Multivariate dynamic mixed frequency density pooling for financial forecasting

Audronė Virbickaitė [*†] Hedibert F. Lopes [‡] Martina Danielova Zaharieva [§]

Abstract

This article investigates the benefits of combining information available from daily and intraday data in financial return forecasting. The two data sources are combined via a density pooling approach, wherein the individual densities are represented as a copula function, and the potentially time-varying pooling weights depend on the forecasting performance of each model. The dependence structure in the daily frequency case is extracted from a standard static and dynamic conditional covariance modeling, and the high frequency counterpart is based on a realized covariance measure. We find that incorporating both high and low frequency information via density pooling provides significant gains in predictive model performance over any individual model and any model combination within the same data frequency. A portfolio allocation exercise quantifies the economic gains by producing investment portfolios with the smallest variance and highest Sharpe ratio.

Keywords: Conditional Value at Risk; Density combination; High Frequency; Realized volatility; Global Minimum Variance.

JEL: C53, C58, C11.

Acknowledgments

The first author is partially supported by the grant PID2020-113192GB-I00/AEI/10.13039/501100011033 (*Mathematical Visualization: Foundations, Algorithms and Applications*); the first and third authors are

[*]Corresponding author, audrone.virbickaite@cunef.edu, Phone: 0034 914 480 892
[†]Department of Quantitative Methods, CUNEF Universidad, Calle Pirineos 55, Madrid, Spain
[‡]Insper Institute of Education and Research, Sao Paulo, Brazil
[§]Department of Quantitative Methods, CUNEF Universidad, Calle Pirineos 55, Madrid, Spain

partially supported by the grant PID2022-138289NB-I00 (*Modeling and forecasting financial returns and macroeconomic variables in data-rich environments: new methods and applications*) from the Spanish State Research Agency (Ministerio de Ciencia e Innovación).

1 Introduction

Since the advent of the availability of high frequency financial data, research on how to use, model and forecast measures extracted from such data has surged (McAleer & Medeiros 2008). As a result, high frequency data based models have proven to be powerful competitors in financial return forecasting to the standard modeling approaches using daily data (Lyócsa et al. 2021). As an alternative to choosing a single modeling approach, some authors have combined the best of both worlds by augmenting low frequency models with high frequency information, see Engle (2002*b*), Ghysels et al. (2004, 2005), Shephard & Sheppard (2010) for univariate modeling, and Noureldin et al. (2012), Bauwens & Xu (2023) for a multivariate approach. Such combinations rely on suitable parametrization, in which the high frequency measure enters the model as an exogenous covariate, thus creating a new class of models.

In contrast to previous research, we combine low and high frequency information not through parameters but through the combination of densities. In particular, we model and forecast the dependence structure of multiple financial returns as a weighted sum of two predictive densities, the first arising from low frequency data and the second arising from high frequency data. Such combinations are also known as opinion pools (the name was first proposed by Stone 1961).

The combination of predictive densities is a recent topic of increasing interest in financial and macroeconomic forecasting. Many of the related literature focuses on proposing a sophisticated combination scheme. For instance, Billio et al. (2013) consider dynamic weights based on the model residuals instead of log scores; Bassetti et al. (2018) assume random combination weights; Opschoor et al. (2017) use scoring rules that focus on a particular area of the density for the computation of the component weights. Furthermore, a model pooling approach is an attractive alternative to model selection when the true model is unknown (O'Doherty et al. 2012). The approach we propose in this paper departs from the strands of literature dealing with forecast combinations based on a number of alternative models that condition on the same information set, as it pools models that arise from parallel theoretical perspectives of understanding volatility. Those are, the volatility as an unobserved process estimated from daily data versus the volatility as an observable quantity extracted from the high frequency data. Given the full availability of financial data at all frequency levels, a density forecast

combination incorporating both perspectives is an obvious and feasible way to improve predictive performance, which we demonstrate by means of two data sets, representing time periods with different market conditions such as the Global Financial Crisis and the Covid 19 pandemic.

Even though models, based on both low- and high frequency data, are essentially aimed at capturing the individual dynamics and the co-dependence structure between the financial returns, they exhibit very different properties. The low frequency data-based models consider the entire series of historical daily data, and the estimated co-volatility processes are usually smooth. In contrast, the high frequency data based models can capture instantaneous changes in co-variation and produce the forecasts accordingly. As noted in Kapetanios et al. (2015), Timmermann (2018), some models might be useful while the markets are in decline, whereas other models might be more informative when the markets are booming. Therefore, the time-varying pooling weights can also indicate whether the preference for one model or another is correlated with the overall market conditions. Such correlation affects an investor's decisions: if an investor anticipates a change in market conditions, they might reconsider the modeling strategy accordingly.

