A Test of Covariance-Matrix Forecasting Methods

Valeriy Zakamulin

Valeriy Zakamulin is a professor of finance at the University of Agder's School of Business and Law in Kristiansand, Norway.
Valeri.Zakamouline@uia.no

The traditional portfolio optimization technique is based on the mean–variance portfolio optimization framework developed by Markowitz [1952]. However, mean–variance portfolio optimization is difficult to implement, due to the challenges associated with forecasting the mean returns. As a result, Markowitz's portfolio optimization technique has never been widely used. Instead, most mutual funds managers focus on uncovering undervalued securities with potentially high expected returns.

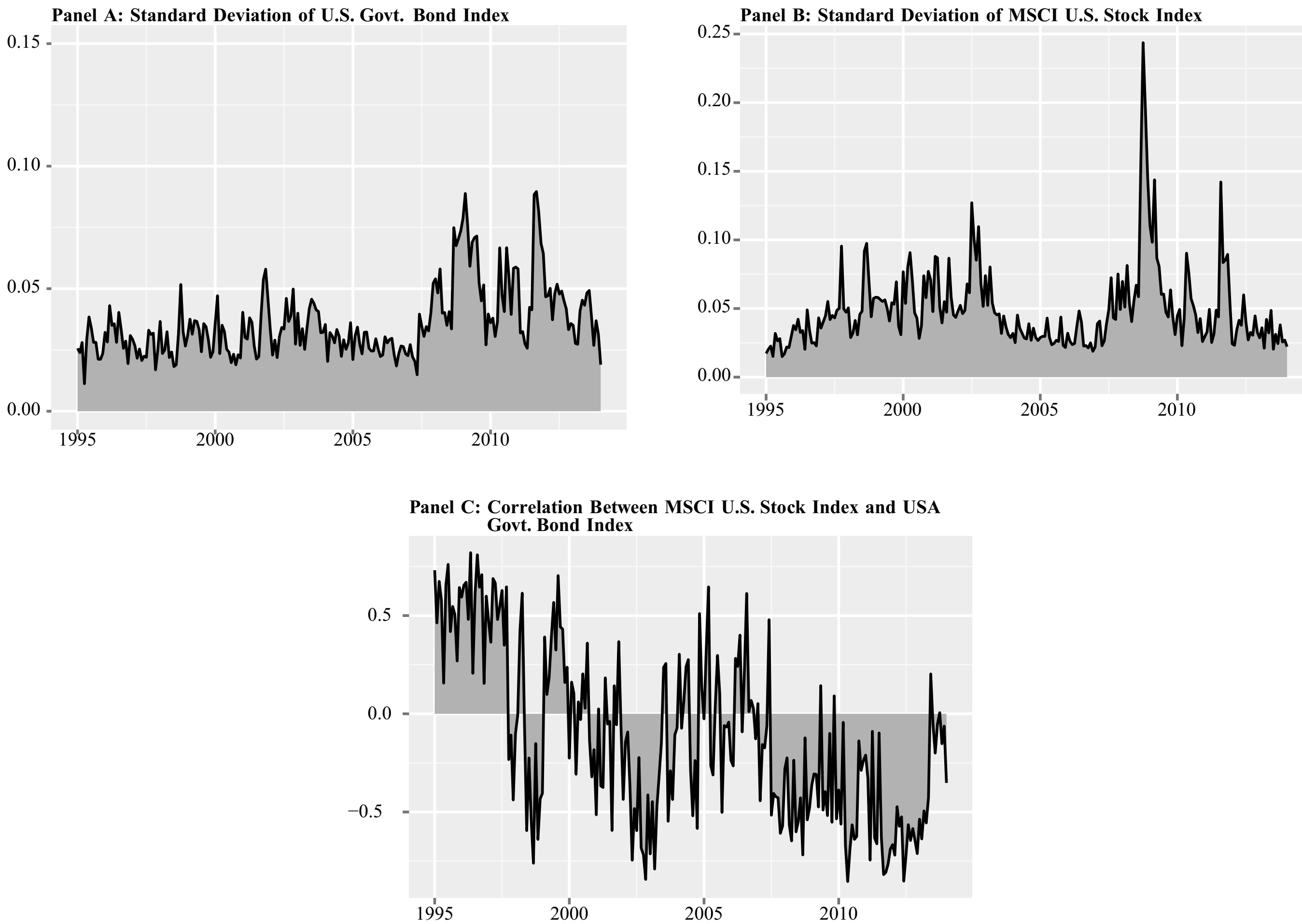
The volatile and turbulent markets of the decade of the 2000s boosted interest in portfolio optimization techniques, with a focus on portfolio risk optimization. These new portfolio optimization methods require only the forecast of the covariance-matrix of returns, without the need to forecast the mean returns. The most popular portfolio risk-optimization techniques are the minimum-variance portfolio (Clarke et al. [2006]; Clarke et al. [2011]), the maximum diversification portfolio (Choueifaty and Coignard [2008]), the risk parity portfolio (Maillard et al. [2010]; Chaves et al. [2011]; Asness et al. [2012]), and a portfolio with volatility targeting (Busse [1999]; Collie et al. [2011]; Butler and Philbrick [2012]; Albeverio et al. [2013]). All of these portfolio risk-optimization techniques were extensively back-tested using historical data and showed superior performance compared to that of mean–variance portfolios and equally weighted portfolios.

Given the increasing popularity of portfolio risk-optimization techniques, in this article we focus on forecasting the covariance-matrix of returns. There is a common consensus that mean returns are notoriously difficult to forecast, whereas the covariance-matrix can be easily forecast using a rolling sample covariance-matrix. A standard approach is to use the monthly rolling n-year covariance-matrix, where n varies from 5 to 20 years, as a forward-looking estimate of the future covariance-matrix (Chan et al. [1999]; DeMiguel et al. [2009]; Duchin and Levy [2009]; Kritzman et al. [2010]). Yet this is a valid approach only if one assumes that the covariance-matrix is either constant over time or varies very slowly over time.

