Bayesian semiparametric multivariate realized GARCH modeling

Accepted for publication in the Journal of Forecasting.

https://doi.org/10.1002/for.3285

Bayesian semiparametric multivariate realized GARCH modeling

Efthimios Nikolakopoulos*

Abstract

This paper introduces a novel Bayesian semiparametric multivariate GARCH framework for modeling returns and realized covariance, as well as approximating their joint unknown conditional density. We extend existing parametric multivariate realized GARCH models by incorporating a Dirichlet Process mixture of countably infinite normal distributions for returns and (inverse-)Wishart distributions for realized covariance. This approach captures time-varying dynamics in higher-order conditional moments of both returns and realized covariance. Our new class of models demonstrates superior out-of-sample forecasting performance, providing significantly improved multiperiod density forecasts for returns and realized covariance, and competitive covariance point forecasts.

Keywords: Multivariate GARCH, Realized covariance, Bayesian nonparametrics, Density forecasts.

*Essex Business School, University of Essex, Wivenhoe Park, Colchester, UK, CO4 3SQ, e.nikolakopoulos@essex.ac.uk

1 Introduction

In this paper, we introduce a novel Bayesian semiparametric multivariate GARCH framework that approximates the joint unknown conditional density of returns and realized covariance (RCov). We extend existing parametric multivariate realized GARCH models by incorporating a Dirichlet Process mixture of countably infinite normal distributions for returns and (inverse-)Wishart distributions for RCov. This enhanced framework captures time-varying dynamics in the higher-order conditional moments of both returns and RCov. Our results demonstrate significant improvements in multiperiod density forecasts for both returns and RCov over current parametric models. Notably, models with the Wishart assumption outperform in forecasting returns densities, while the inverse-Wishart specification provides the best RCov density and point forecasts.

Modeling the time-varying dynamics of asset returns covariance provides forecasting improvements in key financial management areas such as portfolio optimization, asset pricing and risk management. Typically, multivariate extensions of GARCH models (Engle, 1982; Bollerslev, 1986) are used to estimate the conditional covariance.[1] Comprehensive studies on the multivariate GARCH models are from Bauwens et al. (2006), Silvennoinen and Teräsvirta (2009), and Virbickaite et al. (2015). These models are based on low frequency (daily, weekly or monthly) squared returns to construct the conditional covariance.

Realized measures of covariance, based on high frequency returns,[2] have been used to assist the construction of conditional covariance matrices. These are parametric extensions in GARCH-type models by Hansen et al. (2012), Hansen et al. (2014), Archakov et al. (2019) and, Gorgi et al. (2019), in HEAVY-type models by Noureldin et al. (2012), Braione (2016), Opschoor et al. (2018), Sheppard and Xu (2019) and, Bauwens and Xu (2023), in multivariate stochastic volatility by Gouriéroux et al. (2009), Shirota et al. (2017) and, Yamauchi and Omori (2020), and in conditional autoregressive models by Golosnoy et al. (2012) and, Jin and Maheu (2013).

Semiparametric extensions of conditional covariance models, such as the ones by Jensen and Maheu (2013), Zaharieva et al. (2020) and, Maheu and Shamsi Zamenjani (2021) accommodate time-varying returns asymmetry and fat-tails. They also provide better returns density forecasts compared to parametric specifications. In the class of joint returns and RCov models a notable semiparametric approach is by Jin and Maheu (2016). To our knowledge, the multivariate realized GARCH-type (and the HEAVY-type) models have not yet been semiparametrically extended.

The purpose of this paper is to extend the multivariate realized GARCH (MRG) models with a Bayesian semiparametric approach and to examine its impact to forecasts. For computational efficiency, we use the two-parameter (scalar) multivariate GARCH specification of Ding and Engle (2001) in which we include the one period lag of RCov in a multivariate GARCH-X setting (Engle, 2002b) but, any other multivariate GARCH (or HEAVY) specification can be used as well. We approximate the joint returns and RCov underlying density by scaling the returns conditional covariance and the RCov distributional scale matrix with components of infinite support from a Dirichlet Process prior (Ferguson, 1973; Escobar and West, 1995). In our Dirichlet Process mixture (DPM) setting, returns follow a countably infinite mixture of normals, as in Jensen and Maheu (2013) and, Maheu and Shamsi Zamenjani (2021), to capture their empirically observed distributional asymmetry and fat-tails. For RCov we use a countably infinite mixture of Wishart and inverse-Wishart distributions.[3]

Our framework contributes to the literature in several ways. It extends the parametric multivariate realized GARCH models (Hansen et al., 2014; Archakov et al., 2019; Gorgi et al., 2019) to a semiparametric framework. It also extends the model of Jensen and Maheu (2013) to a returns and RCov joint modeling. Our approach is closer to Jin and Maheu (2016) but, with two differences. First, their mixture process is only governed by RCov information while we use both returns and RCov in determining the mixing clusters. Second, we include the returns mean vector in the mixture to capture distributional asymmetry.[4] We find both of these two features of our framework important in forecasting.

[1] Popular specifications are from Bollerslev et al. (1988), Baba et al. (1990), Engle and Kroner (1995) and Engle (2002a).

[2] For instance, daily realized covariance matrices are constructed from high-frequency (e.g., 5-minute) intraday returns.

[3] The inverse-Wishart accommodates fat-tails since its second moment exists only for large degrees of freedom values. See Jin and Maheu (2016).

[4] This feature is also missing from the tails-focused HEAVY model of Opschoor et al. (2018).

We develop several restricted model specifications to test the forecasting impact of our model features. We use the popular equity stocks dataset of Noureldin et al. (2012)[5] and we focus on out-of-sample multiperiod (daily, weekly and monthly) returns and RCov density forecasting. The semiparametric framework provides superior density forecasts compared to parametric specifications that have popular distributional assumptions. We find that using the Wishart assumption for RCov provides in general better returns density forecasts compared to the inverse-Wishart assumption. The latter one results in better RCov density and point forecasts. The new semiparametric models provide out-of-sample portfolio optimization improvements in a global minimum variance portfolio application. This is observed only for daily (one-period) forecasts and not for weekly or monthly ones.

The paper is organized as follows: Section 2 presents the specification of the proposed models and a brief discussion of the estimation steps, Section 3 presents the forecasting process, Section 4 presents the empirical application results and Section 5 concludes. An Appendix details the model estimation algorithm.

2 Multivariate realized GARCH

We start this section with a brief discussion on RCov. Realized measures of covariation are based on high-frequency (intraday) data to estimate the latent low-frequency (daily) covariance matrices. They are multivariate extensions of realized variance measures.[6] For a trading day t, we can have Q high frequency observations for the n-length vector of logarithmic asset prices $p_{t,i},\ i = 1, ..., Q$. Based on these, the high-frequency log returns at time i of day t are calculated as $r_{t,i} = p_{t,i} - p_{t,i-1}$, $r_{t,i} \in \mathbb{R}^n$. The basic nonparametric estimator of RCov, for day t, is calculated as $\text{RCov}_t = \sum_{i=1}^{Q} r_{t,i} r'_{t,i}$, with RCov_t being a $n \times n$ positive definite symmetric matrix where the diagonal elements consist the realized variance estimators and the off-diagonal elements are the realized covariances. Banrdorff-Nielsen and Shephard (2004) show that, if returns are synchronized and free of market microstructure noise, when $Q \to \infty$, RCov_t converges to the quadratic covariation, which is equal to the conditional returns covariance (Andersen et al., 2003).[7]

2.1 Parametric models

In this section we discuss several parametric MRG models with different distributional assumptions that serve as appropriate benchmarks to our proposed semiparametric framework presented in the following section. For the conditional covariance construction we employ the two-parameter (scalar) multivariate GARCH specification of Ding and Engle (2001) with covariance targeting. We extend that to a three-parameter multivariate GARCH-X framework (Engle, 2002b). This parsimonious specification keeps the number of parameters unaffected by the number of data series. Any other multivariate GARCH or HEAVY specification can be used as well.

In these models literature, the standard distributional choice for returns is the multivariate normal and for RCov matrices is the Wishart distribution. Empirically, financial returns have distributions with fat-tails (Richardson and Smith, 1993; Ding and Engle, 2001; Diamantopoulos and Vrontos, 2010; Opschoor et al., 2018, among many others). We employ a fully parametric framework which allows returns to follow the fat-tailed multivariate Student-t distribution and RCov matrices to follow the inverse-Wishart (IW) distribution as suggested by Jin and Maheu (2016). Given the information set $\mathcal{I}_{t-1}$ available at time $t-1$, the model referred to as MRG-t-IW is specified as

$$r_t|\mathcal{I}_{t-1} \sim \text{t}(0, H_t, \zeta), \quad \zeta > 2, \tag{1}$$

$$\text{RCov}_t|\mathcal{I}_{t-1} \sim \text{IW}\left(\nu, (\nu - n - 1)\frac{\zeta}{\zeta - 2} H_t^{1/2} V \left(H_t^{1/2}\right)'\right), \quad \nu > n + 1, \tag{2}$$

$$H_t = \Omega + a\, r_{t-1}(r_{t-1})' + b\, H_{t-1} + c\, \text{RCov}_{t-1}, \tag{3}$$

[5] It is also used by Shirota et al. (2017), Jin et al. (2019), Yamauchi and Omori (2020) among others.

[6] See Andersen and Bollerslev (1998), Barndorff-Nielsen and Shephard (2002), Andersen et al. (2003) and, Barndorff-Nielsen and Shephard (2004).

[7] See Zhang et al. (2005) for the use of *subsampling* for realized measures robust to noise and Barndorff-Nielsen et al. (2008, 2011) for the use of the multivariate realized kernel framework for non-synchronous and noisy returns.

where r_t is a n-length vector of log returns at time t (trading day) for n data series. $\mathrm{t}(0, H_t, \zeta)$ is the multivariate Student-t distribution with zero mean[8], scale the $n \times n$ matrix H_t, with $n(n+1)/2$ unique elements, and degrees of freedom ζ.

By definition the conditional returns covariance from (1) is $\mathrm{Cov}(r_t|\mathcal{I}_{t-1}) = \frac{\zeta}{\zeta-2}H_t$, and the scale matrix, H_t, is constructed from equation (3)[9] in which a, b and c are scalar parameters and Ω is a $n \times n$ positive definite parameter matrix.

The measurement equation (2) models the joint dependence between r_t and RCov_t. RCov_t follows an IW distribution, of dimension n, with ν degrees of freedom and scale matrix $(\nu - n - 1)\frac{\zeta}{\zeta-2}H_t^{1/2}V\left(H_t^{1/2}\right)'$. $H_t^{1/2}$ is the Cholesky decomposition of H_t. From (2), the conditional mean of RCov_t is linked with the returns conditional covariance through $\mathbb{E}(\mathrm{RCov}_t|\mathcal{I}_{t-1}) = \frac{\zeta}{\zeta-2}H_t^{1/2}V\left(H_t^{1/2}\right)'$, with V being an $n \times n$ positive definite parameter matrix which completes the model as it helps to capture deviations between RCov_t and conditional covariance. If there are no deviations, V would be approximately equal to an identity matrix and RCov_t is synonymous to the conditional returns covariance, $\mathbb{E}[\mathrm{RCov}_t|\mathcal{I}_{t-1}] = \frac{\zeta}{\zeta-2}H_t$. The scale matrix of IW distribution is positive definite as long as V and H_t are positive definite and $\nu > n+1$. These conditions are ensured during model estimation.

Compared to the existing parametric multivariate realized GARCH models, the proposed framework has several differences. Unlike models that link realized and conditional covariance elements through linear measurement equations (e.g., Hansen et al., 2012; Archakov et al., 2019), we link RCov_t in (2) to the full conditional covariance matrix. This parsimoniously allows full dependence between the RCov_t and H_t elements. Our approach on that is close to Gorgi et al. (2019). However, they assume that RCov_t follows a Wishart distribution. Jin and Maheu (2016) show that an inverse-Wishart assumption, as in (2), is a better fit for realized covariance matrices, compared to a Wishart one.

We empirically test the model distributional assumptions with alternative specifications. The model with the inverse-Wishart assumption replaced by a Wishart one is referred to as MRG-t-W. This model is specified with (1), (3) and the following RCov measurement equation

$$\mathrm{RCov}_t|\mathcal{I}_{t-1} \sim \mathrm{W}\left(\mathrm{v}, \mathrm{v}^{-1}\frac{\zeta}{\zeta-2}H_t^{1/2}V\left(H_t^{1/2}\right)'\right), \quad \mathrm{v} > n+1, \tag{4}$$

with v being the Wishart degrees of freedom and $\mathrm{v}^{-1}\frac{\zeta}{\zeta-2}H_t^{1/2}V\left(H_t^{1/2}\right)'$ the scale matrix. The conditional mean of RCov_t is linked with H_t in a similar way as in MRG-t-IW. We also test models with the t-distributed innovations substituted by Gaussian ones. They are referred to as MRG-N-IW and MRG-N-W. The latter one serves as an appropriate benchmark since it has the same distributional assumptions as in Gorgi et al. (2019) and, Bauwens and Xu (2023). The models we test are presented in Table 1.

In the following section the parametric models are extended to a semiparametric framework and Bayesian methods are used for their estimation. For paradigm consistency, the parametric models are estimated similarly. The data likelihood is factored as $p(r_t, \mathrm{RCov}_t|\mathcal{I}_{t-1}) = p(r_t|\mathcal{I}_{t-1})p(\mathrm{RCov}_t|\mathcal{I}_{t-1})$ and a set of Markov chain Monte Carlo (MCMC) steps is employed to sample the model parameters from conditional posterior distributions. See the 5ppendix for details.

2.2 Semiparametric models

Empirically returns distributions have fat-tails and asymmetry. Jensen and Maheu (2013) show that a uniformly fat-tails assumption across the dataset is restrictive. Financial time series present unique features which are better approximated by a mixture of distributions. They extend the parametric multivariate GARCH to a semiparametric setting in which returns distribution follows an infinite mixture of normals with a Dirichlet Process prior (Ferguson,

[8]Parsimoniously, the mean vector in (1) is restricted to zero which is common in literature. We have also developed a version with a parametric mean vector, and in our empirical application, the means were estimated to be approximately zero.

[9]It is common in the realized GARCH related literature to not include the ARCH parameter(s) in the construction (Engle, 2002b; Hansen et al., 2012, 2014; Gorgi et al., 2019; Archakov et al., 2019, among many others). We empirically tested this restricted version and we found that it provides inferior forecasts.

1973). Their model, has a two component returns covariance, a GARCH constructed one and a scale matrix from the DPM. They also include the mean vector in the mixture to capture asymmetric effects of the data distribution. They show that the semiparametric framework provides better returns density forecasts than the parametric one. Following we discuss our proposed framework that semiparametrically extends the MRG models.

2.2.1 Multivariate realized GARCH-DPM

We develop several different specifications of MRG models with a DPM mixture. In this section we discuss the general one that nests all the rest. This MRG-DPM model imposes an infinite mixture in: the returns mean vector (M), their conditional covariance component (Λ) and the conditional scale matrix of RCov (V), hence it is referred to as MRG-DPM-MΛV. Several nested versions of it are tested with one or two of the infinite components being parametrically restricted. The name of each model indicates which of the three components are included in the mixture. See Table 1 for a description of each specification.

We describe the model with the inverse-Wishart distributional assumption for RCov and give a brief discussion for the Wishart alternative. The novel semiparametric multivariate realized GARCH model notated as MRG-DPM-MΛV-IW has the following hierarchical form

$$r_t|\mathcal{I}_{t-1}, H_t, m_t, L_t \sim \mathrm{N}\left(m_t, H_t^{1/2} L_t \left(H_t^{1/2}\right)'\right), \tag{5}$$

$$\mathrm{RCov}_t|\mathcal{I}_{t-1}, A_t \sim \mathrm{IW}\left(\nu, (\nu - n - 1) H_t^{1/2} A_t \left(H_t^{1/2}\right)'\right), \tag{6}$$

$$m_t, L_t, A_t|G \overset{iid}{\sim} G, \tag{7}$$

$$G|G_0, \alpha \sim \mathrm{DP}(\alpha, G_0), \tag{8}$$

$$G_0(m_t, L_t, A_t) \equiv \mathrm{N}(m_0, M_0) - \mathrm{IW}(\lambda_0, \Lambda_0) - \mathrm{W}(\mathrm{v}_0, V_0), \tag{9}$$

with $\nu, \lambda_0, \mathrm{v}_0 > n + 1$, and H_t having the functional form as in (3).

Eq.(5)-(9) place an infinite mixture of multivariate normals on the returns distribution and an infinite mixture of inverse-Wisharts on the RCov distribution. The mixing is on the returns mean vector m_t, their covariance component L_t and on the RCov scale matrix A_t. L_t and A_t are positive definite matrices that capture shocks in returns and RCov, respectively, that occur at time t, by scaling the conditionally constructed H_t. That way, we approximate the unknown conditional densities of returns and RCov.

The mixing components are distributed according to the latent G which is nonparametrically modelled with a DP prior. A draw from a DP, $G \sim \mathrm{DP}(\alpha, G_0)$, is almost surely a discrete distribution and has two parameters, the base measure G_0 and the precision $\alpha > 0$. The DP is centred around G_0 since $\mathbb{E}[G] = G_0$ and the precision parameter α determines how close is G to G_0 since $\mathrm{Var}[G] = G_0[1 - G_0]/(\alpha + 1)$.