Each of the individual densities in the pool, namely that modeled with low frequency and that modeled with high frequency data, are constructed by a copula function. Using a copula instead of the complete high-dimensional density is a convenient solution when the focus of the modeling is explicitly on the dependence structure rather than on the individual series dynamics. Modeling the dependence via a copula also has practical advantages. It allows to simplify the assessment of the marginal distributions and avoids dealing with highly parametrized and possibly nonstandard multivariate density functions. Models in which dynamic copula parameters are obtained from daily data are considered a standard approach in the financial times series literature (multivariate GARCH models in Dias & Embrechts 2004, Patton 2006a, Ausín & Lopes 2010; score-driven models in Koopman et al. 2018, Nguyen & Javed 2021; and factor models in Oh & Patton 2017, Opschoor et al. 2021). In contrast, models in which the copula dependence structure is obtained from high frequency data are rather sparse (Salvatierra & Patton 2015, Fengler & Okhrin 2016, Okhrin & Tetereva 2017). Salvatierra & Patton (2015) have modeled the dynamics of the copula parameter as a function of past realized correlations via an autoregressive score-type model, whereas Fengler & Okhrin (2016) and Okhrin & Tetereva (2017) have used multiple univariate AR-type processes for the realized covariances. In this work, we use static and dynamic

multivariate GARCH-type models for the copula parameter arising from daily data, and two models for the copula parameter arising from intra-day data: the Additive Inverse Wishart (AIW) model of Jin & Maheu (2013, 2016), and the Conditional Autoregressive Wishart (CAW) model of Golosnoy et al. (2012).

In order to control for the effects of the particular weighting scheme, we consider four options for density pooling, all based on the log predictive scores (LPS): equally weighted, static (Geweke & Amisano 2011), naïve dynamic (Jore et al. 2010) and dynamic (Del Negro et al. 2016). In the static weighting scheme, the weights are re-balanced daily as a function of the expanding set of the past LPS, converging to a stable equilibrium, hence the name static. In the naïve dynamic scheme, the set of the past LPS is smaller, and only the most recent observations are considered. Finally, in the *fully* dynamic scheme, the weights are latent and are updated as a function of the past weights via an AR-type process. For a general introduction and comprehensive reviews of aggregating probability distributions, readers are referred to Clemen & Winkler (2007), among others. Timmermann (2018) briefly reviews the forecast combinations in the financial econometrics context.

In this article, we rely on a Bayesian estimation approach in three stages. In the first stage, the marginal distributions are estimated. In particular, the daily data are transformed to the unit interval by standardization of the financial returns and application of the probability integral transform. In the second stage, we fit a copula model to the resulting Uniformly distributed data to obtain the joint predictive density for the returns. Finally, the density pooling weights assigned to a low- and to a high frequency model are obtained in order to produce pooled forecasts for the daily log returns.

We illustrate the performance of the proposed method using daily and intraday return data of 14 exchange rates sampled from 2008 to 2023. Empirical results show that pooled models outperform the best individual model in terms of the entire density forecast as well in the left tail. In addition, the density pool shows improvement in predictive performance with respect to mixed frequency models, in this case the DCC-HEAVY model of Bauwens & Xu (2023). An additional empirical application containing ten assets from the DJIA index is available in Appendix E in the Online Supplementary Material and confirms the main results. Finally, we perform a Global Minimum Variance and a Conditional Value-at-Risk portfolio allocation exercise to quantify the economic gains in using the proposed approach. The results illustrate the benefits of pooling by delivering investment portfolios with the

smallest variance and the highest Sharpe ratio values. Overall, we demonstrate that combining information from daily and intraday data sources not only produces superior joint density forecasts but also leads to tangible economic benefits for the investor.

The article is organized as follows. Sections 2 and 3 present the pooled copula model, estimation approach and model evaluation. Sections 4 contains the empirical application, and finally, Section 6 concludes.

