Forecasting Financial Risk in High-Dimensional Time Series: The General Dynamic Factor Approach

Carlos Trucíos

Universidade Estadual de Campinas (UNICAMP).

ctruciosm.github.io
ctrucios@unicamp.br

Table of contents

Basic concepts

- Let $P_{t,i}$ denote the closing price of asset i on day t.

- Let $P_{t,i}$ denote the closing price of asset i on day t.
- The return of asset i at time t is defined as

$$r_{t,i} = (P_{t,i} - P_{t-1,i})/P_{t-1,i} \approx \log(P_{t,i}/P_{t-1,i}).$$

- Let $P_{t,i}$ denote the closing price of asset i on day t.
- The return of asset i at time t is defined as

$$r_{t,i} = (P_{t,i} - P_{t-1,i})/P_{t-1,i} \approx \log(P_{t,i}/P_{t-1,i}).$$

- Let $\mathbf{r}_t = (r_{t,1}, \ldots, r_{t,n})^\top$ denote the n-dimensional vector of returns at time t.

- Let $P_{t,i}$ denote the closing price of asset i on day t.
- The return of asset i at time t is defined as

$$r_{t,i} = (P_{t,i} - P_{t-1,i})/P_{t-1,i} \approx \log(P_{t,i}/P_{t-1,i}).$$

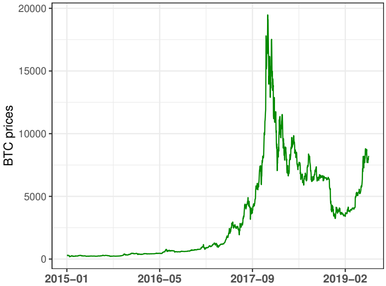
- Let $\mathbf{r}_t = (r_{t,1}, \ldots, r_{t,n})^\top$ denote the n-dimensional vector of returns at time t.

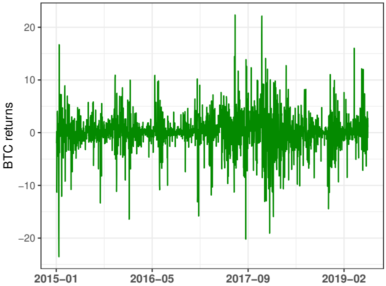
Basic concepts

- Let $P_{t,i}$ denote the closing price of asset i on day t.
- The return of asset i at time t is defined as

$$r_{t,i} = (P_{t,i} - P_{t-1,i})/P_{t-1,i} \approx \log(P_{t,i}/P_{t-1,i}).$$

- Let $\mathbf{r}_t = (r_{t,1}, \ldots, r_{t,n})^\top$ denote the n-dimensional vector of returns at time t.





Risk Measures

Statistical tools used to assess and quantify the risk associated with investments or financial portfolios. They help in understanding potential losses and uncertainty, with common examples including:

- (Co)volatility.
- Value at Risk (VaR).
- Expected Shortfall(ES).
- Etc.

Conditional Covariance Matrix

$$\Sigma_t = \mathbb{V}(\mathbf{r}_t \mid \mathcal{F}_{t-1}).$$

- The diagonal elements $\sigma^2_{t,ii}$ $(i = 1, \ldots, n)$ represent the conditional variances (squared volatilities) of the individual assets.
- The off-diagonal elements $\sigma_{t,ij}$ $(i \neq j)$ represent the conditional covariances (co-volatilities) between asset returns.

Conditional Covariance Matrix

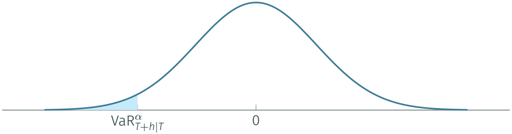
$$\Sigma_t = \mathbb{V}(\mathbf{r}_t \mid \mathcal{F}_{t-1}).$$

- The diagonal elements $\sigma^2_{t,ii}$ $(i = 1, \ldots, n)$ represent the conditional variances (squared volatilities) of the individual assets.
- The off-diagonal elements $\sigma_{t,ij}$ $(i \neq j)$ represent the conditional covariances (co-volatilities) between asset returns.

Forecasting the conditional covariance matrix is crucial for portfolio allocation, risk management, and asset pricing applications (Ardia et al., 2017; Torres and Villena, 2024; Trucíos, 2026).

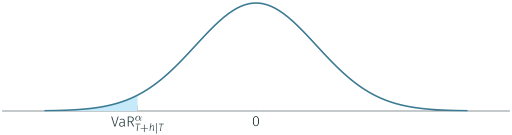
Value-at-Risk (VaR)

Maximum possible loss for a given portfolio at risk level α over a specific time horizon h



Value-at-Risk (VaR)

Maximum possible loss for a given portfolio at risk level α over a specific time horizon h



The VaR h-steps-ahead at risk level α, denoted by $\text{VaR}^{\alpha}_{T+h|T}$, is a number such that

$$\mathbb{P}\left(r_{T+h} \leq \text{VaR}^{\alpha}_{T+h|T}\right) = \alpha$$

Expected Shortfall (ES)

VaR is the maximum loss for given a risk level during a certain period, while the ES is the average loss once this loss overcomes VaR

The ES h-steps-ahead at risk level α, denoted by $\text{ES}^{\alpha}_{T+h|T}$ is given by

$$\text{ES}^{\alpha}_{T+h|T} = \mathbb{E}\left[r_{T+h} \mid r_{T+h} < \text{VaR}^{\alpha}_{T+h|T}\right]$$

Risk measures are important for:

- risk management,
- trading strategies,
- algorithmic decision-making,
- order execution in limit order books,
- etc.

Risk measures are important for:

- risk management,
- trading strategies,
- algorithmic decision-making,
- order execution in limit order books,
- etc.

Typically, beyond its intrinsic importance, Σ_t serves as an input for the estimation of VaR and ES.

Motivation

Given their importance, the development of methodologies for modelling and forecasting (co)volatilities, VaR, and ES is essential.

Motivation

Given their importance, the development of methodologies for modelling and forecasting (co)volatilities, VaR, and ES is essential.

- Σ_t is an input parameter for several portfolio allocation strategies.

Given their importance, the development of methodologies for modelling and forecasting (co)volatilities, VaR, and ES is essential.

- Σ_t is an input parameter for several portfolio allocation strategies.
- Typically, Σ_t serves as an input for the estimation of VaR and ES

Given their importance, the development of methodologies for modelling and forecasting (co)volatilities, VaR, and ES is essential.

