Econometrics and Statistics 27 (2023) 1–15

Contents lists available at ScienceDirect

Econometrics and Statistics

journal homepage: www.elsevier.com/locate/ecosta

Forecasting value-at-risk and expected shortfall in large portfolios: A general dynamic factor model approach

Marc Hallin[a,*], Carlos Trucíos[b]

[a] Department of Mathematics and ECARES, Université libre de Bruxelles, Belgium
[b] Faculty of Business Administration and Accounting, Federal University of Rio de Janeiro (UFRJ), Brazil

ARTICLE INFO

Article history:
Received 14 December 2020
Revised 29 April 2021
Accepted 29 April 2021
Available online 14 May 2021

JEL classification:
C10
C32
C53
C55
G17
G32

Keywords:
Conditional covariance
High-dimensional time series
Large panels
Risk measures
Volatility

ABSTRACT

Beyond their importance from the regulatory policy point of view, Value-at-Risk (VaR) and Expected Shortfall (ES) play an important role in risk management, portfolio allocation, capital level requirements, trading systems, and hedging strategies. However, due to the curse of dimensionality, their accurate estimation and forecast in large portfolios is quite a challenge. To tackle this problem, two procedures are proposed. The first one is based on a filtered historical simulation method in which high-dimensional conditional covariance matrices are estimated via a general dynamic factor model with infinite-dimensional factor space and conditionally heteroscedastic factors; the other one is based on a residual-based bootstrap scheme. The two procedures are applied to a panel with concentration ratio close to one. Backtesting and scoring results indicate that both VaR and ES are accurately estimated under both methods, which both outperform the existing alternatives.

1. Introduction

Value-at-Risk (VaR) and Expected Shortfall (ES) play an essential role in economics, finance, and insurance. They are lying at the core of the regulatory framework settled by the Solvency II regulation, the Swiss Solvency Test, and the Basel II and III accords, all intended to strengthen the financial stability of the banking system. Beyond this, VaR and ES constitute a basic tool for investors, hedge fund managers, traders, and all decision-makers involved in risk management, portfolio allocation, trading desk limits, investment strategies, and the development of trading algorithms. Their accurate estimation and forecasting, thus, is of primary importance.

Several procedures dealing with VaR and ES estimation in univariate and multivariate settings are available in the literature: see Braione and Scholtes (2016), Gao and Zhou (2016), Bayer (2018), Broda et al. (2018), Trucíos et al. (2018), Patton et al. (2019), Taylor (2019), Francq and Zakoian (2020), Trucíos et al. (2020b), ... to quote only a few recent references. In the multivariate framework, most procedures are based on the estimation of the conditional covariance matrix, which in

* Corresponding author.
E-mail address: mhallin@ulb.ac.be (M. Hallin).

https://doi.org/10.1016/j.ecosta.2021.04.006
2452-3062/

a high-dimensional context is quite challenging due to the difficulty to guarantee positive definiteness and the fastly increasing number of parameters—the classical *curse of dimensionality*. As a consequence, traditional multivariate volatility estimation procedures, while remaining useful in small and moderate dimensions, turn out to be helpless or unfeasible in high dimensions.

A high-dimensional context, unfortunately, is the rule rather than the exception in the area, where relevant information typically is scattered among large numbers of time series. Alternative procedures that are able to handle very large panels of financial data thus are badly needed. This need has aroused much interest, and several procedures have been proposed recently to estimate high-dimensional conditional covariance matrices; see Zumbach (2007), Fan et al. (2008), Alessi et al. (2009), Li et al. (2016), Chang et al. (2018), Engle et al. (2019), Trucíos et al. (2019), Pakel et al. (2020), among others.

Recently, Trucíos et al. (2020a) have proposed a high-dimensional conditional covariance matrix estimator based on the general dynamic factor model which does not restrict the factor space to be finite-dimensional. Their procedure, which is based on the two-step approach by Barigozzi and Hallin (2016, 2017, 2020) and the one-sided estimators of Forni et al. (2015, 2017), has shown a good performance in simulations and empirical data, outperforming several of the aforementioned approaches.

Estimating (forecasting) VaRs and ESs in large panels is part of daily practice in financial institutions. The usual approaches are the so-called *historical simulation* method, which does not account for time-varying volatilities and correlations, and the *filtered historical simulation* method, which does (see, for instance; Pérignon and Smith, 2010; Aramonte et al., 2013). The filtered historical simulation method actually applies the historical simulation method to *devolatilized* (or *filtered*) portfolio returns, where devolatilized returns are obtained via a univariate conditional variance estimator or via a multivariate conditional covariance estimator—which, in theory, is much better but once once again runs into the problem of estimating high-dimensional matrices. See, for instance, Giannopoulos and Tunaru (2005), Aramonte et al. (2013), and Gurrola-Perez and Murphy (2015).

In accordance with current practice, we propose an implementation of the filtered historical simulation procedure based on the high-dimensional conditional covariance matrix estimator of Trucíos et al. (2020a) for the estimation of VaR and ES in large portfolios. This filtered historical simulation approach is adopted in conformity with practitioners' familiarity; combining it with our covariance matrix estimator is a natural option in view of the good performance of the latter (as demonstrated in Monte Carlo experiments and empirical applications in Trucíos et al. (2020a)). This filtered historical simulation approach, however, requires the Cholesky factorization and inversion of high-dimensional matrices which, depending on the cross-sectional dimension of the panel, may become problematic. As an alternative, we also propose a computationally loghter residual-based bootstrap approach which only requires the Cholesky factorization and inversion of low-dimensional and diagonal high-dimensional matrices.

These two approaches are compared to alternative state-of-the-art procedures on the basis of a rolling window scheme out-of-sample analysis performed on a panel of 652 stock returns observed over 750 trading days. A backtesting analysis of the results indicates that the two methods we are proposing yield quite similar performance and both significantly outperform the existing ones.

The rest of the paper is organized as follows. In Section 2, we describe the general dynamic factor model and the conditional covariance matrix estimation procedure of Trucíos et al. (2020a). In Section 3 we define the VaR and ES forecasts and describe the proposed filtered historical simulation and residual-based bootstrap forecasting methods. Section 4 introduces the various backtesting tools used to evaluate the accuracy of VaR and ES forecasts. In Section 5, we conduct an empirical out-of-sample comparative analysis of our methods and its competitors. Section 6 concludes.

2. The general dynamic factor model with conditionally heteroscedastic factors

The general dynamic factor model (GDFM) was introduced by Forni et al. (2000) and encompasses most other high-dimensional factor models proposed in the econometric and time series literatures.

Let $\{\mathbf{X}_t := (X_{1t},\ X_{2t}, \ldots, X_{it}, \ldots)',\ t \in \mathbb{Z}\}$, be a double-indexed zero-mean second-order stationary stochastic process, where i is a cross-sectional index and t stand for time. The GDFM is based on the decomposition of X_{it} into a *common* component χ_{it} and an *idiosyncratic* one ξ_{it} driven by *common* ($\mathbf{u}_t$) and *idiosyncratic* (v_{it}) *shocks*, respectively.

Letting $\mathbf{X}_{nt} := \{X_{it} | i = 1, ..., n, t \in \mathbb{Z}\}$, $\boldsymbol{\chi}_{nt} := \{\chi_{it} \mid i = 1, ..., n,\ t \in \mathbb{Z}\}$, and $\boldsymbol{\xi}_{nt} := \{\xi_{it} | i = 1, ..., n, t \in \mathbb{Z}\}$, the GDFM decomposition in vector notation takes the form

$$\mathbf{X}_{nt} = \boldsymbol{\chi}_{nt} + \boldsymbol{\xi}_{nt} = \mathbf{B}_n(L)\mathbf{u}_t, +\mathbf{D}_n(L)\mathbf{v}_{nt}, \quad n \in \mathbb{N}_0, \quad t \in \mathbb{Z} \tag{1}$$

with

$$\mathbf{B}_n(L) := (\mathbf{b}_1(L)...\mathbf{b}_n(L))', \qquad \mathbf{D}_n(L) := \mathrm{diag}(d_1(L)...d_n(L)),$$
$$\mathbf{u}_t := (u_{1t}\ u_{2t} \ldots\ u_{qt})', \qquad \text{and} \qquad \mathbf{v}_{nt} := (v_{1t} \ldots v_{nt})'.$$

where $\{\mathbf{u}_t, \mathbf{v}_{nt}\}$ is orthonormal white noise.

The models considered in Bai and Ng (2002) or Stock and Watson (2002a,b) are particular cases under which $\mathbf{B}_n(L)\mathbf{u}_t$ reduces to $\mathbf{B}_n\mathbf{F}_t$, where $\mathbf{B}_n$ is an $n \times r$ matrix of loadings and the factors $\mathbf{F}_t = (F_{1t}, \ldots, F_{rt})'$ span an r-dimensional factor space ($q \leq r < \infty$). The loadings, in these models, are static; for convenience, call them *static* factor models.

Consistent estimation procedures for the GDFM have been proposed under various assumptions in Forni et al. (2000, 2005) and Forni and Lippi (2011); the most general results are those of Forni et al. (2015, 2017). Contrary to Bai and Ng (2002) or Stock and Watson (2002a,b), their estimation procedure, which we now briefly describe, accommodates infinite-dimensional factor spaces, yet only involves one-sided filters—an essential feature in a forecasting context where "future" observations are not available.

2.1. The GDFM estimation procedure

For an observed panel $\mathbf{X}_{nT}$ with cross-sectional dimension n and sample size T, the GDFM estimation procedure of Forni et al. (2017) proceeds as follows; the assumptions made are the same as in Trucíos et al. (2020a), where we refer to for details.

- **Step 1.** Determine the number q of common shocks via an information criterion, for instance, using (Hallin and Liška, 2007).
- **Step 2.** Randomly reorder the n observed series.
- **Step 3.** Estimate the spectral density matrix of $\mathbf{X}$ by

$$\widehat{\mathbf{\Sigma}}^X_{nT}(\theta) = \frac{1}{2\pi}\sum_{k=-M_T}^{M_T} e^{-ik\theta} K\left(\frac{k}{B_T}\right)\widehat{\mathbf{\Gamma}}^X_k, \quad \theta \in [0, 2\pi]$$

  where $K(\cdot)$ is a kernel function, M_T a truncation parameter, B_T a bandwidth, and $\widehat{\mathbf{\Gamma}}^X_k$ the sample lag-k cross-covariance matrix.
- **Step 4.** Estimate the spectral density matrix of the common components by

$$\widehat{\mathbf{\Sigma}}^\chi_{nT}(\theta) := \widehat{\mathbf{P}}^X_{nT}(\theta)\widehat{\mathbf{\Lambda}}^X_{nT}(\theta)\widehat{\mathbf{P}}^{X*}_{nT}(\theta), \quad \theta \in [0, 2\pi]$$

  where $\widehat{\mathbf{\Lambda}}^X_{nT}(\theta)$ is a $q \times q$ diagonal matrix with diagonal elements the q largest eigenvalues of $\widehat{\mathbf{\Sigma}}^X_{nT}(\theta)$ and $\widehat{\mathbf{P}}^X_{nT}(\theta)$ (with complex conjugate $\widehat{\mathbf{P}}^{X*}_{nT}$) is the corresponding $n \times q$ matrix of eigenvectors.
- **Step 5.** Let $n^* := m(q+1)$ with $m := \left\lceil \frac{n}{q+1} \right\rceil$ and denote by $\widehat{\mathbf{\Sigma}}^\chi_{n^*T}(\theta)$ the empirical $n^* \times n^*$ spectral density matrix computed from $\mathbf{X}_{n^*t}$.