In this framework, the base measure of DP, $G_0(m_t, L_t, A_t)$, in (9) is a normal $-$ inverse-Wishart $-$ Wishart prior. These are well-defined distributions, that are chosen to have $\mathbb{E}[m_t] = 0_n$, $\mathbb{E}[L_t] = I_n$ and $\mathbb{E}[A_t] = I_n$, with 0_n being a vector of zeros and I_n the identity matrix, both of dimension n. If there are no observed shocks, then $m_t = 0_n$, $L_t = I_n$ and $A_t = I_n, \forall t$. This would set the conditional covariance of returns equal to the conditional mean of RCov, $\mathrm{Cov}(r_t|\mathcal{I}_{t-1}) = \mathbb{E}[\mathrm{RCov}_t|\mathcal{I}_{t-1}] = H_t$.

The infinite mixture of the MRG-DPM-MΛV-IW model in (5)-(9) can also be written with the Sethuraman (1994) stick-breaking representation as

$$p(r_t, \mathrm{RCov}_t|\mathcal{I}_{t-1}, H_t, \mathrm{M}, \Lambda, \mathrm{V}, \mathrm{W}) = \sum_{j=1}^{\infty} w_j \, \mathrm{N}\left(r_t \middle| \mu_j, H_t^{1/2} \Lambda_j \left(H_t^{1/2}\right)'\right) \mathrm{IW}\left(\mathrm{RCov}_t \middle| \nu, (\nu - n - 1) H_t^{1/2} \mathrm{V}_j \left(H_t^{1/2}\right)'\right),$$

where $\mathrm{W} = \{w_1, w_2, ...\}$ is the infinite set of mixture weights with $\sum_{j=1}^{\infty} w_j = 1$ and a stick-breaking prior generated as $w_1 = v_1, w_j = v_j \prod_{l=1}^{j-1}(1 - v_l), j > 1, v_j \overset{iid}{\sim} \mathrm{B}(1, \alpha)$, where $\mathrm{B}(.)$ denotes the Beta distribution.

Let $\mathrm{M} = \{\mu_1, \mu_2, ...\}$, $\Lambda = \{\Lambda_1, \Lambda_2, ...\}$ and $\mathrm{V} = \{\mathrm{V}_1, \mathrm{V}_2, ...\}$ denote the unique points of support in G. A given dataset $\{(r_1, \mathrm{RCov}_1), ..., (r_T, \mathrm{RCov}_T)\}$ will be associated with a finite set $\{(m_1, L_1, A_1), ..., (m_T, L_T, A_T)\}$ of draws from G in (7). The DPM permits data clustering under identical mixing components. The model learns from the data and clusters them into identical sets of (m_t, L_t, A_t). This allows pooling data into a finite number of k unique clusters, $\{\mu_j, \Lambda_j, \mathrm{V}_j\}_{j=1}^k$, with $k < T$.

A key mixture component is the DP precision parameter α. This controls the number of unique mixture clusters in order to approximate the unknown data distribution. Ascolani et al. (2023) show that estimating α in a Bayesian fashion makes the DPM cluster consistent. The semiparametric framework nests the parametric models in 2.1. If $\alpha \to 0$, then $w_1 = 1$, $w_j = 0$, $j > 1$, and all the data belong to the same cluster with returns mean vector μ_1, covariance scale Λ_1 and RCov scale component V_1. Hence, MRG-DPM-MΛV-IW would be equivalent to MRG-N-IW, returns would follow a multivariate normal distribution and RCov would follow an inverse-Wishart distribution, instead of a mixture. If $\alpha \to \infty$, then $G \to G_0$, there are as many clusters as data observations and each data observation has its own unique cluster. And if also μ_j is constant $\forall j$, then returns follow a multivariate Student-t distribution.

We also consider the following restricted versions of the MRG-DPM-MΛV-IW model. The first one, MRG-DPM-MΛ-IW, sets $\mathrm{V}_j = V$, $\forall j$, leaving the mixture only in the returns distribution, while the second one, MRG-DPM-Λ-IW, further restricts $\mu_j = 0_n$, $\forall j$, and leaves the mixing only for the returns covariance components, Λ_j.

An alternative distributional framework is also considered. In the general model, notated as MRG-DPM-MΛV-W, returns follow the mixture in (5) while RCov follows a mixture of Wishart distributions as

$$\mathrm{RCov}_t|\mathcal{I}_{t-1}, A_t \sim \mathrm{W}\left(\mathrm{v}, \mathrm{v}^{-1} H_t^{1/2} A_t \left(H_t^{1/2}\right)'\right), \quad \mathrm{v} > n+1. \tag{10}$$

The mixing parameters are distributed as in (7)-(8), with the base measure selected for sampling convenience as $G_0(m_t, L_t, A_t) \equiv \mathrm{N}(m_0, M_0) - \mathrm{IW}(\lambda_0, \Lambda_0) - \mathrm{IW}(v_0, V_0)$. The conditional covariance H_t is constructed as in (3). Similar to MRG-DPM-MΛV-IW, the base measure is selected to have $\mathbb{E}[m_t] = 0_n$, $\mathbb{E}[L_t] = I_n$ and $\mathbb{E}[A_t] = I_n$. We also consider the following restricted versions of this model. MRG-DPM-MΛ-W sets $A_t = V$, $\forall t$, leaving the mixture only in the returns distribution while MRG-DPM-Λ-W further restricts $m_t = 0_n$, $\forall t$, mixing only the returns covariance components. The full list of models is in Table 1.

2.2.2 Estimation

Since this model framework is new we present a brief overview of estimation. The following estimation process is for the MRG-DPM-MΛV-IW. Details for the rest of the models are in the 5ppendix.

In order to make the estimation feasible we use the stick-breaking formulation and the slice sampler by Walker (2007) and Kalli et al. (2011). These truncate the infinite mixture into a finite number k, $k < T$, with the associated unique normal clusters $\{\mu_j, \Lambda_j, \mathrm{V}_j\}_{j=1}^k$. To do so, the parameter space is expanded with the introduction of two latent vectors. The first one is a cluster or state indicator $s_{1:T} = \{s_1, ...s_T\}$ which maps each observation set (r_t, RCov_t) to a cluster j. The second auxiliary slice vector is $u_{1:T} = \{u_1, ...u_T\}$, with $u_t \in (0, 1)$. This helps to truncate the infinite sum into a finite mixture. u_t is defined such that the joint density of $(r_t, \mathrm{RCov}_t, u_t)$, $f(r_t, \mathrm{RCov}_t, u_t|\mathcal{I}_{t-1}, \mathrm{M}, \Lambda, \mathrm{V}, \mathrm{W})$ is equal to

$$\sum_{j=1}^{\infty} \mathbf{1}\{u_t < w_j\}\ \mathrm{N}\left(r_t \middle| \mu_j, H_t^{1/2}\Lambda_j \left(H_t^{1/2}\right)'\right)\ \mathrm{IW}\left(\mathrm{RCov}_t \middle| \nu, (\nu - n - 1) H_t^{1/2} \mathrm{V}_j \left(H_t^{1/2}\right)'\right), \tag{11}$$

where $\mathbf{1}\{.\}$ is an indicator function. Slice variable u_t makes the set $\{u_t < w_j\}_{j=1}^{\infty}$ finite. Integrating (11) over u_t would give the desired density of (r_t, RCov_t). Conditional on $H_{1:T}$, the model likelihood is

$$\begin{aligned} p(r_{1:T}, &\mathrm{RCov}_{1:T}, s_{1:T}, u_{1:T}|H_{1:T}, \mu_{1:k}, \Lambda_{1:k}, \mathrm{V}_{1:k}, w_{1:k}) = \\ &\prod_{t=1}^{T} \mathbf{1}\{u_t < w_{s_t}\}\ \mathrm{N}\left(r_t \middle| \mu_{s_t}, H_t^{1/2}\Lambda_{s_t} \left(H_t^{1/2}\right)'\right)\ \mathrm{IW}\left(\mathrm{RCov}_t \middle| \nu, (\nu - n - 1) H_t^{1/2} \mathrm{V}_{s_t} \left(H_t^{1/2}\right)'\right). \end{aligned} \tag{12}$$

The posterior density of the model parameters is proportional to

$$p(\theta)p(w_{1:k})p(u_{1:T})p(s_{1:T})p(\alpha) \prod_{j=1}^{k} p(\mu_j)p(\Lambda_j)p(\mathrm{V}_j)\ p(r_{1:T}, \mathrm{RCov}_{1:T}, s_{1:T}, u_{1:T}|H_{1:T}, \mu_{1:k}, \Lambda_{1:k}, \mathrm{V}_{1:k}, w_{1:k}), \tag{13}$$

with the likelihood given in (12), $\theta = \{a, b, c, \nu\}$ and k being the smallest positive integer that satisfies the condition $\sum_{j=1}^{k} w_j > 1 - \min(u_{1:T})$. The posterior in (13) does not have a known form. We use a series of MCMC steps to draw from conditional posterior distributions. After $\theta, k, w_{1:k}, s_{1:T}, \mu_{1:k}, \Lambda_{1:k}, \mathrm{V}_{1:k}, \alpha$ are initialized, posterior draws are taken through the following:

1. Sample $\mu_{1:k}, \Lambda_{1:k}, \mathrm{V}_{1:k} | r_{1:T}, \mathrm{RCov}_{1:T}, H_{1:T}, s_{1:T}$.
2. Update $w_{1:k} | s_{1:T}, \alpha$ with a stick-breaking process.
3. Sample the slice vector $u_{1:T} | w_{1:k}, s_{1:T}$.
4. Update k as the smallest integer that satisfies: $\sum_{j=1}^{k} w_j > 1 - \min(u_{1:T})$.
5. Sample $s_{1:T} | r_{1:T}, \mathrm{RCov}_{1:T}, H_{1:T}, \mu_{1:k}, \Lambda_{1:k}, \mathrm{V}_{1:k}, w_{1:k}, u_{1:T}$.
6. Sample $\alpha | T, \kappa$, with $\kappa \leq k$ being the number of clusters in use.
7. Sample $\theta | r_{1:T}, \mathrm{RCov}_{1:T}, \mu_{1:k}, \Lambda_{1:k}, \mathrm{V}_{1:k}, w_{1:k}, s_{1:T}$.

In H_t construction, covariance stationarity is assumed, $\mathbb{E}(H_t) = \bar{\Sigma}$, to target the parameter matrix Ω as

$$\Omega = \bar{\Sigma} \odot \left(\iota\iota' - a - b - c\right), \tag{14}$$

with ι being a n-length vector of ones. Repeating the above steps R times, after R_0 burnin sweeps, gives the posterior draws for inference. Details of the estimation steps are provided in the Appendix.

2.3 Benchmark models

The proposed multivariate realized GARCH models are compared with the following multivariate GARCH specifications that do not include RCov data in the information set. The first benchmark is a parametric two-parameter multivariate GARCH with a multivariate normal distributional assumption for returns, referred to as MG-N, and defined as

$$\begin{aligned} &r_t | r_{1:t-1} \sim \mathrm{N}(0, H_t) \\ &H_t = \Omega + a\, r_{t-1}(r_{t-1})' + b\, H_{t-1}. \end{aligned} \tag{15}$$

The second benchmark has the multivariate-t assumption for the returns distribution and can capture the extreme tails that empirically are observed in financial returns. It is a model proposed by Diamantopoulos and Vrontos (2010) and the benchmark of Jensen and Maheu (2013). A similar specification is also used as a benchmark from Maheu and Shamsi Zamenjani (2021). The model MG-t is defined as

$$r_t | r_{1:t-1} \sim \mathrm{t}(0, H_t, \zeta), \quad \zeta > 2,$$

with ζ being the multivariate-t degrees of freedom and H_t constructed as in (15).

The final benchmark is the semiparametric multivariate GARCH model, developed by Jensen and Maheu (2013), which has no underlying distributional assumption for returns but instead, it uses an infinite mixture of normals to approximate their unknown density. A general specification has also been used by Maheu and Shamsi Zamenjani (2021). The specification used here, and referred to as MG-DPM, is the following

$$r_t | r_{1:t-1}, H_t, L_t \sim \mathrm{N}\left(\mu_t, H_t^{1/2} L_t \left(H_t^{1/2}\right)'\right),$$

with $\mu_t, L_t | G \overset{iid}{\sim} G,\ G | G_0, \alpha \sim \mathrm{DP}(\alpha, G_0)$, base measure $G_0(\mu_t, L_t) \equiv \mathrm{N}(m_0, M_0) - \mathrm{IW}(\nu_0, V_0)$, $\nu_0 > n + 1$ and H_t constructed as in (15). The benchmark models are also presented in Table 1.

Table 1: Models specifications.

Multivariate GARCH:	$H_t = \Omega + a\, r_{t-1}(r_{t-1})' + b\, H_{t-1}$
MG-N:	$r_t \vert r_{1:t-1} \sim \mathrm{N}(0, H_t)$.
MG-t:	$r_t \vert r_{1:t-1} \sim \mathrm{t}(0, H_t, \zeta), \quad \zeta > 2$.
MG-DPM:	$r_t \vert r_{1:t-1}, H_t, L_t \sim \mathrm{N}\left(\mu_t, H_t^{1/2} L_t \left(H_t^{1/2}\right)'\right)$,
	$\mu_t, L_t \vert G \overset{iid}{\sim} G, \quad G \vert G_0, \alpha \sim \mathrm{DP}(\alpha, G_0), \quad G_0(\mu_t, L_t) \equiv \mathrm{N}(0, 0.1 I_n) - \mathrm{IW}(n+10, 9 I_n)$.
Multivariate Realized GARCH:	$H_t = \Omega + a\, r_{t-1}(r_{t-1})' + b\, H_{t-1} + c\, \mathrm{RCov}_{t-1}$
MRG-N-W:	$r_t \vert \mathcal{I}_{t-1} \sim \mathrm{N}(0, H_t), \quad \mathrm{RCov}_t \vert \mathcal{I}_{t-1} \sim \mathrm{W}\left(\mathrm{v}, \mathrm{v}^{-1} H_t^{1/2} V \left(H_t^{1/2}\right)'\right), \quad \mathrm{v} > n+1$.
MRG-N-IW:	$r_t \vert \mathcal{I}_{t-1} \sim \mathrm{N}(0, H_t), \quad \mathrm{RCov}_t \vert \mathcal{I}_{t-1} \sim \mathrm{IW}\left(\nu, (\nu - n - 1) H_t^{1/2} V \left(H_t^{1/2}\right)'\right), \quad \nu > n+1$.
MRG-t-W:	$r_t \vert \mathcal{I}_{t-1} \sim \mathrm{t}(0, H_t, \zeta), \quad \zeta > 2, \quad \mathrm{RCov}_t \vert \mathcal{I}_{t-1} \sim \mathrm{W}\left(\mathrm{v}, \mathrm{v}^{-1} \frac{\zeta}{\zeta-2} H_t^{1/2} V \left(H_t^{1/2}\right)'\right), \quad \mathrm{v} > n+1$.
MRG-t-IW:	$r_t \vert \mathcal{I}_{t-1} \sim \mathrm{t}(0, H_t, \zeta), \quad \zeta > 2, \quad \mathrm{RCov}_t \vert \mathcal{I}_{t-1} \sim \mathrm{IW}\left(\nu, (\nu - n - 1) \frac{\zeta}{\zeta-2} H_t^{1/2} V \left(H_t^{1/2}\right)'\right), \quad \nu > n+1$.
MRG-DPM-Λ-W:	$r_t \vert \mathcal{I}_{t-1}, H_t, L_t \sim \mathrm{N}\left(0, H_t^{1/2} L_t \left(H_t^{1/2}\right)'\right), \quad \mathrm{RCov}_t \vert \mathcal{I}_{t-1} \sim \mathrm{W}\left(\mathrm{v}, \mathrm{v}^{-1} H_t^{1/2} V \left(H_t^{1/2}\right)'\right)$,
	$\mathrm{v} > n+1, \quad L_t \vert G \overset{iid}{\sim} G, \quad G \vert G_0, \alpha \sim \mathrm{DP}(\alpha, G_0), \quad G_0(L_t) \equiv \mathrm{IW}(n+10, 9 I_n)$.
MRG-DPM-Λ-IW:	$r_t \vert \mathcal{I}_{t-1}, H_t, L_t \sim \mathrm{N}\left(0, H_t^{1/2} L_t \left(H_t^{1/2}\right)'\right), \quad \mathrm{RCov}_t \vert \mathcal{I}_{t-1} \sim \mathrm{IW}\left(\nu, (\nu - n - 1) H_t^{1/2} V \left(H_t^{1/2}\right)'\right)$,
	$\nu > n+1, \quad L_t \vert G \overset{iid}{\sim} G, \quad G \vert G_0, \alpha \sim \mathrm{DP}(\alpha, G_0), \quad G_0(L_t) \equiv \mathrm{IW}(n+10, 9 I_n)$.
MRG-DPM-MΛ-W:	$r_t \vert \mathcal{I}_{t-1}, H_t, m_t, L_t \sim \mathrm{N}\left(m_t, H_t^{1/2} L_t \left(H_t^{1/2}\right)'\right), \quad \mathrm{RCov}_t \vert \mathcal{I}_{t-1} \sim \mathrm{W}\left(\mathrm{v}, \mathrm{v}^{-1} H_t^{1/2} V \left(H_t^{1/2}\right)'\right)$,
	$\mathrm{v} > n+1, m_t, L_t \vert G \overset{iid}{\sim} G, \quad G \vert G_0, \alpha \sim \mathrm{DP}(\alpha, G_0), \quad G_0(m_t, L_t) \equiv \mathrm{N}(0, 0.1 I_n) - \mathrm{IW}(n+10, 9 I_n)$.
MRG-DPM-MΛ-IW:	$r_t \vert \mathcal{I}_{t-1}, H_t, m_t, L_t \sim \mathrm{N}\left(m_t, H_t^{1/2} L_t \left(H_t^{1/2}\right)'\right), \quad \mathrm{RCov}_t \vert \mathcal{I}_{t-1} \sim \mathrm{IW}\left(\nu, (\nu - n - 1) H_t^{1/2} V \left(H_t^{1/2}\right)'\right)$,
	$\nu > n+1, \quad m_t, L_t \vert G \overset{iid}{\sim} G, \quad G \vert G_0, \alpha \sim \mathrm{DP}(\alpha, G_0), \quad G_0(m_t, L_t) \equiv \mathrm{N}(0, 0.1 I_n) - \mathrm{IW}(n+10, 9 I_n)$.
MRG-DPM-MΛV-W:	$r_t \vert \mathcal{I}_{t-1}, H_t, m_t, L_t \sim \mathrm{N}\left(m_t, H_t^{1/2} L_t \left(H_t^{1/2}\right)'\right), \quad \mathrm{RCov}_t \vert \mathcal{I}_{t-1}, A_t \sim \mathrm{W}\left(\mathrm{v}, \mathrm{v}^{-1} H_t^{1/2} A_t \left(H_t^{1/2}\right)'\right)$,
	$\mathrm{v} > n+1, \quad m_t, L_t, A_t \vert G \overset{iid}{\sim} G, \quad G \vert G_0, \alpha \sim \mathrm{DP}(\alpha, G_0)$,
	$G_0(m_t, L_t, A_t) \equiv \mathrm{N}(0, 0.1 I_n) - \mathrm{IW}(n+10, 9 I_n) - \mathrm{IW}(n+10, 9 I_n)$.
MRG-DPM-MΛV-IW:	$r_t \vert \mathcal{I}_{t-1}, H_t, m_t, L_t \sim \mathrm{N}\left(m_t, H_t^{1/2} L_t \left(H_t^{1/2}\right)'\right), \quad \mathrm{RCov}_t \vert \mathcal{I}_{t-1}, A_t \sim \mathrm{IW}\left(\nu, (\nu - n - 1) H_t^{1/2} A_t \left(H_t^{1/2}\right)'\right)$,
	$\nu > n+1, \quad m_t, L_t, A_t \vert G \overset{iid}{\sim} G, \quad G \vert G_0, \alpha \sim \mathrm{DP}(\alpha, G_0)$,
	$G_0(m_t, L_t, A_t) \equiv \mathrm{N}(0, 0.1 I_n) - \mathrm{IW}(n+10, 9 I_n) - \mathrm{W}(n+10, (n+10)^{-1} I_n)$.

