

UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Sede Amministrativa: Università degli Studi di Padova

Dipartimento di Scienze Statistiche
Corso di Dottorato di Ricerca in Scienze Statistiche
Ciclo XXXV

Modeling of high-frequency financial data

Coordinatore del Corso: Prof. Nicola Sartori

Supervisore: Prof. Massimiliano Caporin

Dottorando: Marco Girardi

11 January 2023

Abstract

This thesis addresses the exploitation of high-frequency financial data through the modeling of realized covariances, with the purpose of forecasting the covariance matrix of the selected assets. Generally, the analysis of realized covariances is carried out with models using a Wishart distribution whose scale matrix is described by an autoregressive moving average structure. The first chapter introduces a novel distribution used in place of the Wishart, which can be employed to identify a common factor in the assets behaviour, representing the systematic risk in the market. The model is empirically tested in an asset allocation framework with the objective of tracking the market index of reference. Therefore, the purpose of this chapter is not to improve the conventional models based on the Wishart distribution, but rather to expand their usual scope of application.

Another common issue stems from the presence of data retrieved at multiple frequencies and the ensuing need to properly exploit the available information content. The second chapter briefly reviews the literature concerning the modeling mixed-frequency data, and formulates a way to integrate the high-frequency information in a low-frequency-based model, using the former as "views" about the assets. Thus, the approach proposed in this chapter requires the derivation of a posterior equation, which depends upon the distributional assumptions for the high and low-frequency data. The resulting expression for the posterior distribution effectively allows to operate with data retrieved at different frequencies, and therefore to provide an alternative to the common approaches used in this framework.

The final chapter is intended to outline a possible way of exploiting high-frequency data in the field of risk spillover analysis, starting from recent results concerning the structured specifications of conventional models employing low-frequency data.

Sommario

Questa tesi affronta il tema dell'utilizzo dei dati finanziari ad alta frequenza attraverso la modellazione delle covarianze realizzate, al fine di prevedere la matrice di covarianza degli asset selezionati. In genere, l'analisi delle covarianze realizzate viene eseguita tramite modelli che utilizzano una distribuzione Wishart la cui matrice di scala è descritta da una struttura autoregressiva a media mobile. Il primo capitolo introduce una nuova distribuzione usata al posto della Wishart, che può essere utilizzata per identificare un fattore comune nell'andamento degli asset, rappresentante il rischio sistematico del mercato. Il modello viene testato empiricamente in un contesto di asset allocation al fine di replicare l'indice di mercato di riferimento. Quindi, l'obiettivo del capitolo non è di migliorare i modelli classici basati sulla distribuzione Wishart, bensì di espandere il loro tradizionale campo di applicazione.

Un altro problema comune sorge per via della presenza di dati raccolti a frequenze multiple e la conseguente necessità di sfruttare in maniera appropriata il contenuto informativo a disposizione. Il secondo capitolo rivede sinteticamente la letteratura riguardante la modellazione di dati a frequenza mista e formula un modo per integrare l'informazione ad alta frequenza in un modello basato su dati a bassa frequenza, usando la prima come "opinione" riguardo gli asset. Quindi, l'approccio proposto in questo capitolo richiede la derivazione dell'equazione del *posterior*, che dipende dalle assunzioni riguardo la distribuzione dei dati ad alta e bassa frequenza. La risultante espressione per la distribuzione del *posterior* consente di operare efficacemente con dati raccolti a frequenze diverse, e quindi di fornire un'alternativa agli approcci più comuni in questo contesto.

Il capitolo finale intende delineare una possibile modalità di sfruttamento dei dati ad alta frequenza nel campo dell'analisi del *risk spillover*, partendo da recenti risultati relativi alle *structured specifications* di modelli convenzionali che impiegano dati a bassa frequenza.

To Willy

Acknowledgements

I would like to thank my supervisor Massimiliano Caporin for his exceptional support and guidance, even in the most difficult times. I would also like to thank the professors and members of the department for contributing to our education as doctoral students. Finally, I would like to express my gratitude to all my colleagues for helping me along the way.

Contents

List of Figures	xii
List of Tables	xv
Introduction	1
Overview	1
Main contributions of the thesis	3
1 Conditional autoregressive \mathcal{G} model for high-frequency data	5
1.1 State of the art	5
1.2 The conditional autoregressive \mathcal{G} model	7
1.2.1 The \mathcal{G} density	7
1.2.2 Dynamics of δ_t	10
1.2.3 Dynamics of C_t	11
1.2.4 Model estimation	12
1.3 Empirical analysis	13
1.3.1 Data	13
1.3.2 Estimation results	14
1.3.3 Market risk tracking portfolio	19
1.3.4 An active asset allocation application	22
1.4 Conclusions	24
2 Blended frequency model	27
2.1 State of the art and proposal	27
2.2 The Black-Litterman approach	29
2.3 Methodological framework	32
2.3.1 COMFORT PCA model	33
2.3.2 High-frequency data	34
2.3.3 Combining HF and LF data	35
2.3.4 Multi-step forecast	36
2.3.5 Asset allocation strategies and performance evaluation	38
2.4 Empirical analysis	39
2.4.1 Asset allocation strategies and performances	39
2.5 Conclusions	46
3 CAW model and risk spillover analysis	49

3.1	Risk spillover	49
3.2	Reference literature and proposal	50
3.2.1	Structured specification	51
3.2.2	Risk spillover in structured specification	52
3.2.3	Structured CAW specification	53
	Conclusions	55
	Appendix A	59
	Appendix B	63
	Bibliography	69

List of Figures

1.1	Average realized variance and covariance.	14
1.2	Common factor and DJIA variance.	15
1.3	Common factor and DJIA realized variance.	16
1.4	Scatter plots of δ_t versus DJIA variance and realized variance.	16
1.5	Median of realized covariances and fitted values of C_t (off-diagonal).	17
1.6	Quantiles of forecasted C_t (off-diagonal elements).	18
1.7	Median forecasts for δ_t , C_t and $\delta_t C_t$ (off-diagonal elements).	18
1.8	Median forecasts for δ_t , C_t and $\delta_t C_t$ (diagonal elements).	19
1.9	Weight by industry (40 days MA).	20
1.10	Portfolio and DJIA realized variances.	21
1.11	Portfolio net return and DJIA gross return.	22
2.1	Annualized Sharpe ratio.	40
2.2	Annualized Sortino ratio.	41
2.3	Annualized mean and variance of returns (0bp).	41
2.4	Annualized Sharpe ratio (Gaussian model).	42
2.5	Annualized Sortino ratio (Gaussian model).	43
2.6	Annualized return and volatility (0bp).	44
2.7	GMV portfolio performance and drawdown (0bp).	45
2.8	MS portfolio performance and drawdown (0bp).	45

List of Tables

1.1	Summary statistics for realized measures.	14
1.2	Estimated parameters and the corresponding standard errors (sandwich approach).	15
1.3	AIC and BIC for the conditional autoregressive \mathcal{G} model compared to the standard CAW.	17
1.4	Target variance and GMV portfolios.	20
1.5	Sharpe ratios net of transaction costs for market risk mimicking portfolio (passive benchmark) and Active optimal portfolio.	24

Introduction

Overview

Modernization that has taken place over the past decades has led to a continuous evolution of the concepts of availability and speed. Needless to say, how this is evident in everyday reality, from purchases delivered in a matter of hours to the exchange of information in real time. Financial markets have not been excluded from this transformation, with participants under an obligation to stay one step ahead of others, in a constant chase to take advantage of the new information available in the shortest possible time and in the best possible way. An aspect arising from the increased flow of information concerns the exploitation of high-frequency data. Although it can be debated how frequent the observations need to be in order to be considered as properly high-frequency, it is necessary to develop a process that allows to exploit them regardless of the frequency at which they are retrieved. An approach that has gained attention in the field of volatility and covariance modeling consists in the use of realized measures constructed from high-frequency data as drivers of information. The focus, when using such measures, is on predictive models, rather than on obtaining non-parametric measurements of past movements. A widely used model in this area is the Conditional Autoregressive Wishart (CAW) model developed by Golosnoy *et al.* (2012), which uses a Wishart distribution to analyze realized covariances, and assumes an autoregressive moving average structure for the scale matrix of such distribution. The model is suitable for large panels of assets, since it ensures positive-definite covariance matrices without imposing parameter constraints and can be easily estimated by maximum likelihood.

This thesis aims to address the exploitation of high-frequency data for covariance modeling in the field of portfolio allocation. Throughout the dissertation, data collected at a frequency of 1 minute are used as high-frequency data for the construction of realized covariances. Chapter 1 departs from the usual design of the CAW model as it introduces a novel distribution used in place of the Wishart distribution and employed to identify a common factor in the assets behaviour, which constitutes the inherent risk in the market.

The distribution is obtained as the product of a scalar component, distributed as a unit-mean inverse gamma, and a matrix component following a Wishart distribution. The mean is endowed with an autoregressive moving average structure, and the one-step-ahead forecasts of the covariance matrix are employed in a portfolio allocation framework with the aim of tracking the reference index performance by limiting the impact of specific risk. The model is empirically tested in its covariance targeting version with autoregressive and moving average components of order one. The common factor detected by the model is highly correlated with market variance and plays a crucial role in estimating and forecasting conditional covariances. With regard to the index tracking objective, the realized variance of the portfolio ensuing from the empirical application is close to the common factor itself as well as to the realized variance of the market. Additionally, the risk-adjusted performance of the portfolio is comparable to the market index for a level of proportional transaction costs of 15 basis points.

As opposed to the high-frequency-based models, there exist a much more crowded literature concerning the estimation of covariances from daily data. Such approaches are generally used in the field of high-frequency data as benchmarks to assess and test the results of the model under development. Differently from this view, Chapter 2 aims to formulate a way to integrate the high-frequency information within the low-frequency-based model, thus finding an expression for operating with data retrieved at different frequencies. In the literature, data collected at multiple frequencies are commonly processed by aggregating the highest frequency data to reduce all data to the same frequency. Other approaches addressing this practice and preserving frequency differences are proposed by Ghysels *et al.* (2004) and Shephard and Sheppard (2010). Nevertheless, they both jointly model high-frequency and low-frequency data. As opposed to such models, Chapter 2 takes its cue from Black and Litterman (1992), and uses the information from high-frequency data as views about the assets under analysis with the aim of integrating the output of a low-frequency-based model. The mixing process requires the derivation of the posterior equation, which depends upon the distributional assumptions for the high and low-frequency data.

Chapter 3 is intended to outline a possible way of exploiting high-frequency data in the field of risk spillover analysis, starting from recent results concerning the structured specifications of conventional models employing low-frequency data. Thus, the purpose of the chapter is not to provide conclusive results, but rather to provide a research idea that exploits what is presented in the other chapters. Risk spillover refers to the impact that one or more events affecting the risk of an asset have on the risk of other assets or asset classes, either within the same country or across countries. This

network effect has its roots in the financial interconnections that have been persistently increasing among countries, markets, financial institutions and even asset classes, and its implications affect a wide range of financial areas, such as credit risk management, portfolio management, hedging strategies, and financial stability as a whole. Thus, much research has been conducted on this topic, also in order to facilitate the development of various regulatory requirements, such as capital requirements or capital controls. Nevertheless, most of the current literature is limited to the use of low-frequency data. The aim is therefore to take inspiration from the most recent strands of literature about this issue, and provide an enhanced specification obtained through the CAW model, which has the advantage of modeling the covariance matrix using an estimate of it (i.e. the realized covariance matrix) rather than using a latent process.

Main contributions of the thesis

The main contributions of this PhD thesis are related to the exploitation of financial high-frequency data, with emphasis on modeling and forecasting covariances, and are organized into the following chapters. The contents presented in Chapter 1 provide the necessary tools for identifying and isolating a common factor in the assets behaviour, representing the specific risk in the market, and the idiosyncratic risk component. This becomes particularly useful when it comes to understanding the extent to which diversification can reduce the risk of an investment. Alternatively, the model can be exploited as an intermediary linking the quantities retrieved from high frequency and a low-frequency-based model. In this perspective, the output of the former can be used to augment the dynamics of the latter, allowing for a richer and finer depiction. Chapter 1 reveals that the model allows to accurately track the reference market index, by limiting the impact of specific risk.

Chapter 2 contributes to the asset allocation literature by trying to improve the signal used for the computations through the blending process. The assessment is conducted under realistic assumptions concerning the allocation process, that is with transaction costs and monthly rebalancing of the assets. Although it may seem reasonable that increasing the frequency of data used would improve the result, Chapter 2 shows that combining high-frequency and low-frequency data yields superior results, despite the former undoubtedly has the advantage of supplying the finest representation of assets. Additionally, the presence of transaction costs highlights that relying solely on high-frequency data leads to excessive turnover. In such a context, in fact, it might be worth giving up accuracy partially for more stability, to avoid the depletion of any possible

benefit provided by high-frequency data.

Chapter 1

Conditional autoregressive \mathcal{G} model for high-frequency data

1.1 State of the art

Multivariate modeling and forecasting of variances and covariances of asset returns is crucial in many financial activities, although the time series of these quantities are not directly observable. Conventional approaches, based on low-frequency data, model variance-covariance matrices either given past observations, such as multivariate GARCH models (Bollerslev *et al.*, 1988), or treat them as an unobserved stochastic process, such as multivariate stochastic volatility models (Harvey *et al.*, 1994).

More recently, realized measures computed from high-frequency data have proven to be a valuable tool for modeling volatility and covariances. Such measures provide powerful insights if included in the analysis and lead to more accurate estimates compared to conventional models, which are limited to daily-level inputs. They are extremely relevant in capturing sudden movements in the markets, and have an empirical application that spans a wide range of areas, including risk management, asset pricing and portfolio optimization.

Andersen *et al.* (2003) and Barndorff-Nielsen and Shephard (2004) showed that these measures are precise estimates of actual covariation, as they quickly adapt to the market conditions by exploiting intraday information.

Noureldin *et al.* (2012) proposed a multivariate extension of the class of high-frequency-based volatility (HEAVY) models developed by Shephard and Sheppard (2010), which uses high-frequency data to obtain greater predictive ability with respect to the multivariate GARCH model. The main advantage of HEAVY models is the usage of a two-equation system, conditioned on a high-frequency information set, that allows to

deliver precise multi-step forecasts exhibiting short response time in periods of abrupt market changes and short-run momentum effects before reverting.

An alternative approach for covariance modeling was adopted by Golosnoy *et al.* (2012), generalizing the Wishart Autoregressive (WAR) model developed by Gouriéroux *et al.* (2009). The proposed Conditional Autoregressive Wishart (CAW) model assumes an autoregressive moving average structure for the mean of Wishart distribution, allowing both lagged covariance matrices and lagged predictions of the covariance to contribute to the prediction, while in the WAR specification, only lagged covariances are used. WAR and CAW models have been used in several empirical studies focusing on the prediction of realized variances, or in the interpretation of spillovers among assets (see, e.g. the contributions of Bonato *et al.* (2013) and Gribisch *et al.* (2020)). We contribute to this strand of literature on both the methodological and empirical sides. We propose to use a different multivariate distribution, the \mathcal{G} distribution, for realized covariance modeling and forecasting. This novel distribution is obtained as the product of Wishart distribution and a scalar component following a unit-mean inverse gamma. The distribution can be scaled by a dynamically evolving scalar factor and can also account for dynamic changes in the entire covariance by allowing the scale matrix of the original Wishart distribution to be time-varying. The resulting distribution has a notable advantage compared to Wishart distribution, since it allows to identify a common market risk factor which can be implicitly obtained from the information associated with the assets; the market risk factor is associated with the dynamic scalar scaling term. The scalar component can also be removed, simplifying, if needed, the model estimation. To recover the model dynamic components, we suggest the use of a two-step procedure. First, the scalar scaling term is modeled as the conditional mean of a Multiplicative Error Model (MEM); then, the observed realized covariance matrices are filtered from this mutual element and endowed with a dynamic structure equivalent to that adopted in CAW models (Golosnoy *et al.*, 2012), which is particularly useful for prediction purposes. The remaining model parameters are estimated through maximum likelihood accounting for symmetry and positive definiteness of the covariance matrices.

Through an empirical example, we provide some insights into the estimated parameters and show that the common factor extracted from a small set of equities, a subset of the constituents of the Dow Jones Industrial Average (DJIA) index, is highly correlated with market volatility, as proxied by the conditional variance estimated, on a univariate basis, on the DJIA daily index returns or by using the daily realized variance of DJIA. By exploiting the informative content of the common factor, we design a market risk replicating portfolio (a sort of tracking portfolio), which can be used as a benchmark

or as a passive investment portfolio. The risk replicating portfolio exhibits reduced cardinality, compared to the full set of assets used for the common factor estimation, and exhibits superior performances compared to the market index. Additionally, we provide a simple example showing how the passive portfolio can be combined with an active strategy for stock picking, which can lead to performances above those of the reference and naive portfolios.

The rest of Chapter 1 is structured as follows. In section 2, we derive the \mathcal{G} density and describe the model, including its estimation. Section 3 is devoted to the empirical analysis, including the model application in a portfolio allocation framework. Section 4 concludes the main results.

1.2 The conditional autoregressive \mathcal{G} model

Let us consider an observed sequence of realized covariance matrices \mathbf{Z}_t of dimension m with a multiplicative structure, such that $\mathbf{Z}_t = \delta_t X_t \mathbf{Y}_t$, where δ_t is a scalar parameter, X_t follows an inverse-gamma density with unit mean and \mathbf{Y}_t follows a Wishart density with n degrees of freedom and scale matrix C_t/n . For modeling such covariances, we first introduce an explicit expression of the density of \mathbf{Z}_t .

In the multiplicative model framework, Freitas *et al.* (2005) introduced a multivariate extension of the polarimetric distributions available in the literature for describing synthetic aperture radar (SAR) data, proposing the family of polarimetric \mathcal{G}_P laws, obtained as the product of a generalized inverse Gaussian distribution and a multivariate complex Wishart. In this section, we derive a similar distribution, which we refer to as \mathcal{G} , for modeling realized covariances. Afterwards, we move to a conditional setting, discussing the model dynamics, constraints and estimation.

1.2.1 The \mathcal{G} density

Let X be a unitary-mean inverse gamma random variable with shape parameter $-\alpha$ and \mathbf{Y} be a Wishart random variable with n degrees of freedom and scale matrix C/n :

$$f_X(x) = \frac{x^{\alpha-1}}{(-\alpha-1)^\alpha \Gamma(-\alpha)} \exp\left(-\frac{\alpha+1}{x}\right), \quad -\alpha, x > 0, \quad (1.1)$$

and

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{|\mathbf{y}|^{(n-m-1)/2} \exp(-\frac{1}{2} \text{Tr}(nC^{-1}\mathbf{y}))}{2^{nm/2} h(n, m) |C/n|^{n/2}}, \quad (1.2)$$

where m is the dimension of \mathbf{Y} and $h(n, m) = \pi^{m(m-1)/4} \prod_{j=1}^m \Gamma(n/2 + (1-j)/2)$. Equation 1.2 can be equivalently expressed in terms of C , which is the expected value of \mathbf{Y} , as

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{(n/2)^{nm/2} |\mathbf{y}|^{(n-m-1)/2} \exp(-\frac{n}{2} \text{Tr}(C^{-1} \mathbf{y}))}{h(n, m) |C|^{n/2}}. \quad (1.3)$$

Considering equations 1.1 and 1.3, the densities of $\mathbf{\Omega} = X\mathbf{Y}$ and $\mathbf{Z} = \delta\mathbf{\Omega}$, where δ is a scalar, are provided by Theorem 1.1 and Corollary 1.2, respectively.

Theorem 1.1. *Let X be a unitary-mean inverse gamma random variable and \mathbf{Y} be a Wishart random variable with n degrees of freedom and scale matrix C/n , then $\mathbf{\Omega} = X\mathbf{Y}$ follows a $\mathcal{G}(\alpha, C, n)$ density; that is,*

$$f_{\mathbf{\Omega}}(\omega) = \frac{(n/2)^{nm/2} |\omega|^{(n-m-1)/2} \Gamma(nm/2 - \alpha)}{h(n, m) |C|^{n/2} (-\alpha - 1)^\alpha \Gamma(-\alpha)} \times \\ \times \left(\frac{n}{2} \text{Tr}(C^{-1} \omega) - (\alpha + 1) \right)^{\alpha - nm/2}. \quad (1.4)$$

Proof. The density of $\mathbf{\Omega}$ is given by

$$f_{\mathbf{\Omega}}(\omega) = \int_{\mathbb{R}_+} f_{x\mathbf{Y}}(\omega) f_X(x) dx, \quad (1.5)$$

where the scale transformation $x\mathbf{Y}$ is distributed as a Wishart with expected value $x\mathbf{C}$:

$$f_{x\mathbf{Y}}(\omega) = \frac{(n/2)^{nm/2} |\omega|^{(n-m-1)/2} \exp(-\frac{n}{2} \text{Tr}((x\mathbf{C})^{-1} \omega))}{h(n, m) |x\mathbf{C}|^{n/2}}. \quad (1.6)$$

Therefore, the density of $\mathbf{\Omega}(\omega)$ is

$$f_{\mathbf{\Omega}}(\omega) = \frac{(n/2)^{nm/2} |\omega|^{(n-m-1)/2}}{h(n, m) |C|^{n/2} (-\alpha - 1)^\alpha \Gamma(-\alpha)} \times \\ \times \int_{\mathbb{R}_+} x^{\alpha-1-nm/2} \exp\left(x^{-1} \left(-\frac{n}{2} \text{Tr}(C^{-1} \omega) + (\alpha + 1)\right)\right) dx. \quad (1.7)$$

The solution of such an integral was provided by Gradshteyn and Ryzhik (2007) in section 3.326 on page 337. In particular,

$$\int_0^\infty x^a \exp(-\beta x^b) dx = \frac{\Gamma(\gamma)}{b\beta^\gamma}, \quad \gamma = \frac{a+1}{b}, \quad \beta > 0. \quad (1.8)$$

In our case, we have

$$\int_0^\infty y^{-\mathcal{A}} \exp(-\mathcal{B}y^{-1}) dy, \quad (1.9)$$

where $\mathcal{A} = 1 - \alpha + nm/2$ and $\mathcal{B} = \frac{n}{2}\text{Tr}(C^{-1}\omega) - (\alpha + 1)$.