Forni et al. (2015) show that there exists an $n^* \times n^*$ block-diagonal VAR filter $\mathbf{A}_{n^*}(L)$ with diagonal blocks $\mathbf{A}^k(L)$, $k = 1, \dots, m$ such that

$$\mathbf{Y}_{n^*t} := \mathbf{A}_{n^*}(L)\mathbf{X}_{n^*t} = \mathbf{R}_{n^*}\mathbf{u}_t + \mathbf{A}_{n^*}(L)\boldsymbol{\xi}_{n^*t}, \tag{2}$$

where the filtered process $\mathbf{Y}_{n^*t}$ satisfies the assumptions of a static factor model with $r = q$ factors $\mathbf{F}_t = \mathbf{u}_t$, loading matrix $\mathbf{R}_{n^*}$, and idiosyncratic component $\mathbf{A}_{n^*}(L)\boldsymbol{\xi}_{n^*t}$.

- **Step 6.** By inverse Fourier transform of $\widehat{\mathbf{\Sigma}}^\chi_{n^*T}(\theta)$, estimate the autocovariance matrices $\widehat{\mathbf{\Gamma}}^\chi_k$ of the m sub-vectors

$$\boldsymbol{\chi}^k_t = (\chi_{(k-1)(q+1)+1,t} \dots \chi_{k(q+1),t})', \quad k = 1, \dots, m$$

  of dimension $(q+1)$. Based on these, compute, after order identification, the Yule-Walker estimators $\hat{\mathbf{A}}^k(L)$ of the m VAR filters $\mathbf{A}^k(L)$ and stack them into a block-diagonal matrix $\hat{\mathbf{A}}_{n^*}(L)$. Compute

$$\widehat{\mathbf{Y}}_{n^*t} := \hat{\mathbf{A}}_{n^*}(L)\mathbf{X}_{n^*t}.$$

- **Step 7.** Based on the first q standard principal components of $\widehat{\mathbf{Y}}_{n^*t}$, obtain estimates $\widehat{\mathbf{R}_{n^*}\mathbf{u}_t}$ of $\mathbf{R}_{n^*}\mathbf{u}_t$ and, via a Cholesky identification constraint, the estimates $\widehat{\mathbf{R}}_{n^*}$ and $\widehat{\mathbf{u}}_t$ of $\mathbf{R}_{n^*}$ and $\mathbf{u}_t$: our estimate of the impulse-response function is $\widehat{\mathbf{B}}_{n^*}(L) := [\hat{\mathbf{A}}_{n^*}(L)]^{-1}\widehat{\mathbf{R}}_{n^*}$.

To be precise, the Cholesky identification constraint in Step 7 is obtained by considering the $q \times q$ matrix $\mathbf{B}_q(0)$ in model (1), with (i, j) entry $b_{ij}(0)$ and the Cholesky factorization $\mathbf{HH}'$ of $\mathbf{B}_q(0)\mathbf{B}'_q(0)$: $\mathbf{H}$ thus is the lower-triangular matrix with positive diagonal satisfying $\mathbf{HH}' = \mathbf{B}_q(0)\mathbf{B}'_q(0)$. The identifiability constraints consist in imposing

$$\mathbf{u}_t = \mathbf{H}'\mathbf{B}_q^{-1\prime}(0)\mathbf{u}_t \quad \text{and} \quad \mathbf{b}_i(L) = \mathbf{b}_i(L)\mathbf{B}_q^{-1}(0)\mathbf{H}, \quad j = 1, \dots, n.$$

The estimator $\widehat{\mathbf{R}_{n^*}\mathbf{u}_t}$ of $\mathbf{R}_{n^*}\mathbf{u}_t$ in Step 7 actually follows from running, on the filtered $\widehat{\mathbf{Y}}_{n^*t}$ series, the estimation procedures proposed by Bai and Ng (2002) or Stock and Watson (2002a,b) for the common component of the "static" factor model (2); these procedures consist in projecting the observations onto the space spanned by the q first eigenvectors of $\widehat{\mathbf{Y}}_{n^*t}$'s empirical covariance matrix.

- **Step 8.** Repeat Steps 2 through 7 (generating, as in Forni et al. (2017) and Trucíos et al. (2020a), 30 randomly chosen cross-sectional permutations): the final estimates (denoted as $\widehat{\mathbf{R}}_n$, $\widehat{\mathbf{u}}_t$, and $\widehat{\mathbf{B}}_n$) are obtained by averaging the 30 estimates $\widehat{\mathbf{R}}_{n^*}$, $\widehat{\mathbf{u}}_t$, and $\widehat{\mathbf{B}}_{n^*}$ associated with these iterations. Let

$$\widehat{\boldsymbol{\chi}}_{nt} := \widehat{\mathbf{B}}_n(L)\widehat{\mathbf{u}}_t \quad \text{and} \quad \widehat{\boldsymbol{\xi}}_{nt} := \mathbf{X}_{nt} - \widehat{\boldsymbol{\chi}}_{nt}.$$

The consistency of the method is established (see Forni et al., 2017 and Barigozzi and Hallin, 2020) under assumptions (stationarity, the existence of a spectral density matrix, etc.) which we do not reproduce here.

2.2. One-step ahead estimation (forecasting) of conditional covariance matrices

In order to exploit the dependence on the past of the conditional covariance matrix, additional assumptions on the volatilities of the common shocks and the idiosyncratic components are to be made, which are similar to the assumptions in Alessi et al. (2009), Aramonte et al. (2013), and Trucíos et al. (2020a). We assume that the heteroscedastic volatility dynamics of the common shocks are those of an MGARCH process stable by aggregation (examples of MGARCH models stable by aggregation are the BEKK and VECH models. See Chapter 10 of Francq and Zakoian (2019) for details). Additionally, we assume that the heteroscedastic volatility dynamics of each idiosyncratic component can be modelled as a GARCH-type process; the common shocks and idiosyncratic components are assumed to be conditionally uncorrelated.

Under these assumptions and using the results in Forni et al. (2015) which state that the process

$$\mathbf{Y}_{nt} := \mathbf{A}_n(L)\mathbf{X}_{nt} = \mathbf{A}_n(L)\boldsymbol{\chi}_{nt} + \mathbf{A}_n(L)\boldsymbol{\xi}_{nt}$$

admits a *static* factor representation of the form $\mathbf{Y}_{nt} = \mathbf{R}\mathbf{u}_t + \epsilon_t$, Trucíos et al. (2020a) show that the conditional covariance matrix $\mathbf{V}^{\mathbf{X}_n}_{t|t-1}$ of $\mathbf{X}_{nt}$ (conditional on the sigma-field $\mathcal{F}_{t-1}$ generated by the observations $\mathbf{X}_{n,t-1}, \mathbf{X}_{n,t-2}, \ldots$) decomposes into

$$\mathbf{V}^{\mathbf{X}_n}_{t|t-1} = \mathbf{R}_n\mathbf{V}^{\mathbf{u}}_{t|t-1}\mathbf{R}_n' + \mathbf{V}^{\boldsymbol{\xi}_n}_{t|t-1}, \tag{3}$$

where $\mathbf{V}^{\mathbf{u}}_{t|t-1}$ and $\mathbf{V}^{\boldsymbol{\xi}_n}_{t|t-1}$ stand for the conditional covariance matrices of the common shocks and idiosyncratic components, respectively. Equation (3) is obtained by considering that, without loss of generality, the block-diagonal VAR filter

$$\mathbf{A}_n(L) := \begin{bmatrix} \mathbf{A}^1(L) & 0 & \ldots & 0 \\ 0 & \mathbf{A}^2(L) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \mathbf{A}^m(L) \end{bmatrix} \tag{4}$$

can be written as $\mathbf{A}_n(L) = \mathbf{I} - \boldsymbol{\Phi}_1 L - \cdots - \boldsymbol{\Phi}_S L^S$; it follows that $\mathbf{V}^{\mathbf{Y}_n}_{t|t-1} = \mathbf{V}^{\mathbf{X}_n}_{t|t-1}$ and $\mathbf{V}^{\epsilon_n}_{t|t-1} = \mathbf{V}^{\boldsymbol{\xi}_n}_{t|t-1}$ (see Trucíos et al., 2020a, for details). Therefore, an estimator of (3) can be obtained as follows.

- **Step 9.** Using the estimated common shocks ($\hat{\mathbf{u}}$) and idiosyncratic components ($\hat{\boldsymbol{\xi}}_n$) obtained in the previous step, estimate the one-step-ahead conditional covariance matrix of the whole panel $\mathbf{X}_n$ as

$$\widehat{\mathbf{V}}^{\mathbf{X}_n}_{t|t-1} := \widehat{\mathbf{R}}_n\widehat{\mathbf{V}}^{\hat{\mathbf{u}}}_{t|t-1}\widehat{\mathbf{R}}_n' + \widehat{\mathbf{V}}^{\hat{\boldsymbol{\xi}}_n}_{t|t-1}, \tag{5}$$

  where the estimated one-step-ahead conditional covariance matrices $\widehat{\mathbf{V}}^{\hat{\mathbf{u}}}_{t|t-1}$ and $\widehat{\mathbf{V}}^{\hat{\boldsymbol{\xi}}_n}_{t|t-1}$ of the estimated common shocks and idiosyncratic components, are obtained by fitting a q-dimensional MGARCH model and n univariate GARCH-type models, respectively, to the estimated $\hat{\mathbf{u}}_t$ and $\hat{\boldsymbol{\xi}}_n$.

Note that, thanks to the fact that $\widehat{\mathbf{V}}^{\hat{\mathbf{u}}}_{t|t-1}$ is $q \times q$ (with q typically small) and $\widehat{\mathbf{V}}^{\hat{\boldsymbol{\xi}}_n}_{t|t-1}$ is well approximated by a diagonal matrix, computing $\widehat{\mathbf{V}}^{\mathbf{X}_n}_{t|t-1}$ in (5) remains quite feasible in high dimensions. Although the diagonal matrix $\widehat{\mathbf{V}}^{\hat{\boldsymbol{\xi}}_n}_{t|t-1}$ neglects possible idiosyncratic cross-covariances, these cross-covariances are mild (non-pervasive if not null) by definition and do not affect the consistency of (5).

It is important to mention that the estimators of MGARCH models such as the stable by aggregation BEKK and VECH ones are strongly affected by the choice of initial values, suffer from convergence problems, and perform quite poorly (see, for instance, Lien et al., 2002; Manabu, 2015). Fortunately, the DCC estimator (Engle, 2002) shows good performance, even when the data-generating process is not a DCC, and proves to be quite robust against model misspecification (see Chevallier, 2012; Laurent et al., 2012; de Almeida et al., 2018; Trucíos et al., 2020a). Therefore, the DCC model, although it is not stable by aggregation, is used in all subsequent numerical applications: call GDFM-CHF method this procedure.