This assumption does not appear to hold for financial asset returns. On the contrary, there is a large strand of the literature that demonstrates that financial asset returns exhibit heteroskedasticity with volatility clustering. The assumption of constant correlation between financial asset returns also appears to be violated. As a motivation, Exhibit 1 plots the monthly realized standard deviations of returns on the U.S. stock market and bond market indices, as well as the monthly realized correlation coefficient between the returns on these indices. The graphs in this exhibit suggest that both the

EXHIBIT 1

Panels A and B Plot the Monthly Realized Standard Deviations of Returns on the U.S. Stock and Bond Market Indices. Panel C Plots the Monthly Realized Correlation Coefficient between the Returns on these Indices. We Compute the Standard Deviations and Correlation Coefficient Using Daily Returns



standard deviation and the correlation coefficient can change dramatically over the course of a few years. For example, there was a ten-fold increase in the standard deviation of stock market returns over the period 2006 to 2008. During the same period, the standard deviation of bond market returns increased by a factor of four, whereas the correlation coefficient experienced a change from virtually zero to a significantly negative value. Therefore, it is only logical to assume that in forecasting the covariance-matrix, one must take into account the time-varying nature of variances and covariances.

The benefits promised by portfolio risk optimization depend crucially on the accuracy of the covariance-matrix forecast. Yet there has been a shortage of studies evaluating the performance of alternative forecasting methods. The aim of this article is to fill this gap in the literature by comparing the performances of different covariance-matrix forecasting methods. We evaluate the alternative forecasting methods on both a statistical and a practical basis.

For this purpose, we perform three studies. In the first study, we directly evaluate the covariance-matrix forecasting methods by comparing their out-of-sample forecast accuracy. In the second study, we follow Chan et al. [1999], Ledoit and Wolf [2003], and Disatnik and Benninga [2007] and evaluate different methods by examining their ability in out-of-sample tests to minimize the variance of portfolios. In the final study, we evaluate the performance of alternative covariance-matrix forecasting methods by their ability in out-of-sample tests to keep the variance of the minimum-variance portfolio at a pre-specified target level. The performance criterion in the first study is the mean squared forecasting error, whereas in the second and the third studies, we use the mean squared tracking error (MSTE) as the performance criterion.

In all of our studies, employing a fixed data window of 10 years that is rolled over time, we generate out-of-sample forecasts of the covariance-matrix using five distinct methods. The first method, which is most common in practice, uses the rolling sample covariance-matrix as a predictor for the future covariance-matrix. In other words, in this approach we estimate the equally weighted covariance-matrix using a 10-year look-back period. To decrease the estimation error of the sample covariance-matrix, in the second method we use the shrinkage estimation of the covariance-matrix proposed by Ledoit and Wolf [2004]. This method applies the concept of shrinking estimation pioneered by Stein [1956]. The third method of estimating the covariance-matrix is popularized by the RiskMetrics group and uses the exponentially weighted covariance-matrix. This method of forecasting is usually called the exponentially weighted moving average (EWMA) model.

The other two methods of covariance-matrix forecasting employ multivariate generalized autoregressive conditional heteroskedasticity (GARCH) models. Univariate GARCH modeling started with the seminal paper by Bollerslev [1986] and proved to be indispensable in modeling and forecasting time-varying financial asset volatility. Unfortunately, a direct extension of univariate GARCH models to multivariate GARCH models (examples are the VEC-GARCH model of Bollerslev et al. [1988] and the BEKK-GARCH model defined in Engle and Kroner [1995]) suffers from the curse of dimensionality and cannot be used to estimate covariance matrices of many financial assets. With a few financial assets, however, Pojarliev and Polasek [2001], [2003] demonstrate that portfolios based on BEKK-GARCH covariance-matrix forecasts outperform portfolios based on sample covariance-matrix forecasts. In our study, we focus on the relatively recent multivariate GARCH models that can be applied to estimate large covariance matrices. These models are the DCC-GARCH model of Engle [2002] and the GO-GARCH model of van der Weide [2002].

To ensure that our findings are not data-set-specific, in our studies, we use nine empirical data sets. To guarantee that our results are not confined to some particular historical episode, we estimate out-of-sample performances of different covariance forecasting models over a rather long period, from January 1995 to December 2013; this time span covers a series of alternating turbulent and calm stock markets.

Our main findings can be summarized as follows. We find that shrinkage of the sample covariance-matrix reduces neither the forecasting error nor the tracking error of minimum-variance portfolios. In contrast, multivariate GARCH models provide superior covariance-matrix forecasts compared to the forecast based on the rolling sample covariance-matrix. Specifically, switching from the sample covariance-matrix forecast to a multivariate GARCH forecast allows one to reduce the forecasting error and portfolio tracking error by more than 50%. We also find that the simple EWMA covariance-matrix forecast performs only slightly worse than the multivariate GARCH forecast.

The rest of the article is organized as follows. The second section describes our data, the construction of minimum-variance portfolios, and alternative covariance-matrix forecasting methods. The third section presents the results of our empirical studies, and the final section draws conclusions.

DATA AND METHODOLOGY

Data

In our study, we use nine empirical data sets, listed in Exhibit 2. These data sets are similar to the data sets used in studies by DeMiguel et al. [2009] and Kritzman et al. [2010]. The first eight data sets come from the data library of Kenneth French.[1] All of these data sets represent value-weighted portfolios formed using different criteria. For example, the data set of portfolios formed on size consists of returns on 10 stock portfolios sorted by market equity. The portfolios are constructed at the end of each June using the June market equity and NYSE breakpoints. As another example, the data set of industry portfolios consists of returns on 10 industry portfolios in the U.S. The 10 industries considered are consumer non-durables, consumer durables, manufacturing, energy, high-tech, telecommunication, wholesale and retail, health, utilities, and others. The ninth data set covers the seven major

Exhibit 2

List of Data Sets Considered. *N* Denotes the Number of Portfolios in Each Data Set

#	Dataset	N
1	Portfolios Formed on Size	10
2	Portfolios Formed on Book-to-Market	10
3	Portfolios Formed on Momentum	10
4	Portfolios Formed on Size and Book-to-Market	6
5	Industry Portfolios	10
6	Portfolios Formed on Size and Momentum	10
7	Portfolios Formed on Size and Short-Term Reversal	10
8	Portfolios Formed on Size and Long-Term Reversal	10
9	Core Asset Classes	7

asset classes and is obtained from the Thomson Reuters Datastream database. The asset classes in this data set are listed in Exhibit 3. The data for all nine data sets come at a daily frequency and cover the period from January 1, 1986 to December 31, 2013. In addition, we use the 90-day nominal U.S. T-bill rate as a proxy for the risk-free rate of return. The daily T-bill rate is also obtained from the data library of Kenneth French.