- Σ_t is an input parameter for several portfolio allocation strategies.
- Typically, Σ_t serves as an input for the estimation of VaR and ES
- Classical multivariate volatility models are helpful tools for predicting Σ_t in small and moderate dimensions, however, they badly suffer from the "curse of dimensionality".

Some alternatives to estimate the conditional covariance matrix in high dimensional data are:

- Composite likelihood

Motivation

Some alternatives to estimate the conditional covariance matrix in high dimensional data are:

- Composite likelihood
- Shrinkage

Some alternatives to estimate the conditional covariance matrix in high dimensional data are:

- Composite likelihood
- Shrinkage
- Non-parametric approaches

Some alternatives to estimate the conditional covariance matrix in high dimensional data are:

- Composite likelihood
- Shrinkage
- Non-parametric approaches
- Dimension reduction / divide and conquer techniques

Dimension reduction techniques to forecast the conditional covariance matrix:

- Principal components analysis (PCA),
- Independent component analysis (ICA),
- Conditionally uncorrelated components (CUC),
- Dynamic orthogonal components (DOC),
- Principal volatility components (PVC), etc.

However, most dimension reduction techniques are based on a static approach, which are not optimal in a time series context (Hallin et al., 2018)

The general dynamic factor model

Spearman (1904): proposes factor analysis in order to account for the dependencies between several cognitive variables. The result was a unobservable factor ("general ability") explaining the cognitive performance (the IQ concept usually refers to this factor).

Spearman (1904): proposes factor analysis in order to account for the dependencies between several cognitive variables. The result was a unobservable factor ("general ability") explaining the cognitive performance (the IQ concept usually refers to this factor).

In the Spearman's approach, the factor model is characterised by

$$\mathrm{X}_t = \boldsymbol{\chi}_t + \boldsymbol{\varepsilon}_t = \mathrm{Bf}_t + \boldsymbol{\epsilon}_t$$

Spearman (1904): proposes factor analysis in order to account for the dependencies between several cognitive variables. The result was a unobservable factor ("general ability") explaining the cognitive performance (the IQ concept usually refers to this factor).

In the Spearman's approach, the factor model is characterised by

$$\mathbf{X}_t = \boldsymbol{\chi}_t + \boldsymbol{\varepsilon}_t = \mathbf{B}\mathbf{f}_t + \boldsymbol{\epsilon}_t$$

- $\mathbf{X}_1, \cdots, \mathbf{X}_T$ is an iid sample of n-dimensional observations;
- $\mathbf{B}$ is an $n \times r$ matrix of scalar *loadings*;
- $\mathbf{f}_t$ is an iid latent process of r-dimensional *common factors*;
- $\boldsymbol{\epsilon}_t$ is an iid process of n-dimensional *idiosyncratic components*.
- $\mathbb{E}(f_{kt}\epsilon_{it}) = 0,\ k = 1, \cdots, r,\ i = 1, \cdots, n;\ t = 1, \cdots, T.$

Geweke (1977): if factor models were to be used in time series econometrics, some of their assumptions should be relaxed. Factors may be loaded by observations with some lag.

Geweke (1977): if factor models were to be used in time series econometrics, some of their assumptions should be relaxed. Factors may be loaded by observations with some lag.

$$\mathrm{X}_t = \boldsymbol{\chi}_t + \boldsymbol{\varepsilon}_t = \mathrm{B}(L)\mathrm{f}_t + \boldsymbol{\epsilon}_t$$

Geweke (1977): if factor models were to be used in time series econometrics, some of their assumptions should be relaxed. Factors may be loaded by observations with some lag.

$$\mathrm{X}_t = \boldsymbol{\chi}_t + \boldsymbol{\varepsilon}_t = \mathrm{B}(L)\mathrm{f}_t + \boldsymbol{\epsilon}_t$$

- $\mathrm{X}_1, \cdots, \mathrm{X}_T$ is a finite realization of an observed n-dimensional second-order stationary process;
- $\mathrm{B}(\mathrm{L}) = \sum_{\nu=0}^{\infty} B_\nu L^\nu$ is an $n \times r$ matrix of *loadings filters*;
- f_t is an r-dimensional latent second-order stationary process of *factors*;
- $\boldsymbol{\epsilon}_t$ is a n-dimensional second-order stationary mutually orthogonal white noise called *idiosyncratic components*

Chamberlain (1983): thanks to the double asymptotic approach (n and $T \to \infty$), it is possible to relax the mutual orthogonality assumption of $\boldsymbol{\epsilon}_t$

Chamberlain (1983): thanks to the double asymptotic approach (n and $T \to \infty$), it is possible to relax the mutual orthogonality assumption of $\boldsymbol{\epsilon}_t$

$$\mathrm{X}_t^{(n)} = \boldsymbol{\chi}_t^{(n)} + \boldsymbol{\varepsilon}_t^{(n)} = \mathrm{B}^{(n)}\mathrm{f}_t + \boldsymbol{\epsilon}_t^{(n)}$$

Chamberlain (1983): thanks to the double asymptotic approach (n and $T \to \infty$), it is possible to relax the mutual orthogonality assumption of $\boldsymbol{\epsilon}_t$

$$\mathrm{X}_t^{(n)} = \boldsymbol{\chi}_t^{(n)} + \boldsymbol{\varepsilon}_t^{(n)} = \mathrm{B}^{(n)}\mathbf{f}_t + \boldsymbol{\epsilon}_t^{(n)}$$

- $\mathrm{X}_1, \cdots, \mathrm{X}_T$ is a finite realization of an observed n-dimensional second-order stationary process;
- $\mathrm{B}^{(n)}$ is an $n \times r$ matrix of scalar *loadings*;
- $\mathbf{f}_t$ is an r-dimensional latent second-order stationary process of *factors*;
- $\boldsymbol{\epsilon}_t$ is a n-dimensional second-order stationary process with finite covariance matrix (not necessarily diagonal).

- Several authors followed Chamberlain (1983) approach (double asymptotic and static loadings): Stock and Watson (2002), Bai (2003), etc.

General dynamic factor model: the genealogy

- Several authors followed Chamberlain (1983) approach (double asymptotic and static loadings): Stock and Watson (2002), Bai (2003), etc.
- Some others followed Geweke (1977) dynamic exact approach: Engle and Watson (1981), Shumway and Stoffer (1982), Watson and Engle (1983), etc.