3. Value at Risk (VaR) and Expected Shortfall (ES) one-step ahead estimation (forecasting)

The one-step-ahead VaR of a series $\{r_t\}$ of returns at time $T+1$ and risk level α is defined as

$$\text{VaR}^{\alpha}_{T+1} := \inf\{x \in \mathbb{R} : F(x|\mathcal{F}_T) \geq \alpha\},$$

where $x \mapsto F(x|\mathcal{F}_T)$ stands for the distribution function of r_{T+1} conditional on the information $\mathcal{F}_T$ available up to time T: $\text{VaR}^{\alpha}_{T+1}$ thus is r_{T+1}'s α-quantile conditional on $r_T, r_{T-1}, \ldots$. The one-step-ahead ES at risk level α similarly is defined as

$$\text{ES}^{\alpha}_{T+1} := \text{E}[r_{T+1}|r_{T+1} \leq \text{VaR}^{\alpha}_{T+1}, \mathcal{F}_T].$$

In practice, of course, $\text{VaR}^{\alpha}_{T+1}$ and ES^{α}_{T+1} have to be estimated from a finite realization $r_1, \ldots, r_T$ of $\{r_t\}$ or, in case r_t is the revenue of a portfolio (the notation R_t will be used when this is to be emphasized), from a finite realization of the returns of the individual stocks composing the portolio. Since these estimators of a quantity related to time $T+1$ are to be computed at time T, we also call them one-step-ahead *forecasts*.

Many procedures to estimate (forecast) the VaR and ES of a portfolio are available in the literature—see Nieto and Ruiz (2016) or Righi and Ceretta (2015) for recent reviews. The estimation of the conditional covariance matrix of individual stocks plays an important role in all these methods. When the number of stocks is high, however, computing this estimation is a major issue and, due to the curse of dimensionality, a number of methods cannot handle very large portfolios. To overcome this problem, we propose using the *filtered historical simulation* approach along with the estimator of the conditional covariance matrix of returns recently proposed by Trucíos et al. (2020a).

3.1. Filtered Historical Simulation approach

Since its inception by Barone-Adesi et al. (1998, 1999), the so-called *filtered historical simulation* (FHS) method has been widely used to forecast VaRs (see, for instance, Aramonte et al., 2013 or Gurrola-Perez and Murphy, 2015). Its application in the ES estimation/forecast context was initiated by Giannopoulos and Tunaru (2005) and since then has been adopted by both academics and practitioners.

FHS modifies the classical historical simulation method by taking into account the conditional heteroscedastic dynamics of financial returns via the power of conditional volatility models (Barone-Adesi et al., 1998; Giannopoulos and Tunaru, 2005; Christoffersen, 2009).

Unlike the original and widely-applied historical simulation method, which basically estimates VaRs by using the α-quantile of historical returns and and ESs as the average of returns falling below that VaR, FHS first “devolatilizes” the portfolio returns/vector returns by means of their estimated conditional volatilities/co-volatilities and then re-scales back these devolatilized returns via the volatility/co-volatility forecast.

Let $\mathbf{r}_t =: \mathbf{H}_t^{1/2}\boldsymbol{\epsilon}_t$ be the $N-$dimensional vector of individual returns at time t, where $\mathbf{H}_t^{1/2}$ follows from the lower triangular Cholesky factorization of the conditional covariance matrix $\mathbf{H}_t$ (an identification constraint that does not imply any loss of generality) and $\boldsymbol{\epsilon}_t$ an i.i.d. $(0, \mathbf{I}_N)$ random vector. The FHS method proceeds in three steps (coming on top of Steps 1-9 as described in Section 2). Denote by $\widehat{\mathbf{H}}_t$ an estimator of $\mathbf{H}_t$ and by $\widehat{\mathbf{H}}_t^{1/2}\widehat{\mathbf{H}}_t^{1/2\prime}$ its lower triangular Cholesky factorization; any consistent estimator of the conditional covariance matrix can be considered, such as those proposed by Engle (2002), Aramonte et al. (2013), Gurrola-Perez and Murphy (2015), Trucíos et al. (2019) and Engle et al. (2019)—after due Cholesky factorization. However, most of these estimators are unfeasible in high dimensions (Caporin and McAleer, 2014) or fail to exploit the information available on the dependence between individual returns. In view of its computability and excellent performance, the estimator we are recommending here is the GDFM-based one proposed in Trucíos et al. (2020a)—more precisely, the Cholesky square root $\widehat{\mathbf{H}}_t^{1/2}$ of $\widehat{\mathbf{H}}_t := \widehat{\mathbf{V}}^{\mathbf{r}_N}_{t|t-1}$ obtained in Step 9.

- **Step 10.** Using the historical returns $\mathbf{r}_1, \ldots, \mathbf{r}_T$ to obtain the estimators $\widehat{\mathbf{V}}^{\mathbf{r}_N}_{t|t-1}$ of the conditional covariance matrices along Steps 1-9, compute the filtered (or *devolatilized*) return vectors $\hat{\boldsymbol{\epsilon}}_t := \widehat{\mathbf{H}}_t^{-1/2}\mathbf{r}_t$, where $\widehat{\mathbf{H}}_t^{1/2}$ follows from the lower triangular Cholesky factorization of $\widehat{\mathbf{H}}_t := \widehat{\mathbf{V}}^{\mathbf{r}_N}_{t|t-1}$ at time $t = 1, \ldots, T+1$.

  Note that Step 10 requires, for each t, the Cholesky factorization and inversion of an $N \times N$ matrix.

- **Step 11.** Generate a bootstrap sample $\boldsymbol{\epsilon}_1^*, \ldots, \boldsymbol{\epsilon}_B^*$ of size B from the devolatilized return vectors $\hat{\boldsymbol{\epsilon}}_t$, $T = 1, \ldots, T+1$ and construct B one-step-ahead return vectors $\mathbf{r}^{i*}_{T+1} := \widehat{\mathbf{H}}^{1/2}_{T+1}\boldsymbol{\epsilon}_i^*$, still with $\widehat{\mathbf{H}}_t := \widehat{\mathbf{V}}^{\mathbf{r}_N}_{t|t-1}$; for portfolio weights $\boldsymbol{\omega} = (\omega_1, \ldots \omega_N)'$, yields B simulated one-step-ahead portfolio return forecasts $R^{i*}_{T+1} := \boldsymbol{\omega}'\mathbf{r}^{i*}_{T+1}$, $i = 1, \ldots, B$.

- **Step 12.** The one-step-ahead forecasts of $\mathrm{VaR}^{\alpha}_{T+1}$ and $\mathrm{ES}^{\alpha}_{T+1}$ are

$$\widehat{\mathrm{VaR}}^{\alpha}_{\mathrm{T}+1} := \widehat{F}^{-1}_{R^*_{T+1}}(\alpha), \tag{6}$$

  and

$$\widehat{\mathrm{ES}}^{\alpha}_{\mathrm{T}+1} := \sum_{i=1}^{B} \frac{R^{i*}_{T+1}\mathbb{I}\left[R^{i*}_{T+1} < \widehat{\mathrm{VaR}}^{\alpha}_{\mathrm{T}+1}\right]}{\sum_{t=1}^{B}\mathbb{I}\left[R^{i*}_{T+1} < \widehat{\mathrm{VaR}}^{\alpha}_{\mathrm{T}+1}\right]}, \tag{7}$$

  respectively, where $\widehat{F}^{-1}_{R^*_{T+1}}(\alpha)$ is the α-quantile of the empirical distribution of the simulated one-step-ahead portfolio returns $R^{1*}_{T+1}, \ldots, R^{B*}_{T+1}$ and $\mathbb{I}[\,\cdot\,]$ denotes the indicator function.

  Call this the GDFM-CHF method.

3.2. Residual-based Bootstrap approach

The FHS procedure described in Section 3.1 requires the Cholesky decomposition and invertion of T matrices of dimension $N \times N$, which, for large N and depending on the computing resources, could be a heavy or even unfeasible task. To overcome this, we also propose a computationally lighter residual-based bootstrap approach, which consist of bootstrapping devolatilized estimated common shocks and devolatilized estimated idiosyncratic components. Steps 10 and 11 in Section 3.1 then are replaced by the following Steps 10' and 11'.

- **Step 10'.** Using the estimated common shocks and idiosyncratic components resulting from the application of Steps 1-9 to the panel of return vectors, devolatilize them by $\boldsymbol{\nu}_t := \widehat{\mathbf{V}}^{\hat{\mathbf{u}},-1/2}_{t|t-1}\hat{\mathbf{u}}_t$ and $\boldsymbol{\eta}_t := \widehat{\mathbf{V}}^{\hat{\boldsymbol{\xi}}_n,-1/2}_{t|t-1}\hat{\boldsymbol{\xi}}_t$, where $\widehat{\mathbf{V}}^{\hat{\mathbf{u}},1/2}_{t|t-1}$ and $\widehat{\mathbf{V}}^{\hat{\boldsymbol{\xi}}_n,1/2}_{t|t-1}$ are the lower triangular Cholesky factorization of $\widehat{\mathbf{V}}^{\hat{\mathbf{u}}}_{t|t-1}$ and $\widehat{\mathbf{V}}^{\hat{\boldsymbol{\xi}}_n}_{t|t-1}$, respectively.
- **Step 11'.** Generate a bootstrap sample $(\boldsymbol{\nu}^*_1, \boldsymbol{\eta}^*_1), \ldots, (\boldsymbol{\nu}^*_B, \boldsymbol{\eta}^*_B)$ of size B from the devolatilized common shocks and idiosyncratic components $(\boldsymbol{\nu}_t, \boldsymbol{\eta}_t)$; letting $\widehat{\boldsymbol{\chi}}^{i*}_{T+1} := \widehat{\mathbf{B}}_n(L)\hat{\mathbf{u}}^{i*}_{T+1}$ with $\hat{\mathbf{u}}^*_{T+1} := \widehat{\mathbf{V}}^{\hat{\mathbf{u}},1/2}_{T+1|T}\boldsymbol{\nu}^*_i$ and $\hat{\boldsymbol{\xi}}^{i*}_{T+1} := \widehat{\mathbf{V}}^{\hat{\boldsymbol{\xi}}_n,1/2}_{T+1|T}\boldsymbol{\eta}^*_i$, construct B one-step-ahead vector returns $\mathbf{r}^{i*}_{T+1} := \widehat{\boldsymbol{\chi}}^{i*}_{T+1} + \hat{\boldsymbol{\xi}}^{i*}_{T+1}$. This, for portfolio weights $\boldsymbol{\omega} = (\omega_1, \ldots \omega_N)'$, this yields the B simulated one-step-ahead portfolio return forecasts $R^{i*}_{T+1} := \boldsymbol{\omega}'\mathbf{r}^{i*}_{T+1}$ for $i = 1, \ldots, B$.

Step 12 remains unchanged. This procedure—call it GDFM-CHF-Boot—still requires Cholesky factorizations and matrix inversions for all t; the matrices involved, however, now are $q \times q$, where q typically is small; only the high-dimensional $N \times N$ diagonal matrix of estimated idiosyncratic covariances is to be inverted, which, of course, does not cause any dimensionality problem.

4. Backtesting

Backtesting is the widely-used set of statistical tools allowing for the assessment of confidence set and quantile-related estimation accuracy—in this case, the accuracy of VaR and ES forecasting. The empirical backtesting exercise we are conducting in Section 5 below is based on calibration tests and scoring functions, which we first briefly describe.