Notes: This table presents the detailed specifications of all the tested models. The top panel presents the Multivariate GARCH benchmarks and the bottom panel presents the Multivariate Realized GARCH models.

3 Forecasting

In Bayesian nonparametric forecasting, the main focus is on the predictive density. For the parametric models, their posterior draws $\{\theta^{(i)}\}_{i=1}^{R}$ can be used to approximate the predictive returns density as

$$p(r_{t+h} \vert \mathcal{I}_t) \approx \frac{1}{R} \sum_{i=1}^{R} f_n\left(r_{t+h} \vert \mathcal{I}_t, \theta^{(i)}\right),$$

with $f_n(.\vert)$ being the n dimension multivariate p.d.f. (normal or Student's-t) associated with each model's return density assumption and $h = 1, ..., h_{\text{max}}$ being the *out-of-sample* forecast horizon.

The following focuses on the MRG-DPM-MΛV-IW but can be easily modified for the other semiparametric models. The key task when forecasting with DPM models is to integrate out the uncertainty about the future state of the modelled series. Conditional on $\mathcal{I}_t = \{r_{1:t}, \mathrm{RCov}_{1:t}\}$, H_{t+h} can be calculated from (3) and the returns predictive density can be approximated as

$$p_h(r_{t+h}|\mathcal{I}_t, \mu_{1:k}, \Lambda_{1:k}, s_{1:k}, w_{1:k}) \approx \frac{1}{R}\sum_{i=1}^{R} \mathrm{N}\left(r_{t+h}\middle|\mu^{(i)}_{s^{(i)}_{t+h}}, H^{1/2(i)}_{t+h}\Lambda^{(i)}_{s^{(i)}_{t+h}}\left(H^{1/2(i)}_{t+h}\right)'\right), \tag{16}$$

$$\text{where} \quad s^{(i)}_{t+h} = \begin{cases} j, & \text{if} \quad \sum_{l=0}^{j-1} w^{(i)}_l < \phi < \sum_{l=0}^{j} w^{(i)}_l, \\ k^{(i)}+1, & \text{if} \quad \phi \geq \sum_{l=0}^{k^{(i)}} w^{(i)}_l, \end{cases}$$

with $w^{(i)}_o = 0$, $j \leq k^{(i)}$ and $\phi \sim \mathrm{U}(0,1)$. The above means that the future value of $s^{(i)}_{t+h}$ is one of the existing clusters with probability equal to the associated weights and there is a non-zero probability of introducing a new normal cluster $\left(\mu^{(i)}_{k^{(i)}+1}, \Lambda^{(i)}_{k^{(i)}+1}\right)$ from the base measure G_0. The predictive density of RCov_{t+h}, conditional on H_{t+h} and s_{t+h}, can be approximated as

$$p_h(\mathrm{RCov}_{t+h}|\mathcal{I}_t, \mathrm{V}_{1:k}, s_{1:k}, w_{1:k}) \approx \frac{1}{R}\sum_{i=1}^{R} \mathrm{IW}\left(\mathrm{RCov}_{t+h}\middle|\nu^{(i)}, (\nu^{(i)}-n-1)H^{1/2(i)}_{t+h}\mathrm{V}^{(i)}_{s^{(i)}_{t+h}}\left(H^{1/2(i)}_{t+h}\right)'\right). \tag{17}$$

To propagate $H^{(i)}_{t+h}$ for $h > 1$, r_{t+h-1} and RCov_{t+h-1} are simulated conditional on $s^{(i)}_{t+h-1}$.

The predictive density is the key component for the calculation of the predictive likelihood (PL) (Geweke, 1994). This measure provides an *out-of-sample* forecast record of a model for a block of data and hence allows model comparison. The density forecasts are evaluated over an identical set τ (with $1 < \tau < T$) of out-of-sample returns $r_{T-\tau+1:T}$ and RCov matrices $\mathrm{RCov}_{T-\tau+1:T}$. The cumulative log-predictive likelihood log-PL_h for a forecast horizon h is calculated observation-by-observation as

$$\text{log-PL}_h(r_{T-\tau+1:T}|\mathcal{I}_T) = \sum_{t=T-\tau-h}^{T-h} \log\left(p_h(r_{t+h}|\mathcal{I}_t, \mu_{1:k}, \Lambda_{1:k}, s_{1:k}, w_{1:k})\right), \tag{18}$$

where $p_h(r_{t+h}|\mathcal{I}_t, \mu_{1:k}, \Lambda_{1:k}, s_{1:k}, w_{1:k})$ is approximated as in (16). Similarly, the log-predictive likelihood of $\mathrm{RCov}_{T-\tau+1:T}$ is calculated as

$$\text{log-PL}_h(\mathrm{RCov}_{T-\tau+1:T}|\mathcal{I}_T) = \sum_{t=T-\tau-h}^{T-h} \log\left(p_h(\mathrm{RCov}_{t+h}|\mathcal{I}_t, \mathrm{V}_{1:k}, s_{1:k}, w_{1:k})\right), \tag{19}$$

where $p_h(\mathrm{RCov}_{t+h}|\mathcal{I}_t, \mathrm{V}_{1:k}, s_{1:k}, w_{1:k})$ is calculated as in (17). In comparing two models, the difference between their (log-)predictive likelihoods is used to calculate the predictive (log-)Bayes factor. A guide to Bayes factors is provided by Kass and Raftery (1995). According to that, the model with the larger log-PL value is preferred and strongly preferred if the log-Bayes factor value is more than five.

The models' performance on forecasting the conditional returns covariance is measured by the mean squared forecast error (MSFE) of the matrix (Frobenius) norm difference between out-of-sample RCov_{t+h} and the conditional returns covariance

$$\mathrm{MSFE}_h = \frac{1}{T-\tau+1}\sum_{t=T-\tau-h}^{T-h} \left|\left|\mathrm{RCov}_{t+h} - \mathrm{Cov}(r_{t+h}|\mathcal{I}_t)\right|\right|, \tag{20}$$

where $\mathrm{Cov}(r_{t+1}|\mathcal{I}_t)$ is the predictive covariance matrix based on each model's return distributional assumption, $\mathrm{Cov}(r_{t+h}|\mathcal{I}_t) \equiv \mathbb{E}(H_{t+h})$ for multivariate normal and $\mathrm{Cov}(r_{t+h}|\mathcal{I}_t) \equiv \mathbb{E}\left(\frac{\zeta}{\zeta-2}H_{t+h}\right)$ for multivariate-t. For the semiparametric models, the predictive returns covariance is approximated by integrating out next period's parameter and state uncertainty as

$$\mathrm{Cov}(r_{t+h}|\mathcal{I}_t) \approx \frac{1}{R}\sum_{i=1}^{R}\left[H^{1/2(i)}_{t+h}\Lambda^{(i)}_{s^{(i)}_{t+h}}\left(H^{1/2(i)}_{t+h}\right)' + \mu^{(i)}_{s_{t+h}}\left(\mu^{(i)}_{s_{t+h}}\right)'\right] - \left(\frac{1}{R}\sum_{i=1}^{R}\mu^{(i)}_{s_{t+h}}\right)\left(\frac{1}{R}\sum_{i=1}^{R}\mu^{(i)}_{s_{t+h}}\right)'.$$

Table 2: Posterior Means and 95% density interval − SPY/BAC

	MRG-N-W		MRG-t-W		MRG-DPM-Λ-W		MRG-DPM-MΛ-W		MRG-DPM-MΛV-W	
	Mean	95% D.I.	Mean	95% D.I.	Mean	95% D.I.	Mean	95% D.I.	Mean	95% D.I.
$\Omega_{1,1}$	0.004	[0.003, 0.006]	0.003	[0.002, 0.004]	0.005	[0.004, 0.007]	0.005	[0.004, 0.007]	0.005	[0.004, 0.007]
$\Omega_{1,2}$	0.006	[0.004, 0.008]	0.005	[0.004, 0.006]	0.008	[0.006, 0.010]	0.008	[0.006, 0.010]	0.008	[0.005, 0.010]
$\Omega_{2,2}$	0.022	[0.015, 0.029]	0.017	[0.013, 0.023]	0.027	[0.020, 0.035]	0.027	[0.021, 0.035]	0.027	[0.019, 0.035]
a	0.029	[0.022, 0.037]	0.022	[0.016, 0.028]	0.028	[0.020, 0.035]	0.027	[0.020, 0.035]	0.027	[0.019, 0.034]
b	0.611	[0.584, 0.636]	0.603	[0.577, 0.627]	0.604	[0.579, 0.629]	0.604	[0.579, 0.629]	0.599	[0.572, 0.625]
c	0.356	[0.330, 0.383]	0.273	[0.249, 0.298]	0.363	[0.339, 0.389]	0.364	[0.339, 0.389]	0.369	[0.343, 0.397]
ζ			8.014	[6.942, 9.249]						
α					0.663	[0.204, 1.313]	0.282	[0.061, 0.667]	0.261	[0.054, 0.639]
κ					9.232	[4.000, 15.000]	3.464	[2.000, 7.000]	3.153	[2.000, 6.000]
v	10.848	[10.494, 11.201]	10.847	[10.512, 11.184]	10.840	[10.503, 11.182]	10.848	[10.497, 11.197]	11.242	[10.870, 11.627]
$V_{1,1}$	1.036	[1.016, 1.056]	1.034	[0.998, 1.079]	1.032	[1.013, 1.051]	1.032	[1.013, 1.051]		
$V_{1,2}$	−0.011	[−0.026, 0.004]	−0.010	[−0.025, 0.006]	−0.013	[−0.026, 0.001]	−0.013	[−0.026, 0.001]		
$V_{2,2}$	1.042	[1.012, 1.075]	1.041	[0.994, 1.104]	1.029	[1.003, 1.054]	1.028	[1.003, 1.053]		
	MRG-N-IW		**MRG-t-IW**		**MRG-DPM-Λ-IW**		**MRG-DPM-MΛ-IW**		**MRG-DPM-MΛV-IW**	
	Mean	95% D.I.	Mean	95% D.I.	Mean	95% D.I.	Mean	95% D.I.	Mean	95% D.I.
$\Omega_{1,1}$	0.007	[0.005, 0.008]	0.005	[0.004, 0.007]	0.008	[0.007, 0.010]	0.008	[0.007, 0.010]	0.008	[0.007, 0.010]
$\Omega_{1,2}$	0.010	[0.008, 0.012]	0.008	[0.006, 0.010]	0.012	[0.010, 0.014]	0.012	[0.010, 0.014]	0.012	[0.010, 0.014]
$\Omega_{2,2}$	0.035	[0.028, 0.043]	0.028	[0.023, 0.034]	0.042	[0.034, 0.050]	0.042	[0.034, 0.050]	0.042	[0.034, 0.050]
a	0.022	[0.015, 0.029]	0.016	[0.011, 0.022]	0.020	[0.013, 0.027]	0.020	[0.013, 0.027]	0.020	[0.013, 0.028]
b	0.601	[0.575, 0.626]	0.592	[0.567, 0.616]	0.596	[0.570, 0.622]	0.596	[0.570, 0.620]	0.596	[0.569, 0.622]
c	0.371	[0.346, 0.397]	0.284	[0.260, 0.309]	0.376	[0.350, 0.402]	0.377	[0.352, 0.403]	0.377	[0.351, 0.403]
ζ			7.881	[6.771, 9.231]						
α					0.517	[0.153, 1.070]	0.241	[0.049, 0.585]	0.286	[0.058, 0.709]
κ					7.073	[3.000, 12.000]	2.835	[2.000, 5.000]	3.508	[2.000, 7.000]
ν	11.959	[11.581, 12.346]	11.950	[11.575, 12.371]	11.960	[11.592, 12.341]	11.973	[11.614, 12.355]	12.071	[11.682, 12.452]
$V_{1,1}$	1.033	[1.011, 1.055]	1.030	[0.982, 1.090]	1.030	[1.009, 1.051]	1.029	[1.008, 1.051]		
$V_{1,2}$	−0.062	[−0.076, −0.049]	−0.062	[−0.077, −0.045]	−0.065	[−0.077, −0.052]	−0.065	[−0.077, −0.052]		
$V_{2,2}$	1.017	[0.986, 1.050]	1.015	[0.950, 1.107]	1.002	[0.976, 1.028]	1.002	[0.976, 1.028]		

Notes: The results are from 30,000 MCMC posterior draws (after 20,000 burnin sweeps). a, b and c are the GARCH scalar of parameters. ζ is the Student-t degrees of freedom parameter, v is the Wishart and ν is the inverse-Wishart degrees of freedom parameters, α is DPM precision parameter and κ the number of active normal clusters. Parameters in $\{\Omega_{1,1}, \Omega_{1,2}, \Omega_{2,2}\} = \text{vech}(\Omega)$ are calculated with covariance targeting.

4 Empirical application

4.1 Data

The data used are from Noureldin et al. (2012), and have been obtained through Oxford-Man Institute Realized Library[10] (Heber et al., 2009) and consist of two different datasets, one with two assets: Spyder (SPY), the S&P 500 ETF and Bank of America (BAC), and one with 10 Dow Jones Industrial Index stocks. The data stocks are (in the order provided): BAC, JP Morgan (JPM), International Business Machines (IBM), Microsoft (MSFT), Exxon Mobil (XOM), Alcoa (AA), American Express (AXP), Du Pont (DD), General Electric (GE) and Coca Cola (KO). The sample period is from February 1st, 2001 to December 31st, 2009 (2242 days). The data are the open-to-close[11] daily log returns and their realized measures of variance and covariance. The realized measures are calculated from open-to-close 5-minute returns, with subsampling using 1-minute returns for each trading day. Returns have been converted to percentages and RCov matrices have been scaled by 100^2. Summary statistics are in the Appendix.

[10] The Oxford-Man Realized Library was discontinued in 2022.

[11] Following a big part of the relevant literature (Gouriéroux et al., 2009; Golosnoy et al., 2012; Jin and Maheu, 2013, 2016, among many others), we do not incorporate over-night returns information in the RCov construction.

