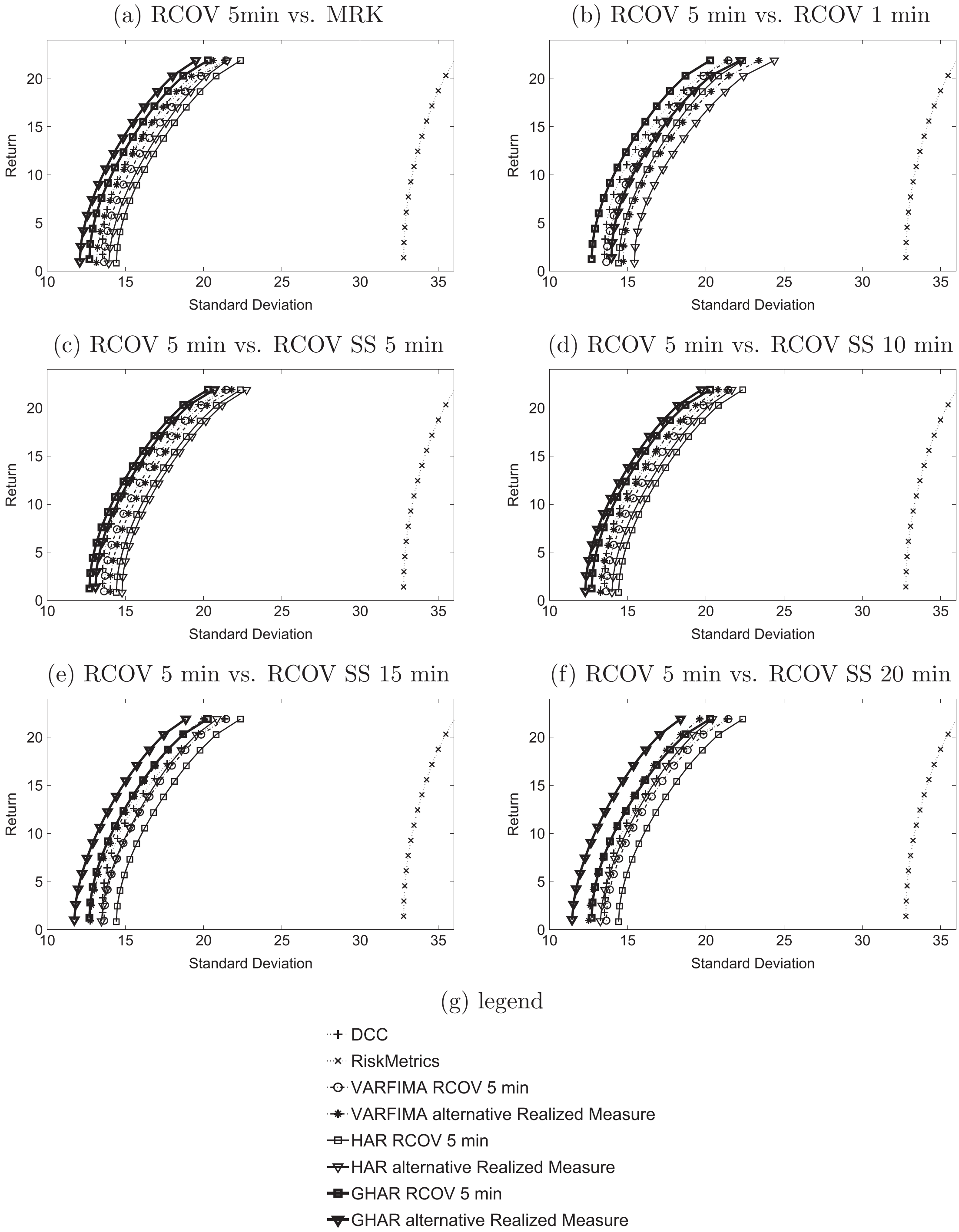

Figure C.3. Efficient frontiers: portfolio of 15 stocks: (a) RCOV 5-minute vs. MRK; (b) RCOV 5-minute vs. RCOV 1-minute; (c) RCOV 5-min vs. RCOV SS 5-minute; (d) RCOV 5-minute vs. RCOV SS 10-minute; (e) RCOV 5-minute vs. RCOV SS 15-minute; (f) RCOV 5-minute vs. RCOV SS 20-minute