4.1. Calibration tests

Calibration tests are hypothesis testing methods aiming at the evaluation of VaR and ES forecasts accuracy. Different calibration tests are constructed taking into account different characteristics expected from the correct VaR and ES forecasts. In all cases, the null hypothesis can be interpreted as “the VaR/ES is correctly specified”, where the term “correct” is defined in various manners. If the null is not rejected, the risk measure estimation procedure is considered adequate.

4.1.1. VaR Calibration tests

Call *violation* the occurrence of a portfolio return smaller than the corresponding estimated VaR.

More precisely, in a rolling window scheme, a sequence of violations is defined as

$$I^\alpha_t := \mathbb{I}[r_t < \widehat{\mathrm{VaR}}^\alpha_t], \qquad t = T+1, \ldots, T+H, \tag{8}$$

where r_t stands for a series of out-of-sample returns—individual stock returns or portfolio returns depending on the context, H stands for the length of the out-of-sample period (equivalently, the number of rolling windows) and T denotes the size of the training sample in the first window.

Let $V_H := \sum_{t=T+1}^{T+H} I^\alpha_t$ denote the number of violations in the out-of sample period: if the VaR forecasts $\widehat{\mathrm{VaR}}^\alpha_{T+1}$ were accurate, the observed proportion V_H/H of violations should be close to the nominal risk level α.

The unconditional coverage (UC) test of Kupiec (1995) is among the most popular methods in backtesting, and addresses the null hypothesis $\mathcal{H}^\alpha_0$ under which the violations (8) constitute an i.i.d. Bernoulli sequence with parameter α. The UC test statistic is usually defined as

$$\mathrm{UC}^\alpha = -2\log[L(\alpha)/L(V_H/H)], \tag{9}$$

where $L(p) = (1-p)^{H-V_H}p^{V_H}$ is the likelihood function of a binomial distribution with parameter p. The null distribution of that statistic is asymptotically χ^2_1 as $H \to \infty$ and chi-square quantiles are usually used as critical values to perform the test.

The UC test, however, is insensitive to alternatives of serial dependence among the violations (such as clustered violations, indicating positive serial dependence).

In order to deal with such alternatives, Christoffersen (1998) developed the so-called conditional coverage (CC) test which is based on a first-order Markov chain model with transition matrix

$$\boldsymbol{\Pi}_1 := \begin{pmatrix} 1-\pi_{01} & \pi_{01} \\ 1-\pi_{11} & \pi_{11} \end{pmatrix}, \tag{10}$$

where π_{01} is the probability of a violation at time $t+1$ conditional on no violation at time t, π_{11} the probability of a violation at time $t+1$ conditional on a violation taking place at time t; under $\mathcal{H}^\alpha_0$ (the violations are i.i.d. Bernoulli with parameter α), one has $\pi_{01} = \pi_{11} = \alpha$. Letting

$$T_{ij} := \sum_{t=T+1}^{T+H-1} \mathbb{I}[I^\alpha_t = i \text{ and } I^\alpha_{t+1} = j], \quad i, j =\in \{0, 1\},$$

the CC test statistic is

$$\mathrm{CC}^\alpha := -2\log\left[L(\alpha)/L_{\widehat{\boldsymbol{\Pi}}_1}\right], \tag{11}$$

where $L(\alpha)$, as in (9), is the likelihood under $\mathcal{H}_0^\alpha$,

$$L_{\Pi_1} := (1-\pi_{01})^{T_{00}}\pi_{01}^{T_{01}}(1-\pi_{11})^{T_{10}}\pi_{11}^{T_{11}},$$

the likelihood under the alternative of a Markov chain with transition matrix (10), and

$$\widehat{\mathbf{\Pi}}_1 := \begin{pmatrix} 1-T_{01}/V_{H-1} & T_{01}/V_{H-1} \\ 1-T_{11}/(H-1-V_{H-1}) & T_{11}/(H-1-V_{H-1}) \end{pmatrix},$$

the maximum likelihood estimator version of $\mathbf{\Pi}_1$.

The dynamic quantile (DQ) test proposed by Engle and Manganelli (2004) is based on the concept of *hits*. A *hit* at time $t \in \{T+1,\dots,T+H\}$ is defined as $\text{Hit}_t^\alpha := I_t^\alpha - \alpha$. Under $\mathcal{H}_0^\alpha$, the sequence of hits is a martingale difference sequence, hence has expected value zero and is uncorrelated with both its own lagged values and the estimated VaR. The DQ test is a Wald-type test of the hypothesis that all coefficients in the linear regression

$$\text{Hit}_t^\alpha = \beta_0 + \sum_{i=1}^{L}\beta_i \text{Hit}_{t-i}^\alpha + \beta_{L+1}\widehat{\text{VaR}}_{t-1}^\alpha + \epsilon_t, \tag{12}$$

are equal to zero. The DQ test statistic is

$$DQ^\alpha := H^{-1}\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}}/\alpha(1-\alpha), \tag{13}$$

where $\hat{\boldsymbol{\beta}} := (\hat\beta_0,\dots,\hat\beta_{L+1})'$ is the vector of OLS estimators of $\boldsymbol{\beta} := (\beta_0,\dots,\beta_{L+1})'$, $\mathbf{X}$ is the $H\times(L+2)$ matrix with columns $(1, \text{Hit}_{t-1}^\alpha,\dots,\text{Hit}_{t-L}^\alpha, \widehat{\text{VaR}}_{t-1}^\alpha)'$ for $t = T+2,\dots,T+H$; the number L of lags is commonly set as $L=4$. Critical values are obtained from the asymptotic (as $H\to\infty$) null distribution of DQ^α, which is chi-square with $L+2$ degrees of freedom.

Finally, based on the fact that the VaR actually is a quantile of the conditional distribution of returns, Gaglianone et al. (2011) proposed the VaR quantile (VQ) regression test in which the null is that the coefficients in the α-quantile regression

$$r_t = \beta_0 + \beta_1\widehat{\text{VaR}}_t^\alpha + \epsilon_t, \tag{14}$$

are $\beta_0 = 0$ and $\beta_1 = 1$. The test statistic is

$$\text{VQ}^\alpha := \hat{\boldsymbol{\theta}}'\mathbf{M}^{-1}\hat{\boldsymbol{\theta}}, \tag{15}$$

where $\hat{\boldsymbol{\theta}} := (\hat\beta_0, \hat\beta_1 - 1)'$ is the OLS estimator of $\boldsymbol{\theta} := (\beta_0, \beta_1-1)'$ and $\mathbf{M}$ is $\hat{\boldsymbol{\theta}}$'s covariance matrix. Critical values are obtained from the asymptotic (as $H\to\infty$) null distribution of VQ^α, which is chi-square with two degrees of freedom.

4.1.2. ES Calibration tests

Turning to expected shortfall, the most common calibration test is the *exceedance residual* (ER) test of McNeil and Frey (2000). It is based on the concept of *standardized residual exceedances* over the VaR, namely,

$$e_t := \frac{r_t - \text{ES}_t^\alpha}{\sigma_t}\mathbb{I}[r_t < \text{VaR}_t^\alpha], \text{ for } t = T+1,\dots,T+H. \tag{16}$$

The null hypothesis $\mathcal{K}_0^\alpha$ to be tested is $\text{E}(e_t|r_t < \text{VaR}_t^\alpha) = 0$, with one-sided alternatives of the form $\text{E}(e_t|r_t < \text{VaR}_t^\alpha) < 0$. Rejecting $\mathcal{K}_0^\alpha$ thus implies that the risk is underestimated (a dangerous situation). The sample counterpart $\hat e_t$ of e_t is obtained by replacing ES_t^α, VaR_t^α, and σ_t by their estimators $\widehat{\text{ES}}_t^\alpha$, $\widehat{\text{VaR}}_t^\alpha$, and $\hat\sigma_t$, yielding the test statistic

$$\text{ER}^\alpha := \sqrt{V_H}m_{\hat e}/s_{\hat e}, \tag{17}$$

where V_H is the number of VaR violations as defined in Section 4.1.1, $m_{\hat e}$ and $s_{\hat e}$ the sample mean and sample standard deviation of $\{\hat e_t | I_t^\alpha = 1\}$—the subsample of $\hat e_t$ values at which violations occur. The null distribution of ER^α is obtained via non-parametric bootstrap.

Another popular calibration test is the *conditional calibration* test (CoC) proposed by Nolde et al. (2017). The idea consists in checking whether the forecasts $\widehat{\text{VaR}}_t^\alpha$ and $\widehat{\text{ES}}_t^\alpha$ are "optimal." Optimality here means satisfying

$$\text{E}(\mathbf{W}(\widehat{\text{VaR}}_t^\alpha, \widehat{\text{ES}}_t^\alpha, r_t)|\mathcal{F}_{t-1}) = \mathbf{0},$$

for some adequate (vector-valued) function $\mathbf{W}$ (see Nolde et al., 2017). Here, we use the $\mathbf{W}$ function

$$\mathbf{W}(R_1, R_2, x) := \begin{pmatrix} \alpha - \mathbb{I}(x < R_1) \\ R_2 - R_1 + \mathbb{I}(x<R_1)(R_1 - x)/\alpha \end{pmatrix}. \tag{18}$$

The CoC test statistic, of the Wald type, is

$$\text{CoC}^\alpha := H\left(\frac{1}{H}\sum_{t=T+1}^{T+H}\mathbf{W}(\widehat{\text{VaR}}_t^\alpha, \widehat{\text{ES}}_t^\alpha, r_t)\right)'\hat{\mathbf{\Omega}}_H^{-1}\left(\frac{1}{H}\sum_{t=T+1}^{T+H}\mathbf{W}(\widehat{\text{VaR}}_t^\alpha, \widehat{\text{ES}}_t^\alpha, r_t)\right), \tag{19}$$

where

$$\widehat{\mathbf{\Omega}}_H := \frac{1}{H}\sum_{t=T+1}^{T+H}\Big(\mathbf{W}(\widehat{\mathrm{VaR}}_t^\alpha, \widehat{\mathrm{ES}}_t^\alpha, r_t)\Big)\Big(\mathbf{W}(\widehat{\mathrm{VaR}}_t^\alpha, \widehat{\mathrm{ES}}_t^\alpha, r_t)\Big)';$$

its (asymptotic, as $H \to \infty$) null distribution is chi-square with two degrees of freedom.

More recently, Bayer and Dimitriadis (2020) proposed yet another approach, the so-called *Expected Shortfall Regression* (ESR) method. That method considers (Dimitriadis et al., 2019; Patton et al., 2019) a double regression model of the form (similar to (14))

$$r_t = \mathbf{V}_t'\boldsymbol{\beta} + u_t^q \quad \text{and} \quad r_t = \mathbf{W}_t'\boldsymbol{\gamma} + u_t^e, \quad t = T+1, \ldots, T+H \tag{20}$$

where the conditional α-quantile of u_t^q given $u_{t-1}^q, u_{t-2}^q, \ldots$ (viz., the level-α VaR_t^α associated with u_t^q) and the level-α expected shortfall ES_t^α associated with u_t^e are zero (the sigma-fields generated by $u_{t-1}^q, u_{t-2}^q, \ldots$ and $u_{t-1}^e, u_{t-2}^e, \ldots$ actually coincide); $\mathbf{V}_t$ and $\mathbf{W}_t$ here are vectors of covariates.