Table 3: Posterior Means and 95% density interval − 10 stocks

	MRG-N-W		MRG-t-W		MRG-DPM-Λ-W		MRG-DPM-MΛ-W		MRG-DPM-MΛV-W	
	Mean	95% D.I.	Mean	95% D.I.	Mean	95% D.I.	Mean	95% D.I.	Mean	95% D.I.
$\Omega_{1,1}$	0.031	[0.028, 0.033]	0.024	[0.022, 0.026]	0.032	[0.030, 0.035]	0.032	[0.030, 0.035]	0.033	[0.030, 0.036]
$\Omega_{1,2}$	0.021	[0.018, 0.022]	0.016	[0.015, 0.017]	0.021	[0.020, 0.023]	0.021	[0.020, 0.023]	0.022	[0.020, 0.024]
$\Omega_{2,2}$	0.026	[0.024, 0.028]	0.020	[0.019, 0.022]	0.027	[0.025, 0.029]	0.027	[0.025, 0.030]	0.028	[0.025, 0.031]
a	0.007	[0.006, 0.008]	0.005	[0.004, 0.006]	0.007	[0.006, 0.008]	0.007	[0.006, 0.008]	0.006	[0.005, 0.007]
b	0.752	[0.745, 0.756]	0.754	[0.750, 0.758]	0.754	[0.749, 0.759]	0.754	[0.750, 0.759]	0.759	[0.754, 0.765]
c	0.236	[0.231, 0.243]	0.172	[0.167, 0.175]	0.234	[0.229, 0.238]	0.234	[0.229, 0.238]	0.229	[0.224, 0.234]
ζ			7.443	[7.112, 7.825]						
α					2.606	[1.619, 3.829]	0.271	[0.069, 0.610]	0.774	[0.328, 1.429]
κ					36.947	[28.000, 48.000]	3.270	[2.000, 5.000]	10.901	[8.000, 16.000]
v	26.992	[26.852, 27.132]	27.012	[26.853, 27.146]	26.993	[26.807, 27.177]	26.994	[26.807, 27.186]	27.538	[27.302, 27.769]
$V_{1,1}$	0.941	[0.930, 0.954]	0.933	[0.918, 0.953]	0.937	[0.925, 0.950]	0.937	[0.924, 0.950]		
$V_{1,2}$	−0.016	[−0.023, −0.008]	−0.015	[−0.023, −0.008]	−0.015	[−0.023, −0.008]	−0.015	[−0.023, −0.008]		
$V_{2,2}$	1.028	[1.017, 1.040]	1.020	[1.006, 1.037]	1.026	[1.014, 1.038]	1.026	[1.014, 1.037]		
	MRG-N-IW		**MRG-t-IW**		**MRG-DPM-Λ-IW**		**MRG-DPM-MΛ-IW**		**MRG-DPM-MΛV-IW**	
	Mean	95% D.I.	Mean	95% D.I.	Mean	95% D.I.	Mean	95% D.I.	Mean	95% D.I.
$\Omega_{1,1}$	0.044	[0.041, 0.045]	0.033	[0.031, 0.035]	0.045	[0.042, 0.048]	0.045	[0.043, 0.048]	0.044	[0.042, 0.048]
$\Omega_{1,2}$	0.029	[0.027, 0.030]	0.022	[0.021, 0.023]	0.030	[0.028, 0.032]	0.030	[0.028, 0.032]	0.029	[0.027, 0.031]
$\Omega_{2,2}$	0.037	[0.035, 0.038]	0.028	[0.026, 0.029]	0.038	[0.036, 0.040]	0.038	[0.036, 0.041]	0.038	[0.035, 0.040]
a	0.007	[0.006, 0.008]	0.005	[0.004, 0.006]	0.006	[0.005, 0.007]	0.006	[0.005, 0.007]	0.005	[0.004, 0.006]
b	0.782	[0.778, 0.786]	0.785	[0.781, 0.791]	0.784	[0.780, 0.789]	0.784	[0.780, 0.789]	0.790	[0.784, 0.795]
c	0.204	[0.200, 0.208]	0.148	[0.142, 0.153]	0.202	[0.198, 0.206]	0.202	[0.197, 0.206]	0.197	[0.192, 0.202]
ζ			7.355	[6.891, 7.881]						
α					2.496	[1.598, 3.580]	0.355	[0.104, 0.759]	0.445	[0.142, 0.920]
κ					35.422	[28.000, 43.000]	4.593	[3.000, 8.000]	5.962	[4.000, 10.000]
ν	28.100	[27.979, 28.355]	28.038	[27.873, 28.296]	28.048	[27.861, 28.238]	28.051	[27.866, 28.236]	28.315	[28.086, 28.536]
$V_{1,1}$	0.828	[0.817, 0.840]	0.814	[0.801, 0.827]	0.823	[0.811, 0.835]	0.822	[0.810, 0.834]		
$V_{1,2}$	−0.062	[−0.069, −0.055]	−0.062	[−0.069, −0.055]	−0.062	[−0.069, −0.055]	−0.062	[−0.069, −0.055]		
$V_{2,2}$	1.013	[1.001, 1.025]	1.001	[0.987, 1.015]	1.011	[0.998, 1.023]	1.010	[0.998, 1.022]		

Notes: The results are from 30,000 MCMC posterior draws (after 20,000 burnin sweeps). a, b and c are the GARCH scalar of parameters. ζ is the Student-t degrees of freedom, v is the Wishart and ν is the inverse-Wishart degrees of freedom parameters, α is DPM precision parameter and κ the number of active normal clusters. Parameters in Ω are calculated with covariance targeting. Here are reported the Ω parameters for the first two series, namely BAC and JPM. The full estimates of matrix V are presented in the Appendix.

4.2 Priors

For most of the parameters we select uninformative but proper priors. GARCH parameters in $a, b,$ and c, have independent normal priors, $\mathrm{N}(0, 100)\mathbf{1}\{a, b, c \geq 0\}$. Posterior draws which do not satisfy positive definite Ω and $H_{1:T}$ are rejected. The inverse-Wishart degrees of freedom parameter has an exponential prior, $\nu \sim \mathrm{Exp}(n + 10)$, with $\mathbb{E}(\nu) = n+10$. The same prior is used for the Wishart degrees of freedom parameter, v. The Student-t degrees of freedom has a uniform prior, $\zeta \sim \mathrm{U}(2, 100)$.

In the DPM base measure, G_0, $\forall j$, the returns mean is set around zero, $\mathbb{E}(\mu_j) = 0_n$, with $m_0 = 0_n$ and $M_0 = 0.1 \times I_n$. The scaling covariance matrices are drawn from $\Lambda_j \sim \mathrm{IW}(\lambda_0, \Lambda_0)$, which centres them around the identity matrix, $\mathbb{E}(\Lambda_j) = I_n$, by setting $\Lambda_0 = (\lambda_0 - n - 1)I_n$ and $\lambda_0 = n + 10$. The RCov mixture components are drawn from $\mathrm{V}_j \sim \mathrm{W}(\mathrm{v}_0, V_0)$ for the IW models and from $\mathrm{V}_j \sim \mathrm{IW}(v_0, V_0)$ for the Wishart ones. In both cases $v_0 = \mathrm{v}_0 = n + 10$ and V_0 is set to values that centre the mixture around identity matrix, $\mathbb{E}(\mathrm{V}_j) = I_n$. See Table 1 for details on each model's base measure. Other values of λ_0, v_0, v_0 and M_0 do not affect the results. The precision parameter of DPM has a Gamma prior, $\alpha \sim \Gamma(2, 8)$, following Jensen and Maheu (2013).

4.3 Posterior estimation results

Results from 30,000 MCMC posterior draws, after 20,000 burnin, are reported in Tables 2 and 3. The covariance persistence H_{t-1} is a key component in the GARCH equation as parameter b has values ranging from 0.592 to 0.611 for SPY/BAC and from 0.752 to 0.790 for the 10 stocks dataset. The lag of RCov is the second most important GARCH structural component, with parameter c values ranging from 0.273 to 0.377 for SPY/BAC and from 0.148 to 0.236 for the 10 stocks. This is expected based on previous findings (Engle, 2002b; Hansen et al., 2012, 2014; Gorgi et al., 2019; Archakov et al., 2019). The ARCH parameter a, which in the literature is often removed from the conditional covariance construction[12], has low and non-zero values ranging from 0.016 to 0.029 for SPY/BAC and from 0.005 to 0.007 for the 10 stocks dataset.

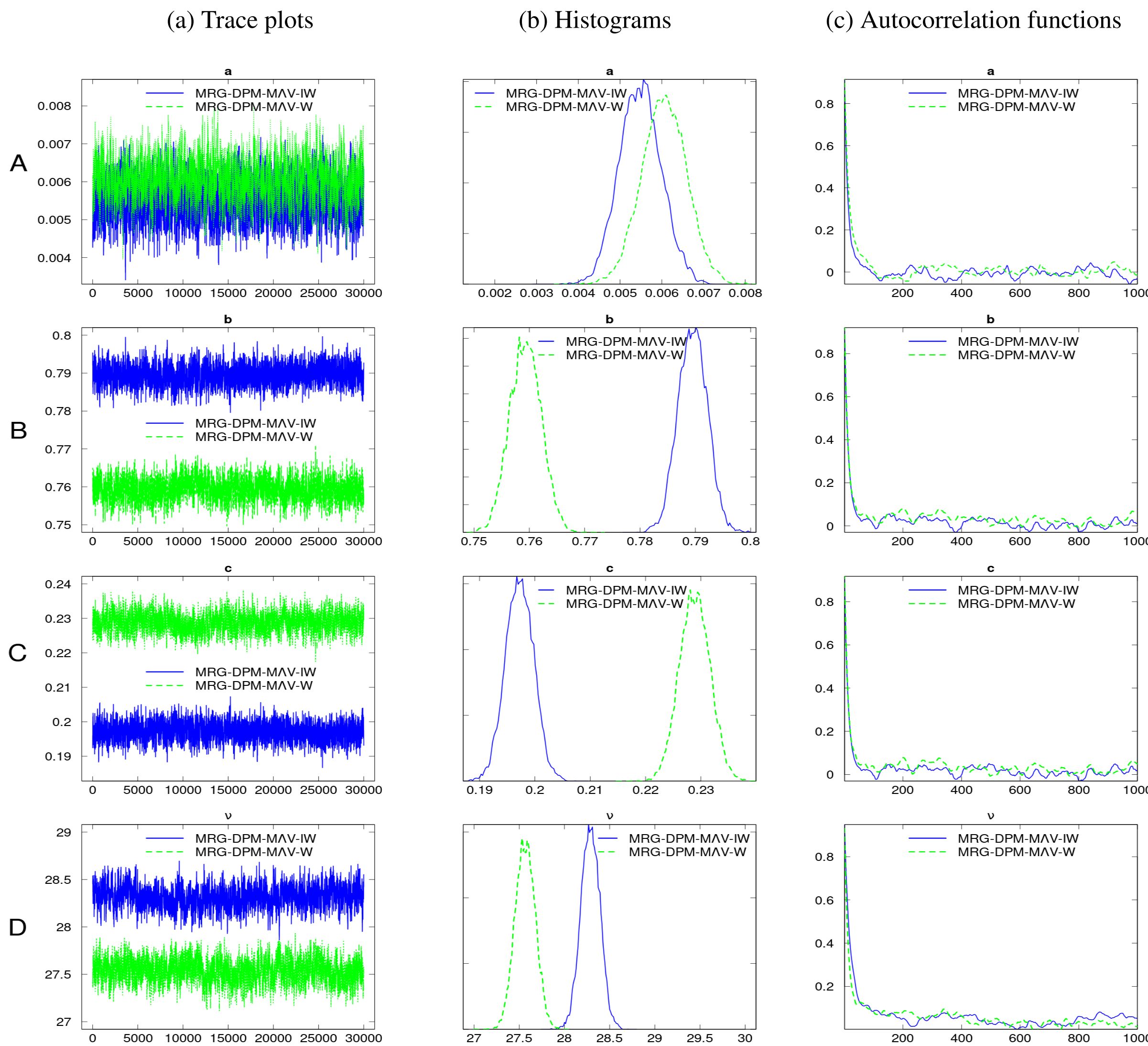


Figure 1: The figure displays (a) trace plots, (b) smoothed histograms and (c) autocorrelation functions for selected parameters of the proposed semiparametric models, for the 10 stocks dataset. Panel A: ARCH parameter a, panel B: GARCH parameter b, panel C: RCov lag parameter c and panel D: (inverse-)Wishart degrees of freedom parameters.

In the RCov measurement equation, the Wishart degrees of freedom parameter is consistently estimated across the models with values around 11 for SPY/BAC and 27 for the 10 stocks. The inverse-Wishart degrees of freedom parameter has values around 12 for SPY/BAC and 28 for the 10 stocks. For all the models and data, the parametric

[12]See Engle (2002b) and Hansen et al. (2012, 2014). We estimated the models without the ARCH parameters and found that their forecasting performance worsened.

scale matrix V, is estimated close to the identity matrix. See the Appendix for the full estimated V matrix for the 10 stocks dataset. The trace and autocorrelation plots of all the above parameters from the general semiparametric models in Figure 1 show good convergence to the stationary posterior.

Regarding the distributional features of the data, the multivariate Student's-t models indicate fat-tails in returns, with degrees of freedom estimates ranging from 7.355 to 8.014 across both datasets. The DPM models suggest that a finite mixture is required to approximate the underlying data distribution. MRG-DPM-MΛ models use, on average, between 2.835 and 4.593 clusters. The more restrictive MRG-DPM-Λ models naturally require a larger number of clusters to approximate the returns distribution. Including RCov in the mixture has a noticeable impact, generally increasing the average number of clusters used in most cases.

(a) SPY/BAC (b) 10 stocks

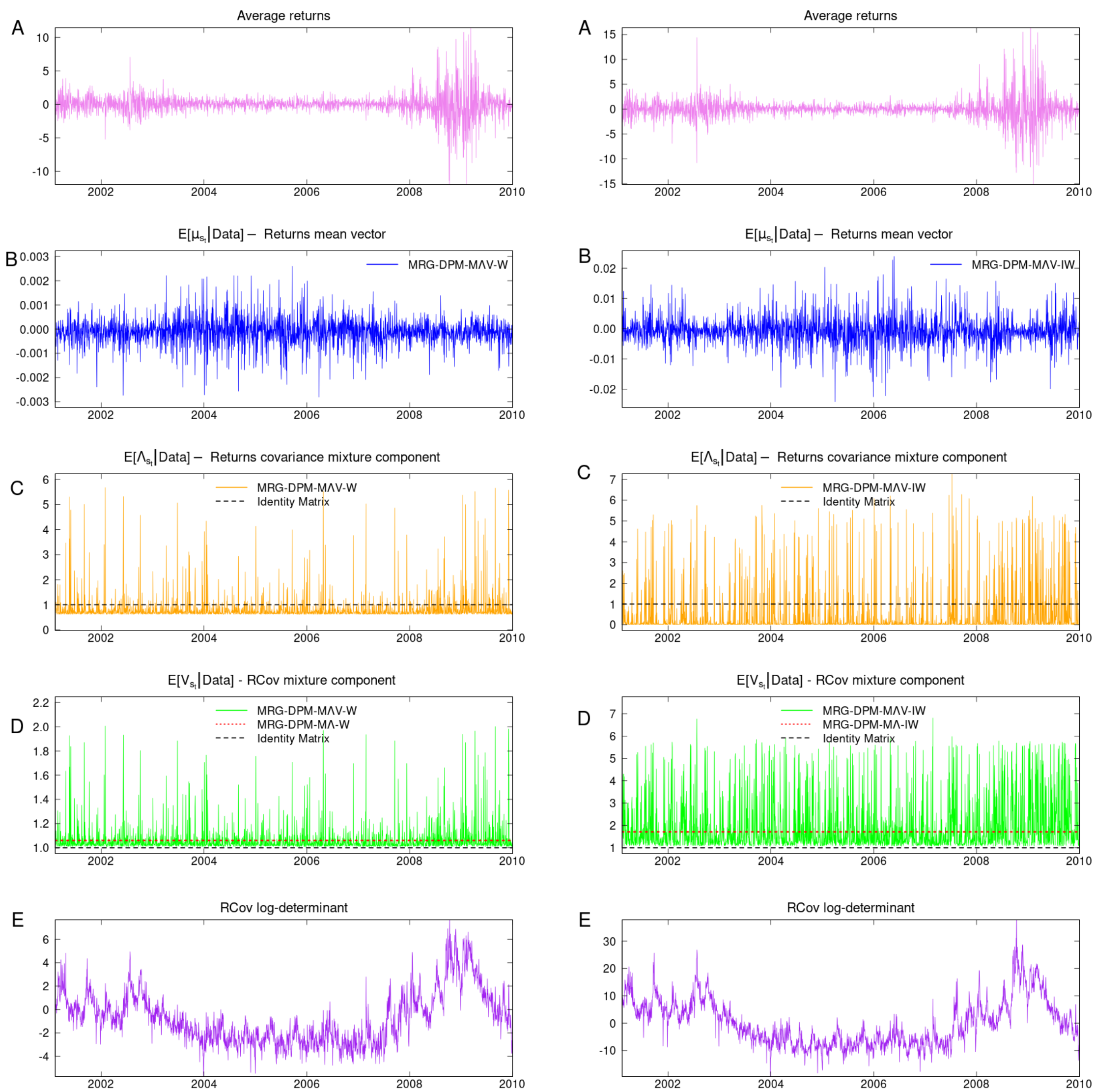


Figure 2: The plots display the posterior estimated nonparametric mixture components against restricted parametric or identity matrix specifications. Panels A: Average returns, panels B: Sum of returns mean vector, $\mathbb{E}\left[\mu_{s_t}|\mathcal{I}_T\right]$, panels C: Determinant of returns covariance mixture component, $\mathbb{E}\left[\Lambda_{s_t}|\mathcal{I}_T\right]$, panels D: Determinant of RCov mixture component matrix, $\mathbb{E}\left[\mathrm{V}_{s_t}|\mathcal{I}_T\right]$, and panels E: log-determinant of RCov matrices.

Figure 2 displays the posterior estimated mixture components from the semiparametric models MRG-DPM-MΛV-W for SPY/BAC and MRG-DPM-MΛV-IW for the 10 stocks, against restricted parametric specifications.