Table C.1. GMVP: portfolio of five stocks

	MRK	RCOV		Subsampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
Cumulative							
DCC	30.50	**30.50**	**30.50**	**30.50**	**30.50**	30.50	30.50
RiskMetrics	40.58	40.58	40.58	40.58	40.58	40.58	40.58
VARFIMA	30.30	33.75	31.80	32.20	30.53	29.41	28.88
GHAR	**30.35**	33.66	31.94	32.21	30.58	**29.40**	**28.81**
HAR	31.16	34.40	32.68	33.01	31.30	30.24	29.74
Annualized							
DCC	17.39	**17.39**	**17.39**	**17.39**	17.39	17.39	17.39
RiskMetrics	23.22	23.22	23.22	23.22	23.22	23.22	23.22
VARFIMA	17.07	18.82	17.84	18.02	17.15	16.54	16.26
GHAR	**17.11**	18.74	17.86	18.04	**17.15**	**16.49**	**16.16**
HAR	17.72	19.32	18.49	18.62	17.73	17.13	16.86

Note: The model with the overall best performance is highlighted; for the given frequency the model with the lowest risk is presented in bold; values represent percentage level of risk; values are scaled by forecasting horizon.

Table C.2. RMSE: portfolio of five stocks

	MRK	RCOV		Subsampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	1.208	1.294	1.375	1.291	1.173	1.107	1.101
RiskMetrics	1.389	1.401	1.431	1.404	1.388	1.384	1.380
VARFIMA	1.153	1.147	1.266	1.173	1.106	1.072	1.078
GHAR	1.287	1.256	1.409	1.307	1.237	1.197	1.205
HAR	1.138	1.133	1.242	1.163	1.091	1.058	1.067

Note: Values are scaled by 10^{-3} and by forecasting horizon; highlighted cells belong to 5% MCS.

Table C.3. GMVP: portfolio of 10 stocks

	MRK	RCOV		Subsampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
Cumulative							
DCC	**22.05**	**22.05**	**22.05**	**22.05**	**22.05**	22.05	22.05
RiskMetrics	42.08	42.08	42.08	42.08	42.08	42.08	42.08
VARFIMA	22.89	26.91	24.17	24.97	23.31	22.09	21.45
GHAR	22.16	26.23	23.55	24.40	22.61	**21.33**	**20.66**
HAR	24.15	27.94	25.42	26.14	24.45	23.25	22.66
Annualized							
DCC	13.03	**13.03**	**13.03**	**13.03**	13.03	13.03	13.03
RiskMetrics	24.40	24.40	24.40	24.40	24.40	24.40	24.40
VARFIMA	13.16	15.13	13.80	14.20	13.32	12.67	12.33
GHAR	**12.56**	14.67	13.28	13.75	**12.76**	**12.06**	**11.69**
HAR	14.01	15.85	14.67	15.01	14.10	13.45	13.14

Note: The model with the overall best performance is highlighted; for the given frequency the model with the lowest risk is presented in bold; values represent percentage level of risk; values are scaled by forecasting horizon.

Table C.4. RMSE: portfolio of 10 stocks

	MRK	RCOV		Subsampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	2.437	2.609	2.687	2.610	2.362	2.260	2.271
RiskMetrics	3.445	3.461	3.455	3.448	3.458	3.487	3.481
VARFIMA	2.139	2.165	2.327	2.208	2.057	2.011	2.041
GHAR	2.605	2.514	2.729	2.607	2.494	2.449	2.491
HAR	2.114	2.110	2.276	2.174	2.026	1.986	2.024

Note: Values are scaled by 10^{-3} and by forecasting horizon; highlighted cells belong to 5% MCS.