General dynamic factor model: the genealogy

- Several authors followed Chamberlain (1983) approach (double asymptotic and static loadings): Stock and Watson (2002), Bai (2003), etc.
- Some others followed Geweke (1977) dynamic exact approach: Engle and Watson (1981), Shumway and Stoffer (1982), Watson and Engle (1983), etc.
- Forni et al. (2000) and Forni and Lippi (2001) adopted a different approach. They combined the double asymptotic of Chamberlain (1983) (which allows for approximate factor models) with the dynamic loading structure introduced by Geweke (1977), leading to the General Dynamic Factor Model (GDFM).

Let $\mathbf{X}_t = (X_{1t}\ X_{2t} \ldots\ X_{nt})'$, $t = 1, ...T$ be a finite realization of a zero-mean double-indexed second-order stationary stochastic process $\mathbf{X} = \{X_{it} : i \in \mathbb{N}, t \in \mathbb{Z}\}$. The GDFM is based on the decomposition

$$
\begin{aligned}
X_{it} &= \chi_{it} + \xi_{it} && (1)\\
\chi_{it} &= b_{i1}(L)u_{1t} + \ldots + b_{iq}(L)u_{qt}, \quad i \in \mathbb{N}, \quad t \in \mathbb{Z}, && (2)
\end{aligned}
$$

where L stands for a lag operator and the unobservable $u_{i,t}$, $\chi_{i,t}$ and $\xi_{i,t}$, stand for the common shocks, common components and idiosyncratic components, respectively.

Under the assumption that the space spanned by the common components is finite-dimensional, the decomposition (1) takes the static form

$$X_{it} = \underbrace{\lambda_{i1}F_{1t} + \ldots + \lambda_{ir}F_{rt}}_{\chi_{it}} + \xi_{it}, \quad r \geq q \tag{3}$$

However, this assumption rules out simple factor-loading patterns (Forni and Lippi, 2011; Forni et al., 2015, 2017) such as

$$X_{i,t} = \overbrace{\underbrace{a_i(1-\alpha_i L)^{-1}u_t}_{a_i(u_t+\alpha_i u_{t-1}+\alpha_i^2 u_{t-2}+\alpha_i^3 u_{t-3}+\ldots)}}^{\chi_{it}} + \xi_{it}. \tag{4}$$

- Forni et al. (2000) proposed a procedure that does not assume that the space spanned by the common components is finite-dimensional. However, is based on a **two-sided filter**, which is not satisfactory for forecasting.

- Forni et al. (2000) proposed a procedure that does not assume that the space spanned by the common components is finite-dimensional. However, is based on a **two-sided filter**, which is not satisfactory for forecasting.
- Forni et al. (2005) proposed a procedure which allows for one-sided filter estimation. However, assume **finite-dimensional** factor space.

- Forni et al. (2000) proposed a procedure that does not assume that the space spanned by the common components is finite-dimensional. However, is based on a **two-sided filter**, which is not satisfactory for forecasting.
- Forni et al. (2005) proposed a procedure which allows for one-sided filter estimation. However, assume **finite-dimensional** factor space.
- Forni et al. (2015, 2017) proposed a procedure which allows **one-sided** filter estimation and **infinite-dimensional** factor space.

Forni et al. (2015, 2017) show that, $\chi_t = (\chi_{1t}\ \chi_{2t} \dots \chi_{nt})'$ admits a block-structure autoregressive representation

$$\mathbf{A}(L)\chi_t = \mathbf{R}u_t. \tag{5}$$

where

$$\mathbf{A}(L) = \begin{bmatrix} \mathbf{A}^1(L) & 0 & \dots & 0 \\ 0 & \mathbf{A}^2(L) & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & \mathbf{A}^m(L) \end{bmatrix} \text{ and } \mathbf{R} = \begin{bmatrix} \mathbf{R}^1 \\ \mathbf{R}^2 \\ \vdots \\ \mathbf{R}^m \end{bmatrix}$$

with $A^k(L)$ being a $(q+1) \times (q+1)$ polynomial matrix with finite degree and R^k a $(q+1) \times q$ matrix, $n = m(q+1)$.

Proposition

Under assumptions in Barigozzi and Hallin (2020) and additionally assuming that u_t and ξ_t are conditionally uncorrelated for any lead and lag, the conditional variance-covariance matrix of $\mathbf{X}_t$ is given by

$$V(\mathbf{X}_t|\mathcal{F}_{t-1}) = \mathbf{R}\, V(\mathbf{u}_t|\mathcal{F}_{t-1})\mathbf{R}' + V(\xi_t|\mathcal{F}_{t-1}). \tag{6}$$

Proposition

Under assumptions in Barigozzi and Hallin (2020) and additionally assuming that u_t and ξ_t are conditionally uncorrelated for any lead and lag, the conditional variance-covariance matrix of $\mathbf{X}_t$ is given by

$$V(\mathbf{X}_t|\mathcal{F}_{t-1}) = \mathbf{R}\, V(\mathbf{u}_t|\mathcal{F}_{t-1})\mathbf{R}' + V(\xi_t|\mathcal{F}_{t-1}). \tag{6}$$

- $u_t \sim$ MGARCH.

Proposition

Under assumptions in Barigozzi and Hallin (2020) and additionally assuming that u_t and ξ_t are conditionally uncorrelated for any lead and lag, the conditional variance-covariance matrix of $\mathbf{X}_t$ is given by

$$V(\mathbf{X}_t|\mathcal{F}_{t-1}) = \mathbf{R}\, V(\mathbf{u}_t|\mathcal{F}_{t-1})\mathbf{R}' + V(\xi_t|\mathcal{F}_{t-1}). \tag{6}$$

- $u_t \sim$ MGARCH.
- The conditional covariance matrix of the idiosyncratic components can be modelled as a full or diagonal matrix, where each conditional variance is modelled independently by a GARCH-type model.

General dynamic factor model: Estimation

- **Step 1.** Determine the number q of common shocks via an information criterion, for instance, using Hallin and Liška (2007).

- **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 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

$$\widehat{\boldsymbol{\Sigma}}^X_{nT}(\theta) = \frac{1}{2\pi} \sum_{k=-M_T}^{M_T} e^{-ik\theta} \underbrace{K\left(\frac{k}{B_T}\right)}_{1-\frac{|k|}{[|\sqrt{T}|]+1}} \widehat{\boldsymbol{\Gamma}}^X_k \quad \theta \in [0, 2\pi] \qquad (7)$$

where $K(\cdot)$ is a kernel function, M_T a truncation parameter, B_T a bandwidth, and $\widehat{\boldsymbol{\Gamma}}^X_k$ the sample lag-k cross-covariance matrix.