The Wald-type test statistic for a null hypothesis of the form $H_0 : \boldsymbol{\gamma} = \boldsymbol{\gamma}_0$ is

$$H(\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0)\widehat{\mathbf{\Omega}}_{\hat{\boldsymbol{\gamma}}}^{-1}(\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0)', \tag{21}$$

where $\hat{\boldsymbol{\gamma}}$ is a root-H consistent estimator of $\boldsymbol{\gamma}$ and $\widehat{\mathbf{\Omega}}_{\hat{\boldsymbol{\gamma}}}$ a consistent estimator of $\hat{\boldsymbol{\gamma}}$'s covariance matrix; its null distribution is asymptotically (as $H \to \infty$) chi-square with q degrees of freedom (q the dimension of $\boldsymbol{\gamma}$, which is either one or two below).

The authors proposed three different versions of the ESR test. Denoting by $\widehat{\mathrm{VaR}}_t^\alpha$ and $\widehat{\mathrm{ES}}_t^\alpha$ the VaRs and ESs estimated in the previous sections, the first version, called the *strict* ESR test, is obtained for $\mathbf{V}_t = \mathbf{W}_t = (1, \widehat{\mathrm{ES}}_t^\alpha)'$. The second version, called *auxiliary* ESR test, is obtained with $\mathbf{V}_t = (1, \widehat{\mathrm{VaR}}_t^\alpha)'$ and $\mathbf{W}_t = (1, \widehat{\mathrm{ES}}_t^\alpha)'$. In both cases, $\boldsymbol{\gamma}_0 = (0, 1)'$ is used. These two versions are intrinsically two-sided, hence provide no information about under- or over-estimation of VaR and ES. A third version of ESR, called *intercept* ESR test, has been proposed to take care of that drawback. The regression equations (20) are replaced with

$$r_t - \widehat{\mathrm{ES}}_t^\alpha = \beta_0 + \beta_1\widehat{\mathrm{ES}}_t^\alpha + u_t^q \quad \text{and} \quad r_t - \widehat{\mathrm{ES}}_t^\alpha = \gamma_1 + u_t^e, \quad t = T+1, \ldots, T+H$$

and the problem consists in testing $H_0 : \gamma_1 \geq 0$ against $H_1 : \gamma_1 < 0$. The test statistic is $H^{1/2}\hat{\gamma}_1/\hat{\Omega}_{\hat{\gamma}_1}$ which under the null is asymptotically standard normal.

Denote by ESR_1, ESR_2, and ESR_3, respectively, these three versions of the ESR method.

4.2. Scoring functions

It often happens that several forecast procedures successfully pass (no rejection) the VaR and ES calibration tests. A criterion then is needed to choose the "best one(s)". This is the purpose of *scoring functions* (also known as *loss functions*): the smaller the scoring function, the better the forecasting procedure.

A scoring function S is called *strictly consistent* for a risk measure $\rho(X)$ if

$$\mathrm{E}(S(\rho(X), X)) < \mathrm{E}(S(r(X), X)) \quad \text{for all} \quad r(X) \neq \rho(X);$$

the risk measure $\rho(X)$ is called *elicitable* if it admits a strictly consistent scoring function.

The VaR is an elicitable risk measure and VaR_t^α admits the strictly consistent scoring function

$$S(\mathrm{VaR}_t^\alpha, r_t) = (\alpha - \mathbb{I}[r_t \leq \mathrm{VaR}_t^\alpha])(G(r_t) - G(\mathrm{VaR}_t^\alpha)), \tag{22}$$

where α is the risk level, $\mathbb{I}[\cdot]$ is the indicator function, and $G(\cdot)$ is a strictly increasing function. A commonly used scoring function for the VaR is the quantile loss (QL) function (González-Rivera et al., 2004), which is obtained when $G(\cdot)$ is the identity function.

Contrary to the VaR, the ES is not elicitable but, as showed by Fissler et al. (2016), the pair (VaR, ES) is. Consistent scoring functions for (VaR_t^α, ES_t^α) take the form

$$\begin{aligned} S\big((\mathrm{VaR}_t^\alpha, \mathrm{ES}_t^\alpha), r_t\big) =& \big(\mathbb{I}[r_t \leq \mathrm{VaR}_t^\alpha] - \alpha\big)G_1(\mathrm{VaR}_t^\alpha) - \mathbb{I}[r_t \leq \mathrm{VaR}_t^\alpha]\,G_1(r_t) \\ &+ G_2(\mathrm{ES}_t^\alpha)\big(\mathrm{ES}_t^\alpha - \mathrm{VaR}_t^\alpha + \mathbb{I}[r_t \leq \mathrm{VaR}_t^\alpha]\big)(\mathrm{VaR}_t^\alpha - r_t)/\alpha + G_3(\mathrm{ES}_t^\alpha) + G_4(r_t), \end{aligned} \tag{23}$$

provided that the functions G_1–G_4 satisfy some conditions for which we refer to Fissler et al. (2016), Fissler and Ziegel (2016), and Taylor (2020).

In Section 5, we are using three distinct scoring functions for a joint evaluation of VaR and ES forecasts. The first one, denoted as FZG, was proposed by Fissler et al. (2016) and is obtained for

$$G_1(x) = x, \quad G_2(x) = e^x/(1+e^x), \quad G_3(x) = \log(1+e^x), \quad \text{and} \quad G_4(x) = \log(2).$$

The second one, denoted as NZ and proposed by Nolde et al. (2017), with

$$G_1(x) = 0, \quad G_2(x) = \sqrt{-x}/2, \quad G_3(x) = -\sqrt{-x}, \quad \text{and} \quad G_4(x) = 0.$$

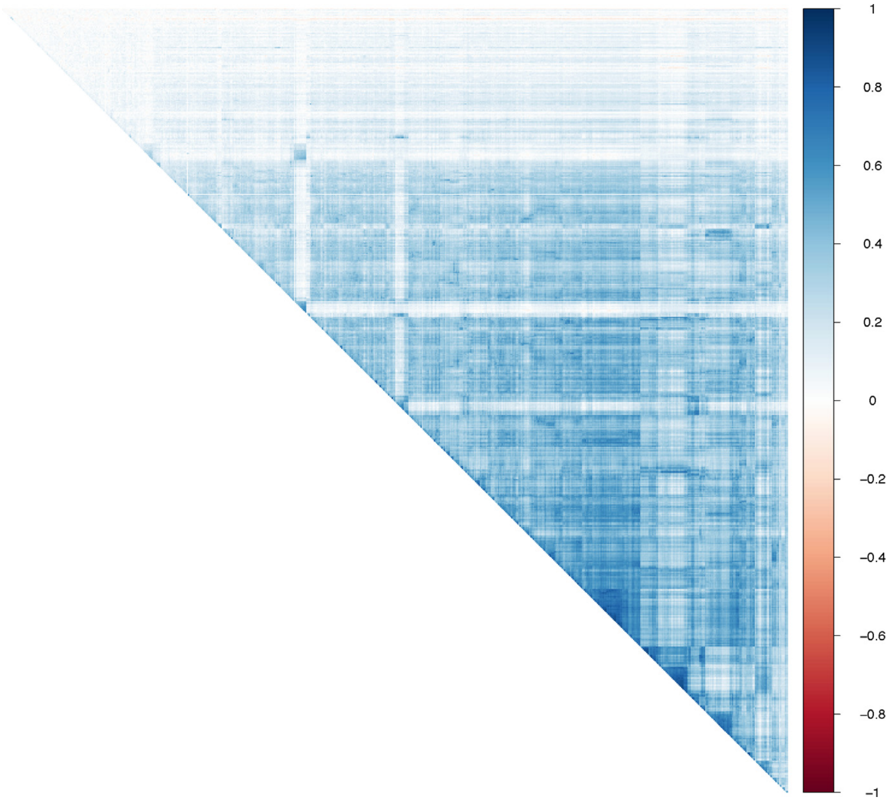

Fig. 1. Sample unconditional contemporaneous correlations between stock returns in Section 5.

The last one, denoted as AL was proposed by Taylor (2019) and is obtained for

$$G_1(x) = 0, \quad G_2(x) = -1/x, \quad G_3(x) = -\log(-x), \quad \text{and} \quad G_4(x) = 1 - \log(1-\alpha).$$

As previously mentioned, a pair (VaR, ES) failing to reject the null hypotheses in calibration tests and achieving the smallest scoring function value is preferable.

5. Empirical results

5.1. Data

Our dataset consists of daily closing prices of stocks used in the composition of the Standard & Poor's 500 Index (S&P 500), the National Association of Securities Dealers Automated Quotations 100 Index (NASDAQ-100), and the NYSE Amex Composite Index (AMEX) from January 3, 2012 to July 1, 2020 ($T = 2136$ observations). Only series with no missing values in the entire sample period were considered, yielding a panel of $N = 655$ stocks.

Daily returns (in percentage) are obtained as $r_{i,t} := 100 \times (P_{i,t}/P_{i,t-1} - 1)$ where $P_{i,t}$ denotes the closing price of the ith stock at time t. Following Engle et al. (2019), if pairs of stock returns are highly correlated (sample correlation larger than 0.95), the one with lower volume is removed from the panel. Using this criterion, the stocks DISCK, GOOGL and LBTYA were removed from the panel. Figure 1 plots the sample correlation among the 652 stock returns over the full sample period, the highest correlation is 0.91 between CMS and XEL stocks while the lowest one is -0,19 between the stocks APT and CCL.

5.2. Out-of-sample analysis

For the sake of simplicity, we throughout consider an equal-weight portfolio re-balanced on a daily basis, with returns $R_t = \sum_{i=1}^{652} \omega_i r_{i,t} = (1/652) \sum_{i=1}^{652} r_{i,t}$. We consider the one-step-ahead forecasts of the VaR and ES of this portfolio at risk levels α =1%, 2.5%, and 5% using the methodologies described in Sections 3.1 and 3.2 over a rolling window of 750 days (with concentration ratio $652/750 = 0.87$): observations 1– 750 thus were used to forecast VaR and ES, as explained in Section 3.1, at time 751, then observations 2–751 to forecast VaR and ES at time 752,..., observations 1386–2135 to forecast VaR and ES at time 2136. This produces a total of 1386 forecasts, from December 29, 2014 through July 1, 2020.

Table 1
One-step-ahead VaR and ES backtesting at 1% (top panel), 2.5% (middle panel), and 5% (bottom panel) risk levels. Out-of-sample period from December 29, 2014 to July 1, 2020 (1386 observations).