Integrating by substitution using $x = y^{-1}$ and then applying formula 3.326 of Gradshteyn and Ryzhik (2007), we get

$$\int_0^\infty x^{\mathcal{A}-2} \exp(-\mathcal{B}x) dx = \frac{\Gamma(\mathcal{A}-1)}{\mathcal{B}^{(\mathcal{A}-1)}}, \quad (1.10)$$

where $\mathcal{A} - 2 = 1 - \alpha + nm/2 - 2 = -\alpha + nm/2 - 1 > 0$ is satisfied by construction given that $-\alpha > 0, m > 1$ and $n > 0$. Moreover, $\mathcal{B} = \frac{n}{2}\text{Tr}(C^{-1}\omega) - (\alpha + 1)$ is positive for $-\alpha > 1$. We rewrite the integral solution as

$$\begin{aligned} \frac{\Gamma(\mathcal{A}-1)}{\mathcal{B}^{(\mathcal{A}-1)}} &= \frac{\Gamma(nm/2 - \alpha)}{\mathcal{B}^{nm/2 - \alpha}} \\ &= \Gamma(nm/2 - \alpha) \mathcal{B}^{\alpha - nm/2}. \end{aligned} \quad (1.11)$$

Therefore, $f_{\Omega}(\omega)$ becomes

$$\begin{aligned} f_{\Omega}(\omega) &= \frac{(n/2)^{nm/2} |\omega|^{(n-m-1)/2} \Gamma(nm/2 - \alpha)}{h(n, m) |C|^{n/2} (-\alpha - 1)^{\alpha} \Gamma(-\alpha)} \times \\ &\quad \times \left(\frac{n}{2} \text{Tr}(C^{-1}\omega) - (\alpha + 1) \right)^{\alpha - nm/2} \end{aligned} \quad (1.12)$$

which is a $\mathcal{G}(\alpha, C, n)$ density. \square

Corollary 1.2. *Let $\Omega \sim \mathcal{G}(\alpha, C, n)$ and δ be a scaling factor, then $\delta\Omega \sim \mathcal{G}(\alpha, \delta C, n)$.*

Proof. Let $\mathbf{Y} \sim W(C, n)$, where C is the expected value of \mathbf{Y} and A is a square matrix of dimension m , then $A\mathbf{Y}A' \sim W(ACA', n)$. If we apply the scale transformation δ to the Wishart density \mathbf{Y} and define $A = \text{diag}(a_1, \dots, a_m)$ with $a_i = \sqrt{\delta}$, $i = 1, \dots, m$, then $\delta\mathbf{Y} = A\mathbf{Y}A' \sim W(\delta C, n)$, and, by Theorem 1.1, the density of $\mathbf{Z} = \delta\Omega$ is

$$\begin{aligned} f_{\mathbf{Z}}(\mathbf{z}) &= \frac{(n/2)^{nm/2} |\mathbf{z}|^{(n-m-1)/2} \Gamma(nm/2 - \alpha)}{h(n, m) |\delta C|^{n/2} (-\alpha - 1)^{\alpha} \Gamma(-\alpha)} \times \\ &\quad \times \left(\frac{n}{2} \text{Tr}((\delta C)^{-1}\mathbf{z}) - (\alpha + 1) \right)^{\alpha - nm/2}, \end{aligned} \quad (1.13)$$

which is a $\mathcal{G}(\alpha, \delta C, n)$ density. \square

Remark $\tilde{X} = \delta X$ follows an inverse gamma with shape $-\alpha$ and mean equal to δ . This allows us to interpret \tilde{X} as a common stochastic market risk factor, whose conditional mean equals δ .

In this work, we use the \mathcal{G} density to characterize the conditional distribution of realized covariance matrices. Namely, we set $\mathbf{Z}_t | \mathcal{I}_{t-1} \sim \mathcal{G}(\alpha, \delta_t C_t, n)$, where \mathcal{I}_{t-1} is the time $t - 1$ information set, and both δ_t and C_t have dynamical representations, which are described in the following sub-sections.

1.2.2 Dynamics of δ_t

The purpose of this component of the conditional density is to extract a common risk factor from a panel of securities. Therefore, the first step is to define how to model the elements of the covariance matrix in a way that allows us to sort out mutual and specific components. If we denote by $z_{ij,t}$ the generic $[ij]$ element of the realized covariance matrix, we can express it in a multiplicative way as $z_{ij,t} = \lambda_t \sigma_{ij,t}$, where λ_t is a common factor for the elements in the matrix and $\sigma_{ij,t}$ is an element-specific factor assuming either positive or negative values. Then, $\log z_{ij,t}^2 = \log \lambda_t^2 + \log \sigma_{ij,t}^2$, and an estimate \tilde{l}_t of $\log \lambda_t^2$ is provided by the arithmetic average:

$$\tilde{l}_t = \frac{\sum_{ij \in \mathbf{S}} \log z_{ij,t}^2}{m(m+1)/2}, \quad (1.14)$$

where \mathbf{S} is the set of ij combinations of the elements in the lower triangular part of the realized covariance matrix, including the elements on the main diagonal. Therefore, the estimate of λ_t can be expressed as

$$\tilde{\lambda}_t = \left(\prod_{ij \in \mathbf{S}} |z_{ij,t}| \right)^{\frac{2}{m(m+1)}}, \quad (1.15)$$

which is a geometric average of $z_{ij,t}$ elements in absolute value. Thus, $\tilde{\lambda}_t$ can be reasonably interpreted as the source of information for the common component.

The scalar δ_t is modeled as the conditional mean of the Multiplicative Error Model (MEM) by Engle (2002b):

$$\delta_t = \theta_1 + \theta_2 \delta_{t-1} + \theta_3 \tilde{\lambda}_{t-1}. \quad (1.16)$$

According to Lee and Hansen (1994), the parameters of equation 1.16 can be estimated with a GARCH software by taking $\tilde{\lambda}^{1/2}$ as dependent variable and setting the mean to zero.

1.2.3 Dynamics of C_t

The conditional mean of \mathbf{Y}_t is endowed with an autoregressive moving average process resembling the CAW model developed by Golosnoy *et al.* (2012). This model allows the predicted covariance matrix to depend on lagged covariance matrices as well as on their lagged predictions. For instance, the mean C_t of a Wishart density at time t can be expressed as a CAW of order (p, q) :

$$C_t = \mathbf{C}\mathbf{C}' + \sum_{i=1}^p \mathbf{A}_i C_{t-i} \mathbf{A}_i' + \sum_{j=1}^q \mathbf{B}_j \mathbf{Z}_{t-j} \mathbf{B}_j', \quad (1.17)$$

where \mathbf{C} is a lower-triangular matrix, and $\mathbf{A}_i, \mathbf{B}_j$ are the parameter matrices. The structure guarantees the symmetry and positive definiteness of the conditional mean C_t without imposing strong parametric restrictions on $(\mathbf{C}, \mathbf{A}_i, \mathbf{B}_i)$.

The number of parameters being estimated in a CAW(p, q) model increases rapidly with the number of assets. In fact, for m assets, the model involves $m(m+1)/2 + (p+q)m^2 + 1$ parameters. A straightforward restriction is to assume \mathbf{A}_i and \mathbf{B}_i to be diagonal matrices; in this case, the number of parameters becomes $m(m+1)/2 + (p+q)m + 1$.

The model can be further simplified by substituting the intercept with a function of the model's long-run covariance obtained as follows. Let $\text{vec}(\cdot)$ denote the operator that stacks all columns of a matrix into a vector, and let $\text{vech}(\cdot)$ and $\text{ivech}(\cdot)$ denote, respectively, the operator that stacks the lower triangular portion (including the diagonal) of a matrix into a vector and the one reversing this operation. Then, the vector representation of the CAW (p, q) is

$$c_t = c + \sum_{i=1}^p \mathcal{A}_i c_{t-i} + \sum_{j=1}^q \mathcal{B}_j z_{t-j}, \quad (1.18)$$

where $c_t = \text{vech}(C_t)$, $z_t = \text{vech}(\mathbf{Z}_t)$, $c = \text{vech}(\mathbf{C}\mathbf{C}')$ and $(\mathcal{A}_i, \mathcal{B}_j)$ are $k \times k$ matrices, with $k = m(m+1)/2$, obtained as

$$\mathcal{A}_i = L_m(\mathbf{A}_i \otimes \mathbf{A}_i) D_m, \quad \mathcal{B}_j = L_m(\mathbf{B}_j \otimes \mathbf{B}_j) D_m, \quad (1.19)$$

where L_m and D_m denote the elimination and duplication matrices, such that $\text{vec}(X) = D_m \text{vech}(X)$ and $\text{vech}(X) = L_m \text{vec}(X)$. Note that z_t can be written as $z_t = \mathbb{E}[z_t | \mathcal{I}_{t-1}] + v_t = c_t + v_t$, with $\mathbb{E}[v_t] = 0$ and $\mathbb{E}[v_s v_t'] = 0$ for all $s \neq t$. Then, the VARMA($\max(p, q), p$)

representation of the CAW(p, q) model is

$$z_t = c + \sum_{i=1}^{\max(p,q)} (\mathcal{A}_i + \mathcal{B}_i) z_{t-i} - \sum_{j=1}^p \mathcal{A}_j v_{t-j} + v_t, \quad (1.20)$$

with $\mathcal{A}_{p+1} = \dots = \mathcal{A}_q = 0$ if $p < q$ and $\mathcal{B}_{q+1} = \dots = \mathcal{B}_p = 0$ if $q < p$. The unconditional mean of z_t is given by

$$\mathbb{E}[z_t] = \left(I_k - \sum_{i=1}^{\max(p,q)} (\mathcal{A}_i + \mathcal{B}_i) \right)^{-1} c, \quad (1.21)$$

and, therefore, the covariance targeting variant is

$$C_t = \tilde{\mathbf{C}} + \sum_{i=1}^p \mathbf{A}_i C_{t-i} \mathbf{A}_i' + \sum_{j=1}^q \mathbf{B}_j \mathbf{Z}_{t-j} \mathbf{B}_j', \quad (1.22)$$

where $\tilde{\mathbf{C}} = \text{ivech}(c)$ has to be positive definite to ensure that C_t is also positive definite. We adopt an analogous approach for modeling the conditional mean of \mathbf{Y} , using diagonal parameter matrices with $p = 1$, $q = 1$ and covariance targeting. The resulting model is

$$C_t = \tilde{\mathbf{C}} + \mathbf{A} C_{t-1} \mathbf{A}' + \mathbf{B} (\mathbf{Z}_{t-1} / \tilde{\lambda}_{t-1}) \mathbf{B}', \quad (1.23)$$

where \mathbf{Z}_{t-1} in equation 1.17 is replaced by $\mathbf{Z}_{t-1} / \tilde{\lambda}_{t-1}$, since realized covariances are filtered by the terms used for the common factor estimation. In the application, the expected value of z_t in equation 1.21 is estimated by $\bar{z} = \text{vech}(\bar{\mathbf{Z}})$, where $\bar{\mathbf{Z}}$ is the sample mean of realized covariances divided by the corresponding $\tilde{\lambda}_{t-1}$.

1.2.4 Model estimation

Initially, the parameters in the δ_t equation are estimated from the data with a GARCH(1,1), using $\tilde{\lambda}^{1/2}$ as dependent variable and setting the mean to zero, according to Lee and Hansen (1994). Subsequently, the parameters $\psi = (\mathbf{A}, \mathbf{B}, \alpha, n)$ are estimated through maximum likelihood. Assuming that $\mathbf{Z}_t \sim \mathcal{G}(\alpha, \delta_t C_t, n)$, the log-likelihood function is

$$\begin{aligned} l_T(\psi) = & \sum_{t=1}^T \left\{ \frac{nm}{2} \log \left(\frac{n}{2} \right) + \frac{n-m-1}{2} \log \left(|\mathbf{z}_t| \right) + \log \left(\Gamma(nm/2 - \alpha) \right) \right. \\ & - \log \left(h(n, m) \right) - \frac{n}{2} \log \left(|\delta_t C_t| \right) - \alpha \log \left(-\alpha - 1 \right) - \log \left(\Gamma(-\alpha) \right) \\ & \left. + \left(\alpha - \frac{nm}{2} \right) \log \left(\frac{n}{2} \text{Tr} \left((\delta_t C_t)^{-1} \mathbf{z}_t \right) - (\alpha + 1) \right) \right\}, \quad (1.24) \end{aligned}$$

and the estimated parameters are

$$\hat{\psi} = \underset{\psi \in \Psi}{\operatorname{argmax}} \hat{l}_T(\psi). \quad (1.25)$$

We assume that given correct model specification and under the standard regularity conditions hold and that the maximum likelihood estimator of the parameter has an asymptotically normal distribution with mean equal to the true parameter and covariance matrix given by the inverse of the Fisher information matrix. However, we understand that evaluating the correctness of model specification is challenging and thus we might interpret the estimator as a quasi-maximum likelihood one, and use the sandwich estimator for the covariance.

Since the model is estimated in two stages, we acknowledge that the standard errors are not the most efficient ones, as demonstrated by Newey and McFadden (1994).

1.3 Empirical analysis

1.3.1 Data

We use high-frequency data at 1-minute level (390 daily observations) of the 26 assets of DJIA with available data for the whole period of analysis. The realized covariance matrices are computed as $Z_t = \sum_{i=1}^I r_{t,i} r'_{t,i}$, where $r_{t,i}$ is the vector of percentage log-returns for the i -th 1-min interval of the t -th trading day.¹ The sample period is 2 January 2003 to 24 June 2020, with 4400 trading days. Table 1.1 is an overview of the realized measures, reporting minimum, median and maximum for three quantiles of every measure (quantiles are computed over time). Apart from a few cases, the realized covariances and correlations are mostly positive. This will reduce the extent to which portfolio risk can be managed by resorting to diversification benefits, as we will show in the following applications. Figure 1.1 plots the time series of the log-average realized variance (top panel) and covariance (bottom panel), showing two main subperiods of greater instability, coinciding with the 2008 financial crises and the 2020 pandemic crises.

¹We are aware that other approaches for estimating the realized covariance exist. However, as our purpose is to illustrate the model properties, we consider the simplest estimator. The comparison of the model fit according to different estimators for the realized covariance is beyond the scope of the present thesis.

TABLE 1.1: Summary statistics for realized measures.

<i>quantiles</i>	Var			Cov			Corr		
	<i>0.05</i>	<i>0.5</i>	<i>0.95</i>	<i>0.05</i>	<i>0.5</i>	<i>0.95</i>	<i>0.05</i>	<i>0.5</i>	<i>0.95</i>
Min.	0.272	0.663	2.969	-0.044	0.159	1.170	-0.059	0.163	0.438
Med.	0.458	1.264	6.463	0.009	0.260	2.362	0.013	0.230	0.527
Max.	0.701	1.890	13.748	0.142	0.642	7.601	0.171	0.489	0.727

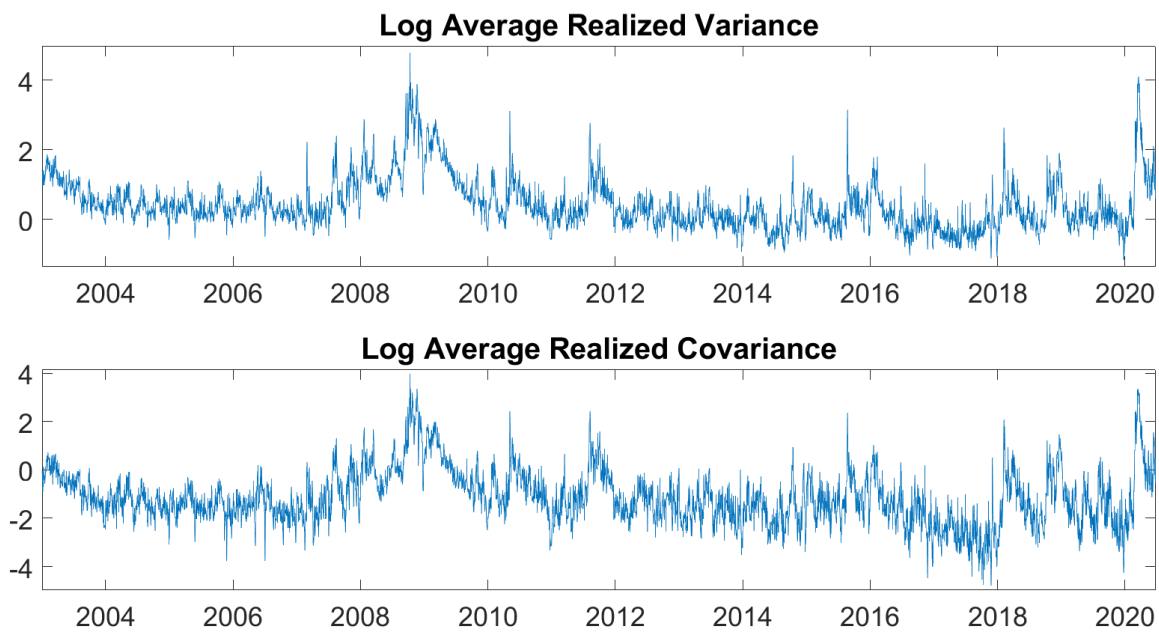


FIGURE 1.1: Average realized variance and covariance.

1.3.2 Estimation results

We first show that δ_t can be interpreted as a proxy for market variance. We fit the model to the complete sample of daily realized covariances to recover the fitted series of the common factor for the whole period. The estimation results are summarized in Table 1.2. As the actual number of assets considered is 26, δ_t would be slightly different if computed on all DJIA assets. Nevertheless, the correlation between the common factor and DJIA variance, where the latter is estimated by fitting an EGARCH(1,1) on daily data, is above 0.9, meaning that δ_t well represents market variance. Figure 1.2 plots δ_t and the variance showing a very close relation. Likewise, Figure 1.3 plots δ_t and

the index realized variance, which is obviously more volatile than the variance obtained from the EGARCH(1,1). Figure 1.4 presents the corresponding scatter plots.

TABLE 1.2: Estimated parameters and the corresponding standard errors (sandwich approach).

Asset	A_{ii}		B_{ii}		Asset	A_{ii}		B_{ii}	
	Value	SE	Value	SE		Value	SE	Value	SE
AAPL	0.8184	0.0116	0.1688	0.0103	JNJ	0.9026	0.0077	0.0871	0.0067
AMGN	0.8692	0.0083	0.1197	0.0073	JPM	0.8595	0.0093	0.1207	0.0074
AXP	0.8647	0.0107	0.1162	0.0083	KO	0.8761	0.0084	0.1066	0.0067
BA	0.8441	0.0170	0.1319	0.0130	MCD	0.8681	0.0103	0.1183	0.0087
CAT	0.8450	0.0097	0.1304	0.0076	MMM	0.8731	0.0081	0.1088	0.0064
CSCO	0.8487	0.0087	0.1334	0.0072	MRK	0.8476	0.0130	0.1305	0.0105
CVX	0.8770	0.0067	0.1107	0.0056	MSFT	0.8448	0.0084	0.1340	0.0068
DIS	0.8165	0.0181	0.1540	0.0140	NKE	0.8412	0.0138	0.1324	0.0107
GS	0.8493	0.0100	0.1276	0.0079	PG	0.8726	0.0083	0.1106	0.0067
HD	0.8394	0.0112	0.1332	0.0087	UNH	0.8312	0.0122	0.1415	0.0094
HON	0.8806	0.0101	0.1029	0.0079	VZ	0.8726	0.0083	0.1132	0.0069
IBM	0.8594	0.0130	0.1189	0.0102	WBA	0.8485	0.0117	0.1309	0.0094
INTC	0.8399	0.0100	0.1385	0.0081	WMT	0.8425	0.0154	0.1370	0.0120
Param.	Value	SE							
n	83.0081	0.1889							
α	4.9829	0.0582							
θ_1	0.0175	0.0023							
θ_2	0.3534	0.0330							
θ_3	0.6293	0.0359							

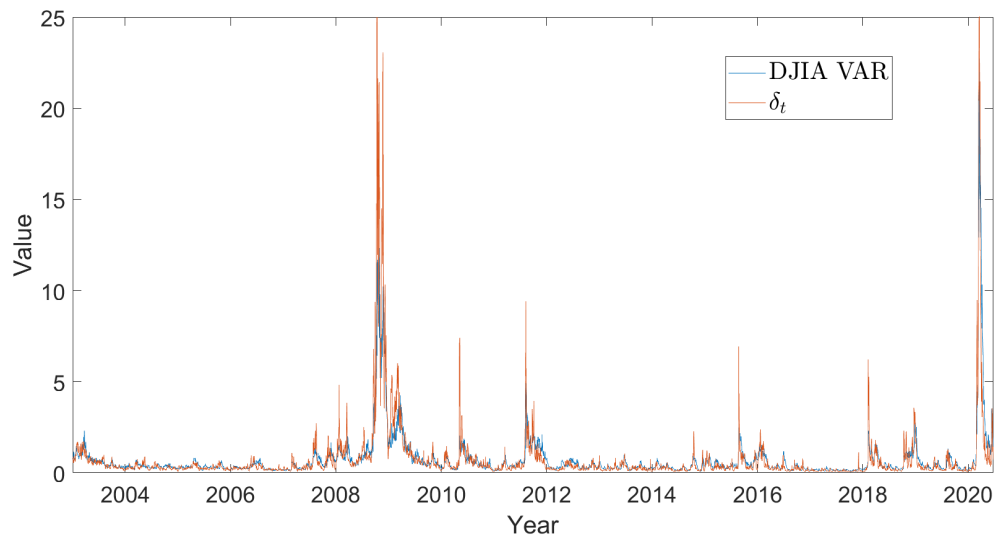


FIGURE 1.2: Common factor and DJIA variance.

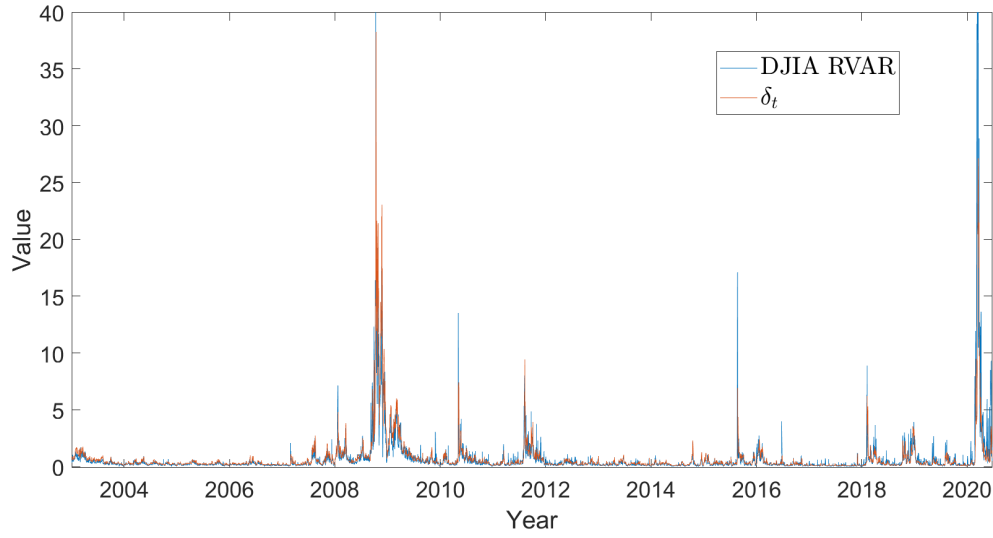


FIGURE 1.3: Common factor and DJIA realized variance.

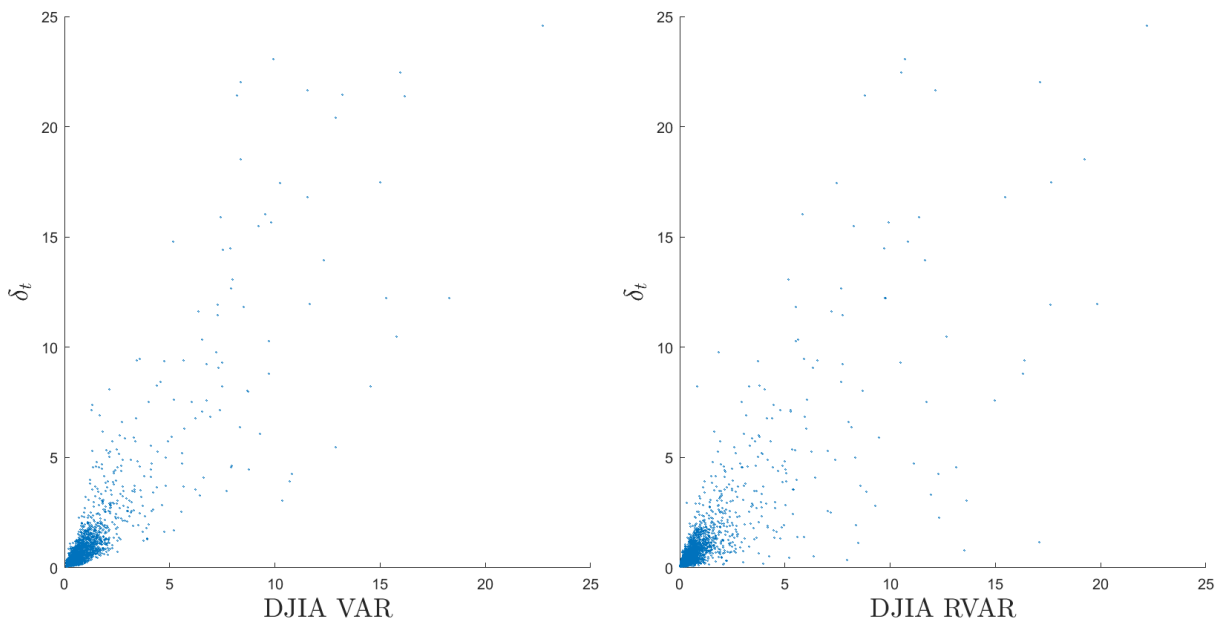


FIGURE 1.4: Scatter plots of δ_t versus DJIA variance and realized variance.