- **Step 4.** Estimate the spectral density matrix of the common components by

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

where $\widehat{\mathbf{\Lambda}}^{\chi}_{nT}(\theta)$ is a $q \times q$ diagonal matrix with diagonal elements the q largest eigenvalues of $\widehat{\mathbf{\Sigma}}^{\chi}_{nT}(\theta)$ and $\widehat{\mathsf{P}}^{\chi}_{nT}(\theta)$ (with complex conjugate $\widehat{\mathsf{P}}^{\chi *}_{nT}$) is the $n \times q$ matrix with the associated eigenvectors.

- **Step 5.** By inverse Fourier transform of $\widehat{\boldsymbol{\Sigma}}^{\chi}_{n^*T}(\theta)$, estimate the autocovariance matrices $\widehat{\boldsymbol{\Gamma}}^{\chi}_{k}$ of the m sub-vectors

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

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

- **Step 5.** By inverse Fourier transform of $\widehat{\boldsymbol{\Sigma}}^{\chi}_{n^*T}(\theta)$, estimate the autocovariance matrices $\widehat{\boldsymbol{\Gamma}}^{\chi}_{k}$ of the m sub-vectors

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

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 $\hat{\mathbf{A}}_{n^*}(L)\mathbf{X}_{n^*t} = \widehat{\mathbf{Y}}_{n^*t} = \widehat{\mathbf{R}}_{n^*}\widehat{\mathbf{u}}_t + \widehat{\varepsilon}_t$
- **Step 6.** 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 $\mathbf{u}_t$; then, an estimate of the impulse-response function is $\widehat{\mathbf{B}}_{n^*}(L) := [\hat{\mathbf{A}}_{n^*}(L)]^{-1}\widehat{\mathbf{R}}_{n^*}$.

- **Step 7.** Repeat Steps 2 through 7 M times: the final estimates (denoted as $\widehat{\mathsf{R}}_n$, $\widehat{\mathsf{u}}_t$, and $\widehat{\mathsf{B}}_n$) are obtained by averaging the estimates $\widehat{\mathsf{R}}_{n*}$, $\widehat{\mathsf{u}}_t$, and $\widehat{\mathsf{B}}_{n*}$ associated with each iteration. Let $\widehat{\chi}_{nt} := \widehat{\mathsf{B}}_n(L)\widehat{\mathsf{u}}_t$ and $\hat{\xi}_{nt} := \mathsf{X}_{nt} - \widehat{\chi}_{nt}$.

- **Step 7.** Repeat Steps 2 through 7 M times: the final estimates (denoted as $\widehat{\mathrm{R}}_n$, $\widehat{\mathrm{u}}_t$, and $\widehat{\mathrm{B}}_n$) are obtained by averaging the estimates $\widehat{\mathrm{R}}_{n*}$, $\widehat{\mathrm{u}}_t$, and $\widehat{\mathrm{B}}_{n*}$ associated with each iteration. Let $\widehat{\chi}_{nt} := \widehat{\mathrm{B}}_n(L)\widehat{\mathrm{u}}_t$ and $\hat{\boldsymbol{\xi}}_{nt} := \mathrm{X}_{nt} - \widehat{\chi}_{nt}$.
- **Step 8.** The estimator of $\mathrm{V}(\mathrm{X}_t|\mathcal{F}_{t-1})$ is given by

$$\widehat{\mathrm{V}}(\mathrm{X}_t|\mathcal{F}_{t-1}) := \widehat{\mathrm{R}}\widehat{\mathrm{V}}(\hat{\mathrm{u}}_t|\mathcal{F}_{t-1})\widehat{\mathrm{R}}' + \widehat{\mathrm{V}}(\hat{\xi}_t|\mathcal{F}_{t-1}). \quad (8)$$

- **Step 7.** Repeat Steps 2 through 7 M times: the final estimates (denoted as $\widehat{\mathbf{R}}_n$, $\widehat{\mathbf{u}}_t$, and $\widehat{\mathbf{B}}_n$) are obtained by averaging the estimates $\widehat{\mathbf{R}}_{n*}$, $\widehat{\mathbf{u}}_t$, and $\widehat{\mathbf{B}}_{n*}$ associated with each iteration. Let $\widehat{\chi}_{nt} := \widehat{\mathbf{B}}_n(L)\widehat{\mathbf{u}}_t$ and $\hat{\boldsymbol{\xi}}_{nt} := \mathbf{X}_{nt} - \widehat{\chi}_{nt}$.
- **Step 8.** The estimator of $\mathrm{V}(\mathbf{X}_t|\mathcal{F}_{t-1})$ is given by

$$\widehat{\mathrm{V}}(\mathbf{X}_t|\mathcal{F}_{t-1}) := \widehat{\mathbf{R}}\widehat{\mathrm{V}}(\hat{\mathbf{u}}_t|\mathcal{F}_{t-1})\widehat{\mathbf{R}}' + \widehat{\mathrm{V}}(\hat{\xi}_t|\mathcal{F}_{t-1}). \tag{8}$$

The consistency of the method is established under some assumptions (stationarity, the existence of a spectral density matrix, etc.). See, Forni et al. (2017), Barigozzi and Hallin (2020), Trucíos et al. (2023).

Monte Carlo experiments

General settings:

- 500 Monte Carlo replications.
- Panel of dimension 60×1000 and 600×700.
- Three DGPs.
  - **DGP1:** Static factor model with $\mathbf{u}_t$ following a bivariate full BEKK and ξ_{it} following univariate GARCH.

General settings:

- 500 Monte Carlo replications.
- Panel of dimension 60×1000 and 600×700.
- Three DGPs.
    - **DGP1:** Static factor model with $\mathbf{u}_t$ following a bivariate full BEKK and ξ_{it} following univariate GARCH.
    - **DGP2:** Dynamic factor model with finite-dimensional factor space, $\mathbf{u}_t$ follows a bivariate full BEKK and ξ_{it} following univariate GARCH.

General settings:

- 500 Monte Carlo replications.
- Panel of dimension 60×1000 and 600×700.
- Three DGPs.
  - **DGP1:** Static factor model with $\mathbf{u}_t$ following a bivariate full BEKK and ξ_{it} following univariate GARCH.
  - **DGP2:** Dynamic factor model with finite-dimensional factor space, $\mathbf{u}_t$ follows a bivariate full BEKK and ξ_{it} following univariate GARCH.
  - **DGP3:** Dynamic factor model with infinite-dimensional factor space, $\mathbf{u}_t$ follows a bivariate full BEKK and ξ_{it} following univariate GARCH.