There are obvious time-varying dynamics in all the mixture components that the restricted versions do not accommodate. The strongest effects are captured in the returns covariance, Λ (panels C) and RCov mixtures, V (panels D). Returns covariance scaling component has values around, but different than, the identity matrix, which is the specification of the parametric models MRG-N-W and MRG-N-IW. These are covariance shocks that the parametric models do not accommodate and capture. The RCov shocks are also captured by the mixture component which has values around the parametric specification V. The deviations between RCov and the conditional covariance are not constant over time. On top of these, the proposed MRG-DPM-MΛV framework captures observed asymmetries in the returns distribution, as seen in the mean vector swifts in panels B of Figure 2. These are effects that in the so far literature of joint returns and RCov flexible models (Jin and Maheu, 2016; Opschoor et al., 2018) are not taken into account.

Further analysis on the DP precision parameter α, which controls the number of clusters used, provides insights into the data distribution. Following closely Jensen and Maheu (2013) we use the Savage-Dickey (Dickey, 1971) density ratio to test the symmetric MRG-DPM-Λ models against the nested MRG-t and MRG-N models, through the DP precision, α.[13] The posterior density of $\mathrm{u} \equiv \alpha/(1+\alpha)$ are in Figure 3. When $\mathrm{u} \to 0$, then the mixture becomes the Gaussian case and when $\mathrm{u} \to 1$, each data point has its own cluster and the mixture becomes the Student's-t case. For both datasets, almost all the probability mass of u is more than 0.1 and less than 0.8, supporting the nonparametric returns mixture against the parametric models. The higher the data dimension, the more concentrated u is around its mode and further away from the Gaussian assumption.

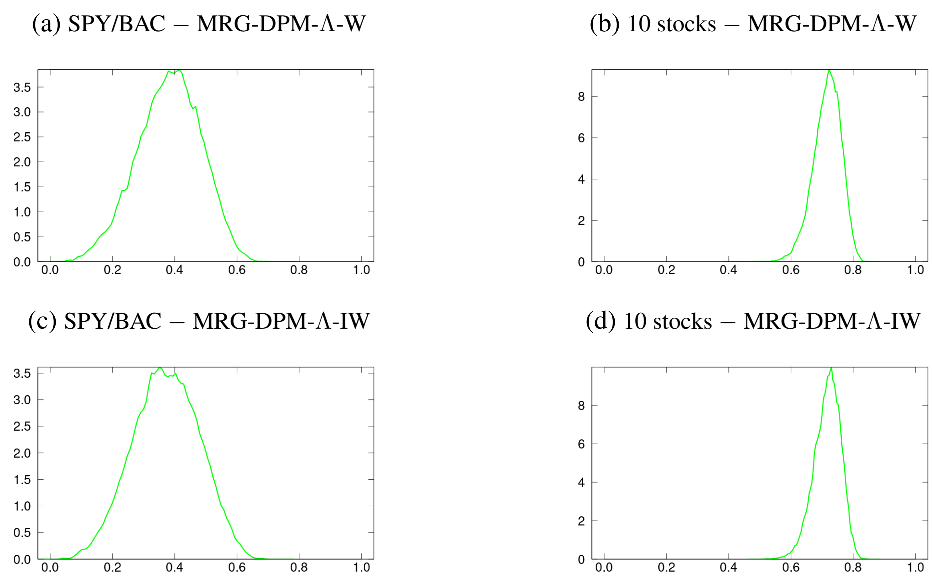


Figure 3: Posterior density of $\mathrm{u} \equiv \alpha/(1+\alpha)$, from the MRG-DPM-$\Lambda$ models. The models are equivalent to MRG-N when $\mathrm{u} = 0$ and equivalent to MRG-t when $\mathrm{u} = 1$.

4.4 Out-of-sample forecasts

Comparison of density forecasts for returns and RCov are presented in Tables 4 and 5, respectively. These report the cumulative log-predictive likelihood for daily ($h = 1$), weekly ($h = 5$) and monthly ($h = 22$) forecast horizons. Log-predictive Bayes factors (Kass and Raftery, 1995) can be formed by subtracting two entries in a column. The forecasts are computed with recursive data expanding posterior estimations for $\tau = 1000$ out-of-sample daily observations from January 12th, 2006 to December 31st, 2009.

From the results presented in Table 4, for both datasets and for all the forecast horizons, the semiparametric models outperform the parametric specifications in returns density forecasting. This indicates time-varying

[13] See also Jensen and Maheu (2014) and Zaharieva et al. (2020).

Table 4: Returns density forecasts.

	SPY/BAC			10 stocks		
	$h=1$	$h=5$	$h=22$	$h=1$	$h=5$	$h=22$
MG-N	−2892.77	−2977.66	−3192.29	−15506.99	−15667.23	−16127.30
MG-t	−2805.70	−2857.53	−2975.95	−14959.10	−15023.16	−15540.49
MG-DPM	−2809.78	−2862.05	−3007.46	−14981.20	−15081.93	−15446.87
MRG-N-W	−2790.07	−2850.24	−2996.01	−14884.22	−15159.45	−15767.67
MRG-N-IW	−2790.02	−2849.39	−3037.67	−14909.86	−15305.58	−16675.90
MRG-t-W	−2750.82	−2817.66	−2976.57	−14655.95	−14890.33	−15539.36
MRG-t-IW	−2752.03	−2823.20	−3008.45	−14639.15	−14849.86	−15443.57
MRG-DPM-Λ-W	−2744.30	−2819.84	−2990.46	−14617.92	−14863.45	−15375.16
MRG-DPM-Λ-IW	−2746.99	−2821.71	−3031.04	−14624.01	−14857.44	−15557.38
MRG-DPM-MΛ-W	−2741.56	−2808.07	−2972.79	**−14598.37**	−14838.14	−15365.78
MRG-DPM-MΛ-IW	−2744.08	−2817.27	−3028.67	−14601.91	−14843.51	−15575.12
MRG-DPM-MΛV-W	**−2741.33**	**−2806.71**	**−2967.25**	−14604.89	−14844.85	**−15356.69**
MRG-DPM-MΛV-IW	−2741.99	−2810.48	−3016.45	−14609.21	**−14830.77**	−15512.95

Notes: This table reports the returns cumulative log-predictive likelihood from (18) for daily ($h = 1$), weekly ($h = 5$) and monthly ($h = 22$) forecast horizons. In the top panel are the parametric multivariate GARCH models, in the middle are the parametric multivariate realized GARCH models and in the bottom panel are the semiparametric ones. Bold indicates the largest value in a column. Forecasting period: 12/01/2006 - 31/12/2009, 1000 out-of-sample forecasts.

higher conditional moments that the parametric specifications cannot accommodate. The most competitive model is MRG-DPM-MΛV-W which gives the best forecasts in four out of six occasions. MRG-DPM-MΛ-W follows closely indicating minor gains of including RCov in the mixture. In general, the models with Wishart RCov modeling outperform their inverse-Wishart alternatives in forecasting returns density. The symmetric returns distributional assumption is too restrictive since the flexible asymmetric models outperform the symmetric parametric and semiparametric specifications. For instance, in the 10 stocks dataset, the MRG-DPM-MΛ-W against the MRG-N-W model has log-Bayes factors values of 285.85 ($h = 1$), 321.31 ($h = 5$) and 401.89 ($h = 22$), and against the MRG-t-W model has log-Bayes factors values of 57.58 ($h = 1$), 52.19 ($h = 5$) and 173.58 ($h = 22$).

Table 5: RCov density forecasts.

	SPY/BAC			10 stocks		
	$h=1$	$h=5$	$h=22$	$h=1$	$h=5$	$h=22$
MRG-N-W	−1771.44	−2400.62	−3098.94	−8822.93	−18170.86	−39815.97
MRG-N-IW	−1472.96	−2181.00	−2978.02	−5705.42	−13390.82	−48802.80
MRG-t-W	−1765.92	−2342.06	−3054.69	−8794.42	−19101.61	−52711.98
MRG-t-IW	−1462.11	−2157.66	−2983.50	−6055.13	−12112.67	−29003.96
MRG-DPM-Λ-W	−1763.29	−2388.65	−3295.46	−8573.97	−17551.01	−39905.00
MRG-DPM-Λ-IW	−1453.91	−2171.52	−3070.55	−6018.42	−12100.25	−28259.70
MRG-DPM-MΛ-W	−1752.18	−2310.61	−3156.96	−8575.77	−17726.47	−40242.11
MRG-DPM-MΛ-IW	−1447.16	−2147.56	−3033.82	−6020.35	−12104.38	−28433.59
MRG-DPM-MΛV-W	−1492.06	−2158.64	**−2977.32**	−4572.98	−12755.76	−33404.41
MRG-DPM-MΛV-IW	**−1397.05**	**−2124.67**	−3008.09	**−4073.62**	**−9954.56**	**−26175.34**

Notes: This table reports the RCov cumulative log-predictive likelihood from (19) for daily ($h = 1$), weekly ($h = 5$) and monthly ($h = 22$) forecast horizons. In the top panel are the parametric multivariate realized GARCH models and in the bottom panel are the semiparametric ones. Bold indicates the largest value in a column. Forecasting period: 12/01/2006 - 31/12/2009, 1000 out-of-sample forecasts.

Figure 4 shows model comparisons for returns daily density forecasts. In panels A, the semiparametric MRG-

(a) SPY/BAC (b) 10 stocks

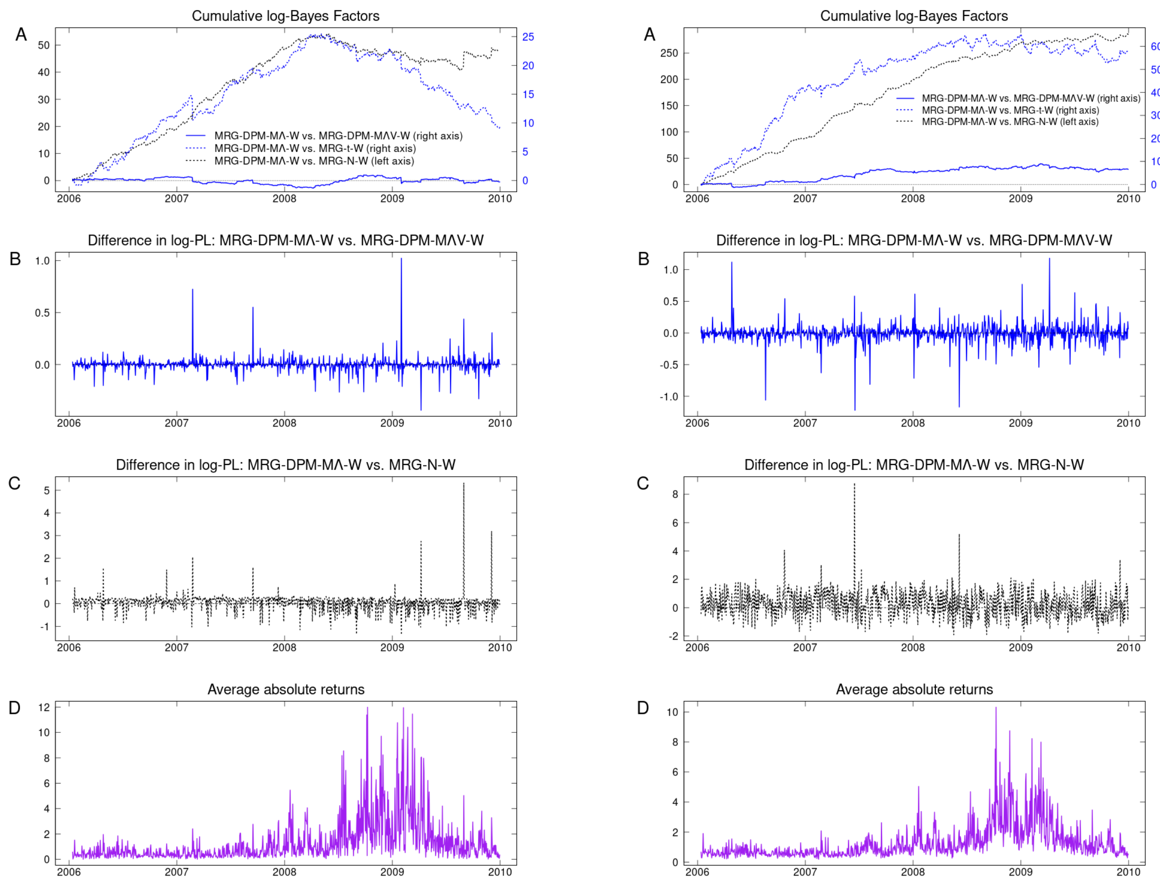


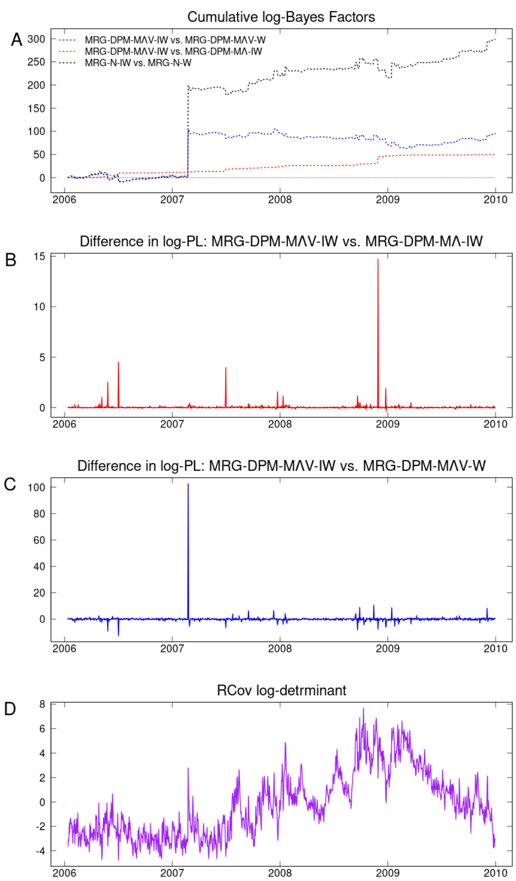
Figure 4: The plots display model comparisons for returns daily ($h = 1$) density forecasts over time. Panels A: Cumulative log-Bayes Factors. Panels B: Difference in log-predictive likelihood of MRG-DPM-MΛ-W vs. MRG-DPM-MΛV-W (mixture vs. non-mixture RCov modeling). Panels C: Difference in log-predictive likelihood of MRG-DPM-MΛ-W vs. MRG-N-W (semiparametric returns modeling vs. multivariate Gaussian assumption). Panels D: Average absolute returns.

DPM-MΛ-W model has ongoing gains against the parametric normal and Student's-t until the first quarter of 2008. After that, the gains level-off and for the SPY-BAC dataset, the tails-focused MRG-t-W model becomes competitive. For the 10 stocks dataset, the inclusion of RCov in the mixture does not improve returns daily density forecasts. For the SPY/BAC it offers marginal gains. In panels B it is shown that RCov mixture provides better returns predictive likelihoods in a few occasions of spiked returns. Compared to the standard parametric MRG-N-W, the semiparametric MRG-DPM-MΛ-W has most of the days higher predictive likelihood and it strongly outperforms in occasions of volatile days during low volatility periods.

The results in Table 5 show that the semiparametric RCov modeling outperforms the parametric specifications. This indicates time-varying higher conditional moments that a parametric specification cannot accommodate. The best model in five out of six occasions is MRG-DPM-MΛV-IW. The inverse-Wishart RCov modeling outperforms the Wishart alternatives in all the daily and weekly forecasts, and in most of the monthly ones. The semiparametric modeling of returns with parametric RCov, outperforms the fully parametric models. For instance, in the 10 stocks dataset, the MRG-DPM-MΛ-IW against the MRG-t-IW has log-Bayes Factors with values of 34.78 ($h = 1$), 8.29 ($h = 5$) and 570.37 ($h = 22$).

Figure 5 shows model comparisons for RCov daily density forecasts. The semiparametric RCov modeling outperforms the parametric specification almost always throughout the forecasting period, in panels B. It provides massive density forecast gains, especially for extreme RCov values, of either high or low volatility. As for the distributional choice, in panels C, the inverse-Wishart, since it accommodates fat-tails, it gives higher predictive likelihood scores than the Wishart in volatile days, such as February 27, 2007. This is also evident in the Gaussian MRG-N models comparison, in panels A. The Wishart distribution has large density forecast gains in the low volatility days, such as the Boxing Day of 2008.

(a) SPY/BAC



(b) 10 stocks

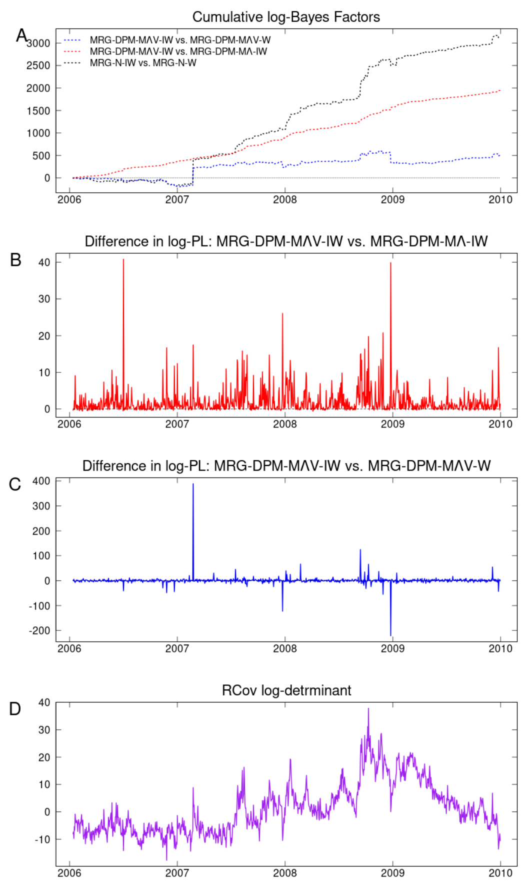


Figure 5: The plots display model comparisons for RCov daily ($h = 1$) density forecasts over time. Panels A: Cumulative log-Bayes Factors. Panels B: Difference in log-predictive likelihood of MRG-DPM-MΛV-IW vs. MRG-DPM-MΛ-IW (mixture vs. non-mixture RCov modeling). Panels C: Difference in log-predictive likelihood of MRG-DPM-MΛV-IW vs. MRG-DPM-MΛV-W (inverse-Wishart vs. Wishart mixture RCov modeling). Panels D: RCov log-determinant.