The crucial role of δ_t also emerges when we focus on its relevance within the conditional mean of realized variances and covariances. Figure 1.5 shows the effect of removing

the common factor from the fitted conditional mean of the realized covariance matrix elements, using the median value of the off-diagonal elements of Z_t and C_t . The latter plays a relatively marginal role in explaining realized covariances, as it exhibits no distinctive movement and has a correlation of 0.02 with the median of the off-diagonal elements of Z_t and $\mathbb{E}[Z_t|\mathcal{I}_{t-1}]$.

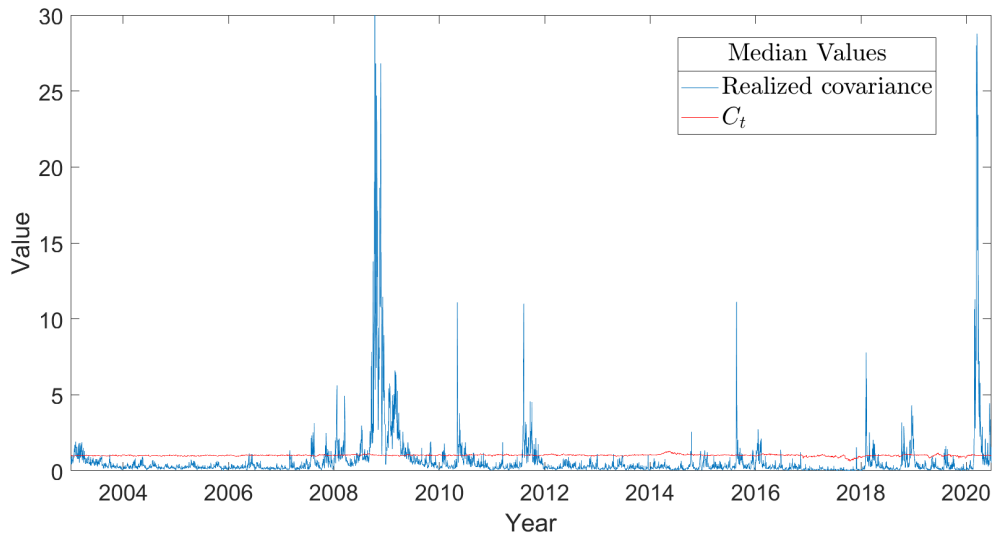


FIGURE 1.5: Median of realized covariances and fitted values of C_t (off-diagonal).

Table 1.3 compares the proposed conditional autoregressive \mathcal{G} model with the baseline CAW specification, using the AIC and BIC.

TABLE 1.3: AIC and BIC for the conditional autoregressive \mathcal{G} model compared to the standard CAW.

	AIC	BIC
\mathcal{G}	-742.6	-378.4
CAW	-677.5	-338.8

Most of the variability of the index dynamics is therefore explained by the common factor, leaving a peripheral role to the C_t component. This is also shown in the following results, obtained as the one-step-ahead forecast after fitting the model in a rolling window approach, using 500 observations for the estimation. Figure 1.6 shows the forecasted values of the off-diagonal elements of C_t for different quantiles. In particular, the

median value floats around 1; thus, the corresponding value of the off-diagonal elements of $\delta_t C_t$ is very close to the common factor itself, as indicated in Figure 1.7. Analogously, Figure 1.8 provides a similar representation considering median values of the diagonal elements of C_t and $\delta_t C_t$. Fluctuations in the conditional variance-covariance matrix are mostly captured by the common factor, which promptly adapts to sudden market changes, while residual movements are explained by the elements in C_t . This is more noticeable for periods marked by extreme instability, indeed we can clearly identify the aftermath of the 2008 financial crisis and the recent turmoil caused by the pandemic.

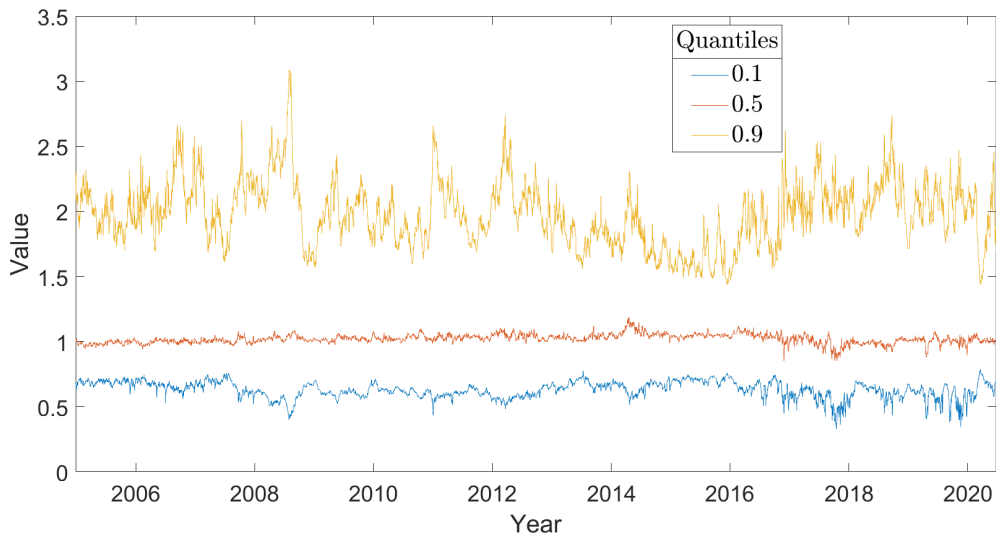


FIGURE 1.6: Quantiles of forecasted C_t (off-diagonal elements).

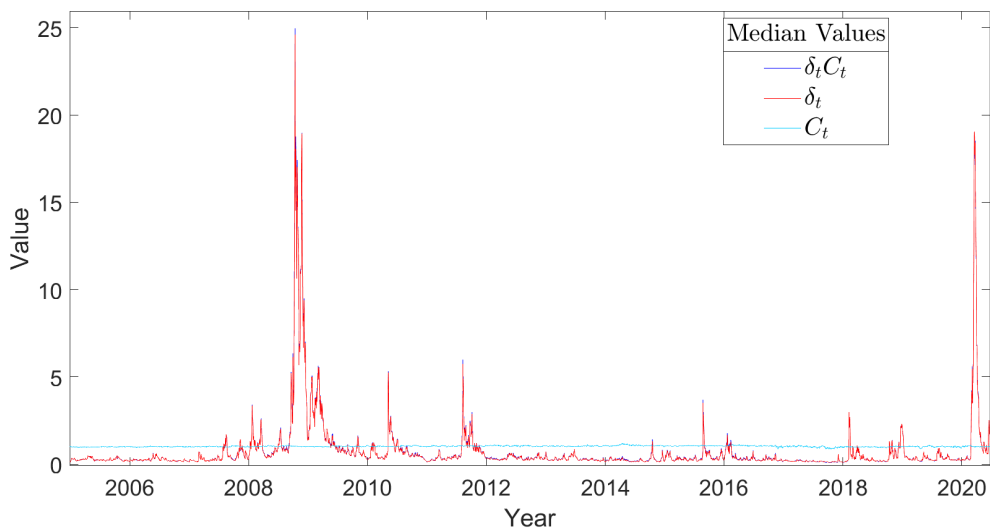


FIGURE 1.7: Median forecasts for δ_t , C_t and $\delta_t C_t$ (off-diagonal elements).

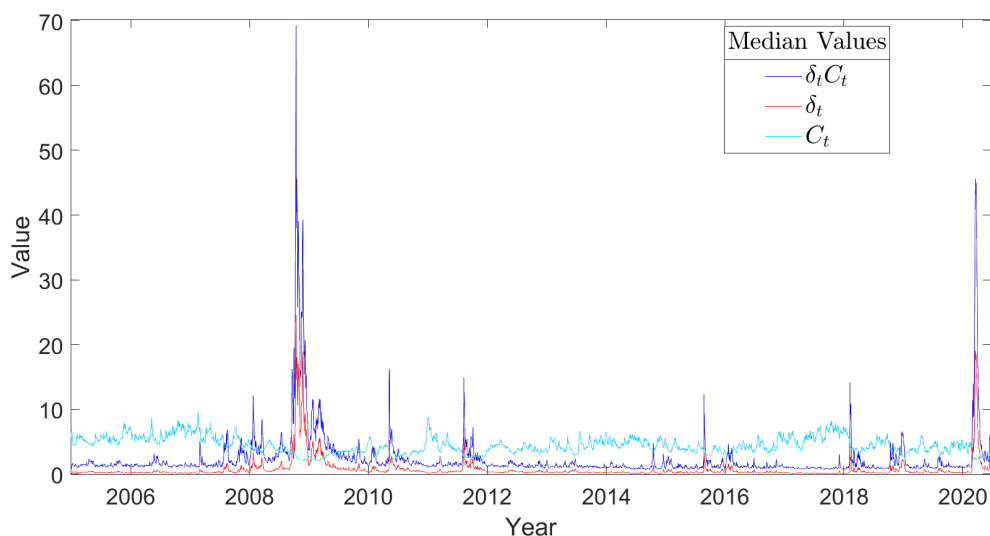


FIGURE 1.8: Median forecasts for δ_t , C_t and $\delta_t C_t$ (diagonal elements).

1.3.3 Market risk tracking portfolio

The interesting property of the common factor to accurately track market variance can be useful in a portfolio allocation framework. As seen before, the model allows to separate out the effect of the common factor from the remainder part; δ_t represents a proxy for the portion of un-diversifiable risk, while C_t embeds the remaining idiosyncratic risk component. It is, therefore, possible to build portfolios whose expected variance equals the common factor itself (i.e. at the systematic level) by imposing $w' C w = 1$, where w is the vector of portfolio weights. To take a realistic perspective, we develop an empirical example in which we use the predicted values of C_t , as mentioned in the previous section, to create a no short selling portfolio with daily rebalancing. Table 1.4 summarizes the performance in terms of profitability and portfolio rotation, compared to the Global Minimum Variance (GMV) and the Equally Weighted (EW) portfolios. In particular, the target variance portfolio delivers superior return with a very low cardinality. To provide a comprehensive representation of portfolio composition, we group companies according to their Industry Classification Benchmark (ICB). This methodology aims to partition companies into 11 broad industry levels. The necessary data are provided by Datastream and point at the companies' ICB. Figure 1.9 displays the portfolio weights using a 40-day moving average.² We can notice greater diversification

²The moving average is introduced for graphical analysis only and is not adopted in the portfolio construction.

before the 2008 financial crisis, even though with moderate relevance of the *Energy*, *Financials*, and *Telecommunications* industries, which tend to become gradually less important, together with *Industrials*, in proximity of the crisis. In the midst of the crisis, the portfolio is almost entirely concentrated on the *Consumer Discretionary* industry. The latter aspect is not observable during the recent pandemic crisis, while we can still notice a reduction in some industries such as *Energy*, *Financials* and *Industrials*.

Considering the aim of the optimization, it is reasonable to compare the portfolio and the index realized variances to assess effectiveness. This is shown graphically in Figure 1.10, where the portfolio properly tracks the index, with a correlation of 0.89.

TABLE 1.4: Target variance and GMV portfolios.

Portfolio	Sharpe	Turnover	Cardinality
Tgt	0.0456	0.128	15.7
GMV	0.0451	0.100	17
EW	0.0451	0.008	26

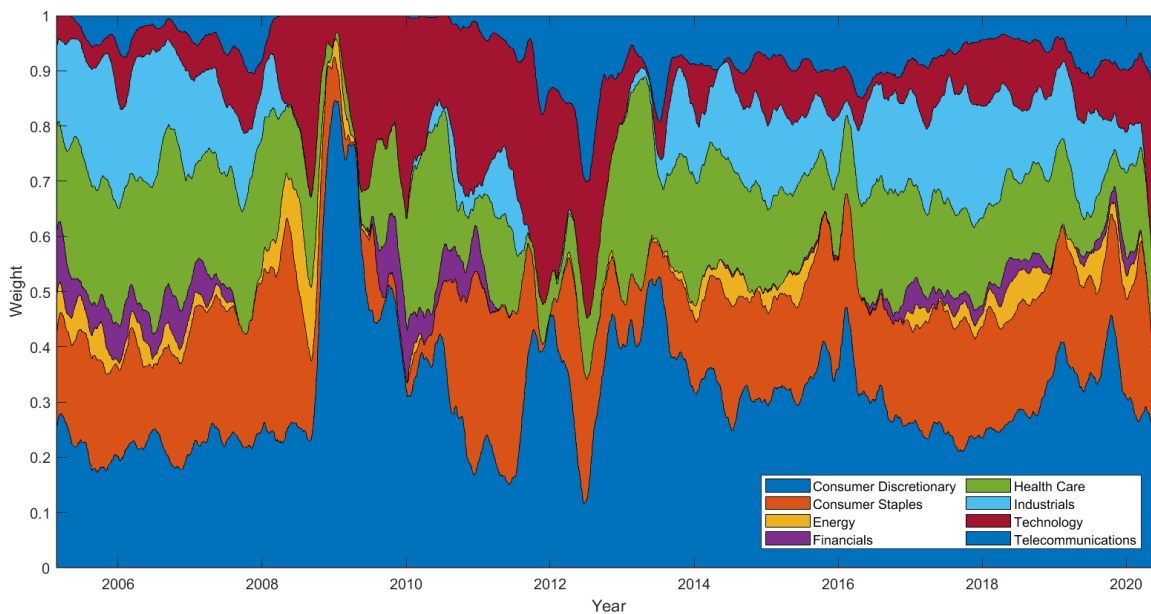


FIGURE 1.9: Weight by industry (40 days MA).

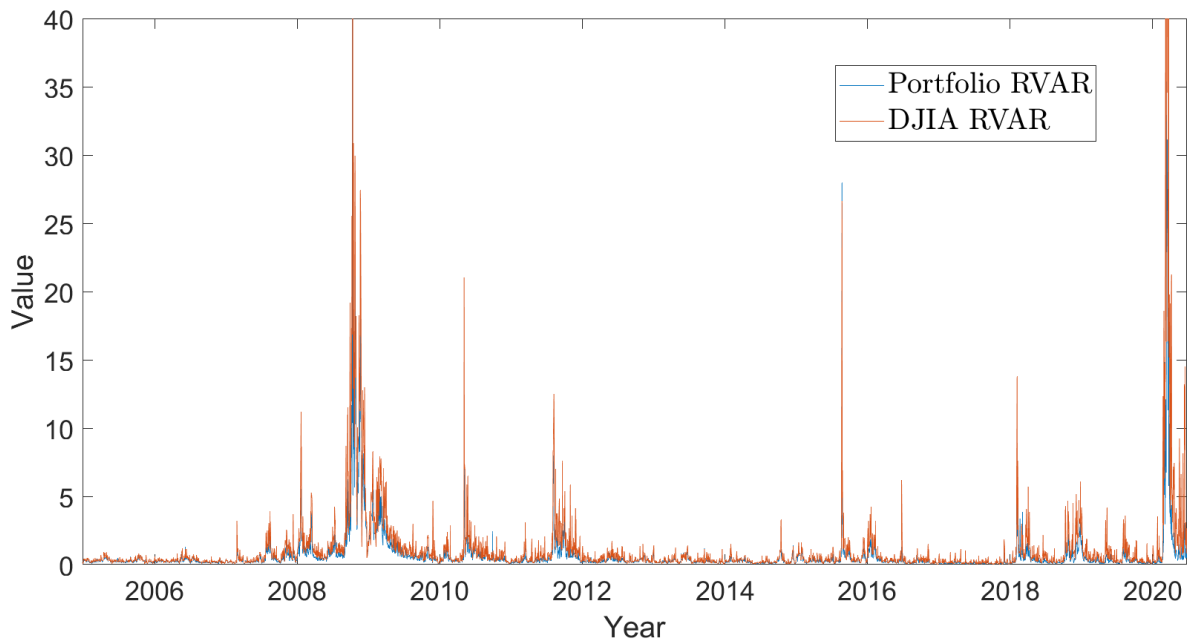


FIGURE 1.10: Portfolio and DJIA realized variances.

For a realistic assessment, the portfolio performance is also compared with the index after deducting the transaction costs, which are assumed to be linear as in DeMiguel *et al.* (2009). Let $c > 0$ be the level of proportional transaction costs expressed in basis points, then the net return of a portfolio P at time $t + 1$ is

$$R_{P,t+1|t}^N = \left(1 - c \sum_{m=1}^M |w_{m,t+1} - w_{m,t+}| \right) (1 + R_{P,t+1|t}^G) - 1, \quad (1.26)$$

where $w_{m,t+}$ is the weight of asset m at time $t + 1$ just before rebalancing. Figure 1.11 plots the cumulated returns for different levels of c . The daily Sharpe ratio of DJIA is approximately 0.024 for the period of interest, which is comparable to the value of 0.026 achieved by the portfolio under proportional transaction costs of 15 basis points.

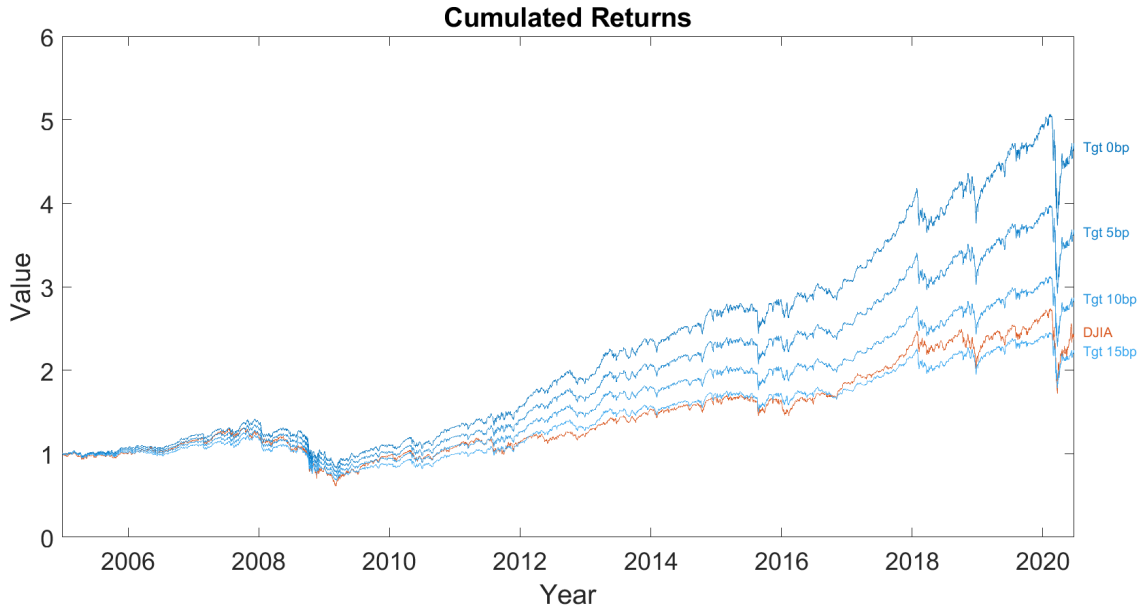


FIGURE 1.11: Portfolio net return and DJIA gross return.

1.3.4 An active asset allocation application

In this section, we provide a further possible empirical analysis involving our model. The proposal starts from one of the features of our modeling strategy, namely the role of δ_t , which tries to mimic the market risk movements. The portfolio with expected variance equal to this common factor, obtained in the previous section, can be, therefore, exploited to set up and handle actively managed portfolio strategies. Treynor and Black (1973) proposed an efficient method which does not require excessive insight or personal beliefs about the market we are operating in, and thus, looks appropriate for the scope. The optimal portfolio mixes a benchmark with actively chosen securities whose weights depend on their abnormal returns and their residual standard deviations. We briefly introduce the essential features and notations of the model. Let the excess rate of return of the i th security at time t over the risk-free rate be

$$R_{i,t} = \alpha_i + \beta_i R_{B,t} + e_{i,t},$$

where $R_{B,t}$ is the excess return of the benchmark index, α_i and β_i are, respectively, the abnormal excess return level and excess return sensitivity to the benchmark, while $e_{i,t}$ is the residual. The position in each security is $w_i = w_i^0 / \sum_{i=1}^n w_i^0$, where $w_i^0 = \alpha_i / \sigma^2(e_i)$. Furthermore, the model establishes the precise share of wealth that is devoted to the

active portfolio:

$$w_A^* = \frac{w_A^0}{1 + (1 - \beta_A)w_A^0},$$

where

$$w_A^0 = \frac{\alpha_A / \sigma^2(e_A)}{E(R_B) / \sigma_B^2},$$

$\alpha_A = \sum_{i=1}^n w_i \alpha_i$, $\beta_A = \sum_{i=1}^n w_i \beta_i$ and $\sigma^2(e_A) = \sum_{i=1}^n w_i^2 \sigma^2(e_i)$. As illustrated, most of the model implications rely on the accuracy of α_i , which are difficult to estimate with sufficient precision. Such uncertainty is not only relevant within the active portfolio, but largely affects its share in the optimal risky portfolio. Additionally, low estimates of $E(R_B)$ (in absolute terms) can exacerbate the problem resulting in completely unrealistic levels of w_A^0 . Different solutions to mitigate the estimation problem have been proposed. He (2007) revisited the model in a Bayesian framework, accounting for the estimation risk of model parameters and simultaneously considering the expected active returns α_i , and the variance–covariance structure of active returns. The solution shrinks the sample estimates of alpha towards the equilibrium-implied value of zero. Pástor and Stambaugh (2000) compared the portfolios selected with three different pricing models used as prior specification. MacKinlay and Pástor (2000) found that factor-based pricing models can be improved when a factor is unobserved, by imposing a link between the alphas and the variance–covariance of returns. Hence, in such a framework, covariances provide information that can be useful in estimating the means. This leads to more precise estimates of expected returns and improved portfolio selection.

This section aims to propose an implementation of our model as a benchmark index and not to improve the effectiveness of the most refined methods in the active portfolio theory; therefore, we relieve the methodology and apply a plain version of the Treynor–Black model. Sticking to the previous section, we impose long-only positions, so $w_i = 0$ for all the securities with a negative α_i , and we impose a 20% cap on each security within the active portfolio, to limit overexposure and mitigate the α_i estimation problem. Furthermore, we set $w_A^* = 0.3$ to avoid making the active portfolio prominent in the operation. For the estimation, we fit a linear regression model considering a rolling window approach using 1000 data points (about 4 years of daily data); thus, the output comprises 2900 daily results, from 16 December 2008 to 24 June 2020. The risk-free rate used is the daily value of the 13-week United States Treasury Bill (recovered from Datastream, with mnemonic FRTBS3M). The results, net of transaction costs, are presented in Table 1.5. The active portfolio becomes increasingly beneficial when the transaction costs rise; thus, the advantage lies in reducing the turnover while preserving a similar return level.

TABLE 1.5: Sharpe ratios net of transaction costs for market risk mimicking portfolio (passive benchmark) and Active optimal portfolio.