Table C.5. GMVP: portfolio of 15 stocks

	MRK	RCOV		Subsampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
Cumulative							
DCC	20.67	**20.67**	**20.67**	**20.67**	**20.67**	20.67	20.67
RiskMetrics	56.60	56.60	56.60	56.60	56.60	56.60	56.60
VARFIMA	21.08	25.00	22.33	23.21	21.56	20.36	19.72
GHAR	**20.21**	24.13	21.54	22.30	20.72	**19.53**	**18.83**
HAR	22.31	26.00	23.46	24.32	22.68	21.49	20.88
Annualized							
DCC	12.56	**12.56**	12.56	**12.56**	12.56	12.56	12.56
RiskMetrics	32.32	32.32	32.32	32.32	32.32	32.32	32.32
VARFIMA	12.32	14.21	12.95	13.38	12.50	11.86	11.53
GHAR	**11.44**	13.55	**12.16**	12.59	**11.70**	**11.04**	**10.66**
HAR	13.19	14.95	13.82	14.19	13.31	12.68	12.37

Note: The model with the overall best performance is highlighted; for the given frequency the model with the lowest risk is presented in bold; values represent percentage level of risk; values are scaled by forecasting horizon.

Table C.6. RMSE: portfolio of 15 stocks

	MRK	RCOV		Subsampled RCOV			
	1 min	1 min	5 min	5 min	10 min	15 min	20 min
DCC	4.141	4.258	4.385	4.323	4.054	3.989	4.010
RiskMetrics	11.806	11.735	11.720	11.719	11.884	11.981	11.961
VARFIMA	3.690	3.542	3.821	3.680	3.496	3.439	3.509
GHAR	4.807	4.514	4.859	4.746	4.571	4.508	4.613
HAR	3.666	3.471	3.767	3.635	3.468	3.424	3.512

Note: Values are scaled by 10^{-3} and by forecasting horizon; highlighted cells belong to 5% MCS.

Table D.1. Descriptive statistics of returns over the period 1 July 2005 to 3 January 2012 APPENDIX D