Table 6 reports the conditional returns covariance mean squared forecast error calculated from (20). The semi-parametric mixture in returns makes their conditional covariance closer to RCov. The best models are the semi-parametric ones that have the inverse-Wishart RCov assumption, namely MRG-DPM-MΛ-IW, MRG-DPM-Λ-IW, followed by MRG-DPM-MΛV-IW. They outperform their Wishart alternatives and the parametric specifications. Including RCov in the mixture does not offer major improvements.

4.5 Global minimum variance portfolios

In this section the models are tested on the economic gains they provide through portfolio optimization. We focus on the Global Minimum Variance portfolio (GMV) problem which is one of the most popular multivariate model applications.[14] A model is preferred if it produces *out-of-sample* portfolios with lower realized variance (see Engle and Colacito, 2006).

[14]It is also used from Engle and Kelly (2012), Jin and Maheu (2013), Jin and Maheu (2016), Opschoor et al. (2018), Archakov et al. (2019), Bauwens and Xu (2023) among many others.

Table 6: Covariance mean squared forecast error.

	SPY/BAC			10 stocks		
	$h=1$	$h=5$	$h=22$	$h=1$	$h=5$	$h=22$
MG-N	8.530	23.021	56.675	24.370	62.525	149.199
MG-t	7.799	21.359	53.623	24.921	63.703	151.243
MG-DPM	7.774	21.413	54.313	22.695	58.948	141.556
MRG-N-W	5.755	18.349	52.652	17.211	48.271	125.884
MRG-N-IW	5.752	18.018	49.367	17.214	47.966	128.520
MRG-t-W	5.770	18.434	53.220	17.101	47.886	126.451
MRG-t-IW	5.764	17.978	48.868	17.260	47.704	125.209
MRG-DPM-Λ-W	5.742	18.184	51.741	17.313	48.438	126.407
MRG-DPM-Λ-IW	5.682	**17.734**	**48.182**	17.051	47.389	124.646
MRG-DPM-MΛ-W	5.769	18.213	51.746	17.202	48.264	126.091
MRG-DPM-MΛ-IW	**5.668**	17.742	48.211	**16.988**	**47.327**	124.663
MRG-DPM-MΛV-W	5.773	18.287	52.319	17.204	48.233	125.840
MRG-DPM-MΛV-IW	5.683	17.755	48.280	17.021	47.386	**124.386**

Notes: This table reports the covariance mean squared forecast error (MSFE) from (20) for daily ($h = 1$), weekly ($h = 5$) and monthly ($h = 22$) forecast horizons. The MSFE is calculated with the use of matrix (Frobenius) norm as: MSFE $= \tau^{-1} \sum_{t=T_0-h}^{T-h} \left|\left|\text{RCov}_{t+h} - \text{Cov}(r_{t+h}|\mathcal{I}_t)\right|\right|$. Bold indicates the lowest value in a column. Forecasting period: 12/01/2006 - 31/12/2009, 1000 out-of-sample forecasts.

Given the information set $\mathcal{I}_t$, available at time t, for n risky assets, investors, who follow a dynamic GMV portfolio strategy, for the forecast horizon h, are interested in accurate forecasts of the covariance matrix, $\text{Cov}(r_{t+h}|\mathcal{I}_t)$. Based on that, they can adjust their GMV portfolio weights for time $t+h$ by solving the following problem

$$\min_{\omega_{t+h|t} \in \mathbb{R}^n} \left\{ \sigma^{2p}_{t+h|t} = (\omega_{t+h|t})' \text{Cov}(r_{t+h}|\mathcal{I}_t) \omega_{t+h|t} \right\}$$
$$\text{s.t.} \quad \iota' \omega_{t+h|t} = 1.$$

where $\omega_{t+h|t}$ is the portfolio weights vector constructed for period $t+h$ given information $\mathcal{I}_t$. The solution to the problem is $\omega^{\text{GMV}}_{t+h|t} = \frac{\text{Cov}(r_{t+h}|\mathcal{I}_t)^{-1}\iota}{(\iota)'\text{Cov}(r_{t+h}|\mathcal{I}_t)^{-1}\iota}$. For the conditional returns covariance, $\text{Cov}(r_{t+h}|\mathcal{I}_t)$, from each model see Section 3.

For the forecast days τ, the sample portfolio returns variance, $\overline{\sigma}^2_p$, of the realized portfolio returns $\{r_{p,t}\}_{t=T-\tau+1}^{T}$, from each model is calculated as $\overline{\sigma}^2_p = \frac{1}{\tau} \sum_{t=T-\tau+1}^{T} \left(r_{p,t} - \bar{r}_p\right)^2$, where $\bar{r}_p$ is the sample mean of $\{r_{p,t}\}_{t=T-\tau+1}^{T}$.

In Table 7, the realized sample standard deviations of the GMV portfolio returns are presented. Capturing the transitory covariance dynamics, with the semiparametric modeling, is important in daily ($h = 1$) out-of-sample portfolio optimization. The semiparametric models perform better than the parametric ones only for the shortest horizon. The best-performing models are the MRG-DPM-MΛ-IW for SPY/BAC and the MRG-DPM-Λ-IW for the 10-stock dataset. Comparing the semiparametric models, including RCov in the mixture improves portfolio optimization only for the monthly forecast horizon. For the daily forecast horizon, the inverse-Wishart RCov models tend to perform better than their Wishart alternatives. For longer forecasting horizons, the covariance persistence becomes more important than the transitory covariance dynamics. The parametric covariance structures, due to their smoothness and stability, stand out in weekly and monthly portfolio construction. The MRG-N-W is the best at the weekly ($h = 5$) horizon for both datasets and at the monthly horizon for SPY/BAC. The parametric MGARCH models are competitive at the monthly horizon ($h = 22$) and perform best for the 10 stocks dataset.

Table 7: Sample standard deviation of realized GMV portfolio returns.

	SPY/BAC			10 stocks		
	$h=1$	$h=5$	$h=22$	$h=1$	$h=5$	$h=22$
MG-N	1.0792	1.0905	1.1264	0.9715	0.9732	0.9931
MG-t	1.0796	1.0917	1.1381	0.9695	0.9707	**0.9896**
MG-DPM	1.0825	1.0964	1.1612	0.9693	0.9799	0.9900
MRG-N-W	1.0714	**1.0716**	**1.0993**	0.9180	**0.9227**	1.0303
MRG-N-IW	1.0716	1.0808	1.2192	0.9172	0.9272	1.1640
MRG-t-W	1.0725	1.0717	1.1042	0.9249	0.9255	1.1400
MRG-t-IW	1.0720	1.0800	1.2249	0.9172	0.9261	1.1358
MRG-DPM-Λ-W	1.0648	1.0735	1.1226	0.9108	0.9364	1.0364
MRG-DPM-Λ-IW	1.0660	1.0725	1.2227	**0.9080**	0.9364	1.0991
MRG-DPM-MΛ-W	1.0654	1.0719	1.1236	0.9089	0.9366	1.0349
MRG-DPM-MΛ-IW	**1.0642**	1.0727	1.2210	0.9093	0.9357	1.0999
MRG-DPM-MΛV-W	1.0655	1.0719	1.1169	0.9096	0.9369	1.0348
MRG-DPM-MΛV-IW	1.0646	1.0725	1.2150	0.9101	0.9338	1.0903

Notes: This table reports the sample standard deviation, $\overline{\sigma}_p$, of the realized GMV portfolio returns for daily $(h = 1)$, weekly $(h = 5)$ and monthly $(h = 22)$ forecast horizons. Bold indicates the lowest value in a column. Forecasting period: 12/01/2006 - 31/12/2009, 1000 out-of-sample forecasts.

5 Conclusions

This paper extends the popular multivariate realized GARCH models to a Bayesian semiparametric framework. We use a countably infinite number of kernels with a Dirichlet prior to approximate the joint unknown density of both returns and RCov matrices. We test several restricted specifications and two distributional assumptions for realized covariance, Wishart and inverse-Wishart.

The empirical application to the dataset of Noureldin et al. (2012) draws many useful forecasting results. First of all, a parametric multivariate normal or Student's-t assumption for returns is restrictive. These specifications are outperformed in returns density forecasts by the new proposed models. The models that capture returns distributional asymmetry and fat-tails perform the best. The Wishart RCov modeling assumption gives better returns density forecasts than the fat-tailed inverse-Wishart alternative. The latter one stands out in RCov density forecasts and the infinite mixture of inverse-Wisharts is strongly preferred. Finally, the new framework provides improvements in covariance point forecasts and economic gains in daily portfolio optimization.

While this work focuses on a multivariate realized GARCH setting, the Bayesian semiparametric extension can be used in a HEAVY setting as well. Regarding the semiparametric modeling, the independent DPM model can be substituted with an infinite hidden Markovian mixture to capture latent regime switches. Both are currently under development.

Acknowledgements

The author thanks the Departmental Editor Bartosz Kurek and two anonymous referees for constructive comments. The author also thanks John M. Maheu, Ronald J. Balvers, Narat Charupat, Arnaud Dufays, Chenxing Li and Ali Gencay Özbekler for helpful comments and suggestions, and the participants of: DeGroote School of Business seminars, Finance and Business Analytics conference, the 1st China Forum of Bayesian Econometrics, and the 14th RCEA Bayesian Econometrics workshop for comments.

Declaration statement

The author declares no conflict of interest.

Funding

No funding was received.

References

Andersen, T. G. and Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, pages 885–905.

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

Archakov, I., Hansen, P. R., and Lunde, A. (2019). A multivariate realized GARCH model. *Working Paper*.

Ascolani, F., Lijoi, A., Rebaudo, G., and Zanella, G. (2023). Clustering consistency with dirichlet process mixtures. *Biometrika*, 110(2):551–558.

Baba, Y., Engle, R. F., Kraft, D. F., and Kroner, K. F. (1990). Multivariate simultaneous generalized ARCH. *Manuscript, University of California, San Diego, Department of Economics*.

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

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2008). Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica*, 76(6):1481–1536.

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

Barndorff-Nielsen, O. E. and Shephard, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(2):253–280.

Barndorff-Nielsen, O. E. and Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, 2(1):1–37.

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

Bauwens, L. and Xu, Y. (2023). DCC-and DECO-HEAVY: Multivariate GARCH models based on realized variances and correlations. *International Journal of Forecasting*, 39(2):938–955.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327.

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.

Braione, M. (2016). A time-varying long run HEAVY model. *Statistics & Probability Letters*, 119:36–44.

Diamantopoulos, K. and Vrontos, I. D. (2010). A student-t full factor multivariate GARCH model. *Computational Economics*, 35(1):63.

Dickey, J. M. (1971). The weighted likelihood ratio, linear hypotheses on normal location parameters. *The Annals of Mathematical Statistics*, pages 204–223.

Ding, Z. and Engle, R. F. (2001). Large scale conditional covariance matrix modeling, estimation and testing.

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. and Colacito, R. (2006). Testing and valuing dynamic correlations for asset allocation. *Journal of Business & Economic Statistics*, 24(2):238–253.

Engle, R. and Kelly, B. (2012). Dynamic equicorrelation. *Journal of Business & Economic Statistics*, 30(2):212–228.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007.

Engle, R. F. and Kroner, K. F. (1995). Multivariate simultaneous generalized ARCH. *Econometric Theory*, pages 122–150.

Escobar, M. D. and West, M. (1995). Bayesian density estimation and inference using mixtures. *Journal of the American Statistical Association*, 90(430):577–588.

Ferguson, T. S. (1973). A Bayesian analysis of some nonparametric problems. *The Annals of Statistics*, pages 209–230.

Geweke, J. (1994). Bayesian comparison of econometric models. Technical report, Working Paper, Federal Reserve Bank of Minneapolis, Minnesota.

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.

Gorgi, P., Hansen, P. R., Janus, P., and Koopman, S. J. (2019). Realized Wishart-GARCH: A score-driven multi-asset volatility model. *Journal of Financial Econometrics*, 17(1):1–32.

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

Hansen, P. R., Huang, Z., and Shek, H. H. (2012). Realized GARCH: a joint model for returns and realized measures of volatility. *Journal of Applied Econometrics*, 27(6):877–906.

Hansen, P. R., Lunde, A., and Voev, V. (2014). Realized beta GARCH: A multivariate GARCH model with realized measures of volatility. *Journal of Applied Econometrics*, 29(5):774–799.

Heber, G., Lunde, A., Shephard, N., and Sheppard, K. (2009). Oxford-Man Institute's Realized Library, version 0.1.

Jensen, M. J. and Maheu, J. M. (2013). Bayesian semiparametric multivariate GARCH modeling. *Journal of Econometrics*, 176(1):3–17.

Jensen, M. J. and Maheu, J. M. (2014). Estimating a semiparametric asymmetric stochastic volatility model with a Dirichlet process mixture. *Journal of Econometrics*, 178:523–538.

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

Jin, X. and Maheu, J. M. (2016). Bayesian semiparametric modeling of realized covariance matrices. *Journal of Econometrics*, 192(1):19–39.

Jin, X., Maheu, J. M., and Yang, Q. (2019). Bayesian parametric and semiparametric factor models for large realized covariance matrices. *Journal of Applied Econometrics*, 34(5):641–660.

Kalli, M., Griffin, J. E., and Walker, S. G. (2011). Slice sampling mixture models. *Statistics and Computing*, 21(1):93–105.

Kass, R. E. and Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90(430):773–795.

Maheu, J. M. and Shamsi Zamenjani, A. (2021). Nonparametric dynamic conditional beta. *Journal of Financial Econometrics*, 19(4):583–613.

Noureldin, D., Shephard, N., and Sheppard, K. (2012). Multivariate high-frequency-based volatility (HEAVY) models. *Journal of Applied Econometrics*, 27(6):907–933.

Opschoor, A., Janus, P., Lucas, A., and Van Dijk, D. (2018). New HEAVY models for fat-tailed realized covariances and returns. *Journal of Business & Economic Statistics*, 36(4):643–657.

Richardson, M. and Smith, T. (1993). A test for multivariate normality in stock returns. *Journal of Business*, pages 295–321.

Sethuraman, J. (1994). A constructive definition of dirichlet priors. *Statistica Sinica*, pages 639–650.

Sheppard, K. and Xu, W. (2019). Factor high-frequency-based volatility (HEAVY) models. *Journal of Financial Econometrics*, 17(1):33–65.

Shirota, S., Omori, Y., Lopes, H. F., and Piao, H. (2017). Cholesky realized stochastic volatility model. *Econometrics and Statistics*, 3:34–59.

Silvennoinen, A. and Teräsvirta, T. (2009). Multivariate garch models. In *Handbook of Financial Time Series*, pages 201–229. Springer.

Virbickaite, A., Ausín, M. C., and Galeano, P. (2015). Bayesian inference methods for univariate and multivariate GARCH models: A survey. *Journal of Economic Surveys*, 29(1):76–96.

Walker, S. G. (2007). Sampling the dirichlet mixture model with slices. *Communications in Statistics — Simulation and Computation®*, 36(1):45–54.

Yamauchi, Y. and Omori, Y. (2020). Multivariate stochastic volatility model with realized volatilities and pairwise realized correlations. *Journal of Business & Economic Statistics*, 38(4):839–855.

Zaharieva, M. D., Trede, M., and Wilfling, B. (2020). Bayesian semiparametric multivariate stochastic volatility with application. *Econometric Reviews*, 39(9):947–970.

Zhang, L., Mykland, P. A., and Aït-Sahalia, Y. (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association*, 100(472):1394–1411.

Appendix

Sampling steps

The following algorithm contains the steps for estimating the MRG-DPM-MΛV-IW model in Section 2.2.2. The algorithm is a hybrid of Gibbs and Metropolis-Hastings draws.