BP	Sharpe Passive	Sharpe Active
0	0.0616	0.0600
5	0.0546	0.0550
10	0.0477	0.0500
15	0.0408	0.0451
20	0.0339	0.0401
30	0.0201	0.0301

1.4 Conclusions

We propose a new multivariate approach for modeling realized covariances of financial assets. The main advantage, compared to standard models for dealing with high-frequency data, is the employment of \mathcal{G} distribution, which allows to separate a common factor embedded in realized covariances and a matrix component. The former captures the inherent risk of the market, while the latter represents the idiosyncratic share of risk. The model is empirically tested in its covariance targeting version with autoregressive and moving average components of order one. We show that the common factor detected by the model is highly correlated with market variance and plays a crucial role in estimating and forecasting conditional covariance. We provide a model application to an asset allocation framework, with the aim of tracking the market by curbing the impact of idiosyncratic risk. The realized variance of the resulting portfolio is close to the common factor itself as well as to the realized variance of the market. Additionally, the risk-adjusted performance of the portfolio is comparable to the market in case proportional transaction costs of 15 basis points are applied. We provide a further extension to the empirical analysis by using the market tracking portfolio as the benchmark to identify companies providing higher than expected return and to exploit such companies in an actively managed portfolio strategy.

For future research, it will be interesting to employ and assess further specifications of the conditional autoregressive \mathcal{G} model as alternative approaches within the frameworks that currently rely on the CAW and WAR models. A possible variant that still addresses the curse of dimensionality problem might include a diagonal intercept in the model instead of using the covariance targeting method, as proposed by Shen *et al.* (2020).

Another relevant element to work on might be to increase the flexibility of the conditional autoregressive \mathcal{G} model by taking a cue from the paper of Yu *et al.* (2017), who developed the generalized CAW model, which possesses features of both CAW and WAR models. A final note should be reserved for the common factor estimation procedure. In this chapter we used the basic MEM to enable fast implementation of the model and its preliminary evaluation, but we might consider applying an extension to improve the accuracy of the estimate, for instance following Brownlees *et al.* (2011).

Chapter 2

Blended frequency model

2.1 State of the art and proposal

An important component of portfolio management, of a possibly large set of assets, is the modeling and forecasting of the variances and covariances. As in the seminal work of Markowitz (1952), the simplest approach is to consider i.i.d. returns and estimate them using sample estimators. However, the two most common stylized facts of financial returns are volatility clustering and heavy-tails. Sample estimator of the covariance matrix captures neither of them, and several multivariate GARCH (MGARCH) models have been proposed to take into account these and other features of returns (see Silvenoinen and Teräsvirta (2009) and references therein). Many of them, however, exhibit limitations due to the sprawl of parameters that need to be estimated, and some of them end up being applicable to a fistful of assets and thus rendering their usefulness in large-scale portfolio optimization futile. Some thrifty specifications include the VECH model of Bollerslev *et al.* (1988), the BEKK model of Engle and Kroner (1995), the constant conditional correlation (CCC) and the dynamic conditional correlation (DCC) models of Bollerslev (1990) and Engle (2002a), respectively.

The MGARCH literature has partly overcome the assumption of multivariate normality of returns. One can use different copula constructions as in Paoletta and Polak (2015a, 2018). However this requires computationally intensive simulations from the estimated copula distribution in high dimensions. A more efficient approach that accounts for asymmetric and heavy-tailed returns is the COMFORT model developed by Paoletta and Polak (2015b). According to the COMFORT specification, the returns are conditionally described by the multivariate generalized hyperbolic (MGHyp) distribution, which allows to include a univariate component common to all the assets, and to endow the conditional dispersion matrix with a GARCH-like structure. Compared to

other non-Gaussian approaches, the estimation via maximum likelihood is feasible for a potentially large numbers of assets owing to an Expectation-Conditional Maximization Either (ECME) algorithm, thus ameliorating the curse of dimensionality. In the context of portfolio optimization, the improved covariance matrix forecasts provided by the broad class of MGARCH models generally lead to more competitive risk-adjusted returns, however, the covariance matrix dynamics of such models are extremely sensitive to new data, leading to frequent pronounced changes in the forecasts and in portfolio weights. Thus, the inclusion of transaction costs allows for a more realistic view of the situation, and highlights the trade-off between the benefits and costs of portfolio rebalancing. An orthogonal specification of the COMFORT model is proposed by Paolella *et al.* (2019) to address this issue. The resulting COMFORT-PCA model allows the dispersion matrix dynamics to be driven by the leading factors in a principal component decomposition, and thus enables to reduce portfolio turnover thanks to the limited evolution of the dispersion matrix. Compared to the original specification, the latter is much more beneficial in real-world applications, where transaction costs impair performances that seem deceptively good.

The aim of this chapter is to exploit data retrieved at high-frequency (HF) as additional information to enhance and refine the performance of the COMFORT-PCA model. The starting point is the model proposed by Black and Litterman (1992), where the authors incorporate some subjective prior views in the analysis. In that context the views represent the personal opinion of an investor about the possible movements in the market, and the model is the statistical way used to quantify and embed the views in the analysis. Among the various extensions of the model, Meucci (2005) and Meucci (2010) propose to rephrase the problem in a more parsimonious way that allows for the inclusion of scenario analysis, by incorporating the views on the future returns and not on the parameters of the distribution. As is often the case, one of the limitations is the Gaussianity assumption of the prior and posterior distribution of returns, while, as mentioned before, we aim to operate with heavy-tailed asymmetric returns. Thus, we derive the posterior distribution under the MGHyp assumption, by incorporating the views from HF data.

We select 100 assets from the S&P 500 and use the proposed approach as well as benchmark approaches to perform a multi-step ahead forecast of the covariance matrix and expected returns. The forecasts are subsequently used in a portfolio allocation analysis aiming to observe how different approaches behave along the efficient frontier: we show that benchmark approaches based solely on high or low-frequency data provide poorer results the closer we move towards the Maximum Sharpe portfolio, while the

blended approach preserves stable results along the frontier.

The rest of this chapter is structured as follows. Section 2.2 provides an overview of the Black-Litterman paradigm. In Section 2.3, we firstly describe the COMFORT-PCA model and the HF data modeling. Then we derive the posterior for the MGHyp distribution and describe the simulation procedure to get multi-step-ahead forecasts using filtered historical simulation. The last part of Section 2.3 concerns the asset allocation procedure and the performance evaluation. Section 2.4 is devoted to the empirical analysis: we start by comparing the proposed approach with the benchmark models, then we replicate the same analysis for the limiting case of the multivariate Gaussian distribution, and finally we highlight how different models behave in two subperiods of the sample. Section 2.5 concludes the main results.

2.2 The Black-Litterman approach

The Black and Litterman model is a portfolio allocation model developed in 1990 at Goldman Sachs by Fischer Black and Robert Litterman (see Black and Litterman, 1990 and Black and Litterman, 1992) which seeks to overcome some of the problems that institutional investors have encountered in applying portfolio theory in practice. The model takes into account the specific opinions (called views) of the investor about asset returns, to obtain an allocation blending the initial inputs with such views. Indeed, it provides a procedure for quantifying the investor's views and their inherent uncertainty, expressed either in absolute or relative terms, and incorporate them into the prediction of the distribution of future returns. The model outputs the estimates of expected returns along with the corresponding precision estimates. In spite of its novelty and success in the field of asset allocation, the Black and Litterman (1992) paper does not provide all derivations and formulas used in the model. Moreover, during the thirty years since the original papers, many authors have published research referring to their model as Black-Litterman, despite being very different from the original one. The survey of Walters *et al.* (2014) provides a chronology of the significant papers contributing to the Black-Litterman literature, as well as the taxonomy of models which have been labeled as Black Litterman. Naturally, the survey contains a complete description of the canonical model including full derivations using both Theil's Mixed Estimation model and Bayes Theory. The various parameters of the model are assessed, along with information on their computation or calibration.

Bevan and Winkelmann (1998) provide details on how they use Black-Litterman as part of their broader asset allocation process at Goldman Sachs, including some

calibrations of the model. They focus on their performance to illustrate how the Black-Litterman framework can be used for designing investment strategies. The Black-Litterman model allows investors to take risk when they have views, with stronger views justifying more risk-taking. The accuracy of the views is then reflected by assessing the performance of the strategy relative to the benchmark.

The research by Mankert (2010) reveals that it is not the Black-Litterman model alone that is rewarding, but rather the combination of model-user situations. Overall, the research indicates the great gap between theory and practice, and the importance of understanding the model in order to keep a critical attitude towards the model itself and its output.

As previously mentioned, many extensions have been proposed since the first release of the model in 1990. Giacometti *et al.* (2007) improve the classical Black-Litterman model by applying more realistic distributions for asset returns and by using alternative risk measures (dispersion-based risk measures, value at risk, conditional value at risk). The results are reported for monthly data and the performance of the models is tested using a rolling window of fixed size along a fixed horizon. The authors show that the incorporation of the investors' views in the model provides information on how the different distributional hypotheses can impact the optimal composition of the portfolio. This work is close to the approach proposed in this thesis because of the non-Gaussian assumption.

A clear exposition of the original Black-Litterman model, including the derivation of the output parameters, is presented by Meucci (2005) and Meucci (2010). Consider a market of N assets whose returns are normally distributed:

$$\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \quad (2.1)$$

While the covariance $\boldsymbol{\Sigma}$ is estimated by exponential smoothing of past returns, $\boldsymbol{\mu}$ is modeled as a random variable whose dispersion represents the possible estimation error. In particular, the original model states that $\boldsymbol{\mu}$ is normally distributed

$$\boldsymbol{\mu} \sim N(\boldsymbol{\pi}, \tau\boldsymbol{\Sigma}), \quad (2.2)$$

where $\boldsymbol{\pi}$ represents the best guess for $\boldsymbol{\mu}$ and $\tau\boldsymbol{\Sigma}$ the uncertainty on this guess. The original Black-Litterman model embraces a market equilibrium approach to set $\boldsymbol{\pi}$. Assuming

there is no estimation error, the reference model becomes

$$\mathbf{X} \sim N(\boldsymbol{\pi}, \boldsymbol{\Sigma}). \quad (2.3)$$

Moreover, assume that all investors maximize a mean-variance trade-off portfolio with unconstrained weights, that is:

$$\mathbf{w}_\lambda \equiv \underset{\mathbf{w}}{\operatorname{argmax}}\{\mathbf{w}'\boldsymbol{\pi} - \lambda\mathbf{w}'\boldsymbol{\Sigma}\mathbf{w}\}. \quad (2.4)$$

The problem can be solved explicitly by setting to zero the derivative with respect to \mathbf{w} of the term in curly brackets, obtaining the relationship between the equilibrium portfolio $\tilde{\mathbf{w}}$, resulting from the average risk-aversion level $\bar{\lambda}$, and $\boldsymbol{\pi}$:

$$\boldsymbol{\pi} \equiv 2\bar{\lambda}\boldsymbol{\Sigma}\tilde{\mathbf{w}}, \quad (2.5)$$

where $\bar{\lambda}$ is exogenous. Thus, the determination of $\boldsymbol{\pi}$ is not influenced by historical information, but rather depends on the market-weighted portfolio $\tilde{\mathbf{w}}$. The Black-Litterman model takes into account the investor's views affecting the expectation $\boldsymbol{\mu}$. A view is therefore a statement concerning the expected performance of an asset or asset class which may clash with the model 2.1. For instance, the investor might claim that an asset class will rise by 10% on an annualized basis, in absolute terms or relatively to another asset class. The K views are represented by a $K \times N$ "pick" matrix \mathbf{P} , whose generic k -th row determines the relative weight of each expected return in the corresponding view. The uncertainty about the views is expressed through the model:

$$\mathbf{P}\boldsymbol{\mu} \sim N(\mathbf{v}, \boldsymbol{\Omega}), \quad (2.6)$$

where \mathbf{v} and $\boldsymbol{\Omega}$ quantify the views and the corresponding uncertainty, respectively. The model also allows the investor to express qualitative views. In this scenario it is convenient to set the entries of \mathbf{v} in terms of the market volatility, that is

$$v_k \equiv (\mathbf{P}\boldsymbol{\pi})_k + \eta_k \sqrt{(\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}')_{k,k}}, \quad k = 1, \dots, K, \quad (2.7)$$

where η_k defines the qualitative view. For instance, one can choose $\eta_k \in \{-\beta, -\alpha, +\alpha, +\beta\}$ with $\alpha = 1$ and $\beta = 2$ to express "very bearish", "bearish", "bullish" and "very bullish" views, respectively. In addition, it is convenient to set

$$\boldsymbol{\Omega} \equiv \frac{1}{c}\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}' \quad (2.8)$$

as in Meucci (2010), where $c \in \{0, \infty\}$ represents the overall level of confidence in the views. Appendix A shows how to obtain the distribution of $\boldsymbol{\mu}$ given the views using Bayes' formula, that is

$$\boldsymbol{\mu}|\mathbf{v}; \boldsymbol{\Omega} \sim N(\boldsymbol{\mu}_{\text{BL}}, \boldsymbol{\Sigma}_{\text{BL}}^{\mu}), \quad (2.9)$$

where

$$\boldsymbol{\mu}_{\text{BL}} \equiv ((\tau\boldsymbol{\Sigma})^{-1} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P})^{-1} ((\tau\boldsymbol{\Sigma})^{-1}\boldsymbol{\pi} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{v}), \quad (2.10)$$

$$\boldsymbol{\Sigma}_{\text{BL}}^{\mu} \equiv ((\tau\boldsymbol{\Sigma})^{-1} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P})^{-1}. \quad (2.11)$$

However, the aim is to find the distribution of \mathbf{X} , which can be easily computed by writing the reference model 2.1 as $\mathbf{X} \stackrel{d}{=} \boldsymbol{\mu} + \mathbf{Z}$, where $\mathbf{Z} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$. The posterior model is therefore

$$\mathbf{X}|\mathbf{v}; \boldsymbol{\Omega} \sim N(\boldsymbol{\mu}_{\text{BL}}, \boldsymbol{\Sigma}_{\text{BL}}), \quad (2.12)$$

where $\boldsymbol{\mu}_{\text{BL}}$ is defined in equation 2.10 and $\boldsymbol{\Sigma}_{\text{BL}}$ follows from equation 2.11 by assuming that $\boldsymbol{\mu}$ and \mathbf{Z} are independent, that is

$$\boldsymbol{\Sigma}_{\text{BL}} = \boldsymbol{\Sigma} + \boldsymbol{\Sigma}_{\text{BL}}^{\mu}. \quad (2.13)$$

Note that equations 2.10 and 2.13 can be written equivalently as

$$\boldsymbol{\mu}_{\text{BL}} = \boldsymbol{\pi} + \tau\boldsymbol{\Sigma}\mathbf{P}'(\tau\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}' + \boldsymbol{\Omega})^{-1}(\mathbf{v} - \mathbf{P}\boldsymbol{\pi}) \quad (2.14)$$

$$\boldsymbol{\Sigma}_{\text{BL}} = (1 + \tau)\boldsymbol{\Sigma} - \tau^2\boldsymbol{\Sigma}\mathbf{P}'(\tau\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}' + \boldsymbol{\Omega})^{-1}\mathbf{P}\boldsymbol{\Sigma}. \quad (2.15)$$

The complete proof is reported in Appendix A.

2.3 Methodological framework

This section is devoted to the methodological procedure adopted to combine data at different frequencies, starting from an introductory part describing how low-frequency and high-frequency data are modeled separately and then moving on to the blending process. Finally, we describe the asset allocation strategies and discuss their assessment.

2.3.1 COMFORT PCA model

The fundamental point of the work is the use of the high-frequency data as a refinement of the low-frequency estimates. This requires a baseline model for the latter. In this chapter we use as benchmark the COMFORT-PCA model developed by Paoletta *et al.* (2019) which uses the multivariate generalized hyperbolic distribution to describe conditional returns. The main advantage is delivered by the inclusion of PCA in the procedure, allowing the dispersion matrix dynamics to be driven only by the leading factors in the decomposition, which are endowed with a univariate GARCH structure, while leaving the secondary factors constant over time. This leads to stable estimates and generates portfolios with much lower turnover and superior risk-adjusted returns net of transaction costs.

Analogously to Paoletta *et al.* (2019), let $\mathbf{Y}_t = (Y_{1,t}, Y_{2,t}, \dots, Y_{K,t})'$, $t = 1, 2, \dots, T$, be the daily return vector of K assets at time t , and let Φ_t be the information set up to time t . The conditional distribution of $\mathbf{Y}_t \mid \Phi_t$ is assumed to be Multivariate Generalized Hyperbolic (MGHyp) with the stochastic representation

$$\begin{aligned} \mathbf{Y}_t \mid \Phi_{t-1} &\stackrel{d}{=} \boldsymbol{\mu} + \boldsymbol{\gamma}G_t + \boldsymbol{\varepsilon}_t, \\ \boldsymbol{\varepsilon}_t &= \sqrt{G_t} \mathbf{H}_t^{1/2} \mathbf{Z}_t, \end{aligned} \quad (2.16)$$

where $\boldsymbol{\mu} = (\mu_1, \dots, \mu_K)' \in \mathbb{R}^K$ and $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_K)' \in \mathbb{R}^K$ are the location and asymmetry vectors, respectively, $\mathbf{H}_t \in \mathbb{R}^{K \times K}$ is a symmetric and positive definite conditional dispersion matrix, $\mathbf{Z}_t \stackrel{\text{iid}}{\sim} \mathcal{N}(\mathbf{0}, \mathbf{I}_K)$ is a multivariate Gaussian random vector, and $G_t \stackrel{\text{iid}}{\sim} \text{GIG}(\lambda, \chi, \psi)$ is a univariate mixing random variable, independent of \mathbf{Z}_t , with generalized inverse Gaussian (GIG) density given by

$$f_{G_t}(g) = \frac{\chi^{-\lambda} (\sqrt{\chi\psi})^\lambda}{2\mathcal{K}_\lambda(\sqrt{\chi\psi})} g^{\lambda-1} \exp\left(-\frac{1}{2}(\chi g^{-1} + \psi g)\right), \quad g > 0, \quad (2.17)$$

where \mathcal{K}_λ is the modified Bessel function of the third kind. The density function of the MGHyp distribution is

$$f_{\mathbf{Y}_t \mid \Phi_{t-1}}(\mathbf{y}) = c d^{\lambda-K/2} \mathcal{K}_{\lambda-K/2}(d) \exp\{(\mathbf{y} - \boldsymbol{\mu})' \mathbf{H}_t^{-1} \boldsymbol{\gamma}\}, \quad (2.18)$$

where

$$\begin{aligned} c &= \frac{(\sqrt{\chi\psi})^{-\lambda} (\psi + \boldsymbol{\gamma}' \mathbf{H}_t^{-1} \boldsymbol{\gamma})^{K/2 - \lambda} \psi^\lambda}{(2\pi)^{K/2} |\mathbf{H}_t|^{1/2} \mathcal{K}_\lambda(\sqrt{\chi\psi})}, \\ d &= [(\chi + (\mathbf{y} - \boldsymbol{\mu})' \mathbf{H}_t^{-1} (\mathbf{y} - \boldsymbol{\mu})) (\psi + \boldsymbol{\gamma}' \mathbf{H}_t^{-1} \boldsymbol{\gamma})]^{1/2}. \end{aligned}$$

The dispersion matrix \mathbf{H}_t can be decomposed as

$$\mathbf{H}_t = \mathbf{Q}\mathbf{\Xi}_t\mathbf{Q}', \quad (2.19)$$

where \mathbf{Q} is the orthogonal matrix of eigenvectors and $\mathbf{\Xi}_t = \text{diag}(\xi_{1,t}, \dots, \xi_{K,t})$ is the diagonal matrix of eigenvalues. The leading factors in $\mathbf{\Xi}_t$ are equipped with a GARCH(1,1) dynamics while the others are kept constant, so only the eigenvalues explaining the largest portion of variation in the data are time-varying.

The model parameters are estimated using an expectation conditional maximization either (ECME) algorithm.

2.3.2 High-frequency data

Realized measures are nonparametric estimators of daily price variation and covariation based on high-frequency data. The simplest of such realized measures is the realized covariance. Andersen *et al.* (2001) discuss the properties of quadratic variation and covariation, suggesting that, under suitable conditions, realized covariance is an unbiased and efficient estimator of return covariance. Based on this idea, we exploit high-frequency data to compute daily realized covariances. Let $\mathbf{r}_{t,i}$ be the vector of returns or log-returns for the i -th 1-min interval of the t -th trading day, then the realized covariance matrix is computed as $\mathbf{Z}_t = \sum_{i=1}^I \mathbf{r}_{t,i}\mathbf{r}'_{t,i}$. The first return of each day is computed on an open-to-close basis (i.e. using the opening and closing prices of the first trading minute) to disregard overnight variation of prices, while the other returns are computed on a close-to-close basis during the day. The aim of this work is to include HF data in a low-frequency-based model, so we use the realized covariance as an approximation of the actual and unobservable covariance matrix of the assets, without using any model to forecast covariances from HF data. Hence, we have that the expected covariance matrix and return vector from HF data are

$$\begin{aligned} \boldsymbol{\Sigma}_{\text{HF},T+1} &= \mathbf{Z}_T \\ \boldsymbol{\mu}_{\text{HF},T+1} &= \frac{1}{W_s} \sum_{t \in S} \sum_{i=1}^I \mathbf{r}_{t,i}, \end{aligned}$$

where the latter corresponds to the average daily log-return, disregarding overnight variations, over the time window $S = \{(T - W_s + 1), \dots, T\}$, of length W_s .

2.3.3 Combining HF and LF data

High-frequency data can be exploited as views about the assets and thus can be used to blend the results of the COMFORT model. However, before specifying the way high-frequency data are included, we need to derive the posterior distribution resulting from the mixing of prior beliefs following a MGHyp distribution and Normal views. According to the COMFORT model, returns follow a normal distribution conditionally on the mixing random variable, that is

$$\mathbf{Y}|G \sim N(\boldsymbol{\mu} + \gamma G, G\mathbf{H}). \quad (2.20)$$

As proposed by Meucci (2005) and Meucci (2010) we incorporate views (in our case from high-frequency data) not on the parameters of the distribution but directly on returns. In particular, we consider views as linear functions of the market $\mathbf{V} \equiv \mathbf{P}\mathbf{Y}$, where \mathbf{P} is a pick matrix characterizing relative views and \mathbf{Y} represents the returns of a broad market. The views \mathbf{V} are a perturbation of the outcome implied by the reference model and as such they are modeled using a conditional distribution, given the realization of the returns vector and of the mixing random variable (in order to accommodate possible common market shocks and their impact on the uncertainty in the views), that is

$$\mathbf{V}|\mathbf{Y} = \mathbf{y}; G = g \sim N(\mathbf{P}\mathbf{y}, g\boldsymbol{\Omega}), \quad (2.21)$$

where $\boldsymbol{\Omega}$ represents the uncertainty about the views. As shown by Meucci (2010), if

$$\mathbf{Y} \sim N(\boldsymbol{\mu}, \mathbf{H}), \quad (2.22)$$

and

$$\mathbf{V}|\mathbf{Y} = \mathbf{y} \sim N(\mathbf{P}\mathbf{y}, \boldsymbol{\Omega}), \quad (2.23)$$

then

$$\mathbf{Y}|\mathbf{V} = \mathbf{v}; \boldsymbol{\Omega} \sim N(\boldsymbol{\mu}_{\text{BL}}, \mathbf{H}_{\text{BL}}), \quad (2.24)$$

where

$$\boldsymbol{\mu}_{\text{BL}} = \boldsymbol{\mu} + \mathbf{H}\mathbf{P}^\top(\mathbf{P}\mathbf{H}\mathbf{P}^\top + \boldsymbol{\Omega})^{-1}(\mathbf{v} - \mathbf{P}\boldsymbol{\mu}), \quad (2.25)$$

$$\mathbf{H}_{\text{BL}} = \mathbf{H} - \mathbf{H}\mathbf{P}^\top(\mathbf{P}\mathbf{H}\mathbf{P}^\top + \boldsymbol{\Omega})^{-1}\mathbf{P}\mathbf{H}. \quad (2.26)$$

In our case, from Equation 2.20 we have that $(\mathbf{Y} - \gamma G)|G \sim N(\boldsymbol{\mu}, G\mathbf{H})$ so we can get the posterior using the results above, considering the views as reported in Equation 2.21. Thus, from Equation 2.20 and 2.21 we obtain that the posterior, given the realization of the mixing random variable, follows a normal distribution with mean and variance equal to

$$\boldsymbol{\mu}_{\text{BL}} = \boldsymbol{\mu} + \mathbf{H}\mathbf{P}^\top(\mathbf{P}\mathbf{H}\mathbf{P}^\top + \boldsymbol{\Omega})^{-1}(\mathbf{v} - \mathbf{P}\boldsymbol{\mu}), \quad (2.27)$$

$$G\mathbf{H}_{\text{BL}} = G\mathbf{H} - G\mathbf{H}\mathbf{P}^\top(\mathbf{P}\mathbf{H}\mathbf{P}^\top + \boldsymbol{\Omega})^{-1}\mathbf{P}\mathbf{H}, \quad (2.28)$$

respectively. This result implies that, without conditioning on the mixing random variable, the vector of returns admits the following decomposition

$$\mathbf{Y} = \boldsymbol{\mu}_{\text{BL}} + \gamma G + \sqrt{G}(\mathbf{H}_{\text{BL}})^{1/2}\mathbf{Z}, \quad (2.29)$$

which is a MGHyp distribution with updated location vector and dispersion matrix. We incorporate the HF data as views in the analysis by blending the outcome of the COMFORT-PCA model, that is we set

$$\begin{aligned} G\boldsymbol{\Omega} &= \boldsymbol{\Sigma}_{\text{HF}} \\ \mathbf{v} &= \boldsymbol{\mu}_{\text{HF}}. \end{aligned}$$

A more detailed derivation of expressions 2.25 and 2.26 is reported in Appendix B.