	AAPL	CVX	GE	GOOG	IBM	JNJ	JPM	KO	MSFT	PFE	PG	T	WFC	WMT	XOM
1 min															
Mean	−0.105	0.086	−0.218	−0.204	0.215	−0.035	−0.036	0.011	−0.040	−0.150	0.133	−0.079	−0.049	−0.018	0.099
Max.	0.046	0.043	0.050	0.022	0.030	0.044	0.032	0.020	0.019	0.032	0.032	0.039	0.042	0.030	0.040
Min.	−0.037	−0.027	−0.032	−0.041	−0.020	−0.033	−0.060	−0.038	−0.025	−0.028	−0.028	−0.038	−0.049	−0.021	−0.034
SD	1.046	0.891	1.088	0.936	0.749	0.559	1.336	0.619	0.833	0.825	0.616	0.834	1.464	0.685	0.827
Skewness	0.037	0.262	0.181	−0.369	0.077	0.394	−0.144	−0.482	−0.063	0.103	−0.066	−0.122	0.068	0.365	−0.188
Kurtosis	37.100	44.353	43.467	34.940	40.526	126.707	41.663	70.490	18.848	30.256	71.747	52.025	40.587	39.587	50.690
5 min															
Mean	−0.629	0.423	−0.994	−1.037	1.177	−0.132	−0.165	0.113	−0.138	−0.745	0.741	−0.353	−0.252	−0.057	0.526
Max.	0.065	0.061	0.052	0.046	0.053	0.032	0.069	0.028	0.030	0.030	0.050	0.034	0.066	0.048	0.053
Min.	−0.048	−0.068	−0.046	−0.069	−0.036	−0.040	−0.068	−0.037	−0.028	−0.038	−0.062	−0.073	−0.077	−0.042	−0.059
SD	2.258	1.916	2.280	2.023	1.580	1.174	2.871	1.297	1.756	1.694	1.316	1.779	3.151	1.478	1.779
Skewness	−0.008	−0.062	0.268	−0.509	0.121	−0.127	0.059	−0.300	−0.091	0.109	−0.460	−0.518	0.095	0.432	−0.098
Kurtosis	28.079	37.935	31.858	39.555	37.413	44.133	35.720	33.284	17.395	18.627	86.196	46.840	34.793	42.507	40.512
10 min															
Mean	−0.960	1.129	−1.740	−1.820	2.856	0.118	−0.611	0.387	0.350	−1.416	1.832	−0.444	−0.827	0.289	1.410
Max.	0.050	0.039	0.052	0.043	0.029	0.030	0.067	0.031	0.029	0.029	0.024	0.038	0.069	0.053	0.051
Min.	−0.079	−0.034	−0.058	−0.073	−0.036	−0.025	−0.102	−0.040	−0.038	−0.027	−0.031	−0.043	−0.092	−0.035	−0.067
SD	3.150	2.630	3.168	2.776	2.169	1.592	3.937	1.793	2.399	2.297	1.765	2.412	4.372	1.999	2.423
Skewness	−0.301	0.229	0.256	−0.373	−0.161	0.310	0.035	−0.321	−0.018	0.244	−0.042	−0.114	0.082	0.435	−0.098
Kurtosis	24.897	15.055	29.864	25.773	20.025	21.991	31.039	27.228	15.157	12.988	20.916	21.321	30.512	23.805	27.665
15 min															
Mean	−1.415	2.229	−2.493	−2.598	4.830	0.250	−0.769	0.643	0.765	−1.986	2.917	−0.370	−1.003	0.712	3.193
Max.	0.058	0.046	0.071	0.049	0.038	0.025	0.113	0.030	0.032	0.039	0.028	0.046	0.099	0.051	0.047
Min.	−0.053	−0.037	−0.070	−0.068	−0.053	−0.024	−0.086	−0.041	−0.042	−0.029	−0.034	−0.053	−0.075	−0.035	−0.037
SD	3.801	3.186	3.877	3.350	2.630	1.946	4.896	2.175	2.921	2.794	2.138	2.951	5.335	2.445	2.925
Skewness	−0.012	0.237	0.161	−0.242	−0.159	0.332	0.314	−0.319	−0.050	0.264	0.086	−0.052	0.421	0.562	0.263
Kurtosis	16.529	15.547	31.899	22.586	21.422	19.781	35.685	22.721	14.785	13.138	21.422	21.883	30.411	20.882	19.090
20 min															
Mean	−1.950	2.227	−4.353	−3.743	5.912	−0.276	−1.903	0.412	0.445	−2.604	3.494	−0.812	−2.259	0.542	3.371
Max.	0.050	0.059	0.062	0.043	0.036	0.034	0.074	0.035	0.034	0.041	0.026	0.053	0.086	0.053	0.069
Min.	−0.048	−0.037	−0.068	−0.118	−0.040	−0.021	−0.102	−0.040	−0.038	−0.029	−0.029	−0.049	−0.080	−0.024	−0.067
SD	4.245	3.608	4.350	3.775	2.939	2.159	5.381	2.446	3.259	3.131	2.380	3.300	6.064	2.741	3.325
Skewness	−0.075	0.252	0.130	−0.867	−0.034	0.420	−0.051	−0.148	−0.044	0.256	0.087	−0.086	0.161	0.510	0.177
Kurtosis	13.207	14.377	26.975	42.831	17.258	17.410	23.170	19.581	12.493	11.697	15.784	19.932	25.774	17.183	25.829

Note: Means are scaled by 10^5; standard deviations are scaled by 10^3.

Authors' biographies:

František Čech is a Ph.D. candidate at the Institute of Economic Studies, Faculty of Social Sciences, Charles University in Prague. His research focuses on computational finance and financial econometrics.

Jozef Baruník is deputy head of the econometric department at UTIA, the Czech Academy of Sciences, and is Assistant Professor at the Institute of Economic Studies, Faculty of Social Sciences, Charles University in Prague. His research has concentrated on the area of financial markets with focus on financial econometrics, high frequency data and volatility models, heterogeneous agent models, and frequency domain econometrics.

František Čech and **Jozef Baruník**, Institute of Economic Studies, Charles University, Opletalova 26, 110 00, Prague, Czech Republic; Institute of Information Theory and Automation, The Czech Academy of Sciences, Pod Vodarenskou Vezi 4, 182 00, Prague, Czech Republic.