Minimum-Variance Portfolios

We assume that the investment universe consists of n risky assets and one risk-free asset. We denote by r_m the $n \times 1$ vector of asset returns for month m and by w_m the $n \times 1$ vector of portfolio weights for month m such that the risky portfolio return and volatility for month m are given by

$$r_{Pm} = \mathrm{w}'_m \mathrm{r}_m, \quad \sigma_{Pm} = \sqrt{\mathrm{w}'_m \Sigma_m \mathrm{w}_m}$$

where Σ_m denotes the $n \times n$ covariance-matrix of asset returns for month m.

The risky portfolio in our study is the global minimum-variance portfolio. We denote by $\hat{\Sigma}_m$ the covariance-matrix forecast for month m made at the end of month m–1. We construct the global minimum-variance portfolio by solving numerically the following quadratic optimization problem with one equality and n inequality constraints:

$$\min_{\mathrm{w}_m} \ \frac{1}{2}\mathrm{w}'_m \hat{\Sigma}_m \mathrm{w}_m, \quad \text{subject to} \quad \mathrm{w}'_m 1 = 1 \quad \text{and} \quad \mathrm{w}_m \geq 0$$

Exhibit 3

List of Core Asset Classes and Their Mnemonics in the Thomson Reuters Datastream Database

#	Asset Class	Mnemonic
1	U.S. Benchmark 30 Year Govt. Bond Index	BMUS30Y
2	MSCI EUROPE Stock Index	MSEROP
3	MSCI Emerging Markets Stock Index	MSEMKF
4	MSCI PACIFIC Stock Index	MSPACF
5	MSCI U.S. Stock Index	MSUSAML
6	S&P U.S. REIT Index	SBBRUSL
7	S&P Goldman Sachs Commodity Index	GSCITOT

where 1 and 0 denote the $n \times 1$ vectors of ones and zeros respectively. Note that the constraint $\mathrm{w}_m \geq 0$ implies that the global minimum-variance portfolio is a long-only portfolio, that is, short sales are prohibited.

The capital allocation consists of investing proportion a_m in the risky portfolio and consequently $1 - a_m$ in the risk-free asset, which provides a return of r_{fm} for month m. Thus, the return on the complete portfolio for month m is given by

$$r_{Cm} = a_m \, r_{Pm} + (1 - a_m) r_{fm}$$

We consider two distinct cases. In the first case, $a_m = 1$, that is, all money is allocated to the risky portfolio only. In the second case, the value of a_m is chosen to keep the volatility of the complete portfolio at a target level denoted by σ^{target}:

$$a_m = \min\left(\frac{\sigma^{\text{target}}}{\hat{\sigma}_{Pm}}, a^{\max}\right)$$

where $\hat{\sigma}_{Pm}$ is the forecast of the portfolio standard deviation for month m and $a^{\max}$ is the maximum acceptable exposure to risk.

Covariance-Matrix Forecasting Methods

Our data comes at a daily frequency, and we assume that the vector of daily asset returns is given by

$$\mathrm{r}_t = \mu + \varepsilon_t$$

where μ is the $n \times 1$ vector of asset mean returns, and ε_t is the $n \times 1$ vector of random disturbances on day t such that

$E[\varepsilon_t] = 0$. The covariance-matrix of asset returns on day t is the $n \times n$ matrix $\Sigma_t = \varepsilon_t \varepsilon_t'$. To compute the composition of the global minimum-variance portfolio for month m, the covariance-matrix of asset returns for month m needs to be forecasted using the information available at the end of month $m - 1$. Below, we review the covariance-matrix forecasting methods employed in this study.

Rolling Historical Covariance

In this method, the covariance-matrix of asset returns on day t (the first day in month m) is forecasted using an equally weighted moving average calculated on a fixed size data window that is rolled over time. Denoting the length of this window by L (also called the length of the look-back period), the covariance-matrix on day t is forecasted using

$$\hat{\Sigma}_t^{\text{Hist}} = \frac{1}{L} \sum_{i=t-L}^{t-1} \varepsilon_i \varepsilon_i'$$

Because our data come at a daily frequency, while the composition of the global minimum-variance portfolio is revised at the month-end, we actually need to forecast the covariance-matrix for the subsequent month. We perform this forecast by multiplying each element of the daily covariance-matrix by the number of days in the subsequent month:

$$\hat{\Sigma}_m^{\text{Hist}} = N_m \hat{\Sigma}_t^{\text{Hist}}$$

where N_m is the number of days in month m.

Rolling Historical Covariance With Shrinkage

The shrinkage estimation of the covariance-matrix is designed to decrease the error in estimation by using an estimator of the form

$$\hat{\Sigma}_t^{\text{Shrink}} = (1 - \delta_t) \hat{\Sigma}_t^{\text{Hist}} + \delta_t \hat{\Sigma}_t^{\text{Target}}$$

where $0 < \delta_t < 1$ is the shrinkage parameter for day t, $\hat{\Sigma}_t^{\text{Hist}}$ is the historical covariance-matrix over a rolling window of length L, and $\hat{\Sigma}_t^{\text{Target}}$ is the shrinkage target, which is also computed over a rolling window of length L. In essence, this estimator shrinks the covariance-matrix toward the shrinkage target. In our study, we use the shrinkage estimator proposed by Ledoit and Wolf [2004], who take the shrinkage target to be the covariance-matrix in the constant correlation model. The covariance-matrix for the subsequent month is forecasted in the same manner as in the previous method:

$$\hat{\Sigma}_m^{\text{Shrink}} = N_m \hat{\Sigma}_m^{\text{Shrink}}$$

Exponentially Weighted Moving Average

This approach to forecasting the covariance-matrix is popularized by the RiskMetrics group. In this approach, the exponentially weighted covariance-matrix is estimated using the following recursive form:

$$\hat{\Sigma}_t^{\text{EWMA}} = (1 - \lambda)\, \varepsilon_{t-1} \varepsilon_{t-1}' + \lambda \hat{\Sigma}_{t-1}^{\text{EWMA}}$$

where $0 < \lambda < 1$ is the decay constant. We follow the recommendations of the RiskMetrics group and estimate the daily EWMA covariance-matrix using $\lambda = 0.97$. We obtain the monthly EWMA covariance-matrix by multiplying each element of the daily EWMA covariance-matrix by the number of days in the subsequent month:

$$\hat{\Sigma}_m^{\text{EWMA}} = N_m \hat{\Sigma}_t^{\text{EWMA}}$$

Dynamic Conditional Correlation GARCH

The dynamic conditional correlation (DCC) GARCH model belongs to the family of multivariate GARCH models. In the DCC-GARCH model, the covariance-matrix is decomposed into the matrix of standard deviations and the correlation matrix as

$$\Sigma_t^{\text{DCC-GARCH}} = D_t R_t D_t$$

where $D_t = \text{diag}(\sigma_{1t}, \ldots, \sigma_{nt})$ is the diagonal matrix containing the standard deviations of asset returns, and $R_t = [\rho_{ijt}]$ is the correlation matrix.