			Calibration tests									Avg. Scoring functions			
		Hits	UC	CC	DQ	VQ	ER	CoC	ESR_1	ESR_2	ESR_3	QL	FZG	NZ	AL
1%	RM2006	0.60	0.161	0.051	0.091	0.331	0.493	0.219	0.344	0.218	0.706	0.038	**0.673**	**1.850**	**2.178**
	DCCc	1.70	0.024	0.000	0.000	0.434	0.404	0.106	0.516	0.505	0.278	**0.039**	**0.670**	**1.863**	**2.181**
	ABC	1.90	0.003	0.001	0.000	0.258	0.024	0.052	0.016	0.043	0.036	**0.042**	**0.698**	**1.991**	**2.344**
	GDFM-CHF	1.40	0.189	0.224	0.266	0.726	0.564	0.493	0.500	0.576	0.180	**0.032**	**0.653**	**1.730**	**2.060**
	GDFM-CHF-Boot	1.20	0.413	0.317	0.209	0.924	0.298	0.705	0.398	0.527	0.120	**0.033**	**0.658**	**1.737**	**2.076**
2.5%	RM2006	1.30	0.002	0.000	0.000	0.001	0.522	0.000	0.009	0.011	0.972	0.075	0.675	1.654	1.975
	DCCc	3.00	0.221	0.003	0.000	0.000	0.087	0.129	0.013	0.165	0.018	0.075	0.666	1.642	1.951
	ABC	3.10	0.166	0.014	0.000	0.420	0.001	0.227	0.016	0.325	0.256	0.077	0.677	1.680	2.007
	GDFM-CHF	2.70	0.689	0.602	0.362	0.437	0.230	0.549	0.421	0.834	0.135	**0.066**	**0.631**	**1.535**	**1.816**
	GDFM-CHF-Boot	3.20	0.088	0.215	0.221	0.324	0.529	0.278	0.406	0.921	0.207	**0.066**	0.636	1.545	1.834
5%	RM2006	2.50	0.000	0.000	0.000	0.000	0.416	0.000	0.000	0.438	0.993	0.123	0.681	1.493	1.797
	DCCc	5.50	0.416	0.000	0.000	0.343	0.037	0.077	0.029	0.035	0.007	**0.121**	0.666	1.473	1.759
	ABC	5.50	0.416	0.024	0.000	0.588	0.000	0.324	0.906	0.010	0.409	**0.119**	0.668	1.478	1.773
	GDFM-CHF	5.60	0.351	0.147	0.207	0.354	0.369	0.405	0.296	0.766	0.072	**0.110**	**0.630**	**1.396**	**1.649**
	GDFM-CHF-Boot	5.90	0.128	0.117	0.338	0.413	0.333	0.225	0.164	0.684	0.055	**0.111**	0.635	1.407	1.665

For the sake of comparison, we are basing our FHS forecasts on various conditional covariance estimators $\widehat{\mathbf{H}}_t$ which were selected for their feasibility in a high-dimensional context (the dimension of the observations here is 652). Specifically, we include the Risk Metrics 2006 methodology (Zumbach, 2007) denoted as RM2006, the dynamic conditional correlation with composite likelihood (Pakel et al., 2020) denoted as DCCc, the conditional covariance estimation using the dynamic factor model approach with finite-dimensional space (Alessi et al., 2009; Aramonte et al., 2013) denoted as ABC; GDFM-CHF stands for the method we are proposing. Whenever univariate volatility models are needed (typically, for the prediction of idiosyncratic volatilities and the marginals of the MGARCH model in the common shocks), the GJR-GARCH model of Glosten et al. (1993) with Student-*t* distribution was used. The GDFM-CHF-Boot procedure described in Section 3.2 is also included in the comparison.

Table 1 reports the results of the out-of-sample VaR and ES backtesting exercise: the *p*-values are reported for the various forecasting methods (RM2006, DCCc, ABC, GDFM-CHF, and GDFM-CHF-Boot) and calibration tests (UC, CC, DQ, and VQ; ER, CoC, ESR_1-ESR_3) described in Section 4.1. Shadowed cells indicate *p*-values larger than 0.05 (non-rejection at nominal level 5%); boldface scoring values indicate the best models (achieving lowest scoring values) identified via the model confidence set approach of (Hansen et al., 2011) at significance level 25%.

All procedures report percentages of violations that are close to the nominal risk levels. Regarding VaR forecasts, the GDFM-CHF and GDFM-Boot methods are doing better than its competitors: at significance level 5%, out of four distinct calibration tests (UC, CC, DQ, VQ), not a single rejection, while all other procedures get rejected at least once, indicating that, irrespective of the calibration method adopted, GDFM-CHF and GDFM-CHF-Boot VaR forecasts are more accurate than RM2006, DCCc, and ABC. Looking at the average QL scoring function, which only takes into account the performance of VaR prediction, the RM2006 is the only procedure that never lies in the model confidence set, while the GDFM-CHF and GDFM-CHF-Boot methods both belong to the model confidence set in all cases; in that respect, this, they perfotm equally well. For ES forecasts, only the GDFM-CHF and the GDFM-CHF-Boot methods are achieving non-rejection in all ES calibration tests (ER, CoC, ESR_1, ESR_2, ESR_3) and for all risk levels considered, confirming the excellent performance of our approaches.

As for scoring function evaluations, irrespective of the choice of the scoring function, the lowest average values are achieved by the GDFM-CHF VaR and ES estimates, closely followed by the GDFM-CHF-Boot ones. In some cases the GDFM-CHF is the only procedure belonging to the set of best models identified by the model confidence set, meaning that the GDFM-CHF procedure is performing best. Overall, these results support evidence of the GDFM-CHF superiority, closely followed by GDFM-CHF-Boot, which in higher dimension could be preferred for being computationally less costly.

Figure 2 plots the out-of-sample equal-weight portfolio returns along with their one-step-ahead VaR forecasts at risk levels $\alpha = 1\%$ (top panel), 2.5% (middle panel), and 5% (bottom panel), for the four methods (RM2006, DCCc, ABC, GDFM-CHF, and GDFM-CHF-Boot). The vertical dotted line stands for March 12, 2020, one day after the COVID-19 virus has been declared a pandemic by the World Health Organisation. Note that, after March 12, 2020, all VaR estimates are strongly affected by the large March 12 shock, the VaR RM2006 estimate being the most severely impacted. As a consequence, the VaR forecasts are larger than should be for quite some time. Figure 3 reports the same information as in Figure 2, but over a smaller out-of-sample COVID-free period ranging from January 2, 2018 to June 28, 2019.

The early February 2018 VaR spike observed in the DCCc, ABC, GDFM-CHF, and GDFM-CHF-Boot column takes place when the stock market peaked in bond rates that were pricing at the risk of inflation (Egan, 2018). In that period the

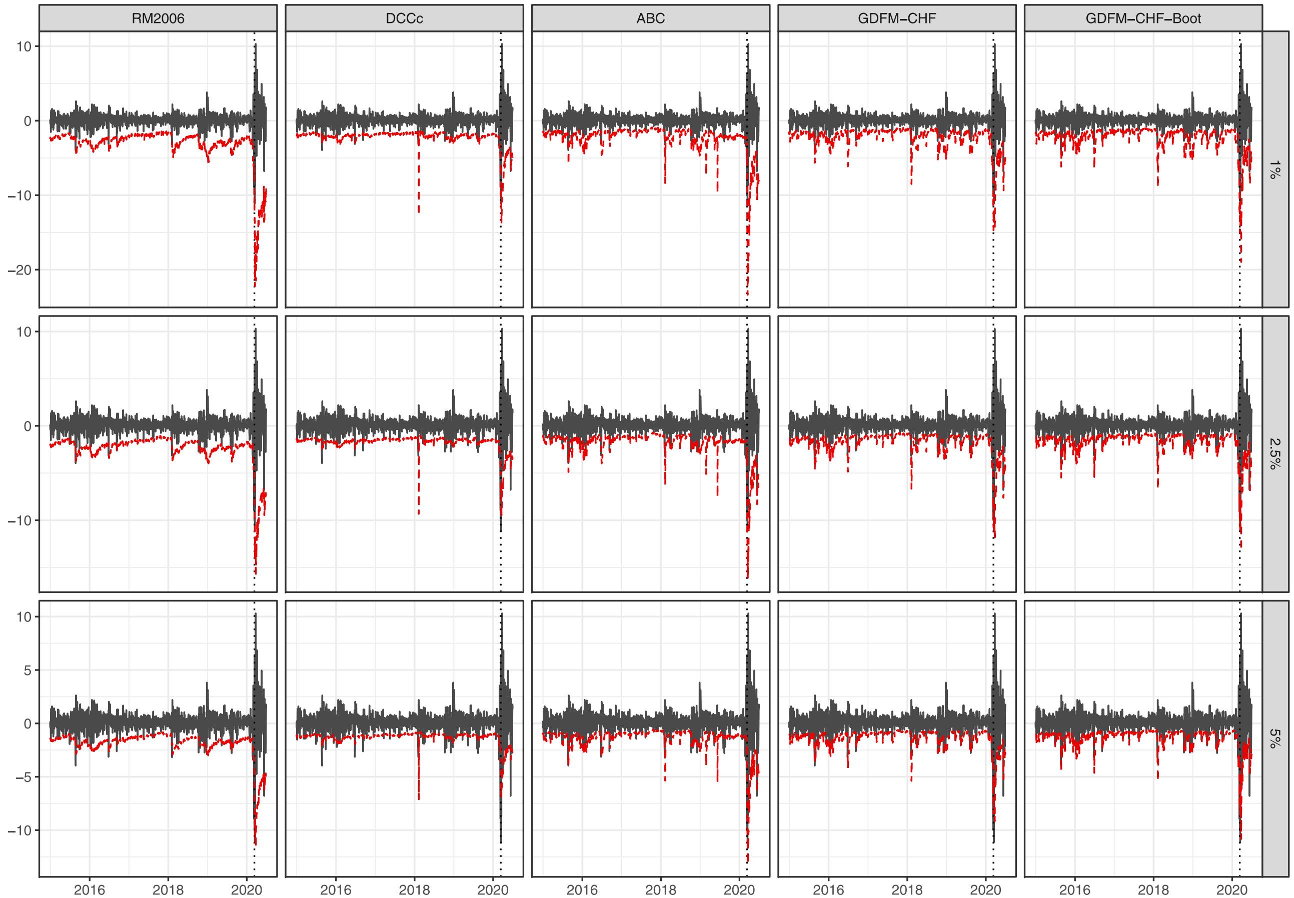

Fig. 2. Out-of-sample portfolio returns (black solid line) and the one-step-ahead forecasts VaR (red dashed line) at risk levels α =1% (top panel), 2.5% (middle panel), and 5% (bottom panel) based on the RM2006, DCCc, ABC, GDFM-CHF, and GDFM-CHF-Boot procedures, respectively, from January 2nd, 2018 through July 1st, 2020.