Two common shocks $u_t = (u_{1t}, u_{2t})'$ generated from a BEKK(1,1,1) model:

$$u_t = Q_t^{1/2} \begin{pmatrix} \eta_{1t} \\ \eta_{2t} \end{pmatrix}, \tag{9}$$

$$Q_t = C_0' C_0 + C_1' u_{t-1} u_{t-1}' C_1 + C_2' Q_{t-1} C_2. \tag{14}$$

where η_{it}, $i = 1, 2$, are i.i.d. innovations generated from $\mathcal{N}(0, 1)$ or a centered and standardized Student-t_5 distribution. In order to guarantee $E(Q_t) = E(u_{t-1} u_{t-1}') = I_q$, we set $C_0' C_0 = I_q - C_1' C_1 - C_2' C_2$.

Loss function

$$L(\hat{\Sigma}_{T+1|T}, \Sigma_{T+1|T}) = \sum_{i=1}^{N} \sum_{j=i}^{N} w(i,j)(\hat{\sigma}_{i,j} - \sigma_{i,j})^2, \tag{10}$$

where $\hat{\sigma}_{i,j}$ and $\sigma_{i,j}$ are the (i,j) elements of $\hat{\Sigma}_{T+1|T}$ and $\Sigma_{T+1|T}$. $w(i,j)$ are weights.

Loss function

$$L(\hat{\Sigma}_{T+1|T}, \Sigma_{T+1|T}) = \sum_{i=1}^{N} \sum_{j=i}^{N} w(i,j)(\hat{\sigma}_{i,j} - \sigma_{i,j})^2, \tag{10}$$

where $\hat{\sigma}_{i,j}$ and $\sigma_{i,j}$ are the (i,j) elements of $\hat{\Sigma}_{T+1|T}$ and $\Sigma_{T+1|T}$. $w(i,j)$ are weights.

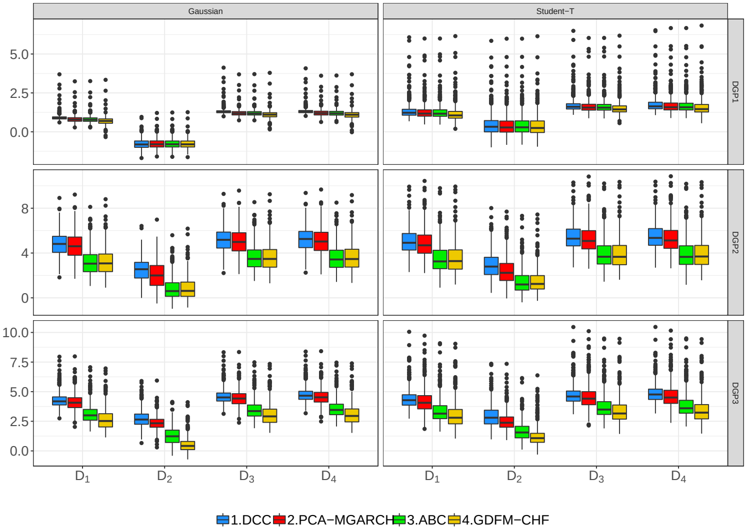
Table 1: Weights $w(i,j)$ in Equation (10), $i = 1, \ldots, N$, $j = i, \ldots, N$

D_1	$w(i,j) = 1 \ \forall i$ and j
D_2	$w(i,j) = 1$ when $i = j$ and zero otherwise
D_3	$w(i,j) = 2$ when $\hat{\sigma}_{i,j}, > h_{i,j}$; 1 otherwise
D_4	$w(i,j) = 2$ when $\hat{\sigma}_{i,j}, < h_{i,j}$; 1 otherwise

Procedures

- PCA-DCC: The static factors are extracted by using PCA and a DCC model is applied on the extracted factors. The idiosyncratic components are modelled by univariate GARCH models.
- DCC: as in Pakel et al. (2020);
- ABC-DCC: dynamic factor extraction under a finite-dimensional factor space assumption as in Alessi et al. (2009);
- GDFM-CHF: Our proposal.

Monte Carlo experiments: N = 60, T = 1000



Empirical Application I

Let H_{T+1} be the conditional covariance matrix of vector returns at time $T+1$ and $\omega = (\omega_1, \cdots, \omega_N)$ be the portfolio weights.

Let H_{T+1} be the conditional covariance matrix of vector returns at time $T+1$ and $\omega = (\omega_1, \cdots, \omega_N)$ be the portfolio weights.

We are interesting in portfolios with minimum risk subject to some constraints, *i.e.*

$$\text{Min:} \quad \omega' H_{T+1} \omega$$

subject to $\omega_i \geq 0$ and $\sum_{i=1}^{N} \omega_i = 1$

Let H_{T+1} be the conditional covariance matrix of vector returns at time $T+1$ and $\omega = (\omega_1, \cdots, \omega_N)$ be the portfolio weights.

We are interesting in portfolios with minimum risk subject to some constraints, *i.e.*

$$\text{Min:} \quad \omega' H_{T+1} \omega$$

subject to $\omega_i \geq 0$ and $\sum_{i=1}^{N} \omega_i = 1$

This portfolio is called **the minimum variance portfolio with no short sales** and is of particular interest for investors.

- **Minimum variance portfolios**
- 656 stocks entering the composition of the S&P 500, NASDAQ100 and AMEX (only series with no missing values were considered).
- Stocks traded from January 2, 2011 through June 29, 2018 (T=1884).
- **Rolling Window scheme:** A window size of 750 days is used for estimation, which represents a concentration ratio of $656/750 = 0.875$; the out-of-sample period was set to 1134 days.

At time $t = 750, \ldots, 1883$ (1134 time points), the one-step ahead conditional covariance matrix is estimated and used to obtain the optimal portfolio allocation weights.

$$\widehat{\boldsymbol{\omega}}_{t+1|t} = \underset{\boldsymbol{\omega}}{\operatorname{argmin}}\, \boldsymbol{\omega}' \widehat{\Sigma}_{t+1|t} \boldsymbol{\omega},$$

subject to the restrictions $\omega_i \geq 0$ and $\sum_{i=1}^{n} \omega_i = 1$.

At time $t = 750, \ldots, 1883$ (1134 time points), the one-step ahead conditional covariance matrix is estimated and used to obtain the optimal portfolio allocation weights.

$$\widehat{\boldsymbol{\omega}}_{t+1|t} = \operatorname*{argmin}_{\boldsymbol{\omega}} \boldsymbol{\omega}' \widehat{\Sigma}_{t+1|t} \boldsymbol{\omega},$$

subject to the restrictions $\omega_i \geq 0$ and $\sum_{i=1}^{n} \omega_i = 1$.