The elements of D_t are modeled as univariate GARCH processes. In our study, we use the DCC-GARCH(1,1). In this method, the (daily) variances of each asset return are modeled as GARCH(1,1), which is given by

$$\sigma_{it}^2 = \alpha_{0i} + \alpha_{1i}\varepsilon_{it-1}^2 + \beta_{1i}\sigma_{it-1}^2 \quad (1)$$

The standardized residuals from the estimation of the GARCH(1,1) models are used to estimate the DCC-GARCH(1,1) correlation specification

$$R_t = Q_t^{*-1}\, Q_t Q_t^{*-1}$$

where

$$Q_t = (1 - a - b)\overline{Q} + a e_{t-1} e'_{t-1} + b Q_{t-1}$$

$e_t = D_t^{-1}\varepsilon_t$ are standardized errors, $\overline{Q} = E[e_t\, e_t']$ is the unconditional covariance-matrix of standardized errors, and Q* is a diagonal matrix with the square root of the diagonal elements of Q_t at the diagonal, $Q^* = \left(\mathrm{diag}(Q_t)\right)^{\frac{1}{2}}$.

We estimate the DCC-GARCH(1,1) model using the daily data and rolling window of L observations. Then we use the estimated parameters to perform the N_m-day ahead forecast of the (daily) covariance matrices. The monthly covariance-matrix is obtained by summing N_m daily covariance matrices:

$$\hat{\Sigma}_m^{\mathrm{DCC-GARCH}} = \sum_{i=0}^{N_m-1} \hat{\Sigma}_{t+i}^{\mathrm{DCC-GARCH}}$$

where $\hat{\Sigma}_{t+i}^{\mathrm{DCC-GARCH}}$ is the covariance-matrix forecast for day $t + i$.

Generalized Orthogonal GARCH

The generalized orthogonal (GO) GARCH model also represents a multivariate generalization of a univariate GARCH model. In the GO-GARCH model, the vector of random disturbances ε_t is modeled as a linear combination of n unobserved independent factors f_t:

$$\varepsilon_t = Z f_t$$

where Z is the $n \times n$ invertible matrix that links the unobserved components with the observed variables. The unobservable components are normalized to have unit variance, such that the unconditional covariance-matrix of ε_t is

$$\overline{\Sigma} = E[\varepsilon_t \varepsilon'_t] = ZZ'$$

The independent factors f_t are modeled as univariate GARCH processes. In our study, we use the GO-GARCH(1,1), where the factors are modeled by a GARCH(1,1) process $H_t = \mathrm{diag}(h_{1t},\ldots,h_{nt})$

$$h_{it}^2 = \alpha_{0i} + \alpha_{1i} f_{it-1}^2 + \beta_{1i} h_{it-1}^2 \quad (2)$$

The conditional covariance-matrix of ε_t is given by

$$\Sigma_t^{\text{GO-GARCH}} = Z H_t Z'$$

Denote by P the $n \times n$ matrix that contains the orthogonal elements of $\overline{\Sigma}$ and by Λ the $n \times n$ diagonal matrix that contains the corresponding eigenvalues, $\Lambda = \mathrm{diag}(\lambda_1,\ldots,\lambda_n)$. These matrices can be estimated because the matrix $\overline{\Sigma}$ can be estimated consistently by $\frac{1}{L}\Sigma_{i=t-L}^{t-1}\,\varepsilon_i\,\varepsilon_t'$. The singular value decomposition of Z yields

$$Z = P\Lambda^{\frac{1}{2}}U'$$

where U is the $n \times n$ orthogonal matrix that can be identified from the structure of the conditional covariance-matrix Σ_t.

Similarly to the DCC-GARCH model, we estimate the GO-GARCH(1,1) model using the daily data and rolling window of L observations. Then the estimated parameters are used to perform the N_m-day ahead forecast of the (daily) covariance matrices. The monthly covariance-matrix is obtained by summing N_m daily covariance matrices:

$$\hat{\Sigma}_m^{\mathrm{GO-GARCH}} = \sum_{i=0}^{N_m-1} \hat{\Sigma}_{t+i}^{\mathrm{GO-GARCH}}$$

EMPIRICAL RESULTS

In this section, we evaluate different covariance-matrix forecasting methods by performing three studies. In the first study, we directly evaluate the covariance-matrix forecasting methods by comparing their forecast accuracy. In the second and the third study, we evaluate the covariance-matrix forecasting methods in practical settings. In particular, we compare the performances of different methods by looking at their ability to optimize the risk of minimum-variance portfolios and minimum-variance portfolios with a volatility target.

In all of our studies, we perform out-of-sample forecasts of the monthly covariance-matrix. That is, the covariance-matrix for month m is forecast based on information available at the end of month $m - 1$. More specifically, our forecasts are based on the rolling-window estimation scheme by using a look-back period of 120 months. The number of days, L, in the look-back period is determined by counting the number of days in the past 120 months. With this choice, our initial in-sample period is from January 1, 1986 to December 31, 1994, and consequently, the performances of different forecasting methods are evaluated over the period from January 1, 1995 to December 31, 2013.