2.3.4 Multi-step forecast

For a realistic application we aim to reduce portfolio rebalancing and increase the holding period. Hence, we have to derive the expression for the forecasts over longer horizons. For the realized covariances we adopt the square root rule, which involves multiplying the last daily realized covariance by the length of the period. On the other hand, for the COMFORT model estimates we adopt a non-parametric approach called filtered historical simulation (FHS) to build up a multi-period conditional distribution of the random vector of returns. This approach was presented by Barone-Adesi *et al.* (1999), who introduced a simulation model based on the historical distribution of returns, thus not imposing any theoretical distribution on the data. The returns are filtered to remove serial correlation and volatility clustering. The resulting standardized returns are independently and identically distributed so the non-parametric bootstrap can be applied. Polak and Ulrych (2021) applied this approach to the baseline version of the COMFORT model, which also includes the mixing random variable.

In this section we extend the approach to the COMFORT-PCA model specified in Equation 2.16. The $(K \times 1)$ standardized residuals vector at time t is

$$\tilde{\mathbf{e}}^{(t)} = \frac{(\mathbf{Y}_t - \hat{\boldsymbol{\mu}} - \hat{\gamma}\mathbf{E}[G_t | \boldsymbol{\Phi}_t])\mathbf{Q}\hat{\boldsymbol{\Sigma}}_t^{-1/2}}{\sqrt{\mathbf{E}[G_t | \boldsymbol{\Phi}_t]}}, \quad (2.30)$$

where $\mathbf{E}[G_t | \boldsymbol{\Phi}_t]$ corresponds to the imputed mixing random variable from the expectation step of the ECME-algorithm. As in Polak and Ulrych (2021), we denote with $\tilde{\mathbf{Z}}$ the $(K \times T)$ matrix of standardized residuals in Equation 2.30. The resulting filtered historical innovations can be drawn randomly with replacement and used as innovations to generate pathways of future returns, although empirical observations depart from the hypothesis of independence and identical distribution (see Barone-Adesi *et al.* (1999)).

Analogously to Polak and Ulrych (2021), we also need to standardize the filtered mixing random variables related to the corresponding standardized residuals for the filtered historical simulation. We use a kernel cumulative distribution function (cdf) estimation to estimate the cdf of the mixing random variable in a non-parametric way. Consider the sample of imputed mixing random variables $\mathbf{E}[G_t | \boldsymbol{\Phi}_t]$, for $t = 1, \dots, T$, and denote with $K_{cdf,T}(\cdot)$ the kernel cdf. Then $K_{cdf,T}(\mathbf{E}[G_t | \boldsymbol{\Phi}_t])$ is the realization of the empirical kernel cdf of the filtered mixing random variable. Additionally, we remind that $(G_{T+1} | \boldsymbol{\Phi}_T) \sim \text{GIG}(\lambda, \chi, \psi)$, so if we denote with $\text{GIG}_{cdf,T}(\cdot)$ and $\text{GIG}_{cdf,T}^{-1}(\cdot)$ the corresponding cdf and its inverse, then the standardized conditional mixing random variables are

$$\tilde{G}^{(t)} = \text{GIG}_{cdf,T}^{-1}(K_{cdf,T}(\mathbf{E}[G_t | \boldsymbol{\Phi}_t])). \quad (2.31)$$

The $(T \times 1)$ vector $\tilde{\mathbf{G}}$ containing the standardized mixing random variables expressed in Equation 2.31 is defined in such a way that each component is related to the corresponding entry of matrix $\tilde{\mathbf{Z}}$.

The filtered historical simulation for the h -step ahead $\mathbf{Y}_{T+h|\boldsymbol{\Phi}_T}^{FHS,(i)}$ is recursively obtained using

$$\mathbf{Y}_{T+h|\boldsymbol{\Phi}_T}^{FHS,(i)} = \hat{\boldsymbol{\mu}} + \hat{\gamma}\tilde{G}_{T+h|\boldsymbol{\Phi}_T} + \sqrt{\tilde{G}_{T+h|\boldsymbol{\Phi}_T}}\mathbf{Q}\hat{\boldsymbol{\Sigma}}_{T+h|\boldsymbol{\Phi}_T}^{1/2}\tilde{\mathbf{Z}}_{T+h|\boldsymbol{\Phi}_T}, \quad i = 1, \dots, B,$$

where B denotes the number of simulations, $\tilde{G}_{T+h|\boldsymbol{\Phi}_T}$ and $\tilde{\mathbf{Z}}_{T+h|\boldsymbol{\Phi}_T}$ are realizations of $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{Z}}$ for a randomly chosen t , and $\hat{\boldsymbol{\Sigma}}_{T+h|\boldsymbol{\Phi}_T}$ is a $(K \times K)$ diagonal matrix such that

$$\hat{\boldsymbol{\Sigma}}_{T+h|\boldsymbol{\Phi}_T} = \begin{cases} \hat{\omega} + \hat{\alpha}(\mathbf{Y}_T - \hat{\boldsymbol{\mu}} - \hat{\gamma}\mathbf{E}[G_T | \boldsymbol{\Phi}_T])^2 + \hat{\beta}(\hat{\boldsymbol{\Sigma}}_{T|\boldsymbol{\Phi}_T}) & \text{for } h = 1 \\ \hat{\omega} + \hat{\alpha}(\mathbf{Y}_{T+h-1|\boldsymbol{\Phi}_T}^{FHS,(i)} - \hat{\boldsymbol{\mu}} - \hat{\gamma}\tilde{G}_{T+h-1|\boldsymbol{\Phi}_T})^2 + \hat{\beta}(\hat{\boldsymbol{\Sigma}}_{T+h-1|\boldsymbol{\Phi}_T}) & \text{for } h > 1 \end{cases},$$

where $\hat{\omega}$, $\hat{\alpha}$ and $\hat{\beta}$ are the GARCH parameters estimated in the COMFORT-PCA model. We can now repeatedly draw sequences of returns and find the h -step conditional cumulative log-return, that is

$$\mathbf{Y}_{T+h|\Phi_T}^{FHS,(i)}(h) = \mathbf{Y}_{T+1|\Phi_T}^{FHS,(i)} + \mathbf{Y}_{T+2|\Phi_T}^{FHS,(i)} + \dots + \mathbf{Y}_{T+h|\Phi_T}^{FHS,(i)}, \quad i = 1, \dots, B, \quad (2.32)$$

and compute the h -step covariance matrix directly from simulated values, for a sufficiently large B .

2.3.5 Asset allocation strategies and performance evaluation

The aim of the analysis is to assess how different models behave when moving along the efficient frontier. We consider as extremes the Global Minimum Variance (GMV) and the Maximum Sharpe (MS) portfolios, and we equally split (in terms of risk) the frontier in the middle to obtain 10 further portfolios. We assume a realistic 4-weeks rebalancing (i.e. 20 days) so the inputs of the mean-variance optimization are the multi-step forecasts obtained as described in Section 2.3.4. Moreover, performances are assessed net of transaction costs, which are assumed to be proportional as in DeMiguel *et al.* (2009). Let $c > 0$ be the level of proportional transaction costs expressed in basis points, then the net return of a portfolio P at time $t + 1$ is

$$\mu_{P,t+h|t}^N = \left(1 - c \sum_{k=1}^K |w_{k,t+h} - w_{k,t+}| \right) (1 + \mu_{P,t+h|t}^G) - 1, \quad (2.33)$$

where $w_{k,t+}$ is the weight of asset m at time $t + h$ just before rebalancing, and $\mu_{P,t+h|t}^N$ and $\mu_{P,t+h|t}^G$ are, respectively, the net and gross portfolio returns at time $t + h$ given the weights selected at time t . Let $\mathbf{1}_K$ be a K -dimensional vector whose elements are 1, Σ and $\boldsymbol{\mu}$ the assets' expected covariance matrix and expected return vector, respectively, then the vector of weights \mathbf{w} is the outcome of the usual optimization problem including the no-short-selling constraint, that is

$$\begin{aligned} \min_{\mathbf{w}} \quad & \mathbf{w}'\Sigma\mathbf{w} \\ \text{s.t.} \quad & \mathbf{w}'\mathbf{1}_K = 1 \\ & w_k \geq 0 \quad \forall k \end{aligned}$$

for the GMV portfolio, while

$$\begin{aligned} \max_{\mathbf{w}} \quad & \frac{\mathbf{w}'\boldsymbol{\mu}}{\sqrt{\mathbf{w}'\boldsymbol{\Sigma}\mathbf{w}}} \\ \text{s.t.} \quad & \mathbf{w}'\mathbf{1}_K = 1 \\ & w_k \geq 0 \quad \forall k \end{aligned}$$

for the MS portfolio assuming a zero risk-free rate, and finally

$$\begin{aligned} \max_{\mathbf{w}} \quad & \mathbf{w}'\boldsymbol{\mu} \\ \text{s.t.} \quad & \sigma_P^2 = \mathbf{w}'\boldsymbol{\Sigma}\mathbf{w} \\ & \mathbf{w}'\mathbf{1}_K = 1 \\ & w_k \geq 0 \quad \forall k \end{aligned}$$

for portfolios with a target variance level, where σ_P^2 is the expected portfolio variance. The obtained portfolios are evaluated based on their Sharpe Ratio and Sortino Ratio after deducing transaction costs, as in Equation 2.33.

2.4 Empirical analysis

We use high-frequency data at 1-minute level (390 daily observations) of the 100 assets with higher capitalization within the S&P 500 index and with available data for the whole period of analysis. The realized covariance matrices Z_t are computed as described in Section 2.3.2 for the period from October 16, 2006 to the end of December 2019.

2.4.1 Asset allocation strategies and performances

For the empirical analysis we adopt a 250-days rolling window approach. We consider the 20-days forecasts obtained from the LF and HF data to get the blended forecasts, as described in the Section 2.3.3. Since we are mixing data at different frequencies and not expressing real views, it is reasonable to set $\mathbf{P} = \mathbf{I}$, where \mathbf{I} is the identity matrix, that is each view is about the single asset and all views have the same magnitude. Additionally, for the empirical analysis we set the number of eigenvalues endowed with a GARCH structure to 3, coherently with the findings of Paoletta *et al.* (2019). Thus, from equation 2.19, $\boldsymbol{\Xi}_t = \text{diag}(\xi_{1,t}, \xi_{2,t}, \xi_{3,t}, \xi_4, \dots, \xi_K)$.

We use the forecasts to build the GMV and MS portfolios and consider 10 further portfolios between these two extremes, as mentioned in Section 2.3.5. The portfolios

are then rebalanced every 20 days, coherently with the forecasts, and evaluated net of transaction costs based on the Sharpe Ratio and Sortino Ratio. The results are shown in Figure 2.1 and 2.2. The difference between the three approaches is minimal for the GMV portfolio, while it intensifies as we move towards the MS portfolio. Specifically, the Sharpe and Sortino ratios are stable across different portfolios when the mixed-frequency approach is considered whilst, conversely, employing a pure high or low-frequency-based method leads to inferior results as we get closer to the MS portfolio. Additionally, this result holds for different levels of transaction costs, although it can be noticed that, as the level increases, the COMFORT-PCA model narrows the gap with the blended approach, whereas the HF model performs substantially worse. Thus, in such a framework, where reducing turnover is essential, the main feature of the COMFORT-PCA model stands out. On the other hand, the accuracy of high-frequency data backfires because of the higher turnover induced.

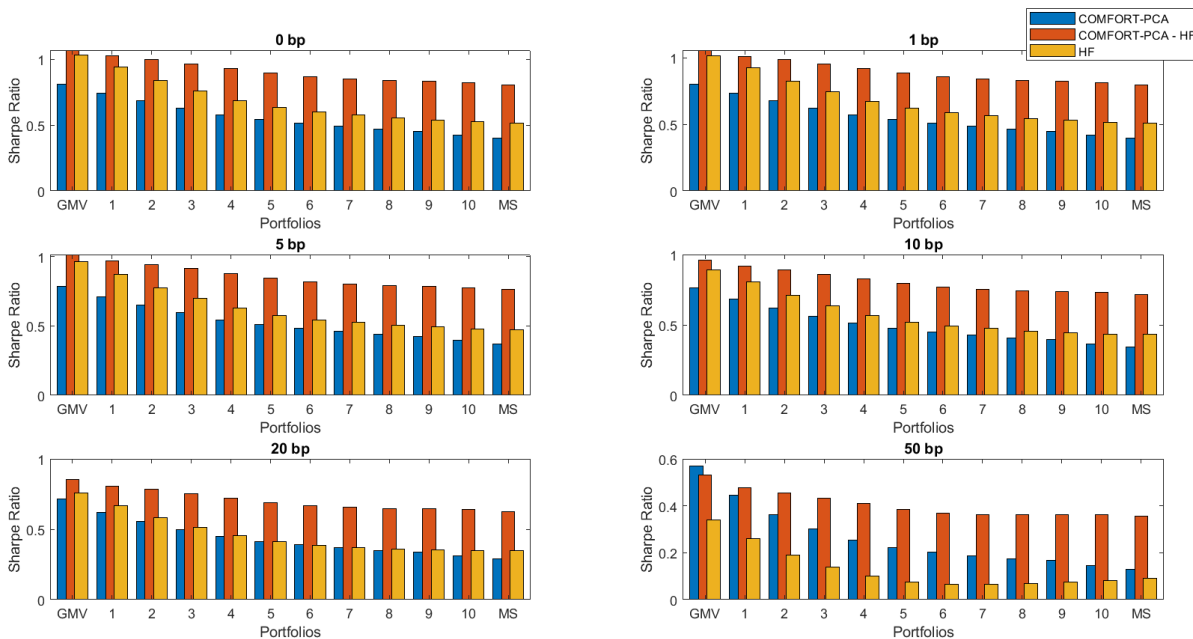


FIGURE 2.1: Annualized Sharpe ratio.

The improvement provided by the blended model is both in terms of higher average return and lower volatility. Figure 2.3 shows the combination of these two effects for the three approaches, considering the zero-transaction costs case. When shifting towards more volatile portfolios, the mixed model maintains a stable return while slightly increasing variance, instead the other approaches have lower annualized return and their

annualized variance increases rapidly. Analogous results are obtained for positive levels of transaction cost, although not displayed for conciseness.

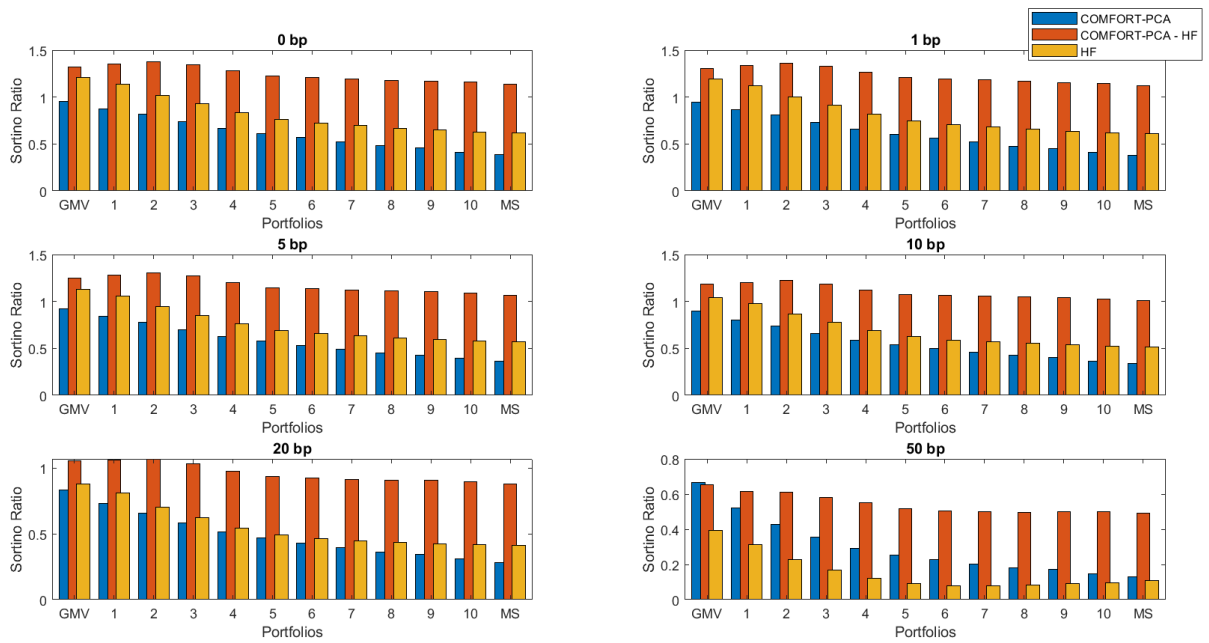


FIGURE 2.2: Annualized Sortino ratio.

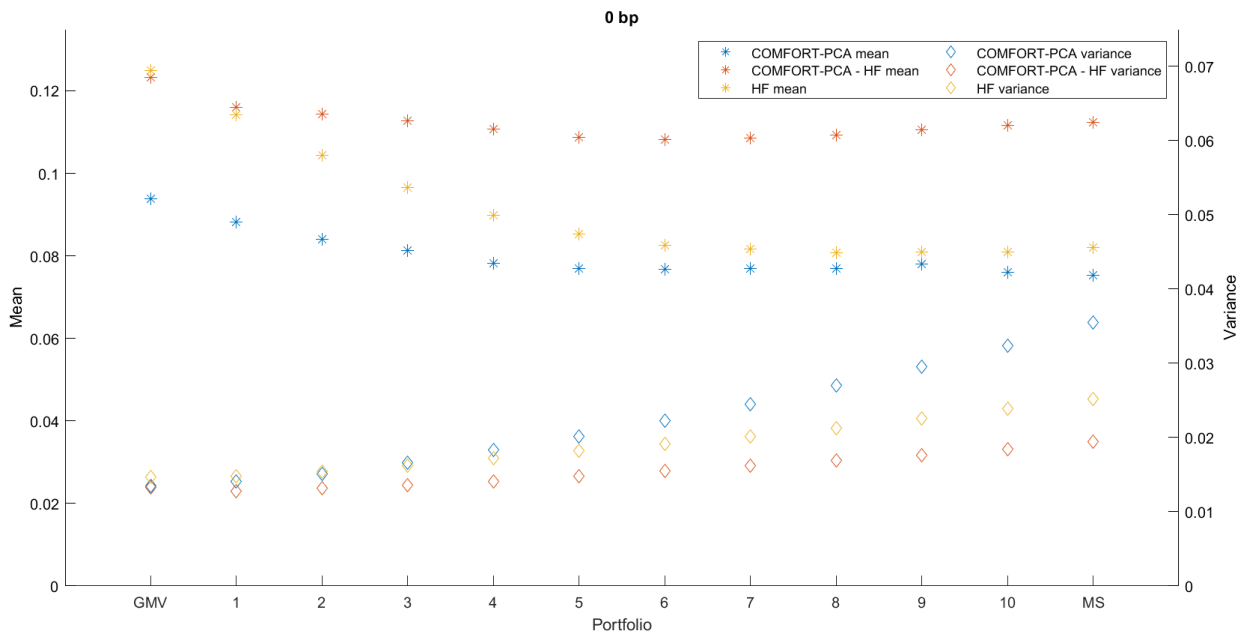


FIGURE 2.3: Annualized mean and variance of returns (0bp).

As a comparison, we use the Gaussian version of the COMFORT-PCA model, that is the limiting case with $\gamma = 0$ and $G_t = 1$ when looking at Equation 2.16. Figure 2.4 and 2.5 show, respectively, the annualized Sharpe and Sortino ratios for this version of the model, highlighting a mild improvement in the blended approach, except for high level of transaction costs. Although the improvement is more restrained, we observe again that using information from different frequencies leads to a superior signal when data are applied in this context. It should be made clear that the Gaussian version of the model is not intended to be a benchmark approach to beat, but rather an alternative application of the blended approach, even though in the research paper by Paoletta *et al.* (2019) the inclusion of asymmetry and of the mixing random variable led to better results in other financial applications.

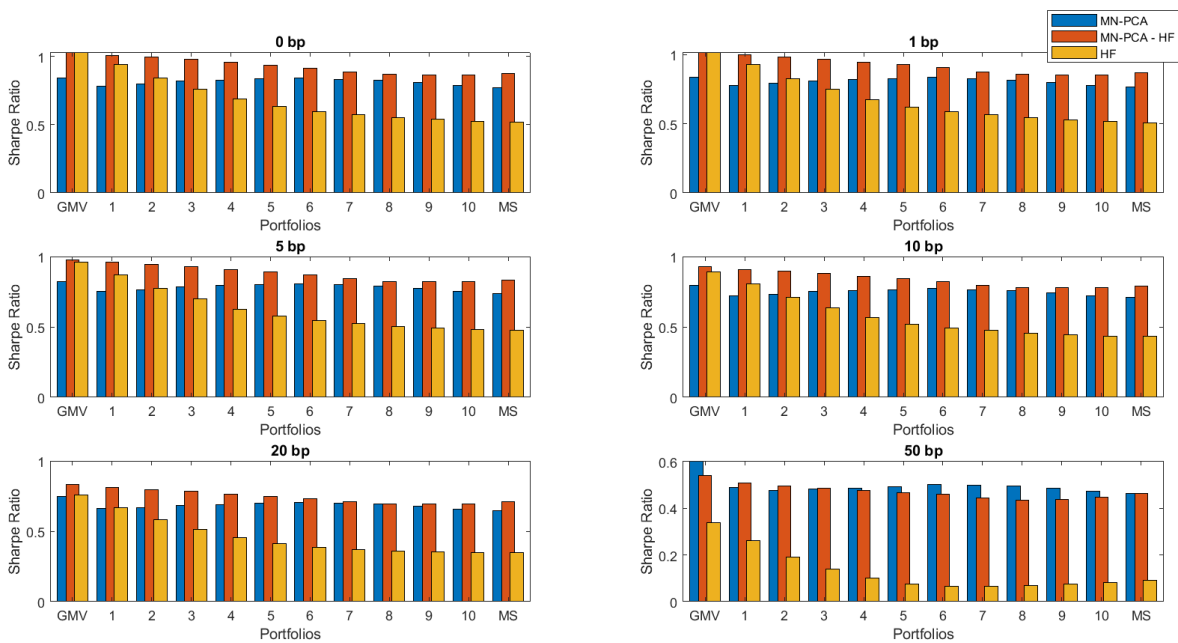


FIGURE 2.4: Annualized Sharpe ratio (Gaussian model).