Then, the resulting (out-of-sample) portfolio return is given by

$$r_{p,t+1} := \sum_{i=1}^{n} \widehat{\omega}_{i;t+1|t} r_{i,t+1}$$

To evaluate the out-of-sample portfolio performance, we use three measures:

- **SD**: Annualized standard deviation (standard deviation of the out-of-sample portfolio returns $\times\sqrt{252}$);
- **IR**: Information ratio (IR $:=$ AV/SD), with AV being the annualized average portfolio (average out-of-sample portfolio returns $\times$ 252);
- **SR**: Sortino ratio.

To evaluate the out-of-sample portfolio performance, we use three measures:

- **SD**: Annualized standard deviation (standard deviation of the out-of-sample portfolio returns $\times\sqrt{252}$);
- **IR**: Information ratio (IR := AV/SD), with AV being the annualized average portfolio (average out-of-sample portfolio returns $\times$ 252);
- **SR**: Sortino ratio.

To assess the statistical differences between our proposal and its competitors we perform the test of Ledoit and Wolf (2011) (for SD) and Ledoit and Wolf (2008) (for IR).

Competitors:

- **1/n**: equal-weighted portfolio strategy advocated by DeMiguel et al. (2009).
- **POET**: the estimator proposed by Fan et al. (2013).
- **NL**: the nonlinear shrinkage estimator of Ledoit and Wolf (2012).
- **RM2006**: the RiskMetrics 2006 methodology of Zumbach (2007).
- **DCC**: the DCC model with composite likelihood of Pakel et al. (2020).
- **OGARCH**: the orthogonal GARCH model of Alexander and Chibumba (1996).
- **GPVC**: generalized principal volatility components of Li et al. (2016).

Competitors:

- **PCA4TS**: principal component analysis for second-order stationary vector time series of Chang et al. (2018).
- **PCA-MGARCH**: as used in our simulation study.
- **ABC**: the procedure of Alessi et al. (2009) based on the general dynamic factor model with finite-dimensional factor space.
- **2GDFM4V**: the two-step GDFM procedure for volatilities proposed by Barigozzi and Hallin (2016, 2017, 2020); does not consider conditional cross-covariances.

Our method outperform all previously mentioned procedures and performs statistically equal than **DCC-NL** and **AFM-DCC-NL**.

Our method outperform all previously mentioned procedures and performs statistically equal than **DCC-NL** and **AFM-DCC-NL**.

- **DCC-NL**: as proposed by Engle et al. (2019); combines the DCC and NL procedures.
- **AFM-DCC-NL**: approximate factor model as in De Nard et al. (2021) but with unobservable common components obtained by classical PCA. The covariance matrix of the idiosyncratic components is estimated using the DCC-NL procedure.

Table 2: Out-of-sample portfolio performance

	SD	IR	SR
1/n	11.5067 (14)	0.5015 (14)	0.6834 (14)
POET	4.6116 (9)	1.2146 (7)	1.6741 (7)
NL	4.6152 (10)	1.0217 (11)	1.4249 (11)
RM2006	4.5446 (7)	1.2327 (6)	1.7241 (6)
DCC	5.9901 (12)	1.1502 (8)	1.6262 (8)
DCC-NL	**3.9358 (2)**	**1.8371 (2)**	2.6227 (2)
OGARCH	4.4551 (5)	1.1051 (10)	1.5616 (10)
GPVC	4.5889 (8)	1.0022 (12)	1.4077 (12)
PCA4TS	4.7256 (11)	1.1364 (9)	1.6032 (9)
PCA-MGARCH	4.4111 (4)	1.4965 (4)	2.0891 (4)
AFM-DCC-NL	**3.9472 (3)**	**1.9764 (1)**	2.8974 (1)
ABC	4.5313 (6)	1.4404 (5)	1.9677 (5)
2GDFM4V	10.2431 (13)	0.7992 (13)	1.0997 (13)
GDFM-CHF-DCC	**3.9205 (1)**	**1.8188 (3)**	2.4109 (3)

VaR and ES

Let $\mathbf{r}_t = \mathbf{H}_t^{1/2}\boldsymbol{\epsilon}_t$, be the $N-$dimensional vector of individual returns at time t

Let $\mathbf{r}_t = \mathbf{H}_t^{1/2}\boldsymbol{\epsilon}_t$, be the $N-$dimensional vector of individual returns at time t

- **Step 10.** Using the historical returns $\mathbf{r}_1, \ldots, \mathbf{r}_T$ to obtain the estimators $\widehat{\mathbf{H}}_t := \widehat{\mathbf{V}}^{\mathbf{r}_N}_{t|t-1}$ 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 $\widehat{\mathbf{H}}_t$ at time $t = 1, \ldots, T+1$.

- **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 $\mathbf{r}_{T+1}^{i*} := \widehat{\mathbf{H}}_{T+1}^{1/2} \boldsymbol{\epsilon}_i^*$; this yields B simulated one-step-ahead portfolio return forecasts $R_{T+1}^{i*} := \boldsymbol{\omega}' \mathbf{r}_{T+1}^{i*}$, for $i = 1, \ldots, B$ for portfolio weights $\boldsymbol{\omega} = (\omega_1, \ldots \omega_N)'$.

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

$$\widehat{\mathrm{VaR}}^{\alpha}_{\mathrm{T+1}} := \widehat{F}^{-1}_{R^{*}_{T+1}}(\alpha), \tag{11}$$

and

$$\widehat{\mathrm{ES}}^{\alpha}_{\mathrm{T+1}} := \sum_{i=1}^{B} \frac{R^{i*}_{T+1}\mathbb{I}\left[R^{i*}_{T+1} < \widehat{\mathrm{VaR}}^{\alpha}_{\mathrm{T+1}}\right]}{\sum_{t=1}^{B} \mathbb{I}\left[R^{i*}_{T+1} < \widehat{\mathrm{VaR}}^{\alpha}_{\mathrm{T+1}}\right]}, \tag{12}$$

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.

Empirical Application II

- 652 stocks used in the composition of the S&P500, NASDAQ-100 and AMEX indexes
- From January 3, 2012 to July 1, 2020 ($T = 2136$ observations)
- A rolling window approach with window size of T= 750 days (concentration ratio of $652/750 = 0.87$).
- 1386 days in the out-of-sample period 8from December 29, 2014 through July 1, 2020).
- In each window, the one-step-ahead VaR and ES is estimated

Alternative models:

- DCC composite likelihood (Pakel et al., 2020)
- RiskMetrics 2006 methodology (Zumbach, 2007)
- ABC (Alessi et al., 2009)

Classical backtesting procedures (calibration tests and scoring functions) and scoring functions are used to evaluate the VaR and ES accuracy.