2 Methodology

In this section, we describe the main contribution of the article: combining information arising from high and low frequency data in the copula modeling framework in order to model and forecast the dependence structure of financial returns. The choice of modeling the joint distribution via copulas, next to being a flexible way of constructing multivariate densities, is also convenient from computational and methodological perspectives. As noted in Opschoor et al. (2021), when the cross-section dimension d is large, specifying and estimating the marginals separately might considerably ease the computational burden. In addition, such an approach enables a focus on modeling the dependence structure explicitly, independently from the marginals. In particular, we are interested in estimating a copula density $c(u_t|\mathcal{M}_{HF}, \mathcal{M}_{LF})$ for Uniformly distributed data. The approach consists of three major tasks:

- Modeling the dynamics of the covariance matrices arising from low frequency data via a model called $\mathcal{M}_{LF}$. In this step, we consider standard specifications for multivariate covolatilities, such as the Static model and the Dynamic Conditional Correlation (DCC) model of Engle (2002*a*).

- Modeling the dynamics of the covariance matrices arising from high frequency data via a model called $\mathcal{M}_{HF}$. We consider the AIW approach of Jin & Maheu (2013, 2016) and the CAW model of Golosnoy et al. (2012).

- Modeling the dynamics of the combination weights ω_t. Here, we consider four options, covering a large part of the variety of linear combination strategies. The schemes are

equally weighted, static (Geweke & Amisano 2011), naïve dynamic (Jore et al. 2010) and dynamic (Del Negro et al. 2016).

We start by defining $r_{i,t}$ as the log returns (in %) for day t and asset i, such that $t = 1, \ldots, T$ and $i = 1, \ldots, d$:

$$r_{i,t} = 100 \times \left(\log \frac{P_{i,t}^{\text{Close}}}{P_{i,t}^{\text{Open}}} \right),$$

where $P_{i,t}^{\text{Close}}$ and $P_{i,t}^{\text{Open}}$ are the opening and the closing prices, respectively.

Next, we present an approach to combine information arising from high and low frequency data for dependence modeling between daily financial returns, by relying on a density combination approach. The linear combination of individual densities is given by:

$$p(r_t) = \sum_{j=1}^{N} \omega_j p(r_t | \mathcal{M}_j), \quad t = 1, \ldots, T,$$

where $r_t = (r_{1,t}, \ldots, r_{d,t})'$ is the d-variate return vector, N is the number of alternative models $\mathcal{M}_j$, ω_j is the combination weight, and $p(r_t|\mathcal{M}_j)$ is the candidate density, originating from different models. The dependence structure between low frequency returns r_t can be modeled by combining a model estimated from daily returns $p(r_t|\mathcal{M}_{LF})$ with a model estimated from high frequency returns $p(r_t|\mathcal{M}_{HF})$.

A convenient way to model the potentially high dimensional joint density $p(r_t|\mathcal{M}_j)$ involves separating the dependence structure from the dynamics of the marginals by using copula functions. The copula approach is a very popular modeling strategy in the financial econometrics context, as it not only reduces the computational burden for large-dimensional datasets but also allows for the construction of non-standard multivariate probability distributions, while including standard specifications as nested cases (Patton 2009, Fan & Patton 2014, Oh & Patton 2017, Opschoor et al. 2021, Nguyen et al. 2024). Furthermore, the treatment of the marginal densities can be substantially simplified by taking advantage of the available *ex post* realized volatility measure defined by the realized variance $RV_{i,t} = \sum_{j=1}^{J} \tilde{r}_{i,t,j}^2$. Here, $\tilde{r}_{i,t,j}$ is an l-minute log-return for day t, and J is the number of l-minute intervals in a trading day (Barndorff-Nielsen & Shephard 2002, Andersen et al. 2003, Barndorff-Nielsen & Shephard 2004). For a review of realized volatility, readers are referred to McAleer & Medeiros (2008). To account for the

time-varying volatility, the log returns are first standardized by the realized volatility measure, then demeaned and standardized by some unconditional standard deviation[1] $z_{i,t} = \left(r_{i,t}/\sqrt{RV_{i,t}} - \mu_i\right)/\sigma_i$. As seen in Andersen et al. (2000, 2001), is it safe to assume that $z_{i,t} \sim N(0,1)$.[2] Finally, we call $u_{i,t} = \Phi_1(z_{i,t})$ the probability integral transform of the $z_{i,t}$, where $\Phi_1(\cdot)$ is a cumulative distribution function for the univariate standard Normal distribution, and the resulting variables are Uniformly distributed $u_{i,t} \overset{iid}{\sim} \mathcal{U}(0,1) \forall i = 1, \ldots, d$ (serially uncorrelated). This approach helps reduce the number of parameters and the computational burden of the estimation procedure. Moreover, using the realized volatility for the standardization step circumvents the inclusion of an additional potential source of estimation error.