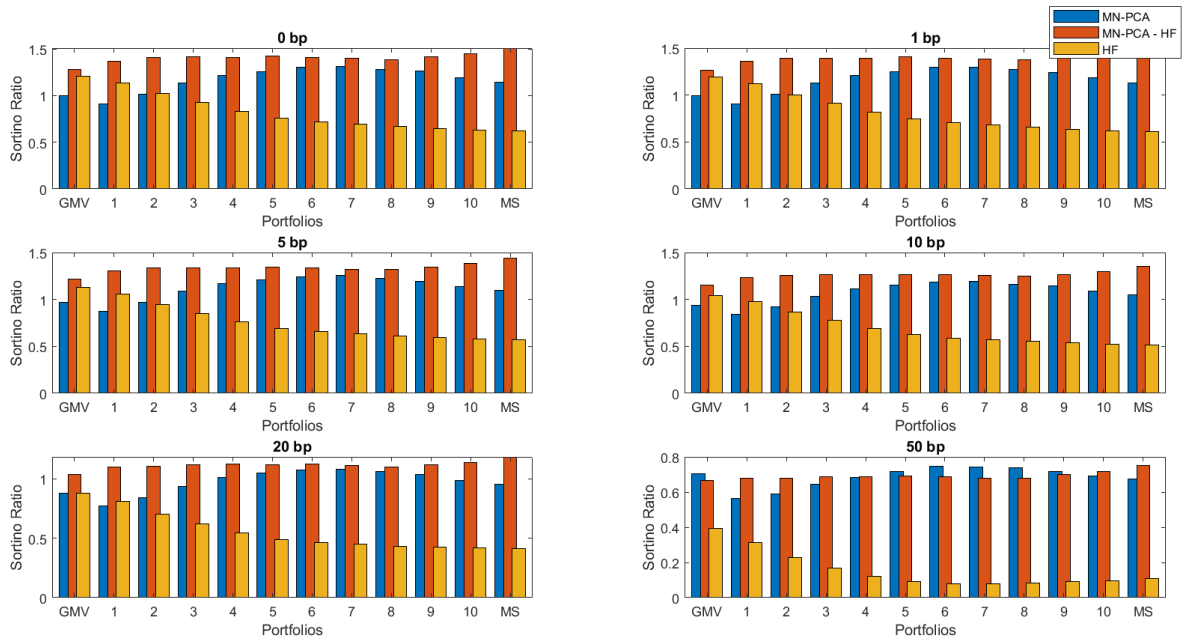


FIGURE 2.5: Annualized Sortino ratio (Gaussian model).

Figure 2.6 shows the results, in terms of annualized return and volatility, obtained with the COMFORT-PCA model and its Gaussian version, compared to the ensuing blended approaches and the high-frequency-based strategy. The blended models provides generally stable returns for all portfolios, at a level above the other approaches, and at the same time they exhibit a moderate increase in volatility while the counterparties are markedly less performing.

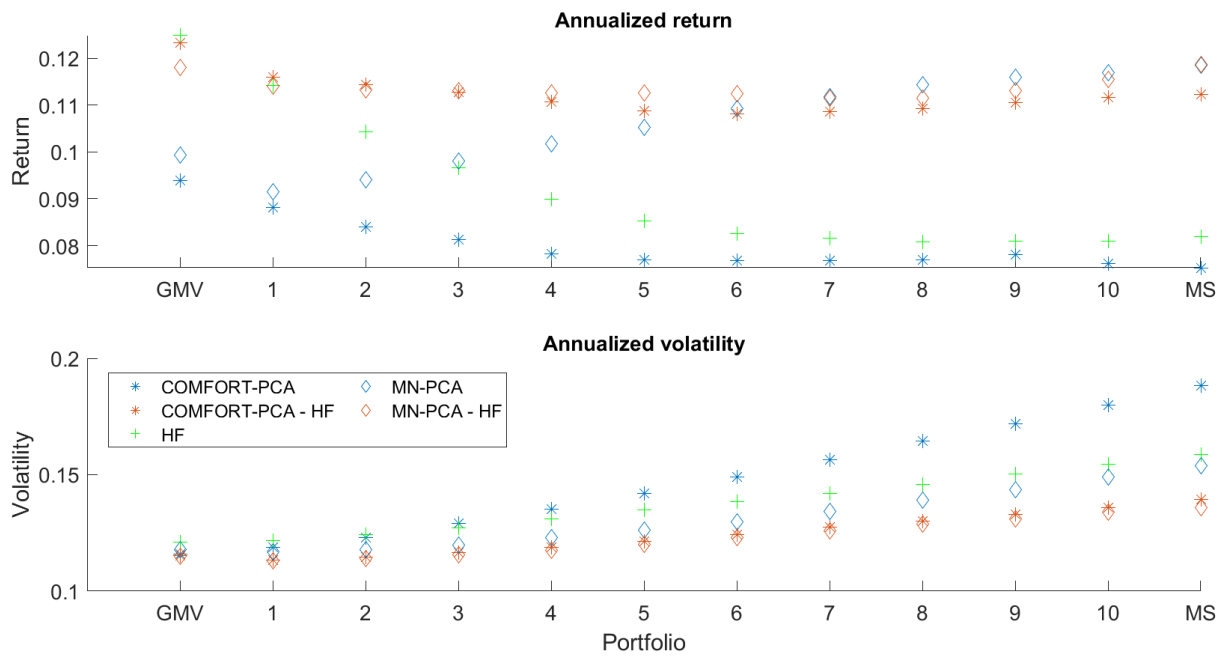


FIGURE 2.6: Annualized return and volatility (0bp).

Another important feature of any approach adopted, concerns the presence of severe drawdowns that persist for a prolonged period. A drawdown measure quantifies the decline in portfolio value from a historical peak. Ideally, it is advisable to have as few drawdowns as possible, and in the event that they happen, to have mild and short declines. Figure 2.7 and 2.8 want to take this measure into consideration for the original version of the COMFORT-PCA. They plot the return and the corresponding drawdown of the GMV and MS portfolios for three approaches, in case of absence of transaction costs. From the plots we can clearly observe the aftermath of the 2008 financial crises, underlined by a very substantial drawdown that lasts for years. As expected, the GMV portfolios exhibit a quicker recovery from the losses and in general the HF and the blended approaches provide close performances. With regard to the MS portfolios we can immediately notice longer and more remarkable drawdowns for all the approaches, compared to the GMV case, and a far better performance of the blended approach over the benchmarks, both in terms of return and in terms of drawdowns, which are milder and shorter for the former.

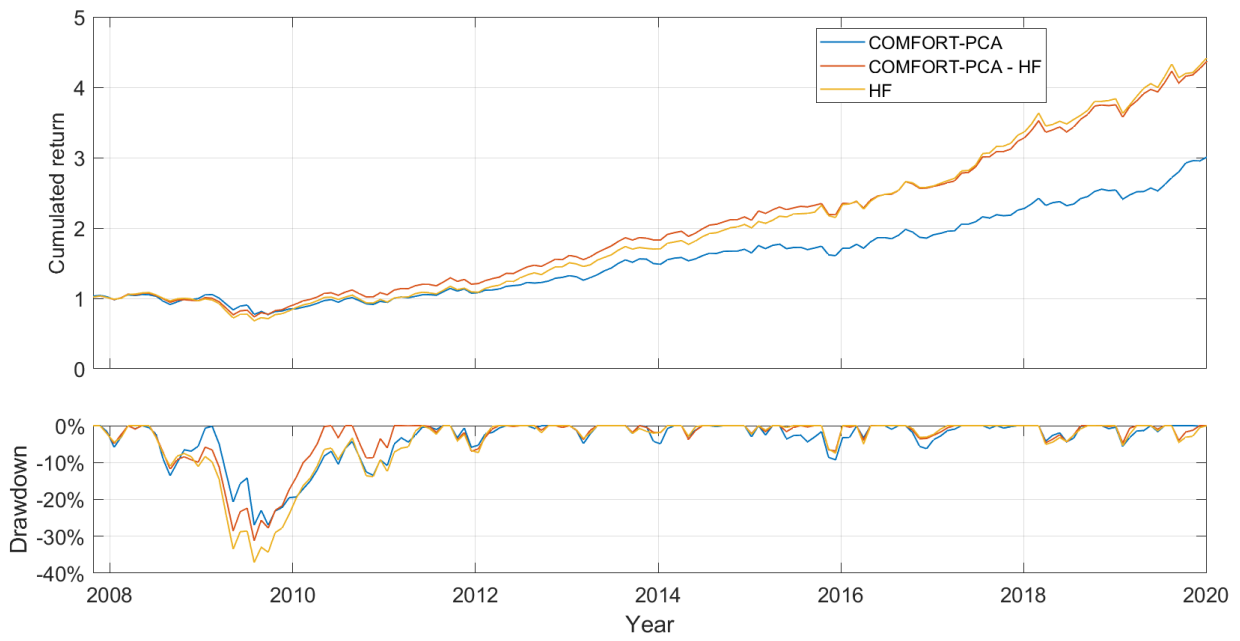


FIGURE 2.7: GMV portfolio performance and drawdown (0bp).

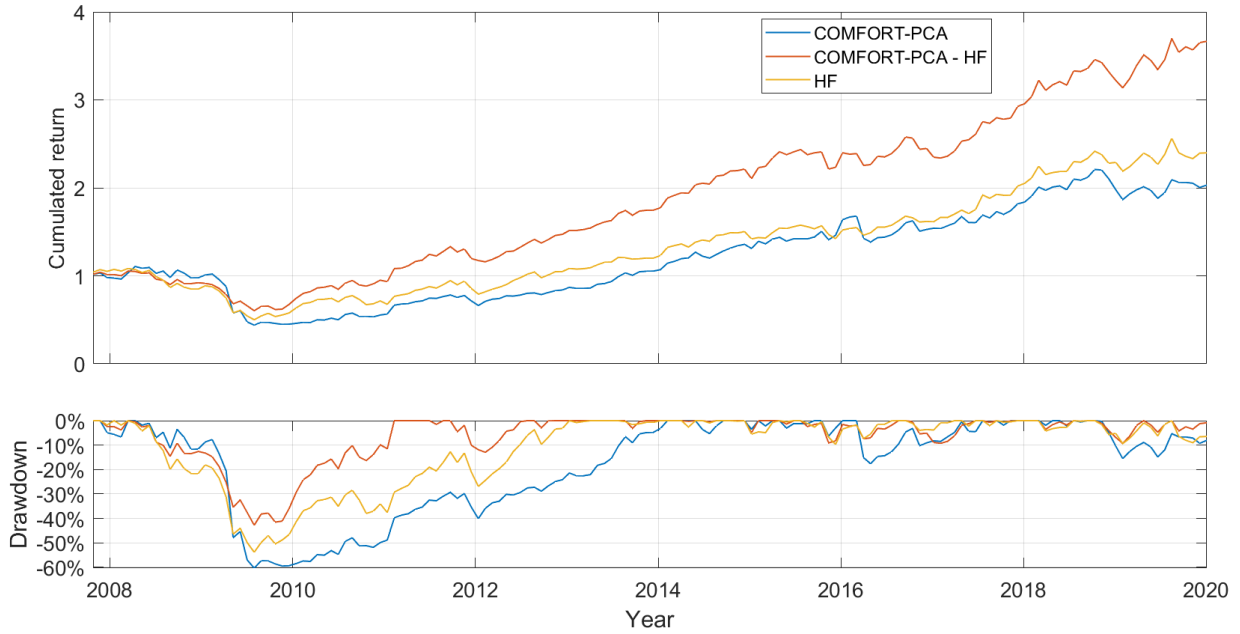


FIGURE 2.8: MS portfolio performance and drawdown (0bp).

2.5 Conclusions

In this chapter we have presented a proficient technique to mix the information modeled at different frequencies while accounting for heavy-tailedness of returns. Compared to many procedures seen in the literature, our approach blends the information after the two typologies of data are modeled separately, obtaining a time efficient result. High-frequency data undoubtedly possess the advantage of supplying the finest representation of assets, however, in such a context it might be worth giving up accuracy partially for more stability. It is not new that Markowitz model is very sensitive to input variation, so minimal changes captured by high-frequency data result in unnecessary turnover and transaction costs that deplete any possible benefit. For sufficiently large levels of transaction costs the high-frequency based approach worsens excessively, due to the high turnover. On the upside, for lower levels the COMFORT-PCA model provides comparable results, despite being grounded in lower frequency data. The proposed blended model provides superior results compared to both alternatives, and its improvements grow along the efficient portfolio frontier. Overall, Sharpe and Sortino ratios remain stable along the efficient frontier, while the alternative models worsen as soon as variance increases. Similar results, although less marked, are obtained using the Gaussian version of the COMFORT-PCA model. The improvement concerns both the return and the risk levels: the former remains stable or slightly increases along the efficient frontier while the counterparties have lower annualized returns, and the latter increases moderately but less markedly compared to the benchmarks. Additionally, the drawdowns of the blended approach are significantly shorter and less pronounced. Thus, the blending process proposed in this chapter shows that mixing information at different frequencies may lead to a finer signal that allows to improve performances in this asset allocation environment.

The work done so far can obviously be deepened and improved. We highlight a possible limitation involving the empirical application of the blending methodology, that is the oversimplifying choice of setting the pick matrix equal to the identity matrix, meaning that each view about the corresponding asset has the same magnitude. This in turn implies that the high-frequency measures (both returns and realized covariances) have the same relevance for each asset, which could easily not be the case. Thus, trying to advance on this side of the application could lead to substantial improvements. Moreover, the high-frequency data are used without any processing, while it might be worth considering possible approaches to model the raw data before the empirical analysis. Finally, we recall that the results depend on the choice of the low-frequency model, that in our case has the characteristic of being extremely parsimonious in terms of

conditional covariance dynamics. This feature is particularly important and desirable as long as transaction costs are taken into account, but it could fail in environments where they are almost negligible and a more accurate covariance dynamics becomes optimal. This implies that other low-frequency models and, more generally, other environment settings could produce different outcomes.

Chapter 3

CAW model and risk spillover analysis

This chapter proposes a possible implementation of the CAW model for the analysis of risk spillover. Thus, the purpose of this chapter is not to provide conclusive results, but rather to outline a research idea that exploits what has been presented so far. Although risk spillover is a much debated topic in finance, most of the current literature is limited to the use of low-frequency data. The aim is therefore to take inspiration from the most recent strands of literature about this issue, and provide an enhanced specification obtained through the CAW model.

3.1 Risk spillover

The current chapter is devoted to the examination of the risk spillover phenomenon, with the aim of proposing possible improvements to the extant literature. The term refers to the proliferation of risk, triggered by one or more events, from one asset to another one, either within the same country or across countries. This network effect has its roots in the financial interconnections that have been persistently increasing among countries, markets, financial institutions and even asset classes, and its implications affect a wide range of financial areas, such as credit risk management, portfolio management, hedging strategies, and financial stability as a whole. Thus, much research has been conducted on this topic, also in order to facilitate the development of various regulatory requirements, such as capital requirements or capital controls.

Reinhart and Rogoff (2008) note that each financial crisis has its own characteristics but at the same time shares some striking similarities with the others, for instance with respect to debt accumulation, growth patterns, and current account deficits. Also,

during crises the financial market volatility generally increases sharply and spills over across markets. This interdependence between financial markets has also been analyzed by King *et al.* (1990) and Forbes and Rigobon (2002), both in terms of returns and return volatilities. Consequently, it is clearly essential to be able to measure and monitor such spillovers, in order to monitor the ongoing market health, detect the early signs of a potential crisis, and eventually track it effectively. To this end, Diebold and Yilmaz (2009) proposed a volatility spillover measure based on forecast error variance decompositions from vector autoregressions, which can be used to quantify spillovers in returns or volatilities across individual assets, portfolios and markets. They found that spillovers have relevant impact and exhibit time-varying intensity, with a strikingly divergent behaviour in the dynamics of return spillovers compared to volatility spillovers. The former display no bursts but just a mild increasing trend, presumably associated with the increasing financial market integration, conversely the latter show no trend but clear bursts associated with crisis events.

The main shortcomings of the approach developed by Diebold and Yilmaz (2009) concerns the variable ordering issue affecting variance decompositions and the evaluation of just total spillovers (from each market to all the other markets). Moreover, the application is limited to the measurement of spillovers across identical assets in different countries. These matters are treated in the paper by Diebold and Yilmaz (2012), where the authors use a generalized vector autoregressive framework in which forecast error variance decompositions are invariant to the variable ordering and include directional volatility spillovers. Additionally, they also evaluate individual-asset spillovers within countries and across asset classes.

Eder and Keiler (2013) quantify and model systemic risk within the financial system using a specific weighting scheme to measure the magnitude of the risk spillover effects. The approach adopted stems from spatial econometrics and allows to decompose the credit spread into a systemic, systematic and idiosyncratic risk premium. Other papers in the spatial econometric literature exploit spatial weights to analyze spillovers (see Billio *et al.* (2022), Blasques *et al.* (2016) and Tonzer (2015) for further details).

3.2 Reference literature and proposal

One of the latest approaches in the framework of risk spillover analysis is proposed by Caporin and Paruolo (2015) and subsequently extended by Billio *et al.* (2021). These works specifically address the curse of dimensionality problem affecting many multivariate volatility models as the cross-sectional dimension increases, through the specification

of a structured parametrization based on weight matrices. The weights are derived from the economic proximity of entities, quantified using economic factors.

The curse of dimensionality problem is often circumvented in literature using diagonal parameter matrices, thus not accounting for covariance spillover and feedback effects. Instead, structured specifications allow for these effects from neighbors, reflecting the factor structure associated. The functioning of structured specifications requires a precise structure derived from economic rationale to outline the association between factors, as opposed to factor volatility models, where factors are not identified. Therefore, structured specifications are easier to interpret.

3.2.1 Structured specification

The structured specifications used in literature adopt weight matrices defined using prior knowledge about the existent connections between the units, and about the way in which distances are converted into weights. The selection of appropriate weight matrices is crucial, and special attention should be given to the approach used to define the right ones. For instance, if the objective is to determine the proximity of some assets, then neighbors can be defined as those assets belonging to the same sector, although a generalization allowing for the presence of covariates can be obtained. Consider the weight matrix definition used in Caporin and Paruolo (2015), that is for a group of n assets and $h = \{0, 1, \dots\}$, let $\mathbf{W}^{(h)}$ be the h -th $n \times n$ weight matrix, whose generic entry $w_{ij}^{(h)}$ indicates the weight, expressed as a real number between 0 and 1, of variable j in the determination of variable i , such that $w_{ii}^{(h)} = 0$ for $i = 1, \dots, n$. From the definition of weight matrix we can define proximity matrices as any matrix of the form

$$\mathbf{\Pi} = \sum_{h=0}^k \text{diag}(\psi^{(h)}) \mathbf{W}^{(h)}, \quad (3.1)$$

where $\psi^{(h)}$ are $n \times 1$ vector of coefficients. The structured specification of a model allows to significantly reduce the number of parameters estimated, for instance consider the BEKK model by Engle and Kroner (1995) for a cross-section of n time series $y_t := (y_{1,t}, \dots, y_{n,t})'$, for $t = 1, \dots, T$. The usual representation employs the vector of demeaned returns $u_t := y_t - \bar{y}$ to predict $\Sigma_t = \mathbb{E}(u_t u_t' | \Phi_{t-1})$, where Φ_{t-1} is information set up to time $t - 1$ and \bar{y} is the sample mean, even though it can also be replaced by

a vector of conditional returns. The resulting structure is

$$\begin{aligned}\Sigma_t &= \mathbf{C} + \mathbf{A}u_{t-1}u'_{t-1}\mathbf{A}' + \mathbf{B}\Sigma_{t-1}\mathbf{B}' \\ u_t &= \Sigma_t^{1/2}\epsilon_t,\end{aligned}\tag{3.2}$$

where ϵ_t is i.i.d. with 0 mean and covariance matrix equal to the identity matrix, and \mathbf{A} , \mathbf{B} and \mathbf{C} are unrestricted $n \times n$ parameter matrices to be estimated, with \mathbf{C} being positive definite. Note that the total number of coefficients in equation 3.2 is $\frac{1}{2}n(n+1) + 2n^2 = O(n^2)$, where $O(n^2)$ is the order in terms of the cross-sectional dimension n . This means that the estimation complexity increases rapidly and becomes unfeasible for a relatively small number of assets (depending on the number of observations available).

Structured specifications allow to curb this problem through the use of proximity matrices, as defined in 3.1. Using a single weight matrix \mathbf{W} , the structured specification of the BEKK model in 3.2 is obtained by setting $\mathbf{C} = \mathbf{S}^{-1}\mathbf{V}\mathbf{S}^{-1}$ and assuming \mathbf{A} , \mathbf{B} , and \mathbf{S} to be:

$$\mathbf{A} = \mathbf{A}_0 + \mathbf{A}_1\mathbf{W}, \quad \mathbf{B} = \mathbf{B}_0 + \mathbf{B}_1\mathbf{W}, \quad \mathbf{S} = \mathbf{I} - \mathbf{S}_1\mathbf{W},\tag{3.3}$$

where \mathbf{I} is the identity matrix, \mathbf{A} , \mathbf{B} and \mathbf{S} are proximity matrices, and $\mathbf{A}_j := \text{diag}(\alpha^{(j)})$, $\mathbf{B}_j := \text{diag}(\beta^{(j)})$ for $j = 0, 1$, $\mathbf{S}_1 := \text{diag}(s^{(1)})$, $\mathbf{V} := \text{diag}(v)$ are $n \times n$ diagonal matrices whose diagonal elements are the corresponding $n \times 1$ parameter vectors $\alpha^{(j)}$, $\beta^{(j)}$, $s^{(1)}$ and v . Note that the specification in 3.3 implies that $\mathbf{W}^{(0)} = \mathbf{I}$. The structured specification $\mathbf{C} = \mathbf{S}^{-1}\mathbf{V}\mathbf{S}^{-1}$ ensures that \mathbf{C} is positive definite, provided the elements of v are positive and \mathbf{S} is invertible. The model presented in structured specification is now estimable for large cross-sections, as the number of parameters in 3.3 reduces to $6n = O(n)$. A refined version of structured specifications is developed in the paper by Billio *et al.* (2021), where a bilateral formulation of proximity matrices is used alongside time-varying weight matrices. Alternatively, one may consider using the more evolved specification presented in Billio *et al.* (2021) and mentioned above.