1. Sample $\mu_{1:k}, \Lambda_{1:k}, \mathrm{V}_{1:k} | r_{1:T}, \mathrm{RCov}_{1:T}, H_{1:T}, s_{1:T}$.

   The state-dependent mean vectors and covariance matrices $\mu_{1:k}, \Lambda_{1:k} | r_{1:T}, H_{1:T}, s_{1:T}$ from linear model conjugate priors with Gibbs sampling as

$$\Lambda_j | r_{1:T}, H_{1:T}, s_{1:T}, \mu_j \sim \mathrm{IW}\left(\lambda_0 + n_j,\ V_0 + \sum_{t:s_t=j} \left(H_t^{-1/2}(r_t - \mu_j)\right)\left(H_t^{-1/2}(r_t - \mu_j)\right)'\right),$$

$$\mu_j | r_{1:T}, H_{1:T}, s_{1:T}, \Lambda_j \sim \mathrm{N}(\tilde{\mu}, \tilde{V}),$$

$$\text{with} \quad \tilde{V} = \left(M_0^{-1} + \sum_{t:s_t=j} (H_t^{-1/2})' \Lambda_j^{-1} H_t^{-1/2}\right)^{-1},$$

$$\tilde{\mu} = \tilde{V}\left(\sum_{t:s_t=j} (H_t^{-1/2})' \Lambda_j^{-1} H_t^{-1/2} r_t + M_0^{-1}\mu_0\right),\ n_j = \sum_{i=1}^{T} \mathbf{1}\{s_t = j\} \text{ and } j = 1, ..., k.$$

   The scale matrices $\mathrm{V}_{1:k} | \nu, \mathrm{RCov}_{1:T}, H_{1:T}, s_{1:T}$ are drawn from a conjugate Wishart prior, with a Gibbs draw as

$$\mathrm{V}_j | \nu, \mathrm{RCov}_{1:T}, H_{1:T}, s_{1:T} \sim \mathrm{W}\left(\mathrm{v}_0 + n_j\nu,\ \left[V_0^{-1} + (\nu - n - 1)\sum_{t:s_t=j}\left(H_t^{1/2}\right)' \mathrm{RCov}_t^{-1} H_t^{1/2}\right]^{-1}\right),$$

   with ν being the inverse-Wishart degrees of freedom parameter from (6).

2. Update the mixture weights in $w_{1:k}|s_{1:T},\alpha$ with a stick-breaking process as

$$v_j|s_{1:T},\alpha \sim \mathrm{B}\left(1+\sum_{t=1}^{T}\mathbf{1}\left\{s_t=j\right\},\alpha+\sum_{t=1}^{T}\mathbf{1}\{s_t>j\}\right),$$

$$w_1=v_1,\quad w_j=v_j\prod_{l=1}^{j-1}(1-v_l),\quad j=2,...,k.$$

3. Update the slice vector $u_{1:T}|w_{1:k},s_{1:T}$ from a uniform draw as: $u_t|w_{1:k},s_{1:T}\sim \mathrm{U}(0,w_{s_t})$.

4. Update the number of mixture clusters k to the smallest positive integer that satisfies: $\sum_{j=1}^{k} w_j > 1-\min(u_{1:T})$. If new clusters are needed to satisfy the inequality, the mixing components are drawn from the base measure.

5. Sample the vector $s_{1:T}|r_{1:T},\mathrm{RCov}_{1:T},H_{1:T},\mu_{1:k},\Lambda_{1:k},\mathrm{V}_{1:k},w_{1:k},u_{1:T}$, from a multinominal distribution with probabilities

$$\begin{aligned}&p(s_t=j|r_{1:T},\mathrm{RCov}_{1:T},H_{1:T},\mu_{1:k},\Lambda_{1:k},\mathrm{V}_{1:k},w_{1:k},u_{1:T})\\&\qquad\propto \mathbf{1}\left\{u_t<w_j\right\}\mathrm{N}\left(r_t\middle|\mu_j,H_t^{1/2}\Lambda_j\left(H_t^{1/2}\right)'\right)\mathrm{IW}\left(\mathrm{RCov}_t\middle|\nu,(\nu-n-1)H_t^{1/2}\mathrm{V}_j\left(H_t^{1/2}\right)'\right),\end{aligned}$$

for $j=1,...,k$. The number of active clusters κ, can be calculated as the ones with at least one assigned data observation.

6. Draw the DPM precision parameter α with a gamma prior $\alpha\sim\Gamma(a_0,b_0)$ following the two steps algorithm of Escobar and West (1995):

   i. draw the random variable $\xi|\alpha,\kappa\sim \mathrm{B}(\alpha+1,T)$.
   ii. sample α from

$$\alpha|\xi\sim\pi_\xi\Gamma(a_0+\kappa,b_0-\log(\xi))+(1-\pi_\xi)\Gamma(a_0+\kappa-1,b_0-\log(\xi)),$$

   with $\frac{\pi_\xi}{1-\pi_\xi}=\frac{a_0+\kappa-1}{T(b_0-\log(\xi))}$.

7. Parameters in $\theta=\{a,b,c,\nu\}$ have the following conditional posterior

$$\begin{aligned}p(\theta|r_{1:T},\mathrm{RCov}_{1:T},\mu_{1:k},\Lambda_{1:k},\mathrm{V}_{1:k},s_{1:k})\propto p(\theta)\prod_{t=1}^{T}\mathrm{N}\left(r_t\middle|\mu_{s_t},H_t^{1/2}\Lambda_{s_t}\left(H_t^{1/2}\right)'\right)\\\times\mathrm{IW}\left(\mathrm{RCov}_t\middle|\nu,(\nu-n-1)H_t^{1/2}\mathrm{V}_{s_t}\left(H_t^{1/2}\right)'\right),\end{aligned}\tag{21}$$

which does not have a standard form. A random-walk MH algorithm is used to take a new draw $\theta^{(i)}$ from the above posterior, with proposal θ' which is from $h(\theta')\sim\mathrm{N}(\theta^{(i-1)},\hat{V}_h)$, with $\hat{V}_h$ the inverse Hessian matrix evaluated at the posterior mode $\hat{\theta}$. $\hat{V}_h$ is computed once at the beginning of estimation. The draw is accepted with probability

$$\min\left\{p(\theta'|\mathcal{I}_T)/p(\theta^{(i-1)}|\mathcal{I}_T),1\right\},$$

where $p(.|\mathcal{I}_T)$ is the posterior in (21). We set the DPM base measure is set such as $\mathbb{E}(\mu_j)=0$, $\mathbb{E}(\Lambda_j)=I_n$, $\mathbb{E}(\mathrm{V}_j)=I_n$, $j=1,2,...$, and we assume that the conditional covariance matrix is stationary and equal to the sample covariance matrix, $\mathbb{E}(H_1)=...=\mathbb{E}(H_T)=\bar{\Sigma}$. For every $\theta^{(i)}$, $\Omega^{(i)}$ is calculated as in (14). Posterior draws that do not result in positive definite $\Omega^{(i)}$ and $H_{1:T}^{(i)}$ are rejected. This rejection rate in our empirical applications is less than 1%. The acceptance rate in the MH algorithm is around $0.35-0.4$.

To estimate the MRG-DPM-MΛV-W model the same algorithm is used with a few alterations. In step 1, $\mu_{1:k}$ and $\Lambda_{1:k}$ are drawn from the same posteriors while the scale matrices $\mathrm{V}_{1:k}|\nu, \mathrm{RCov}_{1:T}, H_{1:T}, s_{1:T}$ are drawn from a conjugate inverse-Wishart prior, with a Gibbs draw as

$$\mathrm{V}_j|\nu, \mathrm{RCov}_{1:T}, H_{1:T}, s_{1:T} \sim \mathrm{IW}\left(\nu_0 + n_j\mathrm{v},\ V_0 + \mathrm{v}\sum_{t:s_t=j} H_t^{-1/2}\mathrm{RCov}_t\left(H_t^{-1/2}\right)'\right),$$

with v being the Wishart degrees of freedom parameter from (10).

In step 5, the vector $s_{1:T}|r_{1:T}, \mathrm{RCov}_{1:T}, H_{1:T}, \mu_{1:k}, \Lambda_{1:k}, \mathrm{V}_{1:k}, w_{1:k}, u_{1:T}$, is drawn from a multinominal distribution with probabilities

$$\begin{aligned} &p(s_t = j|r_{1:T}, \mathrm{RCov}_{1:T}, H_{1:T}, \mu_{1:k}, \Lambda_{1:k}, \mathrm{V}_{1:k}, w_{1:k}, u_{1:T}) \\ &\qquad \propto \mathbf{1}\{u_t < w_j\}\,\mathrm{N}\left(r_t\middle|\mu_j, H_t^{1/2}\Lambda_j\left(H_t^{1/2}\right)'\right)\mathrm{W}\left(\mathrm{RCov}_t\middle|\mathrm{v}, \mathrm{v}^{-1}H_t^{1/2}\mathrm{V}_j\left(H_t^{1/2}\right)'\right), \end{aligned}$$

for $j = 1, ..., k$.

In step 7, parameters in $\theta = \{a, b, c, \mathrm{v}\}$ have the following conditional posterior

$$\begin{aligned} p(\theta|r_{1:T}, \mathrm{RCov}_{1:T}, \mu_{1:k}, \Lambda_{1:k}, \mathrm{V}_{1:k}, s_{1:k}) \propto p(\theta)\prod_{t=1}^{T}\ &\mathrm{N}\left(r_t\middle|\mu_{s_t}, H_t^{1/2}\Lambda_{s_t}\left(H_t^{1/2}\right)'\right) \\ \times\ &\mathrm{W}\left(\mathrm{RCov}_t\middle|\mathrm{v}, \mathrm{v}^{-1}H_t^{1/2}\mathrm{V}_{s_t}\left(H_t^{1/2}\right)'\right), \end{aligned}$$

and a random-walk MH sampler is used to sample from it.

For the MRG-DPM-MΛ models, the above algorithms are used and parametric matrices V are samples from the same posteriors as V_js by setting n_j equal to the number of data observations T and for $t = 1, ..., T$. For the MRG-DPM-Λ models is imposed the restriction of $\mu_j = 0_n, \forall j$.

For the parametric models MRG-t-W and MRG-t-IW the parameters in vector $\theta = \{a, b, c, \zeta, \nu\}$ and V have the following posterior density

$$p(\theta, V|\mathcal{I}_T) \propto p(\theta)p(V)\prod_{t=1}^{T}\ \mathrm{t}(r_t|0, H_t, \zeta)\ \mathrm{F}\left(\mathrm{RCov}_t\middle|H_t, V, \nu\right),$$

with F(.) being the p.d.f. of either Wishart or inverse-Wishart. To draw from it we employ the following two-step MCMC algorithm:

1. sample $\theta^{(i)}|\mathcal{I}_T, V^{(i)}$ with a random-walk MH algorithm with proposal $h(\theta') \sim \mathrm{N}(\theta^{(i-1)}, \hat{V}_h)$. For every $\theta^{(i)}$, $\Omega^{(i)}$ is targeted as $\Omega^{(i)} = \bar{\Sigma} \odot \left(\iota\iota' - \frac{\zeta^{(i)}}{\zeta^{(i)}-2}a^{(i)} - b^{(i)} - \frac{\zeta^{(i)}}{\zeta^{(i)}-2}c^{(i)}\right)$, by assuming covariance stationarity, $\mathbb{E}(H_t) = \bar{\Sigma}$, and $\mathbb{E}(V) = I_n$. $\hat{V}_h$ the inverse Hessian matrix evaluated at the posterior mode $\hat{\theta}$. $\hat{V}_h$ is computed once at the beginning of estimation. Only posterior draws that result in positive definite $\Omega^{(i)}$ and $H_{1:T}^{(i)}$ are accepted.
2. sample $V^{(i)}|\theta^{(i)}, \mathrm{RCov}_{1:T}, H_{1:T}$ with a Gibbs sampler from (inverse-)Wishart conjugate priors.

For the MRG-N models, $\theta = \{a, b, c, \nu\}$, and the above algorithm is used, with $\frac{\zeta}{\zeta-2} \equiv 1$, to sample from $p(\theta, V|\mathcal{I}_T) \propto p(\theta)p(V)\prod_{t=1}^{T}\ \mathrm{N}(r_t|0, H_t)\ \mathrm{F}\left(\mathrm{RCov}_t\middle|H_t, V, \nu\right)$.

Summary statistics and posterior estimation tables.

SPY/BAC

	Sample returns		Sample covariance		RCov sample mean	
	Skewness	Kurtosis	SPY	BAC	SPY	BAC
SPY	−0.12	9.67	1.10		1.12	
BAC	0.33	21.81	1.62	5.71	1.46	5.46

10 stocks

	Sample returns		Sample returns covariance										Mean of RCov matrices									
	Skewness	Kurtosis	BAC	JPM	IBM	MSFT	XOM	AA	AXP	DD	GE	KO	BAC	JPM	IBM	MSFT	XOM	AA	AXP	DD	GE	KO
BAC	0.33	21.81	5.46										5.71									
JPM	0.58	17.05	3.05	5.06									3.77	4.83								
IBM	0.01	6.32	1.23	1.33	1.93								1.21	1.54	1.69							
MSFT	0.25	6.16	1.36	1.48	1.12	2.45							1.29	1.63	1.16	2.06						
XOM	−0.19	11.64	1.18	1.26	0.87	0.97	2.07						1.19	1.28	0.78	0.87	1.67					
AA	−0.68	9.89	1.97	1.94	1.19	1.30	1.37	4.94					2.40	2.30	1.29	1.32	1.38	4.44				
AXP	0.32	11.23	2.52	2.46	1.26	1.37	1.20	1.85	4.42				3.29	3.12	1.42	1.49	1.29	2.28	4.35			
DD	0.03	7.28	1.50	1.52	0.98	1.07	1.04	1.67	1.46	2.53			1.67	1.77	0.98	1.03	0.96	1.76	1.73	2.11		
GE	0.22	10.96	1.89	1.86	1.13	1.27	1.13	1.63	1.76	1.29	3.20		2.22	2.23	1.16	1.23	1.03	1.77	2.07	1.36	2.68	
KO	0.11	6.92	0.85	0.91	0.66	0.71	0.66	0.83	0.85	0.73	0.82	1.41	0.74	0.87	0.55	0.68	0.61	0.80	0.89	0.63	0.76	1.13

Table 8: Summary statistics. Sample period is 1/2/2001 - 31/12/2009 (2242 trading days). Returns sample means are close to zero.

Table 9: Posterior Means and 95% density interval of parametric matrix V from the Wishart models - 10 stocks.