The dependence structure of the resulting probability integral transforms can be easily modeled by using copulas. To define a copula, we consider a collection of random variables $Y_1, \ldots, Y_d$ with corresponding distribution functions $F_i(y_i) = P[Y_i \leq y_i]$ for $i = 1 \ldots, d$ and a joint distribution function $H(y_1, \ldots, y_d) = P[Y_1 \leq y_1, \ldots, Y_d \leq y_d]$. Then, according to a theorem by Sklar (1959), a copula C exists such that

$$H(y_1, \ldots, y_d) = C(F_1(y_1), \cdots, F_d(y_d)),$$

and it is unique if the variables are continuous. Furthermore, based on the conditional copula definition by Patton (2006b) Sklar's theorem is extended to the time series case by defining $Y_t = (Y_{1,t}, \ldots, Y_{d,t})'$ and $\mathcal{F}_{t-1}$ being an information set:

$$H(y_t|\mathcal{F}_{t-1}) = C(F_1(y_{1,t}|\mathcal{F}_{t-1}), \cdots, F_d(y_{d,t}|\mathcal{F}_{t-1})|\mathcal{F}_{t-1}),$$

where $F(\cdot|\mathcal{F}_{t-1})$ and $H(\cdot|\mathcal{F}_{t-1})$ are the corresponding conditional distribution functions and C is the conditional copula of Y_t given $\mathcal{F}_{t-1}$. The joint density $h(y_{1,t}, \ldots, y_{d,t}|\mathcal{F}_{t-1})$ is then a product of individual marginal densities $f_i(y_{i,t}|\mathcal{F}_{t-1})$ and a copula density: $h(y_{1,t}, \ldots, y_{d,t}|\mathcal{F}_{t-1}) = c(F_1(y_{1,t}|\mathcal{F}_{t-1}), \cdots, F_d(y_{d,t}|\mathcal{F}_{t-1})) \cdot \prod_{i=1}^{d} f_i(y_{i,t}|\mathcal{F}_{t-1})$. In other words, the dependence structure can be separated

[1] σ_i is a scaling factor that allows the standard deviation of the returns to deviate from the RV measure; see Jin & Maheu (2013, 2016).

[2] Andersen et al. (2000, 2001) have found that the distributions of the returns scaled by realized standard deviations are approximately Gaussian.

from the marginals. Copulas are defined in the unit hypercube $[0,1]^d$, where d is the dimension of the data, and all univariate marginals are Uniformly distributed. For a detailed treatment of copulas and areas of applications, readers are referred to McNeil et al. (2005), Nelsen (2006), Patton (2012), Joe (2015). For the ease of notation in the following we omit the conditioning on the information set $\mathcal{F}_{t-1}$.

In this article, we use Gaussian and t copulas, because they are available in high dimensions ($d > 2$), and their implementation is straightforward. Gaussian copulas, although widely used, do not allow for fat-tailed co-dependence — an assumption that can be relaxed by using the t copula.

Call $u_t = (u_{1,t}, \ldots, u_{d,t})'$ the collection of Uniformly distributed data at time t. The d-variate Gaussian copula has the following distribution and density functions (Joe 2015):

$$
\begin{aligned}
C(u_t|\boldsymbol{R}) &= \Phi_d(\Phi_1^{-1}(u_{1,t}), \ldots, \Phi_1^{-1}(u_{d,t})|\boldsymbol{R}), \\
c(u_t|\boldsymbol{R}) &= \frac{\phi_d(\Phi_1^{-1}(u_{1,t}), \ldots, \Phi_1^{-1}(u_{d,t})|\boldsymbol{R})}{\prod_{i=1}^{d} \phi_1(\Phi_1^{-1}(u_{i,t}))}.
\end{aligned}
$$

Here, $\Phi_d(\cdot|\boldsymbol{R})$ and $\phi_d(\cdot|\boldsymbol{R})$ are a d-variate standard Normal distribution and density functions with a correlation matrix $\boldsymbol{R}$. The d-variate t copula has the following distribution and density functions (Joe 2015):

$$
\begin{aligned}
C(u_t|\boldsymbol{R}, \eta) &= T_{d,\eta}(T_{1,\eta}^{-1}(u_{1,t}), \ldots, T_{1,\eta}^{-1}(u_{d,t})|\boldsymbol{R}), \\
c(u_t|\boldsymbol{R}, \eta) &= \frac{t_{d,\eta}(T_{1,\eta}^{-1}(u_{1,t}), \ldots, T_{1,\eta}^{-1}(u_{d,t})|\boldsymbol{R})}{\prod_{i=1}^{d} t_{1,\eta}(T_{1,\eta}^{-1}(u_{i,t}))}.
\end{aligned}
$$