Covariance-Matrix Forecast

In this study, we evaluate the covariance-matrix forecast accuracy provided by each distinct method of forecasting. Let $\Sigma_m = [\sigma_{ij,m}]$ denote the monthly realized covariance-matrix for month m and $\hat{\Sigma}_m = [\hat{\sigma}_{ij,m}]$ denote a forecast of Σ_m made at the end of month $m - 1$. The month m squared forecast error (SFE) for single covariance is given by $(\sigma_{ij,m} - \hat{\sigma}_{ij,m})^2$. To avoid double accounting of covariance forecast errors, we compute the month m total SFE for covariance-matrix Σ_m by summing the SFEs for single covariances in the upper-triangular part of matrix Σ_m, including the main diagonal:

$$SFE_m = \sum_{i=1}^{n} \sum_{j=1}^{i} \left(\sigma_{ij,m} - \hat{\sigma}_{ij,m}\right)^2$$

The mean squared forecast error (MSFE) is computed as

$$MSFE = \frac{1}{M} \sum_{m=1}^{M} SFE_m$$

where M is the number of months in the out-of-sample period.

Exhibit 4 reports the MSFE produced by different covariance-matrix forecasting methods for each data set separately and the average MSFEs over all data sets. Exhibit 5 plots the average of MSFEs over all data sets for different forecasting methods. The results reported in Exhibit 4 demonstrate that the rolling historical covariance method and the rolling historical covariance with shrinkage method provide identical forecast accuracy. That is, similarly to Disatnik and Benninga [2007], we find that shrinkage does not produce a better forecast of the covariance-matrix for the subsequent month than does the rolling historical covariance. In contrast, as compared with the two first methods of forecasting, the remaining three methods reduce the MSFE by approximately 50%. The differences in forecast accuracy between these three methods are marginal. In particular, the methods based on the multivariate GARCH models produce slightly better forecasts than the method that uses EWMA, and the forecast produced by the DCC-GARCH model is marginally better than that of the GO-GARCH model.

The Global Minimum-Variance Portfolio

In this study, we evaluate the performance of alternative covariance-matrix forecasting methods by their ability in out-of-sample tests to minimize the variance of portfolios. We measure performance using the MSTE, which is computed as follows. Denote by $\hat{w}_m$ the ex ante vector of weights of the global minimum-variance portfolio for month m and by w_m the *ex post* vector of weights of the global minimum-variance portfolio for month m. The former is determined on the basis of $\hat{\Sigma}_m$ provided by each distinct method of forecasting at the end of month $m - 1$, and the latter is determined on the basis of Σ_m, which is the realized covariance-matrix for month m. Further denote by σ_{Pm}^{oos} the out-of-sample (i.e., realized) monthly standard deviation of the minimum-variance portfolio and by σ_{Pm}^{is} the in-sample monthly standard deviation of the minimum-variance portfolio. These standard deviations are given by

$$\sigma_{Pm}^{oos} = \sqrt{\hat{w}_m' \Sigma_m \hat{w}_m}, \quad \sigma_{Pm}^{is} = \sqrt{w_m \Sigma_m w_m}$$

The month m squared tracking error (STE) is defined as the squared difference between the in-sample and the out-of-sample standard deviations of the minimum-variance portfolio

$$STE_m = \left(\sigma_{Pm}^{is} - \sigma_{Pm}^{oos}\right)^2$$

The MSTE is computed as

EXHIBIT 4

Forecast Accuracy Over the Period 1995–2013 vs. the Covariance-Matrix Forecasting Method

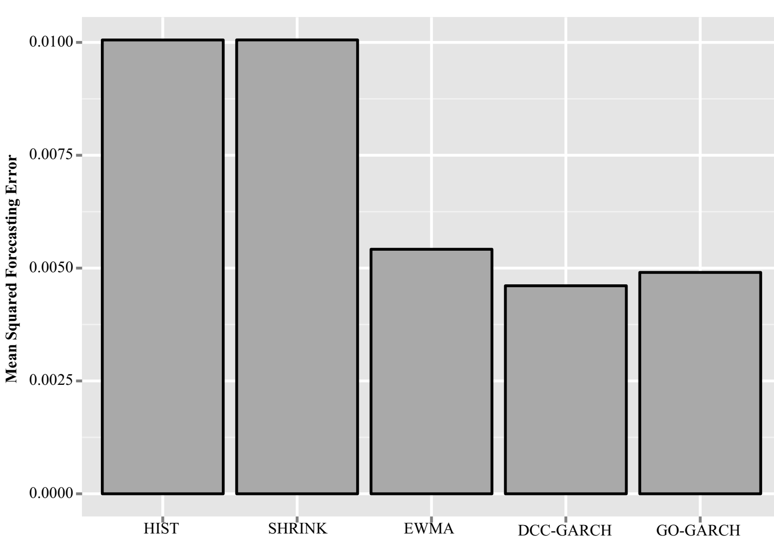
This exhibit reports the MSFE produced by different covariance-matrix forecasting methods. HIST denotes the rolling historical covariance; SHRINK denotes the rolling historical covariance with shrinkage; EWMA denotes the EWMA covariance; DCC-GARCH denotes the DCC-GARCH; GO-GARCH denotes the GO-GARCH. In all methods, the length of the rolling estimation window is 120 months.

Dataset	Covariance-Matrix Forecasting Method				
	HIST	SHRINK	EWMA	DCC-GARCH	GO-GARCH
Portfolios Formed on Size	0.0069	0.0069	0.0040	0.0033	0.0028
Portfolios Formed on Book-to-Market	0.0089	0.0080	0.0042	0.0035	0.0043
Portfolios Formed on Momentum	0.0149	0.0149	0.0069	0.0061	0.0070
Portfolios Formed on Size and Book-to-Market	0.0081	0.0081	0.0042	0.0036	0.0040
Industry Portfolios	0.0081	0.0081	0.0052	0.0045	0.0043
Portfolios Formed on Size and Momentum	0.0149	0.0149	0.0069	0.0061	0.0070
Portfolios Formed on Size and Short-Term Reversal	0.0079	0.0079	0.0048	0.0040	0.0039
Portfolios Formed on Size and Long-Term reversal	0.0153	0.0153	0.0093	0.0081	0.0074
Core Assets	0.0065	0.0065	0.0032	0.0023	0.0032
Averages Over All Datasets	0.0101	0.0101	0.0054	0.0046	0.0049

EXHIBIT 5

Average Forecast Accuracy Over the Period 1995–2013 vs. the Covariance-Matrix Forecasting Method

For each distinct method of forecasting, this exhibit plots the average of the MSFEs over all data sets. HIST denotes the rolling historical covariance; SHRINK denotes the rolling historical covariance with shrinkage; EWMA denotes the EWMA covariance; DCC-GARCH denotes the DCC-GARCH; GO-GARCH denotes the GO-GARCH. In all methods, the length of the rolling estimation window is 120 months.



$$MSTE = \frac{1}{M}\sum_{m=1}^{M} STE_m$$

where M is the number of months in the out-of-sample period.