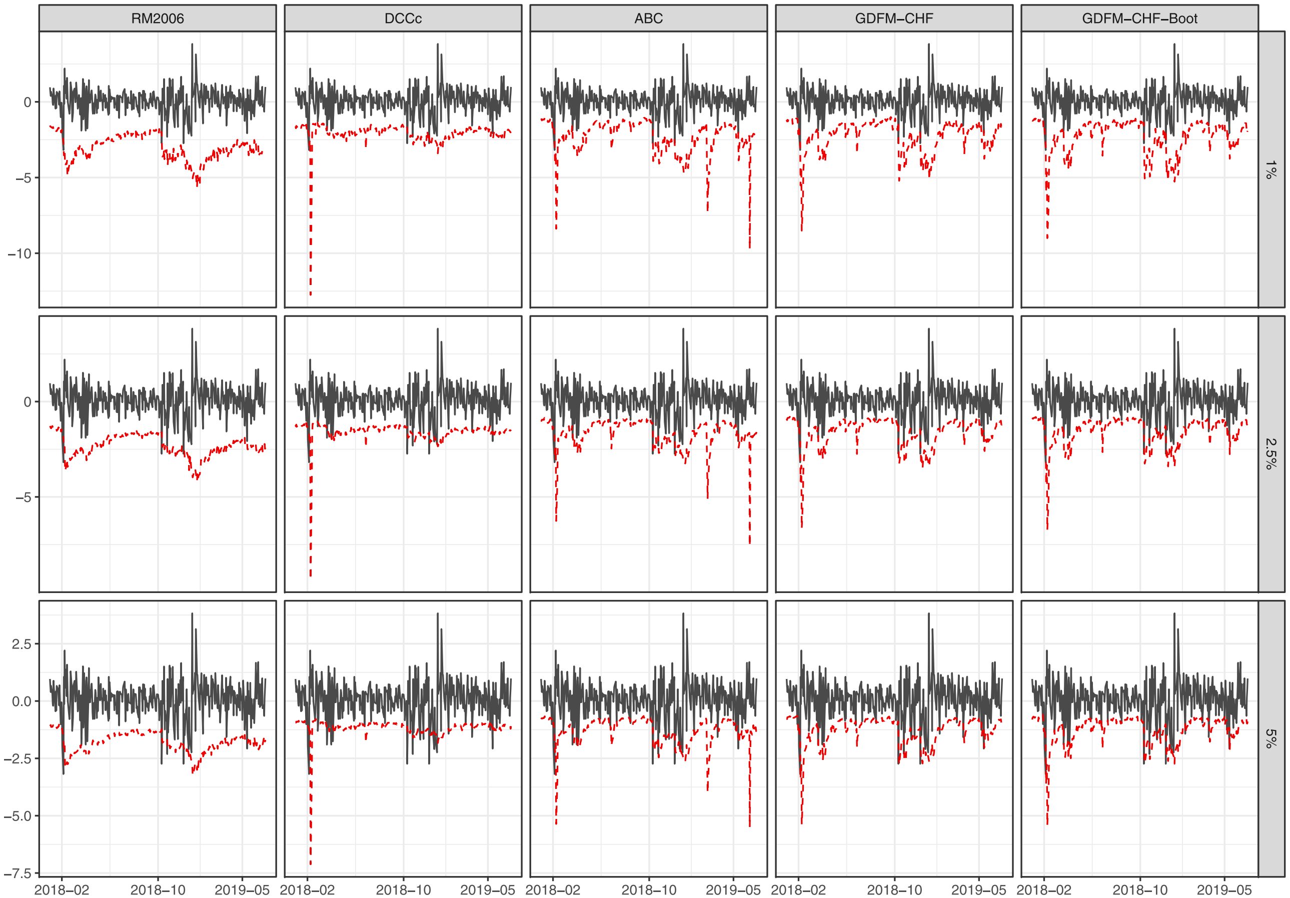

Fig. 3. Out-of-sample portfolio returns (black solid line) and the one-step-ahead VaR forecasts (red dashed line) at risk levels $\alpha = 1\%$ (top panel), 2.5% (middle panel), and 5% (bottom panel) based on the RM2006, DCCc, ABC, GDFM-CHF, and GDFM-CHF-Boot procedures, respectively, from January 2, 2018 through June 28, 2019.

Table 2
One-step-ahead VaR and ES backtesting at 1% (top panel), 2.5% (middle panel), and 5% (bottom panel) risk levels. Out-of-sample period from December 29, 2014 to March 11, 2020.

			Calibration tests									Avg. Scoring functions			
		Hits	UC	CC	DQ	VQ	ER	CoC	ESR_1	ESR_2	ESR_3	QL	FZG	NZ	AL
1%	RM2006	0.70	0.229	0.069	0.108	0.285	0.493	0.383	0.266	0.158	0.741	0.032	**0.664**	1.743	**2.093**
	DCCc	1.50	0.075	0.000	0.000	0.153	0.398	0.256	0.171	0.135	0.168	**0.032**	**0.660**	**1.737**	**2.085**
	ABC	1.80	0.007	0.005	0.000	0.149	0.135	0.076	0.027	0.004	0.010	**0.035**	**0.687**	**1.834**	**2.214**
	GDFM-CHF	1.40	0.197	0.221	0.255	0.537	0.689	0.491	0.514	0.654	0.407	**0.028**	**0.643**	**1.635**	**1.971**
	GDFM-CHF-Boot	1.10	0.604	0.328	0.565	0.743	0.350	0.838	0.555	0.537	0.298	**0.028**	**0.650**	**1.649**	**1.997**
2.5%	RM2006	1.20	0.001	0.000	0.000	0.000	0.331	0.000	0.091	0.005	0.819	0.063	0.659	1.559	1.891
	DCCc	2.90	0.362	0.010	0.000	0.625	0.129	0.288	0.033	0.061	0.079	0.062	0.648	1.541	1.863
	ABC	3.10	0.158	0.046	0.000	0.501	0.007	0.211	0.021	0.043	0.002	0.064	0.656	1.567	1.898
	GDFM-CHF	2.60	0.823	0.559	0.539	0.930	0.324	0.832	0.626	0.886	0.183	**0.056**	**0.612**	**1.447**	**1.730**
	GDFM-CHF-Boot	3.10	0.158	0.357	0.343	0.566	0.599	0.422	0.418	0.800	0.170	0.056	0.618	1.457	1.750
5%	RM2006	2.40	0.000	0.000	0.000	0.000	0.374	0.000	0.000	0.000	1.000	0.103	0.653	1.404	1.711
	DCCc	5.30	0.655	0.000	0.000	0.031	0.051	0.170	0.019	0.023	0.036	0.101	0.639	1.386	1.675
	ABC	5.60	0.347	0.045	0.001	0.614	0.002	0.248	0.414	0.258	0.157	0.101	0.639	1.387	1.677
	GDFM-CHF	5.30	0.568	0.104	0.078	0.677	0.404	0.724	0.619	0.869	0.188	**0.093**	**0.604**	**1.319**	**1.569**
	GDFM-CHF-Boot	5.50	0.413	0.241	0.509	0.493	0.278	0.622	0.359	0.821	0.096	**0.093**	**0.606**	**1.322**	**1.578**

VIX Volatility Index doubled, reflecting the investors' sentiment. As expected, volatility forecasts also are impacted, yielding largely overestimated VaR forecasts.

On the other hand, the 2019 VaR spikes in the ABC column are due to the presence of extreme observation in the in-sample period, which cause bad estimation of the common shocks. Although the influence of outliers in the GDFM estimation process when the space spanned by the common factor is finite have not been analyzed yet, our own analysis (not shown here) in such case reveals that the common shocks present large peaks and spikes at some points of the in-sample period and very small volatility in the remaining periods. This common shock behaviour is very similar to that observed in the supplementary material of Trucíos et al. (2020) when dealing with outliers in the GDFM with infinite-dimensional factor space. The common shock behaviour in the ABC procedure, thus, can be explained by the presence of outliers in the in-sample period.

The consequences of neglecting the possible presence of additive outliers in the volatility estimation process in most of the high-dimensional conditional covariance matrix estimators proposed recently have not been analyzed yet, but it is widely known that additive outliers badly affect volatility estimation in other contexts (univariate and moderate-dimensional multivariate). Therefore, the fact that extreme observations can affect the performance of estimators is hardly surprising. We refer to Boudt et al. (2013), Grané and Veiga (2014), Hotta and Trucíos (2018), Trucíos et al. (2017, 2018) and the references therein for a discussion of the influence of outliers on other volatility models.

The poor performance of alternatives procedures is not only due to the COVID-19 period (starting on March 12, 2020). Table 2 reports the same backtesting exercise as Table 1 but using out-of-sample data until March 11, 2020 (one day before the largest spike observed in the out-of-sample period). The results in all scenarios are qualitatively the same as in Table 1, confirming the superiority of the GDFM-CHF and GDFM-CHF-Boot methods over the alternative approaches.

Due to the short COVID-19 period in our sample (77 trading days), results solely for this period are not reported since the power of the statistical tests over such a short observation period could lead to misleading conclusions.

In our application, we have considered an equal-weight portfolio re-balanced on a daily basis. This re-balancing strategy implies transaction costs which may be nonnegligible and potentially make this strategy unattractive for some investors. Nevertheless, portfolios based on different strategies such as minimum variance or maximum diversification as well as portfolios with monthly or yearly re-balancing—the so-called *lazy portfolios*—also can be considered with no methodological changes but for replacing $\boldsymbol{\omega} = (1/N, \ldots, 1/N)$ with the desired portfolio weights. Note that we only use the portfolio weights to build $R^{i*}_{T+1} := \boldsymbol{\omega}'\mathbf{r}^{i*}_{T+1}$, so that $\boldsymbol{\omega}$ can be any desired combination of portfolio weights without affecting the VaR and ES estimation procedure.

6. Conclusions

In this paper, we deal with one-step-ahead risk measures (VaR and ES) forecasts in high-dimensional portfolios. We propose two methods. The first one is a filtered historical simulation method combined with the conditional covariance estimator of Trucíos et al. (2020a), itself based on the general dynamic factor model with infinite-dimensional factor space of (Forni et al., 2015; 2017). The second one is building on a residual-based bootstrap scheme, which is about 15% faster than the first proposal and, contrary to the first one, does not require the factorization and inversion of high-dimensional matrices.

These two procedures and a few selected alternatives are empirically evaluated on a large portfolio of 652 stocks returns observed over 750 trading days via a comprehensive backtesting exercise based on calibration tests and scoring functions. Our results indicate that, both for VaR and ES, our procedures are outperforming their competitors—including the ABC method based on a dynamic factor model with finite-dimensional factor space (a *static* factor model). Our filtered historical simulation method appears to be best. However, as the dimension increases, it gets numerically more costly, in which case our residual-based bootstrap proposal, which yields nearly as accurate forecasts, is the best choice. Whenever possible, thus, we suggest to use the filtered historical simulation method along with the conditional covariance matrix estimator of Trucíos et al. (2020a). When this becomes numerically unfeasible, we suggest the residual-based bootstrap alternative.

Finally, some high-dimensional conditional covariance matrix estimators seem to be more affected by extreme observations and outliers than some others. A detailed analysis of this fact as well as robust versions of those estimators are the subject of our current research.

Acknowledgment

We are grateful to two anonymous referees for a very careful reading of our manuscript. Their thoughtful remarks and suggestions have brought a number of improvements over the original version, including the residual-based bootstrap alternative to the filtered historical simulation approach. The authors also thank the Centre for Applied Research on Econometrics, Finance and Statistics (CAREFS), Brazil for their computational support. Part of this research was carried out while the second author was Visiting Professor at the Center for Mathematics, Computing and Cognition—Federal University of ABC (UFABC), Brazil, whose hospitality is gratefully acknowledged.