Here, $T_{1,\eta}$, $T_{d,\eta}(\cdot|\boldsymbol{R})$, $t_{1,\eta}$ and $t_{d,\eta}(\cdot|\boldsymbol{R})$ are the univariate and d-variate t distribution and density functions with degrees of freedom parameter $\eta > 0$ and correlation matrix $\boldsymbol{R}$. When $\eta \to \infty$, the t copula becomes a Gaussian copula.

Another important reason exists to focus on Gaussian copula, at least for the high frequency model. We make use of the fact that the variance-covariance (or correlation) matrix, estimated from the de-meaned and standardized log returns $z_{i,t}$ (given that they are approximately Normally distributed), is equivalent to the copula parameter $\boldsymbol{R}$. This result is valid for only a Gaussian copula with standard Normal marginals and is a result of Hoeffding's lemma and Sklar's theorem; details are described in Fengler & Okhrin (2016). In other words, the Gaussian copula function enables one-to-one mapping

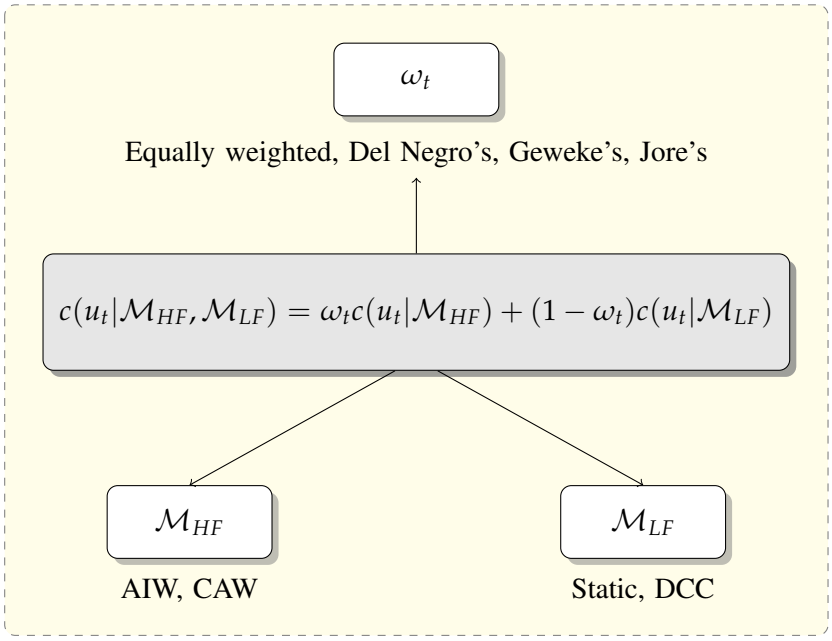

Figure 1: Model components combining high ($\mathcal{M}_{HF}$) and low ($\mathcal{M}_{LF}$) frequency models via density pooling with time varying weights ω_t.

between the copula dependence parameter and the linear dependence measure. This result is relevant to our work because it enables use of the realized correlation, obtained from high frequency data, as a copula parameter $\boldsymbol{R}$.

Finally, given our proposed framework, the resulting density to describe daily dependence structure at time t can be written in terms of a copula density pool:

$$c(u_t|\mathcal{M}_{LF},\mathcal{M}_{HF}) = \omega_t c(u_t|\mathcal{M}_{HF}) + (1-\omega_t)c(u_t|\mathcal{M}_{LF}), \tag{1}$$

where $\mathcal{M}_{HF}$ and $\mathcal{M}_{LF}$ present the models estimated by using high and low frequency data. Note that, when assuming the same marginal models in both joint densities, the density pool simplifies to a product of individual marginal densities and the copula density pool in Eq. (1). Therefore, the following sections focus on the high- and low frequency correlation matrix modeling, which correspond to $\mathcal{M}_{HF}$ and $\mathcal{M}_{LF}$. Figure 1 summarizes the main model components and the approach used at each phase.