Exhibit 6 reports the out-of-sample performance of alternative covariance forecasting methods when tracking the variance of the global minimum-variance portfolio over the period 1995 to 2013. Specifically, this exhibit reports the portfolio MSTEs produced by different covariance-matrix forecasting methods. Exhibit 7 plots the average of MSTEs over all data sets. The results reported in Exhibit 6 suggest that both the rolling historical covariance method and the rolling historical covariance with shrinkage method provide virtually identical MSTEs. That is, shrinkage does not reduce the tracking error of the minimum-variance portfolios. As opposed to the two first methods of forecasting, the remaining three methods allow the average MSTE to be reduced by 50% to 65%. The forecast obtained using the DCC-GARCH model produces the least MSTE. Somewhat surprisingly, the use of the EWMA model produces a smaller MSTE than the use of the much more advanced GO-GARCH model,

EXHIBIT 6

The Portfolio MSTE Over the Period 1995–2013 vs. the Covariance-Matrix Forecasting Method

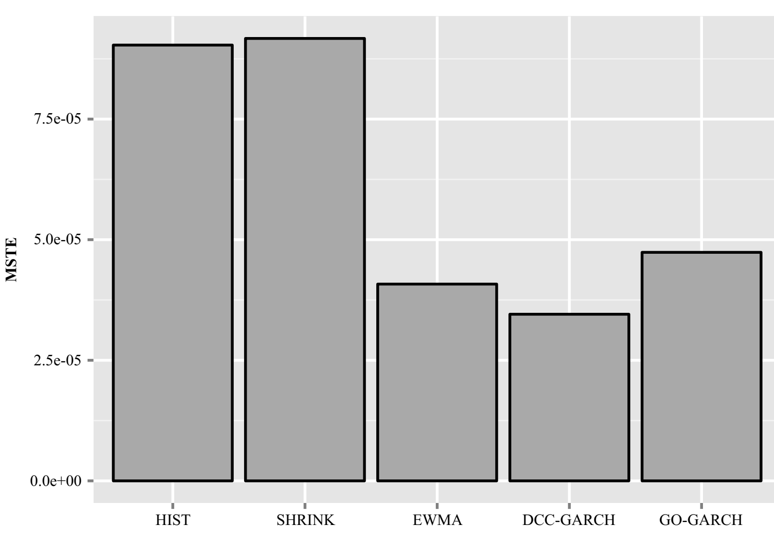
This exhibit reports the out-of-sample performance of alternative covariance forecasting methods when tracking the variance of the global minimum-variance portfolio. HIST denotes the rolling historical covariance, SHRINK denotes the rolling historical covariance with shrinkage, EWMA denotes the EWMA covariance, DCC-GARCH denotes the DCC-GARCH, GO-GARCH denotes the GO-GARCH. In all methods, the length of the rolling estimation window is 120 months.

	Covariance-Matrix Forecasting Method				
Dataset	**HIST**	**SHRINK**	**EWMA**	**DCC-GARCH**	**GO-GARCH**
Portfolios Formed on Size	5.37e-05	7.23e-05	1.22e-05	2.16e-05	1.90e-05
Portfolios Formed on Book-to-Market	1.18e-04	1.15e-04	2.35e-05	2.43e-05	2.80e-05
Portfolios Formed on Momentum	1.24e-04	1.23e-04	4.98e-05	4.23e-05	8.50e-05
Portfolios Formed on Size and Book-to-Market	1.16e-04	1.16e-04	2.46e-05	1.86e-05	1.69e-05
Industry Portfolios	6.27e-05	6.21e-05	2.68e-05	3.24e-05	3.25e-05
Portfolios Formed on Size and Momentum	1.24e-04	1.23e-04	4.98e-05	4.23e-05	8.50e-05
Portfolios Formed on Size and Short-Term Reversal	5.34e-05	5.45e-05	3.08e-05	3.66e-05	5.00e-05
Portfolios Formed on Size and Long-Term Reversal	1.21e-04	1.18e-04	1.28e-04	7.27e-05	8.13e-05
Core Assets	4.07e-05	4.09e-05	2.14e-05	2.01e-05	2.87e-05
Averages Over All Datasets	9.03e-05	9.17e-05	4.08e-05	3.45e-05	4.74e-05

EXHIBIT 7

Average Out-of-Sample Performance of Alternative Covariance Forecasting Methods when Tracking the Variance of the Global Minimum-Variance Portfolio Over the Period 1995–2013

For each distinct method of forecasting, this exhibit plots the average MSTEs over all data sets. HIST denotes the rolling historical covariance, SHRINK denotes the rolling historical covariance with shrinkage, EWMA denotes the EWMA covariance, DCC-GARCH denotes the DCC-GARCH; GO-GARCH denotes the GO-GARCH. In all methods, the length of the rolling estimation window is 120 months.



but the differences between their MSTEs are marginal.

The Global Minimum-Variance Portfolio with a Volatility Target

In the last study, we evaluate the performance of alternative covariance-matrix forecasting methods based on their ability in out-of-sample tests to keep the variance of the global minimum-variance portfolio at a prespecified target level. The performance of these methods is also measured by the MSTE. To implement this portfolio construction method, we first define the composition of the global minimum-variance portfolio for month m on the basis of the covariance-matrix forecast provided at the end of month $m-1$. The forecasted standard deviation of the minimum-variance portfolio for month m is given by

$$\hat{\sigma}_{Pm} = \sqrt{\hat{\mathrm{w}}'_m \hat{\Sigma}_m \hat{\mathrm{w}}_m}$$

where $\hat{\mathrm{w}}_m$ is the ex ante vector of weights of the global minimum-variance portfolio for month m, and $\hat{\Sigma}_m$ is the covariance-matrix forecast

EXHIBIT 8

The Portfolio MSTE Over the Period 1995–2013 vs. the Covariance-Matrix Forecasting Method