Calibration tests:

Test	Proposed by	Used to evaluate
Unconditional coverage (UC)	Kupiec (1995)	VaR
Conditional coverage (CC)	Christoffersen (1998)	VaR
Dynamic quantile (DQ)	Engle and Manganelli (2004)	VaR
VaR quantile regression (VQ)	Gaglianone et al. (2011)	VaR
Exceedance residuals (ER)	McNeil and Frey (2000)	ES and VaR
Conditional calibration (CoC)	Nolde et al. (2017)	ES and VaR
Exceedance shortfall regression (ESR)	Bayer and Dimitriadis (2020)	ES

Table 3: Calibration tests used to evaluate VaR and ES accuracy.

$$H_0 : \text{The VaR/ES is correctly specified}$$

Scoring functions:

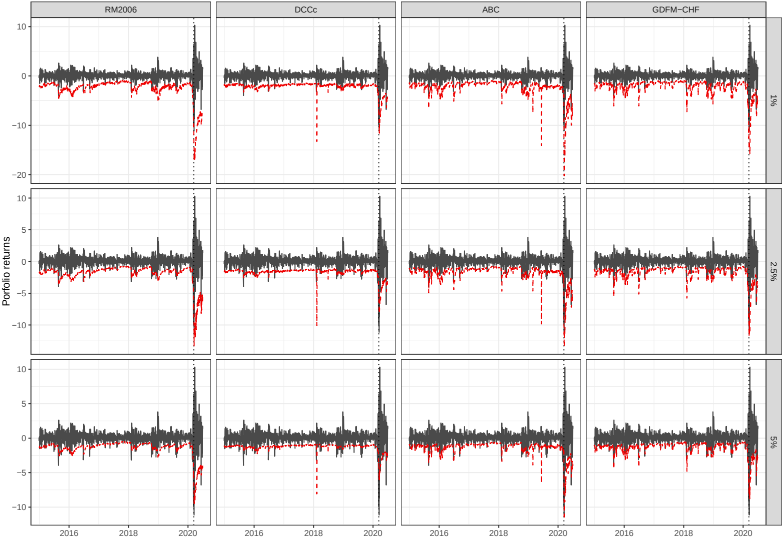
For the VaR:

$$S(\mathrm{VaR}_{\mathrm{t}}^{\alpha}, r_t) = (\alpha - \mathbb{I}[r_t \leq \mathrm{VaR}_{\mathrm{t}}^{\alpha}])(G(r_t) - G(\mathrm{VaR}_{\mathrm{t}}^{\alpha})), \tag{13}$$

For the VaR and ES (jointly):

$$\begin{aligned} S\big((\mathrm{VaR}_t^{\alpha}, \mathrm{ES}_t^{\alpha}), r_t\big) = & \big(\mathbb{I}[r_t \leq \mathrm{VaR}_t^{\alpha}] - \alpha\big) G_1(\mathrm{VaR}_t^{\alpha}) - \mathbb{I}[r_t \leq \mathrm{VaR}_t^{\alpha}]\, G_1(r_t) \\ & + G_2(\mathrm{ES}_t^{\alpha})\big(\mathrm{ES}_t^{\alpha} - \mathrm{VaR}_t^{\alpha} + \mathbb{I}[r_t \leq \mathrm{VaR}_t^{\alpha}]\big)(\mathrm{VaR}_t^{\alpha} - r_t)/\alpha \\ & + G_3(\mathrm{ES}_t^{\alpha}) + G_4(r_t), \end{aligned} \tag{14}$$

Empirical Application



		Hits	UC	CC	DQ	VQ	ER	CoC	ESR_1	ESR_2	ESR_3
			Calibration tests								
1%	RM2006	0.60	0.161	0.051	0.091	0.331	0.493	0.219	0.344	0.218	0.706
	DCCc	1.70	0.024	0.000	0.000	0.434	0.404	0.106	0.516	0.505	0.278
	ABC	1.90	0.003	0.001	0.000	0.258	0.024	0.052	0.016	0.043	0.036
	GDFM-CHF	1.40	0.189	0.224	0.266	0.726	0.564	0.493	0.500	0.576	0.180
2.5%	RM2006	1.30	0.002	0.000	0.000	0.001	0.522	0.000	0.009	0.011	0.972
	DCCc	3.00	0.221	0.003	0.000	0.000	0.087	0.129	0.013	0.165	0.018
	ABC	3.10	0.166	0.014	0.000	0.420	0.001	0.227	0.016	0.325	0.256
	GDFM-CHF	2.70	0.689	0.602	0.362	0.437	0.230	0.549	0.421	0.834	0.135
5%	RM2006	2.50	0.000	0.000	0.000	0.000	0.416	0.000	0.000	0.438	0.993
	DCCc	5.50	0.416	0.000	0.000	0.343	0.037	0.077	0.029	0.035	0.007
	ABC	5.50	0.416	0.024	0.000	0.588	0.000	0.324	0.906	0.010	0.409
	GDFM-CHF	5.60	0.351	0.147	0.207	0.354	0.369	0.405	0.296	0.766	0.072

Table 4: One-step-ahead VaR and ES backtesting (calibration tests).

		Avg. Scoring functions			
		QL	FZG	NZ	AL
1%	RM2006	0.038	0.673	1.850	2.178
	DCCc	0.039	0.670	1.863	2.181
	ABC	0.042	0.698	1.991	2.344
	GDFM-CHF	0.032	0.653	1.730	2.060
2.5%	RM2006	0.075	0.675	1.654	1.975
	DCCc	0.075	0.666	1.642	1.951
	ABC	0.077	0.677	1.680	2.007
	GDFM-CHF	0.066	0.631	1.535	1.816
5%	RM2006	0.123	0.681	1.493	1.797
	DCCc	0.121	0.666	1.473	1.759
	ABC	0.119	0.668	1.478	1.773
	GDFM-CHF	0.110	0.630	1.396	1.649

Table 5: One-step-ahead VaR and ES backtesting (scoring functions). Shaded cells stand for the smallest scoring function.