References

Alessi, L., Barigozzi, M., Capasso, M., 2009. Estimation and forecasting in large datasets with conditionally heteroskedastic dynamic common factors. Working Paper Series 1115. European Central Bank, Frankfurt am Main, Germany.
de Almeida, D., Hotta, L.K., Ruiz, E., 2018. MGARCH models: Trade-off between feasibility and flexibility. International Journal of Forecasting 34 (1), 45–63.
Aramonte, S., del Giudice Rodriguez, M., Wu, J., 2013. Dynamic factor value-at-risk for large heteroskedastic portfolios. Journal of Banking & Finance 37 (11), 4299–4309.
Bai, J., Ng, S., 2002. Determining the number of factors in approximate factor models. Econometrica 70 (1), 191–221.
Barigozzi, M., Hallin, M., 2016. Generalized dynamic factor models and volatilities: recovering the market volatility shocks. Econometrics Journal 19, 33–60.
Barigozzi, M., Hallin, M., 2017. Generalized dynamic factor models and volatilities: estimation and forecasting. Journal of Econometrics 201, 307–321.
Barigozzi, M., Hallin, M., 2020. Generalized dynamic factor models and volatilities: Consistency, rates, and prediction intervals. Journal of Econometrics 216, 4–34.
Barone-Adesi, G., Bourgoin, F., Giannopoulos, K., 1998. Don't look back. Risk 11, 100–104.
Barone-Adesi, G., Giannopoulos, K., Vosper, L., 1999. Var without correlations for portfolios of derivative securities. Journal of Futures Markets 19 (5), 583–602.
Bayer, S., 2018. Combining value-at-risk forecasts using penalized quantile regressions. Econometrics and Statistics 8, 56–77.
Bayer, S., Dimitriadis, T., 2020. Regression based expected shortfall backtesting. Journal of Financial Econometrics doi:10.1093/jjfinec/nbaa013.
Boudt, K., Danielsson, J., Laurent, S., 2013. Robust forecasting of dynamic conditional correlation garch models. International Journal of Forecasting 29 (2), 244–257.
Braione, M., Scholtes, N.K., 2016. Forecasting value-at-risk under different distributional assumptions. Econometrics 4 (1), 3.
Broda, S.A., Krause, J., Paolella, M.S., 2018. Approximating expected shortfall for heavy-tailed distributions. Econometrics and Statistics 8, 184–203.
Caporin, M., McAleer, M., 2014. Robust ranking of multivariate GARCH models by problem dimension. Computational Statistics & Data Analysis 76, 172–185.
Chang, J., Guo, B., Yao, Q., et al., 2018. Principal component analysis for second-order stationary vector time series. The Annals of Statistics 46 (5), 2094–2124.
Chevallier, J., 2012. Time-varying correlations in oil, gas and CO^2 prices: an application using BEKK, CCC, and DCC-MGARCH models. Applied Economics 44 (32), 4257–4274.
Christoffersen, P., 2009. Value–at–Risk Models. In: Andersen, T.G., Davis, R.A., Kreiß, J.-P., Mikosch, T.V. (Eds.), Handbook of financial time series. Springer, pp. 753–766.
Christoffersen, P.F., 1998. Evaluating interval forecasts. International Economic Review 39 (4), 841–862.
Dimitriadis, T., Bayer, S., et al., 2019. A joint quantile and expected shortfall regression framework. Electronic Journal of Statistics 13 (1), 1823–1871.
Egan, M., 2018. February was an insane month for the stock market. Retrieved from https://money.cnn.com/2018/02/28/investing/stock-market-february-dow-jones/index.htmlon November 12, 2020..
Engle, R., 2002. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. Journal of Business & Economic Statistics 20 (3), 339–350.
Engle, R.F., Ledoit, O., Wolf, M., 2019. Large dynamic covariance matrices. Journal of Business & Economic Statistics 37 (2), 363–375.
Engle, R.F., Manganelli, S., 2004. CAViaR: Conditional autoregressive value at risk by regression quantiles. Journal of Business & Economic Statistics 22 (4), 367–381.
Fan, J., Wang, M., Yao, Q., 2008. Modelling multivariate volatilities via conditionally uncorrelated components. Journal of the Royal Statistical Society Series B (Statistical Methodology) 70 (4), 679–702.
Fissler, T., Ziegel, J., 2016. Higher order elicitability and Osbands principle. The Annals of Statistics 44 (4), 1680–1707.
Fissler, T., Ziegel, J.F., Gneiting, T., 2016. Expected Shortfall is jointly elicitable with Value at Risk – implications for backtesting. Risk 5 (1), 8–16.
Forni, M., Hallin, M., Lippi, M., Reichlin, L., 2000. The generalized dynamic-factor model: Identification and estimation. Review of Economics and Statistics 82 (4), 540–554.
Forni, M., Hallin, M., Lippi, M., Reichlin, L., 2005. The generalized dynamic factor model: one-sided estimation and forecasting. Journal of the American Statistical Association 100 (471), 830–840.
Forni, M., Hallin, M., Lippi, M., Zaffaroni, P., 2015. Dynamic factor models with infinite-dimensional factor spaces: One-sided representations. Journal of Econometrics 185 (2), 359–371.
Forni, M., Hallin, M., Lippi, M., Zaffaroni, P., 2017. Dynamic factor models with infinite-dimensional factor space: asymptotic analysis. Journal of Econometrics 199 (1), 74–92.
Forni, M., Lippi, M., 2011. The general dynamic factor model: One-sided representation results. Journal of Econometrics 163 (1), 23–28.
Francq, C., Zakoian, J., 2019. GARCH Models: Structure, Statistical Inference and Financial Applications, 2nd Wiley, West Sussex.
Francq, C., Zakoian, J.-M., 2020. Virtual historical simulation for estimating the conditional VaR of large portfolios. Journal of Econometrics 217 (2), 356–380.

Gaglianone, W.P., Lima, L.R., Linton, O., Smith, D.R., 2011. Evaluating value-at-risk models via quantile regression. Journal of Business & Economic Statistics 29 (1), 150–160.
Gao, C.-T., Zhou, X.-H., 2016. Forecasting VaR and ES using dynamic conditional score models and skew student distribution. Economic Modelling 53, 216–223.
Giannopoulos, K., Tunaru, R., 2005. Coherent risk measures under filtered historical simulation. Journal of Banking & Finance 29 (4), 979–996.
Glosten, L.R., Jagannathan, R., Runkle, D.E., 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. The Journal of Finance 48 (5), 1779–1801.
González-Rivera, G., Lee, T.-H., Mishra, S., 2004. Forecasting volatility: A reality check based on option pricing, utility function, value-at-risk, and predictive likelihood. International Journal of Forecasting 20 (4), 629–645.
Grané, A., Veiga, H., 2014. Outliers, GARCH-type models and risk measures: A comparison of several approaches. Journal of Empirical Finance 26, 26–40.
Gurrola-Perez, P., Murphy, D., 2015. Filtered historical simulation Value-at-Risk models and their competitors. Bank of England Working Paper.
Hallin, M., Liška, R., 2007. Determining the number of factors in the general dynamic factor model. Journal of the American Statistical Association 102 (478), 603–617.
Hansen, P.R., Lunde, A., Nason, J.M., 2011. The model confidence set. Econometrica 79 (2), 453–497.
Hotta, L.K., Trucíos, C., 2018. Inference in (M) GARCH models in the presence of additive outliers: Specification, estimation, and prediction. In: Lavor, C., Gomes, F.A. (Eds.), Advances in Mathematics and Applications. Springer, pp. 179–202.
Kupiec, P., 1995. Techniques for verifying the accuracy of risk measurement models. The Journal of Derivatives 3 (2), 73–84.
Laurent, S., Rombouts, J.V., Violante, F., 2012. On the forecasting accuracy of multivariate GARCH models. Journal of Applied Econometrics 27 (6), 934–955.
Li, W., Gao, J., Li, K., Yao, Q., 2016. Modeling multivariate volatilities via latent common factors. Journal of Business & Economic Statistics 34 (4), 564–573.
Lien, D., Tse, Y.K., Tsui, A.K., 2002. Evaluating the hedging performance of the constant-correlation GARCH model. Applied Financial Economics 12 (11), 791–798.
Manabu, A., 2015. Initial values on quasi-maximum likelihood estimation for BEKK multivariate GARCH models. Soka Economic Studies Quarterly 44 (1), 45–52.
McNeil, A.J., Frey, R., 2000. Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. Journal of Empirical Finance 7 (3-4), 271–300.
Nieto, M.R., Ruiz, E., 2016. Frontiers in VaR forecasting and backtesting. International Journal of Forecasting 32 (2), 475–501.
Nolde, N., Ziegel, J.F., et al., 2017. Elicitability and backtesting: Perspectives for banking regulation. The Annals of Applied Statistics 11 (4), 1833–1874.
Pakel, C., Shephard, N., Sheppard, K., Engle, R.F., 2020. Fitting vast dimensional time-varying covariance models. Journal of Business & Economic Statistics (Forthcoming) 1–17.
Patton, A.J., Ziegel, J.F., Chen, R., 2019. Dynamic semiparametric models for Expected Shortfall (and Value-at-Risk). Journal of Econometrics 211 (2), 388–413.
Pérignon, C., Smith, D.R., 2010. The level and quality of Value-at-Risk disclosure by commercial banks. Journal of Banking & Finance 34 (2), 362–377.
Righi, M.B., Ceretta, P.S., 2015. A comparison of expected shortfall estimation models. Journal of Economics and Business 78, 14–47.
Stock, J.H., Watson, M.W., 2002. Forecasting using principal components from a large number of predictors. Journal of the American Statistical Association 97 (460), 1167–1179.
Stock, J.H., Watson, M.W., 2002. Macroeconomic forecasting using diffusion indexes. Journal of Business & Economic Statistics 20 (2), 147–162.
Taylor, J.W., 2019. Forecasting value at risk and expected shortfall using a semiparametric approach based on the asymmetric Laplace distribution. Journal of Business & Economic Statistics 37 (1), 121–133.
Taylor, J.W., 2020. Forecast combinations for value at risk and expected shortfall. International Journal of Forecasting 36 (2), 428–441.
Trucíos, C., Hotta, L.K., Pereira, P.L.V., 2019. On the robustness of the principal volatility components. Journal of Empirical Finance 52, 201–219.
Trucíos, C., Hotta, L.K., Ruiz, E., 2017. Robust bootstrap forecast densities for GARCH returns and volatilities. Journal of Statistical Computation and Simulation 87 (16), 3152–3174.
Trucíos, C., Hotta, L.K., Ruiz, E., 2018. Robust bootstrap densities for dynamic conditional correlations: implications for portfolio selection and value-at-risk. Journal of Statistical Computation and Simulation 88 (10), 1976–2000.
Trucíos, C., Mazzeu, J.H.G., Hallin, M., Hotta, L.K., Valls Pereira, P.L., Zevallos, M., 2020. Forecasting conditional covariance matrices in high-dimensional time series: a general dynamic factor approach. Available at SSRN 3399782.
Trucíos, C., Tiwari, A.K., Alqahtani, F., 2020. Value-at-risk and expected shortfall in cryptocurrencies portfolio: a vine copula–based approach. Applied Economics 52 (24), 2580–2593.
Trucíos, C., Mazzeu, J.H., Hotta, L.K., Valls Pereira, P.L., Hallin, M., 2020. Robustness and the general dynamic factor model with infinite-dimensional space: Identification, estimation, and forecasting. International Journal of Forecasting (in press) doi:10.1016/j.ijforecast.2020.09.013.
Zumbach, G. O., 2007. A gentle introduction to the RM2006 methodology. Working Paper available at SSRN: https://ssrn.com/abstract=1420183.