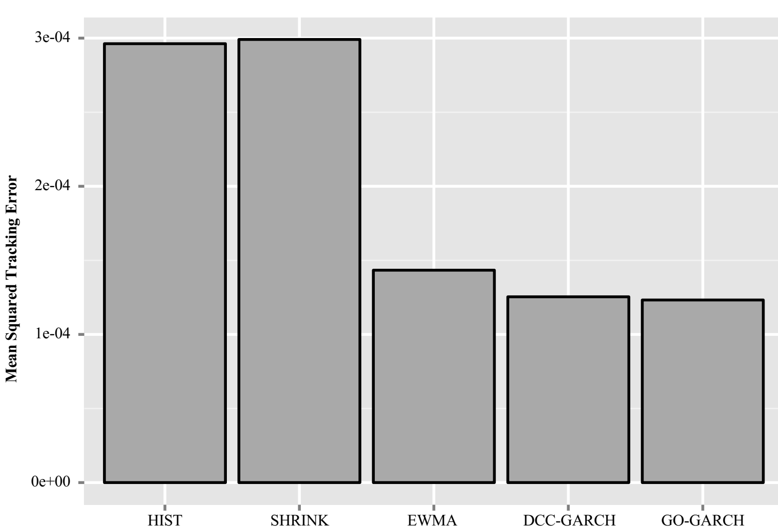
This exhibit reports the out-of-sample performance of alternative covariance forecasting methods in keeping the variance of the global minimum-variance portfolio at a target level. The annualized target volatility of the minimum-variance portfolios is 10%. HIST denotes the rolling historical covariance, SHRINK denotes the rolling historical covariance with shrinkage, EWMA denotes the EWMA covariance, DCC-GARCH denotes the DCC-GARCH; GO-GARCH denotes the GO-GARCH. In all methods, the length of the rolling estimation window is 120 months.

	Covariance-Matrix Forecasting Method				
Dataset	**HIST**	**SHRINK**	**EWMA**	**DCC-GARCH**	**GO-GARCH**
Portfolios Formed on Size	3.61e-04	3.66e-04	1.83e-04	1.59e-04	1.62e-04
Portfolios Formed on Book-to-Market	3.38e-04	3.41e-04	1.50e-04	1.28e-04	1.22e-04
Portfolios Formed on Momentum	2.99e-04	3.01e-04	1.44e-04	1.26e-04	1.20e-04
Portfolios Formed on Size and Book-to-Market	3.70e-04	3.75e-04	1.59e-04	1.48e-04	1.50e-04
Industry Portfolios	2.85e-04	2.83e-04	1.39e-04	1.19e-04	1.18e-04
Portfolios Formed on Size and Momentum	2.99e-04	3.01e-04	1.44e-04	1.26e-04	1.20e-04
Portfolios Formed on Size and Short-Term Reversal	2.80e-04	2.85e-04	1.48e-04	1.28e-04	1.20e-04
Portfolios Formed on Size and Long-Term Reversal	2.90e-04	2.95e-04	1.43e-04	1.34e-04	1.25e-04
Core Assets	1.44e-04	1.43e-04	8.15e-05	6.13e-05	7.20e-05
Averages Over All Datasets	2.96e-04	2.99e-04	1.43e-04	1.25e-04	1.23e-04

EXHIBIT 9

Average Out-of-Sample Performance of Alternative Covariance Forecasting Methods in Keeping the Variance of the Global Minimum-Variance Portfolio at a Target Level

The out-of-sample period is 1995–2013, and the annualized target volatility of the minimum-variance portfolios is 10%. For each distinct method of forecasting, this exhibit plots the average MSTE over all data sets. HIST denotes the rolling historical covariance, SHRINK denotes the rolling historical covariance with shrinkage, EWMA denotes the EWMA covariance, DCC-GARCH denotes the DCC-GARCH; GO-GARCH denotes the GO-GARCH. In all methods, the length of the rolling estimation window is 120 months.



for month m. Then, given the volatility target σ^{target}, the global minimum-variance portfolio is mixed with the risk-free asset to keep the volatility of the complete portfolio at the target level. Specifically, the weight of the minimum-variance portfolio in the complete portfolio is determined as

$$a_m = \min\left(\frac{\sigma^{\text{target}}}{\hat{\sigma}_{Pm}}, a^{\max}\right)$$

Given the daily returns generated by the minimum-variance portfolio with a volatility target, we compute the portfolio out-of-sample (i.e., realized) monthly standard deviation, which is denoted by σ_{Cm}^{oos}. The month m STE and the MSTE are computed as

$$STE_m = \left(\sigma_{Cm}^{\text{oos}} - \sigma^{\text{target}}\right)^2, \quad MSTE = \frac{1}{M}\sum_{m=1}^{M} STE_m$$

where M is the number of months in the out-of-sample period. In our study, the annual volatility target, σ^{target}, is set at 10% and the maximum risk exposure, $a^{\max}$, is set at 150%.

Exhibit 8 reports the out-of-sample performance (over the period 1995 to 2013) of alternative covariance forecasting methods in keeping the variance of the global minimum-variance portfolio at a target level. Specifically, this exhibit reports the portfolio MSTEs produced by different covariance-matrix forecasting methods. Exhibit 9 plots the average MSTEs over all data sets. The results for the portfolio MSTEs reported in Exhibit 8 agree with those presented in the previous subsections. In brief, there is no difference between the MSTEs produced by the rolling historical covariance method and those produced by the rolling historical covariance with shrinkage method. The other three methods of forecasting reduce the MSTE by 50% to 60%.

CONCLUSIONS

Our research showed that there are large differences between the alternative covariance-matrix forecasting methods. For those practitioners who routinely use the sample covariance-matrix as a forward-looking estimate for the future covariance-matrix, our empirical results imply that there are potentially huge benefits from adopting multivariate GARCH models. Our study revealed that, as compared with the sample covariance-matrix forecast, the multivariate GARCH models allow one to decrease the covariance-matrix forecasting error by 50%. In addition, our study demonstrated that portfolios constructed using multivariate GARCH forecasts deliver the best performance in terms of risk optimization and portfolio tracking error. In particular, switching from the sample covariance-matrix forecast to a multivariate GARCH forecast let one reduce the tracking error of minimum-variance portfolios by more than 50%. The same degree of improvement is obtained in the tracking error of minimum-variance portfolios with a volatility target. We tested two multivariate GARCH models, DCC-GARCH and GO-GARCH, and in our tests, the performance of the DCC-GARCH model was found to be marginally better than that of the GO-GARCH model.