3.2.2 Risk spillover in structured specification

The structured BEKK specification solves the challenge of reducing the number of parameters in the original BEKK model while still accounting for covariance spillover, in contrast to the diagonal BEKK specification. The spillover effect can be observed by looking at the term $\mathbf{A}u_{t-1}u'_{t-1}\mathbf{A}'$ in 3.2, from which we get

$$\mathbf{A}u_{t-1} = \text{diag}(\alpha^{(0)})u_{t-1} + \text{diag}(\alpha^{(1)})\mathbf{W}u_{t-1}.\tag{3.4}$$

Consider now the two terms in 3.4 and focus on the i -th element of $\mathbf{A}u_{t-1}$, that is $\alpha_i^{(0)}u_{i,t-1}$ and $\alpha_i^{(1)}w'_i u_{t-1}$. The former only includes the corresponding lagged term u_{t-1} , while the latter delivers the spatial effect from neighbors, where w'_i is the i -th row of \mathbf{W} , and $w'_i u_{t-1}$ is proportional to the average of u_{t-1} for stocks of the same sector, hence representing the spillover effects from other stocks in the same sector of i . Overall, the term $\mathbf{A}u_{t-1}u'_{t-1}\mathbf{A}'$ contains both diagonal and spillover effects from the same sector. A similar reasoning applies to the $\mathbf{B}\Sigma_{t-1}\mathbf{B}'$ term in 3.2. Let e_i be the i -th column of \mathbf{I} , then $(\Sigma_t)_{ij}$ depends on

$$\begin{aligned} (\mathbf{B}\Sigma_{t-1}\mathbf{B}')_{ij} &= \beta_i^{(0)}\beta_j^{(0)}(\Sigma_{t-1})_{ij} + \beta_j^{(0)}\beta_i^{(1)}(w'_i\Sigma_{t-1}e_j) + \beta_i^{(0)}\beta_j^{(1)}(e'_i\Sigma_{t-1}w_j) \\ &\quad + \beta_i^{(1)}\beta_j^{(1)}(w'_i\Sigma_{t-1}w_j). \end{aligned} \quad (3.5)$$

The first term represents a diagonal effect, while the last three are feedback effects from the covariances of i and j with their neighbors. For a more detailed explanation of the terms, as well as of the properties and advantages of structured specifications, see Caporin and Paruolo (2015). Additionally, the authors extensively discuss the determination of weight matrices with one or more classification criteria, or even using covariates such as market value, book-value, earnings/price, dividend yield and others. However, the intent of this chapter is simply to outline an attractive research topic connected to what has already been done in the previous two chapters, and not to review in detail the relevant literature. Therefore, all these issues will not be deepened further.

3.2.3 Structured CAW specification

The proposal presented in this chapter is to use a structured CAW specification for the analysis of risk spillover. The main contribution is not in terms of improved specification, as the approach presented is reliant upon the existent settings regarding proximity matrices, but rather relates to the belief that more precise results can be obtained through the use of high-frequency data. Apart from being more accurate, high-frequency data provide an additional advantage to the CAW, which lies in the fact that it can model the covariance matrix using an estimate of it (i.e. the realized covariance matrix) rather than using a latent process as in the BEKK specification. Consider the usual formulation of the CAW model, as presented in Chapter 1, that is

$$C_t = \mathbf{C}\mathbf{C}' + \sum_{i=1}^p \mathbf{A}_i C_{t-i} \mathbf{A}'_i + \sum_{j=1}^q \mathbf{B}_j \mathbf{Z}_{t-j} \mathbf{B}'_j, \quad (3.6)$$

where C_t is the conditional mean of the Wishart distribution used by the model, \mathbf{Z}_{t-j} is the lagged realized covariance matrix, and \mathbf{C} , \mathbf{A}_i , \mathbf{B}_j are the parameter matrices, with \mathbf{C} being lower-triangular. As already mentioned, the number of parameters increases rapidly with the number of assets. In fact, for n assets and using the basic specification with $p = q = 1$, the model involves $\frac{1}{2}n(n+1) + 2n^2 = O(n^2)$ parameters, where $O(n^2)$ is the order in terms of the cross-sectional dimension. Adopting diagonal parameter matrices, as done in Chapter 1, would not allow for covariance spillover and feedback effects. The resort to a structured specification, instead, ensures an easier estimation while accounting for these effects. The necessary proximity matrices can be defined as done in the literature of reference, that is similarly to 3.3, and the same reasoning applies to the weight matrices, since they are defined based on classification criteria or using covariates not depending upon the frequency of sampling. A reasonable alternative approach that needs to be evaluated consists in the use of a bilateral specification for proximity matrices coupled with time-varying weight matrices, as proposed in Billio *et al.* (2021).

The objective is to exploit the informative content of high-frequency data in a framework that is usually confined to the use of daily observations, with the expectation of getting more accurate analysis of risk spillover. Thus, given the need for high-frequency data, the scope of application is constrained to analyses for which such data can be retrieved. Although still to be defined, it is reasonable to consider assessing risk spillover among a group of assets using different classification criteria, such as sector membership or through the use of covariates, in order to understand which are most appropriate in defining proximity matrices. Should computational advantages be observed from embedding realized covariances in the modeling, then the work could be further extended to other models such as realized DCC, WAR, or the conditional autoregressive \mathcal{G} presented in the first chapter.

Conclusions

Discussion

Financial high-frequency data have a key role in modern analysis because of their extraordinarily informative content, which can be exploited for multiple purposes. This thesis focused on the modeling of such data with special emphasis on the evaluation and management of risk. Moreover, the whole work is based on realized covariances, that is the high-frequency data are not treated as they are, but rather the data are handled to obtain daily measures of risk. With these considerations in mind, it is easier to debate the contributions delivered to the literature.

The first chapter primarily aimed to introduce a novel distribution in the framework of financial high-frequency data. The rationale is to enable the identification and separation of the two leading risk components, that is systematic risk and specific risk. Compared to standard approaches for modeling realized covariances, such as the CAW model, the employment of the \mathcal{G} distribution allows to separate a common factor embedded in realized covariances and a matrix component. The former captures the inherent risk of the market, while the latter represents the idiosyncratic share of risk. The model is tested in a restricted version to reduce the estimated parameters, but more sophisticated version are pretty straightforward to obtain. The common factor detected by the model is highly correlated with market variance and plays a crucial role in estimating and forecasting conditional covariance. At the same time, the complementary component of the forecasted conditional covariance matrix, representing the idiosyncratic risk, can be easily limited in empirical applications, depending on the objective of the analysis.

The second chapter deals with the trade-off problem between accuracy and stability in the asset allocation framework. When both low and high-frequency data are available, it is obvious that the primary and most direct choice is to rely on the data supplying the finest representation of assets. Nevertheless, in the context of asset allocation, where rebalances are infrequent and transaction costs deplete potential profits, it might be worth giving up accuracy partially for more stability. Minimal variations in the

inputs often lead to substantial turnover in the portfolio, so the extreme precision and promptness of high-frequency data can backfire under certain conditions. In fact, the empirical analysis in chapter 2 confirms that portfolios relying exclusively on high-frequency data lead to excessive portfolio turnover. Conversely, the analysis based on a low-frequency-based model accounts for the necessity of stability although obviously failing to capture the informative content of high-frequency data. The proposed blended model allows to find an appealing trade-off between the strategies, providing superior overall results and therefore highlighting that mixing information at different frequencies may lead to a finer signal for the asset allocation environment.

Future directions of research

There are several possible extensions and improvements toward which the work done so far can be directed. As for the first chapter, it would be interesting to employ and assess further specifications of the conditional autoregressive \mathcal{G} model as alternative approaches within the frameworks that currently rely on the CAW and WAR models. A possible variant that still addresses the curse of dimensionality problem might include a diagonal intercept in the model instead of using the covariance targeting method, as proposed by Shen *et al.* (2020). Also, it would be attractive to increase the flexibility of the conditional autoregressive \mathcal{G} model by taking a cue from the generalized CAW model, developed by Yu *et al.* (2017).

With regard to the work done in Chapter 2 we can underline a limitation involving the empirical application of the blending methodology, that is the oversimplifying choice of setting $\mathbf{P} = \mathbf{I}$, meaning that each view about the corresponding asset has the same magnitude. This in turn implies that the high-frequency measures (both returns and realized covariances) have the same relevance for each asset, which could easily not be the case. Thus, trying to advance on this side of the application could lead to substantial improvements. Moreover, the high-frequency data are used without any processing, while it might be worth considering possible approaches to model the raw data before the empirical analysis. Also, we recall that the results depend on the choice of the low-frequency model, that in our case has the characteristic of being extremely parsimonious in terms of conditional covariance dynamics. This feature is particularly important and desirable as long as transaction costs are taken into account, but it could fail in environments where they are almost negligible and a more accurate covariance dynamics becomes optimal. This implies that other low-frequency models and, more generally, other environment settings could produce different outcomes.

Another topic that may be worth evaluating as a possible area of research that goes to extend what has been done so far, is that covered in Chapter 3. As explained earlier, it involves using a structured CAW specification in the study of risk spillover. The objective is to exploit the informative content of high-frequency data in a framework that is usually confined to the use of daily observations, with the expectation of getting a more accurate analysis of risk spillover. In addition, unlike the BEKK and the other models relying on low-frequency data, the CAW has the advantage of using an observed estimator of covariance in the process.

Appendix A

Consider the pdf of

$$\boldsymbol{\mu} \sim \text{N}(\boldsymbol{\pi}, \tau\boldsymbol{\Sigma}), \quad (\text{A.1})$$

that is

$$f_{\boldsymbol{\mu}}(\boldsymbol{\mu}) \equiv \frac{|\tau\boldsymbol{\Sigma}|^{-\frac{1}{2}}}{(2\pi)^{\frac{N}{2}}} e^{-\frac{1}{2}(\boldsymbol{\mu}-\boldsymbol{\pi})'(\tau\boldsymbol{\Sigma})^{-1}(\boldsymbol{\mu}-\boldsymbol{\pi})}, \quad (\text{A.2})$$

and the normal model associating uncertainty to the K views

$$\mathbf{P}\boldsymbol{\mu} \sim \text{N}(\mathbf{v}, \boldsymbol{\Omega}), \quad (\text{A.3})$$

then we can write A.3 as

$$\mathbf{v} \stackrel{d}{=} \mathbf{P}\boldsymbol{\mu} + \mathbf{Z}, \quad (\text{A.4})$$

where $\mathbf{Z} \sim \text{N}(\mathbf{0}, \boldsymbol{\Omega})$. Thus, \mathbf{v} can be modeled as a random variable \mathbf{V} whose distribution, conditioned on the realization of $\boldsymbol{\mu}$ is:

$$\mathbf{V}|\boldsymbol{\mu} \sim \text{N}(\mathbf{P}\boldsymbol{\mu}, \boldsymbol{\Omega}). \quad (\text{A.5})$$

Therefore, the pdf is

$$f_{\mathbf{V}|\boldsymbol{\mu}}(\mathbf{v}) \equiv \frac{|\boldsymbol{\Omega}|^{-\frac{1}{2}}}{(2\pi)^{\frac{K}{2}}} e^{-\frac{1}{2}(\mathbf{v}-\mathbf{P}\boldsymbol{\mu})'\boldsymbol{\Omega}^{-1}(\mathbf{v}-\mathbf{P}\boldsymbol{\mu})}. \quad (\text{A.6})$$

The posterior of $\boldsymbol{\mu}$ given \mathbf{V} is derived using the Bayes rule

$$f_{\boldsymbol{\mu}|\mathbf{v}}(\boldsymbol{\mu}) = \frac{f_{\boldsymbol{\mu},\mathbf{V}}(\boldsymbol{\mu}, \mathbf{v})}{f_{\mathbf{V}}(\mathbf{v})} = \frac{f_{\mathbf{V}|\boldsymbol{\mu}}(\mathbf{v})f_{\boldsymbol{\mu}}(\boldsymbol{\mu})}{\int f_{\mathbf{V}|\boldsymbol{\mu}}(\mathbf{v})f_{\boldsymbol{\mu}}(\boldsymbol{\mu}) d\boldsymbol{\mu}} \quad (\text{A.7})$$

by proving that the numerator can be written as the required conditional pdf $f_{\mathbf{v}|\boldsymbol{\mu}}$ multiplied by a factor which is not of our interest.

$$\begin{aligned} f_{\boldsymbol{\mu}, \mathbf{v}}(\boldsymbol{\mu}, \mathbf{v}) &= f_{\mathbf{v}|\boldsymbol{\mu}}(\mathbf{v})f_{\boldsymbol{\mu}}(\boldsymbol{\mu}) \\ &\propto |\tau\boldsymbol{\Sigma}|^{-\frac{1}{2}}|\boldsymbol{\Omega}|^{-\frac{1}{2}}e^{-\frac{1}{2}[(\boldsymbol{\mu}-\boldsymbol{\pi})'(\tau\boldsymbol{\Sigma})^{-1}(\boldsymbol{\mu}-\boldsymbol{\pi})+(\mathbf{v}-\mathbf{P}\boldsymbol{\mu})'\boldsymbol{\Omega}^{-1}(\mathbf{v}-\mathbf{P}\boldsymbol{\mu})]} \end{aligned} \quad (\text{A.8})$$

This expression coincides with the formula B.6 used to derive the market version of the Black-Litterman model, after substituting $\boldsymbol{\mu}$ with \mathbf{X} , $\boldsymbol{\pi}$ with $\boldsymbol{\mu}$, and $\tau\boldsymbol{\Sigma}$ with $\boldsymbol{\Sigma}$. Therefore, by using the same approach, we obtain

$$\boldsymbol{\mu}|\mathbf{v}; \boldsymbol{\Omega} \sim N(\boldsymbol{\mu}_{\text{BL}}, \boldsymbol{\Sigma}_{\text{BL}}^{\boldsymbol{\mu}}), \quad (\text{A.9})$$

where

$$\boldsymbol{\mu}_{\text{BL}} \equiv ((\tau\boldsymbol{\Sigma})^{-1} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P})^{-1} ((\tau\boldsymbol{\Sigma})^{-1}\boldsymbol{\pi} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{v}), \quad (\text{A.10})$$

$$\boldsymbol{\Sigma}_{\text{BL}}^{\boldsymbol{\mu}} \equiv ((\tau\boldsymbol{\Sigma})^{-1} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P})^{-1}. \quad (\text{A.11})$$

Using the identity

$$(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{C}\mathbf{A}^{-1}\mathbf{B} - \mathbf{D})^{-1}\mathbf{C}\mathbf{A}^{-1}, \quad (\text{A.12})$$

where \mathbf{A} and \mathbf{D} are invertible matrices, \mathbf{B} and \mathbf{C} are generic and conformable matrices, we can express A.10 as

$$\begin{aligned} \boldsymbol{\mu}_{\text{BL}} &\equiv ((\tau\boldsymbol{\Sigma})^{-1} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P})^{-1} ((\tau\boldsymbol{\Sigma})^{-1}\boldsymbol{\pi} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{v}) \\ &= ((\tau\boldsymbol{\Sigma}) - (\tau\boldsymbol{\Sigma})\mathbf{P}'(\mathbf{P}(\tau\boldsymbol{\Sigma})\mathbf{P}' + \boldsymbol{\Omega})^{-1}\mathbf{P}(\tau\boldsymbol{\Sigma})) ((\tau\boldsymbol{\Sigma})^{-1}\boldsymbol{\pi} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{v}) \\ &= \boldsymbol{\pi} + (\tau\boldsymbol{\Sigma})\mathbf{P}'(\boldsymbol{\Omega}^{-1} - (\mathbf{P}(\tau\boldsymbol{\Sigma})\mathbf{P}' + \boldsymbol{\Omega})^{-1}\mathbf{P}(\tau\boldsymbol{\Sigma})\mathbf{P}'\boldsymbol{\Omega}^{-1})\mathbf{v} \\ &\quad - (\tau\boldsymbol{\Sigma})\mathbf{P}'(\mathbf{P}(\tau\boldsymbol{\Sigma})\mathbf{P}' + \boldsymbol{\Omega})^{-1}\mathbf{P}\boldsymbol{\pi} \\ &= \boldsymbol{\pi} + (\tau\boldsymbol{\Sigma})\mathbf{P}'(\mathbf{P}(\tau\boldsymbol{\Sigma})\mathbf{P}' + \boldsymbol{\Omega})^{-1}(\mathbf{v} - \mathbf{P}\boldsymbol{\pi}), \end{aligned} \quad (\text{A.13})$$

where the last step follows from the observation that

$$\boldsymbol{\Omega}^{-1} - (\mathbf{P}(\tau\boldsymbol{\Sigma})\mathbf{P}' + \boldsymbol{\Omega})^{-1}\mathbf{P}(\tau\boldsymbol{\Sigma})\mathbf{P}'\boldsymbol{\Omega}^{-1} = (\mathbf{P}(\tau\boldsymbol{\Sigma})\mathbf{P}' + \boldsymbol{\Omega})^{-1}, \quad (\text{A.14})$$

obtained by left-multiplying both sides by $(\mathbf{P}(\tau\boldsymbol{\Sigma})\mathbf{P}' + \boldsymbol{\Omega})$.

Similarly, using A.12 we can write 2.13 as

$$\begin{aligned}\boldsymbol{\Sigma}_{\text{BL}} &\equiv \boldsymbol{\Sigma} + ((\tau\boldsymbol{\Sigma})^{-1} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P})^{-1} \\ &= (1 + \tau)\boldsymbol{\Sigma} - (\tau\boldsymbol{\Sigma})\mathbf{P}'(\mathbf{P}(\tau\boldsymbol{\Sigma})\mathbf{P}' + \boldsymbol{\Omega})^{-1}\mathbf{P}(\tau\boldsymbol{\Sigma}).\end{aligned}\tag{A.15}$$

Appendix B

Derivation of the equations 2.25 and 2.26 for the Black and Litterman model with views directly applied on returns. Consider a market with N assets whose distribution is

$$\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \quad (\text{B.1})$$

and the K views

$$\mathbf{V}|\mathbf{x} \sim N(\mathbf{P}\mathbf{x}, \boldsymbol{\Omega}), \quad (\text{B.2})$$

then

$$f_{\mathbf{X}}(\mathbf{x}) \equiv \frac{|\boldsymbol{\Sigma}|^{-\frac{1}{2}}}{(2\pi)^{\frac{N}{2}}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}, \quad (\text{B.3})$$

$$f_{\mathbf{V}|\mathbf{x}}(\mathbf{v}) \equiv \frac{|\boldsymbol{\Omega}|^{-\frac{1}{2}}}{(2\pi)^{\frac{K}{2}}} e^{-\frac{1}{2}(\mathbf{v}-\mathbf{P}\mathbf{x})'\boldsymbol{\Omega}^{-1}(\mathbf{v}-\mathbf{P}\mathbf{x})}. \quad (\text{B.4})$$

The posterior of the market given the views is derived using the Bayes' rule

$$f_{\mathbf{X}|\mathbf{v};\boldsymbol{\Omega}}(\mathbf{x}) = \frac{f_{\mathbf{X},\mathbf{v}}(\mathbf{x}, \mathbf{v})}{f_{\mathbf{V}}(\mathbf{v})} = \frac{f_{\mathbf{V}|\mathbf{x}}(\mathbf{v})f_{\mathbf{X}}(\mathbf{x})}{\int f_{\mathbf{V}|\mathbf{x}}(\mathbf{v})f_{\mathbf{X}}(\mathbf{x}) dx}, \quad (\text{B.5})$$

from which

$$f_{\mathbf{V}|\mathbf{x}}(\mathbf{v})f_{\mathbf{X}}(\mathbf{x}) \propto |\boldsymbol{\Sigma}|^{-\frac{1}{2}}|\boldsymbol{\Omega}|^{-\frac{1}{2}}e^{-\frac{1}{2}[(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})+(\mathbf{v}-\mathbf{P}\mathbf{x})'\boldsymbol{\Omega}^{-1}(\mathbf{v}-\mathbf{P}\mathbf{x})]}. \quad (\text{B.6})$$

The expression in square brackets in B.6 can be written as

$$\begin{aligned}
[\dots] &= (\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) + (\mathbf{v} - \mathbf{P}\mathbf{x})' \boldsymbol{\Omega}^{-1} (\mathbf{v} - \mathbf{P}\mathbf{x}) & (B.7) \\
&= \mathbf{x}' \boldsymbol{\Sigma}^{-1} \mathbf{x} - 2\mathbf{x}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} + \boldsymbol{\mu}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} + \mathbf{v}' \boldsymbol{\Omega}^{-1} \mathbf{v} - 2\mathbf{x}' \mathbf{P}' \boldsymbol{\Omega}^{-1} \mathbf{v} + \mathbf{x}' \mathbf{P}' \boldsymbol{\Omega}^{-1} \mathbf{P} \mathbf{x} \\
&= \mathbf{x}' (\boldsymbol{\Sigma}^{-1} + \mathbf{P}' \boldsymbol{\Omega}^{-1} \mathbf{P}) \mathbf{x} - 2\mathbf{x}' (\boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} + \mathbf{P}' \boldsymbol{\Omega}^{-1} \mathbf{v}) + \boldsymbol{\mu}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} + \mathbf{v}' \boldsymbol{\Omega}^{-1} \mathbf{v}.
\end{aligned}$$

Let's define

$$\boldsymbol{\mu}_{\text{BL}} \equiv (\boldsymbol{\Sigma}^{-1} + \mathbf{P}' \boldsymbol{\Omega}^{-1} \mathbf{P})^{-1} (\boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} + \mathbf{P}' \boldsymbol{\Omega}^{-1} \mathbf{v}), \quad (B.8)$$

which is equivalent to

$$(\boldsymbol{\Sigma}^{-1} + \mathbf{P}' \boldsymbol{\Omega}^{-1} \mathbf{P}) \boldsymbol{\mu}_{\text{BL}}(\mathbf{v}) \equiv (\boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} + \mathbf{P}' \boldsymbol{\Omega}^{-1} \mathbf{v}), \quad (B.9)$$

Using B.9, equation B.7 can be expressed as

$$\begin{aligned}
[\dots] &= \mathbf{x}' (\boldsymbol{\Sigma}^{-1} + \mathbf{P}' \boldsymbol{\Omega}^{-1} \mathbf{P}) \mathbf{x} - 2\mathbf{x}' (\boldsymbol{\Sigma}^{-1} + \mathbf{P}' \boldsymbol{\Omega}^{-1} \mathbf{P}) \boldsymbol{\mu}_{\text{BL}} & (B.10) \\
&\quad + \boldsymbol{\mu}_{\text{BL}}' (\boldsymbol{\Sigma}^{-1} + \mathbf{P}' \boldsymbol{\Omega}^{-1} \mathbf{P}) \boldsymbol{\mu}_{\text{BL}} - \boldsymbol{\mu}_{\text{BL}}' (\boldsymbol{\Sigma}^{-1} + \mathbf{P}' \boldsymbol{\Omega}^{-1} \mathbf{P}) \boldsymbol{\mu}_{\text{BL}} \\
&\quad + \boldsymbol{\mu}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} + \mathbf{v}' \boldsymbol{\Omega}^{-1} \mathbf{v} \\
&= (\mathbf{x} - \boldsymbol{\mu}_{\text{BL}})' (\boldsymbol{\Sigma}^{-1} + \mathbf{P}' \boldsymbol{\Omega}^{-1} \mathbf{P}) (\mathbf{x} - \boldsymbol{\mu}_{\text{BL}}) + \alpha,
\end{aligned}$$

where

$$\alpha \equiv \boldsymbol{\mu}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} + \mathbf{v}' \boldsymbol{\Omega}^{-1} \mathbf{v} - \boldsymbol{\mu}_{\text{BL}}' (\boldsymbol{\Sigma}^{-1} + \mathbf{P}' \boldsymbol{\Omega}^{-1} \mathbf{P}) \boldsymbol{\mu}_{\text{BL}}. \quad (B.11)$$