Conclusions

Conclusions

- GDFM is a powerful tool for modelling and forecasting high-dimensional time series.
- Based on the GDFM, we propose a new procedure for forecasting the conditional covariance matrix in high-dimensional time series with good finite sample properties and consistency results, is proposed (Trucíos et al., 2023).
- A filtered historical simulation method with the conditional covariance estimator of Trucíos et al. (2023) to estimate the VaR and ES in large portfolios is suggested (Hallin and Trucíos, 2023).
- Our results are quite competitive with state-of-the-art alternatives.

References

Alessi, L., Barigozzi, M., and 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.

Alexander, C. O. and Chibumba, A. (1996). Multivariate orthogonal factor GARCH. *University of Sussex Discussion Papers in Mathematics*.

Ardia, D., Bolliger, G., Boudt, K., and Gagnon-Fleury, J.-P. (2017). The impact of covariance misspecification in risk-based portfolios. *Annals of Operations Research*, 254(1):1–16.

Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica*, 71(1):135–171.

Barigozzi, M. and Hallin, M. (2016). Generalized dynamic factor models and volatilities: recovering the market volatility shocks. *The Econometrics Journal*, 19(1):C33–C60.

Barigozzi, M. and Hallin, M. (2017). Generalized dynamic factor models and volatilities: estimation and forecasting. *Journal of Econometrics*, 201(2):307–321.

Barigozzi, M. and Hallin, M. (2020). Generalized dynamic factor models and volatilities: Consistency, rates, and prediction intervals. *Journal of Econometrics*.

Bayer, S. and Dimitriadis, T. (2020). Regression based expected shortfall backtesting. *Journal of Financial Econometrics*.

Chamberlain, G. (1983). Funds, factors, and diversification in arbitrage pricing models. *Econometrica*, pages 1305–1323.

Chang, J., Guo, B., and Yao, Q. (2018). Principal component analysis for second-order stationary vector time series. *The Annals of Statistics*, 46(5):2094–2124.

Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review*, 39(4):841–862.

De Nard, G., Ledoit, O., and Wolf, M. (2021). Factor models for portfolio selection in large dimensions: The good, the better and the ugly. *Journal of Financial Econometrics*, 19(2):236–257.

DeMiguel, V., Garlappi, L., and Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *The Review of Financial Studies*, 22(5):1915–1953.

Engle, R. and Watson, M. (1981). A one-factor multivariate time series model of metropolitan wage rates. *Journal of the American Statistical Association*, 76(376):774–781.

Engle, R. F., Ledoit, O., and Wolf, M. (2019). Large dynamic covariance matrices. *Jounal of Business & Economic Statistics*, (doi: 10.1080/07350015.2017.1345683).

Engle, R. F. and Manganelli, S. (2004). CAViaR: Conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics*, 22(4):367–381.

Fan, J., Liao, Y., and Mincheva, M. (2013). Large covariance estimation by thresholding principal orthogonal complements. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 75(4):603–680.

Forni, M., Hallin, M., Lippi, M., and 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., and 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., and 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., and Zaffaroni, P. (2017). Dynamic factor models with infinite-dimensional factor space: asymptotic analysis. *Journal of Econometrics*, 199(1):74–92.

Forni, M. and Lippi, M. (2001). The generalized dynamic factor model: representation theory. *Econometric Theory*, 17(6):1113–1141.

Forni, M. and Lippi, M. (2011). The general dynamic factor model: One-sided representation results. *Journal of Econometrics*, 163(1):23–28.

Gaglianone, W. P., Lima, L. R., Linton, O., and Smith, D. R. (2011). Evaluating value-at-risk models via quantile regression. *Journal of Business & Economic Statistics*, 29(1):150–160.

Geweke, J. (1977). The dynamic factor analysis of economic time series. *Latent variables in socio-economic models*.

Hallin, M., Hörmann, S., and Lippi, M. (2018). Optimal dimension reduction for high-dimensional and functional time series. *Statistical Inference for Stochastic Processes*, 21(2):385–398.

Hallin, M. and Liška, R. (2007). Determining the number of factors in the general dynamic factor model. *Journal of the American Statistical Association*, 102(478):603–617.

Hallin, M. and Trucíos, C. (2023). Forecasting value-at-risk and expected shortfall in large portfolios: A general dynamic factor model approach. *Econometrics and Statistics*, 27:1–15.

Kupiec, P. (1995). Techniques for verifying the accuracy of risk measurement models. *The Journal of Derivatives*, 3(2):73–84.

Ledoit, O. and Wolf, M. (2008). Robust performance hypothesis testing with the sharpe ratio. *Journal of Empirical Finance*, 15(5):850–859.

Ledoit, O. and Wolf, M. (2011). Robust performances hypothesis testing with the variance. *Wilmott*, 2011(55):86–89.

Ledoit, O. and Wolf, M. (2012). Nonlinear shrinkage estimation of large-dimensional covariance matrices. *The Annals of Statistics*, pages 1024–1060.

Li, W., Gao, J., Li, K., and Yao, Q. (2016). Modeling multivariate volatilities via latent common factors. *Journal of Business & Economic Statistics*, 34(4):564–573.

McNeil, A. J. and 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.

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., and Engle, R. F. (2020). Fitting vast dimensional time-varying covariance models. *Journal of Business & Economic Statistics*, pages 1–17.

Shumway, R. H. and Stoffer, D. S. (1982). An approach to time series smoothing and forecasting using the em algorithm. *Journal of Time Series Analysis*, 3(4):253–264.

Spearman, C. (1904). General intelligence, objectively determined and measured. *The American Journal of Psychology*, 15(2):201–292.

Stock, J. H. and Watson, M. W. (2002). Forecasting using principal components from a large number of predictors. *Journal of the American statistical association*, 97(460):1167–1179.

Torres, R. and Villena, M. (2024). On the empirical performance of different covariance-matrix forecasting methods. *Neural Computing and Applications*, 36(16):9503–9524.

Trucíos, C. (2026). Hierarchical risk clustering versus traditional risk-based portfolios: an empirical out-of-sample comparison. *Empirical Economics*. Forthcoming.

Trucíos, C., Mazzeu, J. H. G., Hallin, M., Hotta, L. K., Valls Pereira, P. L., and Zevallos, M. (2023). Forecasting conditional covariance matrices in high-dimensional time series: a general dynamic factor approach. *Journal of Business and Economic Statistics*.

Watson, M. W. and Engle, R. F. (1983). Alternative algorithms for the estimation of dynamic factor, mimic and varying coefficient regression models. *Journal of Econometrics*, 23(3):385–400.

Zumbach, G. O. (2007). A gentle introduction to the RM2006 methodology. *Working Paper available at SSRN: https://ssrn.com/abstract=1420183*.