There is a common impression that the shrinkage of the sample covariance-matrix improves the accuracy of the covariance-matrix forecast. The results obtained previously by Disatnik and Benninga [2007] revealed that this impression is wrong. Our study confirms that shrinkage is a computationally intensive method that reduces neither the forecasting error nor the tracking error of minimum-variance portfolios. Our results showed that the simple EWMA covariance-matrix forecast performs only slightly worse than the multivariate GARCH forecast. Therefore, those practitioners who find that multivariate GARCH models are cumbersome to implement are advised to adopt the EWMA method, which is computationally simple and fast.

ENDNOTES

The author thanks Steen Koekebakker, an anonymous referee, and the participants at the 2014 Paris Financial Management Conference for their insightful comments. The author also thanks Alexios Ghalanos for developing the R package "rmgarch," which was used to estimate the multivariate GARCH models and perform forecasting. The usual disclaimer applies.

[1]http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

REFERENCES

Albeverio, S., V. Steblovskaya, and K. Wallbaum. "Investment Instruments With Volatility Target Mechanism." *Quantitative Finance*, Vol. 13, No. 10 (2013), pp. 1519-1528.

Asness, C., A. Frazzini, and L. Pedersen. "Leverage Aversion and Risk Parity." *Financial Analysts Journal*, Vol. 68, No. 1 (2012), pp. 47-59.

Bollerslev, T. "Generalized Autoregressive Conditional Heteroskedasticity." *Journal of Econometrics*, Vol. 31, No. 3 (1986), pp. 307-327.

Bollerslev, T., R. Engle, and J. Wooldridge. "A Capital Asset Pricing Model with Time-Varying Covariances." *Journal of Political Economy*, Vol. 96, No. 1 (1988), pp. 116-131.

Busse, J. "Volatility Timing in Mutual Funds: Evidence from Daily Returns." *Review of Financial Studies*, Vol. 12, No. 5 (1999), pp. 1009-1041.

Butler, A., and M. Philbrick. "Volatility Management for Better Absolute and Risk-Adjusted Performance." White paper, Macquarie Private Wealth Inc., 2012.

Chan, L., J. Karceski, and J. Lakonishok. "On Portfolio Optimization: Forecasting Covariances and Choosing the Risk Model." *Review of Financial Studies*, Vol. 12, No.5 (1999), pp. 937-974.

Chaves, D., J. Hsu, F. Li, and O. Shakernia. "Risk Parity Portfolio versus Other Asset Allocation Heuristic Portfolios." *The Journal of Investing*, Vol. 20, No. 1 (2011), pp. 108-108.

Choueifaty, Y., and Y. Coignard. "Toward Maximum Diversification." *The Journal of Portfolio Management*, Vol. 35, No. 1 (2008), pp. 40-51.

Clarke, R., H. de Silva, and S. Thorley. "Minimum-Variance Portfolios in the U.S. Equity Market." *The Journal of Portfolio Management*, Vol. 33, No. 1 (2006), pp. 10-24.

——. "Minimum-Variance Portfolio Composition." *The Journal of Portfolio Management*, Vol. 37, No. 2 (2011), pp. 31-45.

Collie, R., M. Sylvanus, and M. Thomas. "Volatility-Responsive Asset Allocation." White paper, Russell Investments, 2011.

DeMiguel, V., L. Garlappi, and R. Uppal. "Optimal versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy?" *Review of Financial Studies*, Vol. 22, No. 5 (2009), pp. 1915-1953.

Disatnik, D., and S. Benninga. "Shrinking the Covariance-Matrix." *The Journal of Portfolio Management*, Vol. 33, No. 4 (2007), pp. 55-63.

Duchin, R., and H. Levy. "Markowitz versus the Talmudic Portfolio Diversification Strategies." *The Journal of Portfolio Management*, Vol. 35, No. 2 (2009), pp. 71-74.

Engle, R. "Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models." *Journal of Business and Economic Statistics*, Vol. 20, No. 3 (2002), pp. 339-350.

Engle, R., and K. Kroner. "Multivariate Simultaneous Generalized ARCH." *Econometric Theory*, Vol. 11, No. 1 (1995), pp. 122-150.

Kritzman, M., S. Page, and D. Turkington. "In Defense of Optimization: The Fallacy of 1/N." *Financial Analysts Journal*, Vol. 66, No. 2 (2010), pp. 31-39.

Ledoit, O., and M. Wolf. "Improved Estimation of the Covariance-matrix of Stock Returns with an Application to Portfolio Selection." *Journal of Empirical Finance*, Vol. 10, No. 5 (2003), pp. 603 - 621.

——. "Honey, I Shrunk the Sample Covariance-Matrix." *The Journal of Portfolio Management*, Vol. 30, No. 4 (2004), pp. 110-119.

Maillard, S., T. Roncalli, and J. Teiletche. "The Properties of Equally Weighted Risk Contribution Portfolios." *The Journal of Portfolio Management*, Vol. 36, No. 4 (2010), pp. 60-70.

Markowitz, H. "Portfolio Selection." *Journal of Finance*, Vol. 7, No. 1 (1952), pp. 77-91.

Pojarliev, M., and W. Polasek. "Applying Multivariate Time Series Forecasts for Active Portfolio Management." *Financial Markets and Portfolio Management*, Vol. 15, No. 2 (2001), pp. 201-211.

——. "Portfolio Construction by Volatility Forecasts: Does the Covariance Structure Matter?" *Financial Markets and Portfolio Management*, Vol. 17, No. 1 (2003), pp. 103-116.

Stein, C. "Inadmissibility of the Usual Estimator for the Mean of a Multivariate Normal Distribution." In *Proceedings of the 3rd Berkeley Symposium on Mathematical Statistics and Probability*, pp. 197-206. Berkeley: University of California Press, 1956.

van der Weide, R. "GO-GARCH: A Multivariate Generalized Orthogonal GARCH Model." *Journal of Applied Econometrics*, Vol. 17, No. 5 (2002), pp. 549-564.