	MRG-N-W		**MRG-t-W**		**MRG-DPM-Λ-W**		**MRG-DPM-MΛ-W**	
	Mean	95% D.I.	Mean	95% D.I.	Mean	95% D.I.	Mean	95% D.I.
$V_{1,1}$	0.941	[0.930, 0.954]	0.933	[0.918, 0.953]	0.937	[0.925, 0.950]	0.937	[0.924, 0.950]
$V_{1,2}$	−0.016	[−0.023, −0.008]	−0.015	[−0.023, −0.008]	−0.015	[−0.023, −0.008]	−0.015	[−0.023, −0.008]
$V_{1,3}$	0.020	[0.012, 0.028]	0.021	[0.012, 0.029]	0.021	[0.013, 0.029]	0.021	[0.013, 0.029]
$V_{1,4}$	0.014	[0.006, 0.022]	0.014	[0.006, 0.022]	0.014	[0.006, 0.022]	0.014	[0.006, 0.022]
$V_{1,5}$	0.011	[0.003, 0.019]	0.011	[0.003, 0.019]	0.011	[0.004, 0.019]	0.011	[0.003, 0.019]
$V_{1,6}$	−0.007	[−0.015, 0.001]	−0.007	[−0.014, 0.001]	−0.007	[−0.015, 0.001]	−0.007	[−0.015, 0.001]
$V_{1,7}$	−0.025	[−0.033, −0.017]	−0.025	[−0.033, −0.017]	−0.025	[−0.034, −0.017]	−0.025	[−0.034, −0.018]
$V_{1,8}$	0.013	[0.005,0.020]	0.013	[0.004, 0.021]	0.013	[0.005, 0.021]	0.013	[0.005, 0.021]
$V_{1,9}$	−0.003	[−0.011, 0.005]	−0.003	[−0.011, 0.005]	−0.003	[−0.011, 0.005]	−0.003	[−0.011, 0.005]
$V_{1,10}$	0.019	[0.011,0.027]	0.019	[0.011, 0.028]	0.020	[0.012, 0.028]	0.020	[0.012, 0.028]
$V_{2,2}$	1.028	[1.017, 1.040]	1.020	[1.006, 1.037]	1.026	[1.014, 1.038]	1.026	[1.014, 1.037]
$V_{2,3}$	−0.010	[−0.019, −0.002]	−0.010	[−0.019, −0.002]	−0.010	[−0.019, −0.002]	−0.010	[−0.019, −0.002]
$V_{2,4}$	−0.005	[−0.013, 0.003]	−0.005	[−0.013, 0.003]	−0.005	[−0.013, 0.003]	−0.005	[−0.013, 0.003]
$V_{2,5}$	0.003	[−0.005, 0.011]	0.003	[−0.005, 0.011]	0.003	[−0.005, 0.011]	0.003	[−0.006, 0.011]
$V_{2,6}$	−0.001	[−0.009, 0.007]	−0.001	[−0.009, 0.007]	−0.001	[−0.009, 0.007]	−0.001	[−0.009, 0.007]
$V_{2,7}$	−0.014	[−0.022, −0.005]	−0.014	[−0.022, −0.005]	−0.014	[−0.022, −0.005]	−0.014	[−0.022, −0.005]
$V_{2,8}$	−0.001	[−0.010, 0.007]	−0.001	[−0.010, 0.007]	−0.001	[−0.010, 0.007]	−0.001	[−0.010, 0.007]
$V_{2,9}$	−0.001	[−0.010, 0.007]	−0.001	[−0.010, 0.007]	−0.001	[−0.010, 0.007]	−0.001	[−0.010, 0.007]
$V_{2,10}$	0.003	[−0.005, 0.011]	0.003	[−0.005, 0.011]	0.003	[−0.005, 0.011]	0.003	[−0.005, 0.011]
$V_{3,3}$	1.032	[1.021, 1.044]	1.025	[1.010, 1.042]	1.031	[1.019, 1.043]	1.031	[1.019, 1.042]
$V_{3,4}$	−0.012	[−0.020, −0.003]	−0.012	[−0.020, −0.003]	−0.012	[−0.020, −0.004]	−0.012	[−0.020, −0.003]
$V_{3,5}$	−0.001	[−0.009, 0.007]	−0.001	[−0.009, 0.007]	−0.001	[−0.010, 0.007]	−0.001	[−0.009, 0.007]
$V_{3,6}$	−0.001	[−0.009, 0.007]	−0.001	[−0.009, 0.007]	−0.001	[−0.010, 0.007]	−0.001	[−0.010, 0.007]
$V_{3,7}$	0.001	[−0.007, 0.010]	0.001	[−0.007, 0.010]	0.001	[−0.007, 0.010]	0.001	[−0.007, 0.010]
$V_{3,8}$	−0.002	[−0.010, 0.006]	−0.002	[−0.011, 0.006]	−0.002	[−0.011, 0.006]	−0.002	[−0.011, 0.006]
$V_{3,9}$	−0.003	[−0.011, 0.006]	−0.003	[−0.011, 0.006]	−0.003	[−0.011, 0.006]	−0.003	[−0.011, 0.006]
$V_{3,10}$	0.001	[−0.007, 0.009]	0.001	[−0.007, 0.009]	0.001	[−0.007, 0.009]	0.001	[−0.008, 0.009]
$V_{4,4}$	1.044	[1.033, 1.056]	1.037	[1.020, 1.055]	1.043	[1.031, 1.055]	1.043	[1.031, 1.054]
$V_{4,5}$	−0.002	[−0.010, 0.007]	−0.002	[−0.010, 0.007]	−0.002	[−0.010, 0.007]	−0.002	[−0.010, 0.007]
$V_{4,6}$	0.001	[−0.008, 0.009]	0.001	[−0.007, 0.009]	0.001	[−0.008, 0.009]	0.001	[−0.008, 0.009]
$V_{4,7}$	0.001	[−0.007, 0.010]	0.001	[−0.007, 0.010]	0.001	[−0.007, 0.010]	0.001	[−0.007, 0.010]
$V_{4,8}$	−0.001	[−0.009, 0.008]	−0.001	[−0.009, 0.008]	−0.001	[−0.009, 0.007]	−0.001	[−0.009, 0.008]
$V_{4,9}$	−0.002	[−0.010, 0.006]	−0.002	[−0.010, 0.006]	−0.002	[−0.010, 0.006]	−0.002	[−0.010, 0.006]
$V_{4,10}$	−0.005	[−0.014, 0.003]	−0.005	[−0.013, 0.003]	−0.005	[−0.014, 0.003]	−0.005	[−0.014, 0.003]
$V_{5,5}$	1.050	[1.038, 1.062]	1.042	[1.026, 1.060]	1.048	[1.036, 1.060]	1.048	[1.036, 1.060]
$V_{5,6}$	−0.004	[−0.012, 0.004]	−0.004	[−0.012, 0.004]	−0.004	[−0.012, 0.005]	−0.004	[−0.012, 0.004]
$V_{5,7}$	0.000	[−0.008, 0.008]	0.000	[−0.008, 0.008]	0.000	[−0.008, 0.008]	0.000	[−0.008, 0.008]
$V_{5,8}$	0.000	[−0.008, 0.009]	0.000	[−0.008, 0.009]	0.000	[−0.008, 0.009]	0.000	[−0.009, 0.008]
$V_{5,9}$	0.002	[−0.007, 0.010]	0.002	[−0.007, 0.010]	0.002	[−0.007, 0.010]	0.002	[−0.007, 0.010]
$V_{5,10}$	−0.002	[−0.010, 0.007]	−0.002	[−0.010, 0.006]	−0.002	[−0.010, 0.006]	−0.002	[−0.010, 0.006]
$V_{6,6}$	1.064	[1.052, 1.076]	1.056	[1.040, 1.074]	1.062	[1.050, 1.074]	1.062	[1.050, 1.074]
$V_{6,7}$	−0.003	[−0.011, 0.006]	−0.003	[−0.011, 0.005]	−0.003	[−0.011, 0.006]	−0.003	[−0.011, 0.005]
$V_{6,8}$	−0.007	[−0.016, 0.001]	−0.007	[−0.016, 0.001]	−0.007	[−0.016, 0.001]	−0.007	[−0.016, 0.001]
$V_{6,9}$	−0.002	[−0.010, 0.007]	−0.002	[−0.010, 0.007]	−0.002	[−0.010, 0.007]	−0.002	[−0.010, 0.007]
$V_{6,10}$	0.002	[−0.006, 0.011]	0.002	[−0.006, 0.010]	0.002	[−0.006, 0.010]	0.002	[−0.006, 0.011]
$V_{7,7}$	1.054	[1.042, 1.066]	1.045	[1.031, 1.062]	1.051	[1.039, 1.063]	1.051	[1.039, 1.063]
$V_{7,8}$	−0.003	[−0.011, 0.006]	−0.003	[−0.011, 0.006]	−0.003	[−0.011, 0.006]	−0.003	[−0.011, 0.006]
$V_{7,9}$	−0.002	[−0.010, 0.007]	−0.002	[−0.010, 0.007]	−0.002	[−0.010, 0.007]	−0.002	[−0.010, 0.007]
$V_{7,10}$	0.002	[−0.006, 0.010]	0.002	[−0.006, 0.011]	0.002	[−0.006, 0.010]	0.002	[−0.006, 0.011]
$V_{8,8}$	1.083	[1.071, 1.095]	1.075	[1.058, 1.094]	1.081	[1.069, 1.094]	1.081	[1.069, 1.093]
$V_{8,9}$	−0.003	[−0.012, 0.005]	−0.003	[−0.012, 0.005]	−0.003	[−0.012, 0.005]	−0.003	[−0.012, 0.005]
$V_{8,10}$	0.000	[−0.008, 0.009]	0.000	[−0.008, 0.009]	0.000	[−0.009, 0.009]	0.000	[−0.009, 0.009]
$V_{9,9}$	1.073	[1.060, 1.085]	1.064	[1.049, 1.082]	1.070	[1.058, 1.083]	1.070	[1.058, 1.082]
$V_{9,10}$	0.000	[−0.009, 0.008]	0.000	[−0.009, 0.008]	0.000	[−0.009, 0.008]	0.000	[−0.009, 0.008]
$V_{10,10}$	1.074	[1.062, 1.086]	1.066	[1.050, 1.084]	1.072	[1.060, 1.084]	1.072	[1.060, 1.084]

Notes: The results are from 30,000 MCMC posterior draws (after 20,000 burnin sweeps). The matrix elements are presented in vector-horizontal form, $\{V_{1,1}, ..., V_{10,10}\} = \text{vech}(V)$.

Table 10: Posterior Means and 95% density interval of parametric matrix V from the inverse-Wishart models - 10 stocks.

	MRG-N-IW		**MRG-t-IW**		**MRG-DPM-Λ-IW**		**MRG-DPM-MΛ-IW**	
	Mean	95% D.I.	Mean	95% D.I.	Mean	95% D.I.	Mean	95% D.I.
$V_{1,1}$	0.828	[0.817, 0.840]	0.814	[0.801, 0.827]	0.823	[0.811, 0.835]	0.822	[0.810, 0.834]
$V_{1,2}$	−0.062	[−0.069, −0.055]	−0.062	[−0.069, −0.055]	−0.062	[−0.069, −0.055]	−0.062	[−0.069, −0.055]
$V_{1,3}$	0.000	[−0.007, 0.007]	0.001	[−0.006, 0.008]	0.001	[−0.006, 0.008]	0.001	[−0.006, 0.009]
$V_{1,4}$	0.002	[−0.006, 0.009]	0.002	[−0.005, 0.010]	0.002	[−0.005, 0.010]	0.003	[−0.005, 0.010]
$V_{1,5}$	0.001	[−0.007, 0.008]	0.001	[−0.007, 0.008]	0.001	[−0.006, 0.008]	0.001	[−0.006, 0.009]
$V_{1,6}$	−0.019	[−0.026, −0.012]	−0.019	[−0.027, −0.012]	−0.019	[−0.026, −0.012]	−0.019	[−0.026, −0.012]
$V_{1,7}$	−0.051	[−0.058, −0.043]	−0.051	[−0.058, −0.044]	−0.051	[−0.058, −0.044]	−0.051	[−0.058, −0.044]
$V_{1,8}$	0.001	[−0.006, 0.009]	0.002	[−0.006, 0.009]	0.002	[−0.006, 0.009]	0.002	[−0.006, 0.009]
$V_{1,9}$	−0.029	[−0.036, −0.021]	−0.028	[−0.036, −0.021]	−0.029	[−0.036, −0.021]	−0.029	[−0.036, −0.021]
$V_{1,10}$	0.021	[0.014, 0.029]	0.021	[0.014, 0.029]	0.022	[0.014, 0.029]	0.022	[0.014, 0.029]
$V_{2,2}$	1.013	[1.001, 1.025]	1.001	[0.987, 1.015]	1.011	[0.998, 1.023]	1.010	[0.998, 1.022]
$V_{2,3}$	−0.030	[−0.038, −0.022]	−0.030	[−0.038, −0.022]	−0.030	[−0.038, −0.022]	−0.030	[−0.038, −0.022]
$V_{2,4}$	−0.020	[−0.028, −0.011]	−0.020	[−0.028, −0.012]	−0.020	[−0.028, −0.012]	−0.020	[−0.028, −0.012]
$V_{2,5}$	−0.011	[−0.019, −0.003]	−0.011	[−0.019, −0.003]	−0.011	[−0.019, −0.003]	−0.011	[−0.020, −0.003]
$V_{2,6}$	−0.011	[−0.019, −0.003]	−0.011	[−0.019, −0.003]	−0.011	[−0.019, −0.003]	−0.011	[−0.019, −0.003]
$V_{2,7}$	−0.021	[−0.029, −0.013]	−0.021	[−0.029, −0.013]	−0.021	[−0.029, −0.013]	−0.021	[−0.030, −0.013]
$V_{2,8}$	−0.005	[−0.014, 0.003]	−0.005	[−0.013, 0.003]	−0.005	[−0.013, 0.003]	−0.005	[−0.014, 0.003]
$V_{2,9}$	−0.017	[−0.025, −0.009]	−0.017	[−0.025, −0.009]	−0.017	[−0.026, −0.009]	−0.017	[−0.026, −0.009]
$V_{2,10}$	0.003	[−0.005, 0.012]	0.003	[−0.005, 0.012]	0.003	[−0.005, 0.012]	0.003	[−0.005, 0.012]
$V_{3,3}$	1.049	[1.037, 1.061]	1.038	[1.023, 1.052]	1.048	[1.036, 1.061]	1.048	[1.035, 1.060]
$V_{3,4}$	−0.036	[−0.044, −0.027]	−0.036	[−0.044, −0.027]	−0.036	[−0.044, −0.028]	−0.036	[−0.045, −0.028]
$V_{3,5}$	−0.022	[−0.030, −0.013]	−0.022	[−0.030, −0.014]	−0.022	[−0.031, −0.014]	−0.022	[−0.030, −0.014]
$V_{3,6}$	−0.013	[−0.021, −0.004]	−0.013	[−0.021, −0.004]	−0.013	[−0.021, −0.004]	−0.013	[−0.021, −0.004]
$V_{3,7}$	−0.019	[−0.028, −0.012]	−0.019	[−0.027, −0.012]	−0.020	[−0.028, −0.011]	−0.020	[−0.028, −0.011]
$V_{3,8}$	−0.007	[−0.016, 0.001]	−0.007	[−0.016, 0.001]	−0.007	[−0.016, 0.001]	−0.007	[−0.016, 0.001]
$V_{3,9}$	−0.019	[−0.027, −0.010]	−0.019	[−0.027, −0.010]	−0.019	[−0.027, −0.011]	−0.019	[−0.027, −0.010]
$V_{3,10}$	−0.011	[−0.020, −0.002]	−0.011	[−0.020, −0.003]	−0.011	[−0.020, −0.003]	−0.011	[−0.020, −0.003]
$V_{4,4}$	1.109	[1.096, 1.122]	1.097	[1.082, 1.112]	1.108	[1.094, 1.121]	1.107	[1.094, 1.121]
$V_{4,5}$	−0.016	[−0.024, −0.007]	−0.016	[−0.024, −0.007]	−0.016	[−0.025, −0.007]	−0.016	[−0.025, −0.007]
$V_{4,6}$	−0.005	[−0.013, 0.004]	−0.005	[−0.013, 0.004]	−0.005	[−0.013, 0.004]	−0.005	[−0.013, 0.004]
$V_{4,7}$	−0.007	[−0.015, 0.001]	−0.007	[−0.015, 0.001]	−0.007	[−0.015, 0.001]	−0.007	[−0.016, 0.001]
$V_{4,8}$	−0.008	[−0.016, 0.001]	−0.008	[−0.017, 0.001]	−0.008	[−0.017, 0.001]	−0.008	[−0.017, 0.001]
$V_{4,9}$	−0.012	[−0.020, −0.003]	−0.012	[−0.020, −0.003]	−0.012	[−0.020, −0.003]	−0.012	[−0.020, −0.003]
$V_{4,10}$	−0.016	[−0.025, −0.008]	−0.016	[−0.025, −0.008]	−0.016	[−0.025, −0.008]	−0.016	[−0.025, −0.008]
$V_{5,5}$	1.116	[1.102, 1.129]	1.103	[1.088, 1.119]	1.115	[1.101 , 1.128]	1.114	[1.101, 1.128]
$V_{5,6}$	−0.016	[−0.024, −0.007]	−0.016	[−0.024, −0.007]	−0.016	[−0.024, −0.007]	−0.016	[−0.024, −0.007]
$V_{5,7}$	−0.004	[−0.012, 0.005]	−0.004	[−0.012, 0.004]	−0.004	[−0.012, 0.005]	−0.004	[−0.012, 0.005]
$V_{5,8}$	−0.013	[−0.022, −0.004]	−0.013	[−0.022, −0.004]	−0.013	[−0.022, −0.004]	−0.013	[−0.022, −0.005]
$V_{5,9}$	−0.012	[−0.021, −0.003]	−0.012	[−0.020, −0.003]	−0.012	[−0.021, −0.003]	−0.012	[−0.021, −0.003]
$V_{5,10}$	−0.016	[−0.025, −0.007]	−0.016	[−0.025, −0.007]	−0.016	[−0.025, −0.007]	−0.016	[−0.025, −0.007]
$V_{6,6}$	1.102	[1.089, 1.115]	1.090	[1.075, 1.106]	1.102	[1.089 , 1.115]	1.101	[1.088, 1.115]
$V_{6,7}$	−0.007	[−0.015, 0.001]	−0.007	[−0.015, 0.001]	−0.007	[−0.015, 0.001]	−0.007	[−0.015, 0.001]
$V_{6,8}$	−0.013	[−0.022, −0.005]	−0.013	[−0.022, −0.005]	−0.013	[−0.022, −0.005]	−0.013	[−0.022, −0.005]
$V_{6,9}$	−0.012	[−0.020, −0.003]	−0.012	[−0.020, −0.003]	−0.012	[−0.020, −0.003]	−0.012	[−0.020, −0.003]
$V_{6,10}$	0.003	[−0.006, 0.012]	0.003	[−0.006, 0.012]	0.003	[−0.006, 0.012]	0.003	[−0.006, 0.012]
$V_{7,7}$	1.030	[1.018, 1.042]	1.017	[1.002, 1.032]	1.027	[1.015, 1.040]	1.027	[1.014, 1.039]
$V_{7,8}$	0.007	[−0.001, 0.016]	0.007	[−0.001, 0.016]	0.007	[−0.001, 0.016]	0.007	[−0.001, 0.016]
$V_{7,9}$	−0.007	[−0.015, 0.001]	−0.007	[−0.015, 0.001]	−0.007	[−0.015, 0.001]	−0.007	[−0.015, 0.001]
$V_{7,10}$	0.006	[−0.003, 0.014]	0.006	[−0.003, 0.014]	0.006	[−0.003, 0.014]	0.006	[−0.003, 0.014]
$V_{8,8}$	1.134	[1.120 , 1.147]	1.121	[1.106, 1.137]	1.132	[1.119, 1.146]	1.132	[1.119, 1.146]
$V_{8,9}$	−0.012	[−0.020, −0.003]	−0.012	[−0.020, −0.003]	−0.012	[−0.020, −0.003]	−0.012	[−0.020, −0.003]
$V_{8,10}$	−0.003	[−0.011, 0.006]	−0.003	[−0.011, 0.006]	−0.003	[−0.012, 0.006]	−0.003	[−0.012, 0.006]
$V_{9,9}$	1.098	[1.085, 1.111]	1.085	[1.069, 1.100]	1.095	[1.082, 1.109]	1.095	[1.082, 1.108]
$V_{9,10}$	−0.007	[−0.015, 0.002]	−0.006	[−0.015, 0.002]	−0.006	[−0.015, 0.002]	−0.006	[−0.015, 0.002]
$V_{10,10}$	1.159	[1.145, 1.173]	1.146	[1.130, 1.162]	1.158	[1.144, 1.171]	1.157	[1.143, 1.171]

Notes: The results are from 30,000 MCMC posterior draws (after 20,000 burnin sweeps). The matrix elements are presented in vector-horizontal form, $\{V_{1,1}, ..., V_{10,10}\} = \text{vech}(V)$.