Using definition B.8, α can be expressed as

$$\begin{aligned}
\alpha &= \boldsymbol{\mu}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} + \mathbf{v}' \boldsymbol{\Omega}^{-1} \mathbf{v} & (B.12) \\
&\quad - (\boldsymbol{\mu}' \boldsymbol{\Sigma}^{-1} + \mathbf{v}' \boldsymbol{\Omega}^{-1} \mathbf{P}) (\boldsymbol{\Sigma}^{-1} + \mathbf{P}' \boldsymbol{\Omega}^{-1} \mathbf{P})^{-1} (\boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} + \mathbf{P}' \boldsymbol{\Omega}^{-1} \mathbf{v}) \\
&= \boldsymbol{\mu}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} + \mathbf{v}' \boldsymbol{\Omega}^{-1} \mathbf{v} \\
&\quad - \boldsymbol{\mu}' \boldsymbol{\Sigma}^{-1} (\boldsymbol{\Sigma}^{-1} + \mathbf{P}' \boldsymbol{\Omega}^{-1} \mathbf{P})^{-1} \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} \\
&\quad - \mathbf{v}' \boldsymbol{\Omega}^{-1} \mathbf{P} (\boldsymbol{\Sigma}^{-1} + \mathbf{P}' \boldsymbol{\Omega}^{-1} \mathbf{P})^{-1} \mathbf{P}' \boldsymbol{\Omega}^{-1} \mathbf{v} \\
&\quad + 2\boldsymbol{\mu}' \boldsymbol{\Sigma}^{-1} (\boldsymbol{\Sigma}^{-1} + \mathbf{P}' \boldsymbol{\Omega}^{-1} \mathbf{P})^{-1} \mathbf{P}' \boldsymbol{\Omega}^{-1} \mathbf{v} \\
&= \mathbf{v}' [\boldsymbol{\Omega}^{-1} - \boldsymbol{\Omega}^{-1} \mathbf{P} (\boldsymbol{\Sigma}^{-1} + \mathbf{P}' \boldsymbol{\Omega}^{-1} \mathbf{P})^{-1} \mathbf{P}' \boldsymbol{\Omega}^{-1}] \mathbf{v} \\
&\quad + 2\mathbf{v}' \boldsymbol{\Omega}^{-1} \mathbf{P} (\boldsymbol{\Sigma}^{-1} + \mathbf{P}' \boldsymbol{\Omega}^{-1} \mathbf{P})^{-1} \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} \\
&\quad + \boldsymbol{\mu}' [\boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1} (\boldsymbol{\Sigma}^{-1} + \mathbf{P}' \boldsymbol{\Omega}^{-1} \mathbf{P})^{-1} \boldsymbol{\Sigma}^{-1}] \boldsymbol{\mu}.
\end{aligned}$$

Using the identity

$$(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{C}\mathbf{A}^{-1}\mathbf{B} - \mathbf{D})^{-1}\mathbf{C}\mathbf{A}^{-1}, \quad (\text{B.13})$$

where \mathbf{A} and \mathbf{D} are invertible matrices, \mathbf{B} and \mathbf{C} are generic and conformable matrices, we can express

$$\mathbf{\Omega}^{-1} - \mathbf{\Omega}^{-1}\mathbf{P}(\mathbf{\Sigma}^{-1} + \mathbf{P}'\mathbf{\Omega}^{-1}\mathbf{P})^{-1}\mathbf{P}'\mathbf{\Omega}^{-1} = (\mathbf{\Omega} + \mathbf{P}\mathbf{\Sigma}\mathbf{P}')^{-1}. \quad (\text{B.14})$$

Additionally, let's define

$$\tilde{\mathbf{v}} \equiv -(\mathbf{\Omega} + \mathbf{P}\mathbf{\Sigma}\mathbf{P}')\mathbf{\Omega}^{-1}\mathbf{P}(\mathbf{\Sigma}^{-1} + \mathbf{P}'\mathbf{\Omega}^{-1}\mathbf{P})^{-1}\mathbf{\Sigma}^{-1}\boldsymbol{\mu} \quad (\text{B.15})$$

which can be written as

$$(\mathbf{\Omega} + \mathbf{P}\mathbf{\Sigma}\mathbf{P}')^{-1}\tilde{\mathbf{v}} \equiv -\mathbf{\Omega}^{-1}\mathbf{P}(\mathbf{\Sigma}^{-1} + \mathbf{P}'\mathbf{\Omega}^{-1}\mathbf{P})^{-1}\mathbf{\Sigma}^{-1}\boldsymbol{\mu}. \quad (\text{B.16})$$

Using equations B.14 and B.16, α can be expressed as

$$\begin{aligned} \alpha &= \mathbf{v}'(\mathbf{\Omega} + \mathbf{P}\mathbf{\Sigma}\mathbf{P}')^{-1}\mathbf{v} - 2\mathbf{v}'(\mathbf{\Omega} + \mathbf{P}\mathbf{\Sigma}\mathbf{P}')^{-1}\tilde{\mathbf{v}} \\ &\quad + \tilde{\mathbf{v}}'(\mathbf{\Omega} + \mathbf{P}\mathbf{\Sigma}\mathbf{P}')^{-1}\tilde{\mathbf{v}} - \tilde{\mathbf{v}}'(\mathbf{\Omega} + \mathbf{P}\mathbf{\Sigma}\mathbf{P}')^{-1}\tilde{\mathbf{v}} \\ &\quad + \boldsymbol{\mu}'(\mathbf{\Sigma}^{-1} - \mathbf{\Sigma}^{-1}(\mathbf{\Sigma}^{-1} + \mathbf{P}'\mathbf{\Omega}^{-1}\mathbf{P})^{-1}\mathbf{\Sigma}^{-1})\boldsymbol{\mu} \\ &= (\mathbf{v} - \tilde{\mathbf{v}})'(\mathbf{\Omega} + \mathbf{P}\mathbf{\Sigma}\mathbf{P}')^{-1}(\mathbf{v} - \tilde{\mathbf{v}}) + \phi, \end{aligned} \quad (\text{B.17})$$

where

$$\begin{aligned} \phi &\equiv \boldsymbol{\mu}'(\mathbf{\Sigma}^{-1} - \mathbf{\Sigma}^{-1}(\mathbf{\Sigma}^{-1} + \mathbf{P}'\mathbf{\Omega}^{-1}\mathbf{P})^{-1}\mathbf{\Sigma}^{-1})\boldsymbol{\mu} \\ &\quad - \tilde{\mathbf{v}}'(\mathbf{\Omega} + \mathbf{P}\mathbf{\Sigma}\mathbf{P}')^{-1}\tilde{\mathbf{v}} \end{aligned} \quad (\text{B.18})$$

does not depend on \mathbf{x} or \mathbf{v} . Substituting equation B.17 in B.10 we get

$$\begin{aligned} [\dots] &= (\mathbf{x} - \boldsymbol{\mu}_{\text{BL}})'(\mathbf{\Sigma}^{-1} + \mathbf{P}'\mathbf{\Omega}^{-1}\mathbf{P})(\mathbf{x} - \boldsymbol{\mu}_{\text{BL}}) \\ &\quad + (\mathbf{v} - \tilde{\mathbf{v}})'(\mathbf{\Omega} + \mathbf{P}\mathbf{\Sigma}\mathbf{P}')^{-1}(\mathbf{v} - \tilde{\mathbf{v}}) + \phi, \end{aligned} \quad (\text{B.19})$$

therefore the joint probability B.6 becomes

$$\begin{aligned}
f_{\mathbf{V}|\mathbf{X}}(\mathbf{v})f_{\mathbf{X}}(\mathbf{x}) &\propto |\boldsymbol{\Sigma}|^{-\frac{1}{2}}|\boldsymbol{\Omega}|^{-\frac{1}{2}}e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_{\mathbf{BL}})'(\boldsymbol{\Sigma}^{-1}+\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P})(\mathbf{x}-\boldsymbol{\mu}_{\mathbf{BL}})} \\
&\quad e^{-\frac{1}{2}(\mathbf{v}-\tilde{\mathbf{v}})'(\boldsymbol{\Omega}+\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}')^{-1}(\mathbf{v}-\tilde{\mathbf{v}})} \\
&= |\boldsymbol{\Sigma}^{-1} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}|^{\frac{1}{2}}e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_{\mathbf{BL}})'(\boldsymbol{\Sigma}^{-1}+\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P})(\mathbf{x}-\boldsymbol{\mu}_{\mathbf{BL}})} \\
&\quad |\boldsymbol{\Omega} + \mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'|^{-\frac{1}{2}}e^{-\frac{1}{2}(\mathbf{v}-\tilde{\mathbf{v}})'(\boldsymbol{\Omega}+\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}')^{-1}(\mathbf{v}-\tilde{\mathbf{v}})},
\end{aligned} \tag{B.20}$$

where the last equality follows from

$$\begin{aligned}
|\boldsymbol{\Sigma}||\boldsymbol{\Omega}||\boldsymbol{\Sigma}^{-1} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}| &= |\boldsymbol{\Sigma}(\boldsymbol{\Sigma}^{-1} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P})||\boldsymbol{\Omega}| \\
&= |\mathbf{I} + \boldsymbol{\Sigma}\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}||\boldsymbol{\Omega}| \\
&= |\mathbf{I} + \boldsymbol{\Omega}^{-1}\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'||\boldsymbol{\Omega}| \\
&= |\boldsymbol{\Omega}(\mathbf{I} + \boldsymbol{\Omega}^{-1}\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}')| \\
&= |\boldsymbol{\Omega} + \mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'|,
\end{aligned} \tag{B.21}$$

which in turn follows from the identity

$$|\mathbf{I}_m + \mathbf{B}\mathbf{C}| \equiv |\mathbf{I}_n + \mathbf{C}\mathbf{B}|, \tag{B.22}$$

with \mathbf{B} and \mathbf{C} being an $m \times n$ and $n \times m$ matrices, respectively. From the joint probability B.20, and using B.8 and B.15, we get

$$f_{\mathbf{V}|\mathbf{X}}(\mathbf{v})f_{\mathbf{X}}(\mathbf{x}) \propto f_{\mathbf{X}|\mathbf{V}}(\mathbf{x}|\mathbf{v})g(\mathbf{v}), \tag{B.23}$$

where

$$f_{\mathbf{X}|\mathbf{V}}(\mathbf{x}|\mathbf{v}) \propto |\boldsymbol{\Sigma}^{-1} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}|^{\frac{1}{2}}e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_{\mathbf{BL}})'(\boldsymbol{\Sigma}^{-1}+\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P})(\mathbf{x}-\boldsymbol{\mu}_{\mathbf{BL}})}. \tag{B.24}$$

Thus,

$$\mathbf{X}|\mathbf{V} \sim \mathcal{N}(\boldsymbol{\mu}_{\mathbf{BL}}, \boldsymbol{\Sigma}_{\mathbf{BL}}), \tag{B.25}$$

with

$$\boldsymbol{\Sigma}_{\mathbf{BL}} \equiv (\boldsymbol{\Sigma}^{-1} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P})^{-1}. \tag{B.26}$$

The last step is to reshuffle the expressions for $\boldsymbol{\mu}_{\mathbf{BL}}$ and $\boldsymbol{\Sigma}_{\mathbf{BL}}$, in B.8 and B.26, to get equations 2.25 and 2.26, respectively. Consider the identity B.13, and notice that B.8

can be written as

$$\begin{aligned}
\boldsymbol{\mu}_{\text{BL}} &\equiv (\boldsymbol{\Sigma}^{-1} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P})^{-1}(\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{v}) & (\text{B.27}) \\
&= (\boldsymbol{\Sigma} - \boldsymbol{\Sigma}\mathbf{P}'(\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}' + \boldsymbol{\Omega})^{-1}\mathbf{P}\boldsymbol{\Sigma}) (\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{v}) \\
&= \boldsymbol{\mu} + \boldsymbol{\Sigma}\mathbf{P}' (\boldsymbol{\Omega}^{-1} - (\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}' + \boldsymbol{\Omega})^{-1}\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'\boldsymbol{\Omega}^{-1}) \mathbf{v} \\
&\quad - \boldsymbol{\Sigma}\mathbf{P}'(\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}' + \boldsymbol{\Omega})^{-1}\mathbf{P}\boldsymbol{\mu}
\end{aligned}$$

Finally, notice that

$$\boldsymbol{\Omega}^{-1} - (\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}' + \boldsymbol{\Omega})^{-1}\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'\boldsymbol{\Omega}^{-1} = (\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}' + \boldsymbol{\Omega})^{-1} \quad (\text{B.28})$$

by left multiplying both sides by $(\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}' + \boldsymbol{\Omega})$, so we can write B.27 as

$$\boldsymbol{\mu}_{\text{BL}} = \boldsymbol{\mu} + \boldsymbol{\Sigma}\mathbf{P}'(\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}' + \boldsymbol{\Omega})^{-1}(\mathbf{v} - \mathbf{P}\boldsymbol{\mu}). \quad (\text{B.29})$$

Similarly, we can express B.26 as

$$\begin{aligned}
\boldsymbol{\Sigma}_{\text{BL}} &\equiv (\boldsymbol{\Sigma}^{-1} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P})^{-1} & (\text{B.30}) \\
&= \boldsymbol{\Sigma} - \boldsymbol{\Sigma}\mathbf{P}'(\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}' + \boldsymbol{\Omega})^{-1}\mathbf{P}\boldsymbol{\Sigma}.
\end{aligned}$$

Bibliography

- Andersen, T. G., Bollerslev, T., Diebold, F. X. and Labys, P. (2001) The distribution of realized exchange rate volatility. *Journal of the American statistical association* **96**(453), 42–55.
- Andersen, T. G., Bollerslev, T., Diebold, F. X. and Labys, P. (2003) Modeling and forecasting realized volatility. *Econometrica* **71**(2), 579–625.
- Barndorff-Nielsen, O. E. and Shephard, N. (2004) Econometric analysis of realized co-variation: High frequency based covariance, regression, and correlation in financial economics. *Econometrica* **72**(3), 885–925.
- Barone-Adesi, G., Giannopoulos, K. and Vosper, L. (1999) Var without correlations for portfolios of derivative securities. *Journal of Futures Markets* **19**(5), 583–602.
- Bevan, A. and Winkelmann, K. (1998) Using the black-litterman global asset allocation model: three years of practical experience. *Fixed Income Research* pp. 1–19.
- Billio, M., Caporin, M., Frattarolo, L. and Pelizzon, L. (2021) Networks in risk spillovers: A multivariate garch perspective. *Econometrics and Statistics* .
- Billio, M., Caporin, M., Panzica, R. and Pelizzon, L. (2022) The impact of network connectivity on factor exposures, asset pricing, and portfolio diversification. *International Review of Economics & Finance* .
- Black, F. and Litterman, R. (1990) Asset allocation: combining investor views with market equilibrium. *Goldman Sachs Fixed Income Research* **115**.
- Black, F. and Litterman, R. (1992) Global portfolio optimization. *Financial analysts journal* **48**(5), 28–43.
- Blasques, F., Koopman, S. J., Lucas, A. and Schaumburg, J. (2016) Spillover dynamics for systemic risk measurement using spatial financial time series models. *Journal of Econometrics* **195**(2), 211–223.

- Bollerslev, T. (1990) Modelling the coherence in short-run nominal exchange rates: a multivariate generalized arch model. *The review of economics and statistics* pp. 498–505.
- Bollerslev, T., Engle, R. F. and Wooldridge, J. M. (1988) A capital asset pricing model with time-varying covariances. *Journal of political Economy* **96**(1), 116–131.
- Bonato, M., Caporin, M. and Rinaldo, A. (2013) Risk spillovers in international equity portfolios. *Journal of Empirical Finance* **24**, 121–137.
- Brownlees, C. T., Cipollini, F. and Gallo, G. M. (2011) Multiplicative error models. *Available at SSRN 1852285* .
- Caporin, M. and Paruolo, P. (2015) Proximity-structured multivariate volatility models. *Econometric Reviews* **34**(5), 559–593.
- 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.
- Diebold, F. X. and Yilmaz, K. (2009) Measuring financial asset return and volatility spillovers, with application to global equity markets. *The Economic Journal* **119**(534), 158–171.
- Diebold, F. X. and Yilmaz, K. (2012) Better to give than to receive: Predictive directional measurement of volatility spillovers. *International Journal of forecasting* **28**(1), 57–66.
- Eder, A. and Keiler, S. (2013) Cds spreads and systemic risk-a spatial econometric approach. *Bundesbank Discussion Paper* **20**.
- Engle, R. (2002a) Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics* **20**(3), 339–350.
- Engle, R. (2002b) New frontiers for arch models. *Journal of Applied Econometrics* **17**(5), 425–446.
- Engle, R. F. and Kroner, K. F. (1995) Multivariate simultaneous generalized arch. *Econometric theory* **11**(1), 122–150.
- Forbes, K. J. and Rigobon, R. (2002) No contagion, only interdependence: measuring stock market comovements. *The journal of Finance* **57**(5), 2223–2261.

- Freitas, C. C., Frery, A. C. and Correia, A. H. (2005) The polarimetric \mathcal{G} distribution for sar data analysis. *Environmetrics: The official journal of the International Environmetrics Society* **16**(1), 13–31.
- Ghysels, E., Santa-Clara, P. and Valkanov, R. (2004) The midas touch: Mixed data sampling regression models .
- Giacometti, R., Bertocchi, M., Rachev, S. T. and Fabozzi, F. J. (2007) Stable distributions in the black–litterman approach to asset allocation. *Quantitative Finance* **7**(4), 423–433.
- Golosnoy, V., Gribisch, B. and Liesenfeld, R. (2012) The conditional autoregressive wishart model for multivariate stock market volatility. *Journal of Econometrics* **167**(1), 211–223.
- Gouriéroux, C., Jasiak, J. and Sufana, R. (2009) The wishart autoregressive process of multivariate stochastic volatility. *Journal of Econometrics* **150**(2), 167–181.
- Gradshteyn, I. S. and Ryzhik, I. M. (2007) *Table of integrals, series, and products*. Elsevier.
- Gribisch, B., Hartkopf, J. P. and Liesenfeld, R. (2020) Factor state–space models for high-dimensional realized covariance matrices of asset returns. *Journal of Empirical Finance* **55**, 1–20.
- Harvey, A., Ruiz, E. and Shephard, N. (1994) Multivariate stochastic variance models. *The Review of Economic Studies* **61**(2), 247–264.
- He, Z. L. (2007) Incorporating alpha uncertainty into portfolio decisions: A bayesian revisit of the treynor–black model. *Journal of Asset Management* **8**(3), 161–175.
- King, M. A., Sentana, E. and Wadhvani, S. (1990) Volatiltiy and links between national stock markets.
- Lee, S.-W. and Hansen, B. E. (1994) Asymptotic theory for the garch (1, 1) quasi-maximum likelihood estimator. *Econometric theory* **10**(1), 29–52.
- MacKinlay, A. C. and Pástor, L. (2000) Asset pricing models: Implications for expected returns and portfolio selection. *The Review of financial studies* **13**(4), 883–916.
- Mankert, C. (2010) *The Black-Litterman model: Towards its use in practice*. Ph.D. thesis, KTH.

- Markowitz, H. (1952) Portfolio selection, *journal of finance*, march .
- Meucci, A. (2005) *Risk and asset allocation*. Volume 1. Springer.
- Meucci, A. (2010) The black-litterman approach: Original model and extensions. *Shorter version in, THE ENCYCLOPEDIA OF QUANTITATIVE FINANCE, Wiley* .
- Newey, W. K. and McFadden, D. (1994) Large sample estimation and hypothesis testing. *Handbook of econometrics* **4**, 2111–2245.
- Noureldin, D., Shephard, N. and Sheppard, K. (2012) Multivariate high-frequency-based volatility (heavy) models. *Journal of Applied Econometrics* **27**(6), 907–933.
- Paolella, M. S. and Polak, P. (2015a) ALRIGHT: Asymmetric LaRge-Scale (I)GARCH with Hetero-Tails. *International Review of Economics and Finance* **40**, 282–297.
- Paolella, M. S. and Polak, P. (2015b) Comfort: A common market factor non-gaussian returns model. *Journal of Econometrics* **187**(2), 593–605.
- Paolella, M. S. and Polak, P. (2018) COBra: Copula-Based Portfolio Optimization. In *Studies in Computational Intelligence: Predictive Econometrics and Big Data*, ed. N. C. Vladik Kreinovich, Songsak Sriboonchitta. Springer.
- Paolella, M. S., Polak, P. and Walker, P. S. (2019) A non-elliptical orthogonal garch model for portfolio selection under transaction costs. *Swiss Finance Institute Research Paper* (19-51).
- Pástor, L. and Stambaugh, R. F. (2000) Comparing asset pricing models: an investment perspective. *Journal of Financial Economics* **56**(3), 335–381.
- Polak, P. and Ulrych, U. (2021) Dynamic currency hedging with ambiguity. *Swiss Finance Institute Research Paper* (21-60).
- Reinhart, C. M. and Rogoff, K. S. (2008) Is the 2007 us sub-prime financial crisis so different? an international historical comparison. *American Economic Review* **98**(2), 339–44.
- Shen, K., Yao, J. and Li, W. K. (2020) Forecasting high-dimensional realized volatility matrices using a factor model. *Quantitative Finance* **20**(11), 1879–1887.
- Shephard, N. and Sheppard, K. (2010) Realising the future: forecasting with high-frequency-based volatility (heavy) models. *Journal of Applied Econometrics* **25**(2), 197–231.

-
- Silvennoinen, A. and Teräsvirta, T. (2009) Multivariate garch models. In *Handbook of financial time series*, pp. 201–229. Springer.
- Tonzer, L. (2015) Cross-border interbank networks, banking risk and contagion. *Journal of Financial Stability* **18**, 19–32.
- Treynor, J. L. and Black, F. (1973) How to use security analysis to improve portfolio selection. *The journal of business* **46**(1), 66–86.
- Walters, C. *et al.* (2014) The black-litterman model in detail. *Available at SSRN 1314585*
.
- Yu, P. L., Li, W. K. and Ng, F. C. (2017) The generalized conditional autoregressive wishart model for multivariate realized volatility. *Journal of Business & Economic Statistics* **35**(4), 513–527.

Marco Girardi

CURRICULUM VITAE

Contact Information

University of Padova
Department of Statistical Sciences
via Cesare Battisti, 241-243
35121 Padova. Italy.

e-mail: marco.girardi.6@phd.unipd.it; marco.girardi94@gmail.com (personal)

Current Position

Since October 2019; (expected completion: January 2023)

PhD Student in Statistical Sciences, University of Padova.

Thesis title: Modeling of high-frequency financial data

Supervisor: Prof. Massimiliano Caporin

Research interests

- Financial times series analysis
- Covariance modeling
- Asset allocation

Education

October 2016 – December 2018

Master's degree (laurea magistrale) in Economics and Finance.

University of Padova, Department of Economics and Management

Supervisor: Prof. Massimiliano Caporin

Final mark: 110/110 with honors

September 2013 – July 2016

Bachelor's degree (laurea triennale) in International Economics (Economia Internazionale).

University of Padova, Department of Economics and Management

Supervisor: Prof. Gianfranco Tuset

Final mark: 110/110 with honors

Work experience

2019 – 2019

Prometeia.

Wealth and asset management.

2018 – 2018

Accenture.

Banking and financial services.

2016 – 2016

Fischer Italia.

Accounting and HR.

Awards and Scholarship

2018

CFASI Fund Management Challenge, first place. Portfolio management competition among Italian universities.

2017 – 2018

Mille e una lode. Merit scholarship, University of Padova.

Computer skills

- Matlab, proficient user
- R, intermediate user
- LaTeX, proficient user
- Refinitiv Eikon and Datastream, intermediate user
- Bloomberg, intermediate user
- IBM SPSS Statistics, intermediate user

Language skills

Italian: native; English: fluent (reading, listening, writing), good (speaking).

Publications

Working papers

Caporin, M., Girardi, M. (2022). Conditional autoregressive \mathcal{G} model for common factor detection in the stock market. *Currently under revision.*

Girardi, M., Polak, P. (2022). COMFORT-PCA with views from high-frequency data. *Currently under revision.*

Conference presentations

Caporin, M., Girardi, M. (2022). Conditional autoregressive \mathcal{G} model for common factor detection in the stock market. (invited) *Computational and Financial Econometrics (CFE)*, London, England, 17 – 19 December 2022.

Caporin, M., Girardi, M. (2022). The conditional autoregressive Wishart- \mathcal{G} model. *SIS 2022 - The 51st Scientific Meeting of the Italian Statistical Society*, Caserta, Italy, 22 – 24 June 2022.

Teaching experience

March 2022

Tirocinio formativo (LaTeX course)
Master's degree in statistical sciences
Teaching assistance, 2 hours
University of Padova
Instructor: Prof. Costantina Bruno

September 2021

Tirocinio formativo (LaTeX course)
Master's degree in statistical sciences
Teaching assistance, 2 hours
University of Padova
Instructor: Prof. Costantina Bruno