2.1 Low frequency covariance modeling

Next, we consider several standard approaches to model the dynamics of the correlation matrix that arises from the low frequency (daily) data. Call Ω a variance-covariance matrix of the observed standard Normally distributed standardized returns $z_t = (z_{1,t}, \ldots, z_{d,t})'$. Then the corresponding correlation matrix is $\boldsymbol{R} = (\text{diag}\Omega)^{-1/2}\Omega(\text{diag}\Omega)^{-1/2}$, with $\text{diag}(X)$ denoting the diagonal matrix obtained by setting to zero all the off-diagonal elements of the squared matrix X. We start with the most straightforward way to measure dependence, by using a sample correlation matrix. The dependence between u_t is modeled either by fitting a Gaussian or t copula with a static correlation matrix $\boldsymbol{R}$, estimated given the daily data up to time t. Such an approach is referred to as Static and is our benchmark specification.

Another possible model, which is dynamic, is the DCC model of Engle (2002*a*):

$$\Omega_t = \overline{\Omega} \odot (\iota\iota' - B_1 - B_2) + B_1 \odot z_{t-1}z'_{t-1} + B_2 \odot \Omega_{t-1}, \tag{2}$$

where $\odot$ is the Hadamard product of two equally sized matrices (element-by-element multiplication); ι is a vector of ones; parameter matrices B_1, B_2 are re-parametrized as rank-1 matrices $B_1 = b_1 b'_1$ and $B_2 = b_2 b'_2$ with b_1 and b_2 being a $d \times 1$ parameter vectors; and $\overline{\Omega}$ is a sample variance covariance matrix. This result corresponds to the time-varying correlation matrix $\boldsymbol{R}_t = (\text{diag}\Omega_t)^{-1/2}\Omega_t(\text{diag}\Omega_t)^{-1/2}$. Naturally, the model choice for daily variance-covariance matrix is not limited to the models outlined above. For extensive reviews of existing multivariate volatility models, readers are referred to Asai et al. (2006), Bauwens et al. (2006), Silvennoinen & Teräsvirta (2009), among others.

2.2 High frequency covariance modeling

As mentioned before, Fengler & Okhrin (2016) have shown that the Gaussian copula's parameter $\boldsymbol{R}_t$ can be estimated by using the correlation matrix of the original data (log returns in our case). In the high frequency data setting, the correlation matrix of the returns can be estimated via Rcor_t, a realized correlation measure, obtained from intraday data (Noureldin et al. 2012): $\mathbf{R}_t \equiv \text{Rcor}_t = (\text{diag}\,\text{Rcov}_t)^{-1/2}\text{Rcov}_t(\text{diag}\,\text{Rcov}_t)^{-1/2}$, where Rcov_t is a realized covariance measure. Modeling the dynamics of the realized covariance matrices is a notoriously difficult task because of the high dimensions and positive-definite restrictions on the matrices. An efficient method is to model the dynamics

of the realized variance-covariance matrices directly by using Wishart distributions (Gourieroux et al. 2009, Jin & Maheu 2013, 2016). In Jin & Maheu (2013, 2016), the scale matrix in the Wishart distribution follows either an additive or a multiplicative component structure, and the authors have found that the additive structure performs better. Such additive models capture strong persistence in the covariances and fat-tailed distributions of the returns. They have compared their proposed model with multiple other models, such as Cholesky-VARFIMA from Chiriac & Voev (2011), the Wishart autoregressive model from Gourieroux et al. (2009), vec-MGARCH from Ding & Engle (2001) and DCC from Engle (2002*a*). The additive Wishart model has been found to produce superior density forecasts for all forecast horizons.

Next, we present the additive component model, introduced in Jin & Maheu (2013, 2016). Consider a sequence of realized covariance matrices Rcov_t of dimension $d \times d$, $t = 1, \ldots, T$. The additive component inverse Wishart AIW(L) model is given by:

$$\begin{aligned} \text{Rcov}_t &\sim \mathcal{IW}((\nu - d - 1)V_t, \nu), \\ V_t &= B_0 + \sum_{j=1}^{L} B_j \odot \Gamma_{t-1,l_j}, \\ B_j &= b_j b_j', \quad j = 1, \ldots, L, \\ \Gamma_{t-1,l_j} &= 1/l_j \sum_{i=1}^{l_j} \text{Rcov}_{t-i}. \end{aligned} \tag{3}$$

Here, $\mathcal{IW}(A, b)$ is the inverse-Wishart distribution with scale matrix A and degrees of freedom b. We set $l_1 = 1$, and further l_js indicates how many past observations are used to form a component Γ_{t-1,l_j} and L is the number of autoregressive components. B_0 is a symmetric positive definite matrix and is set to $B_0 = (\iota\iota' - B_1 - \ldots - B_K) \odot \overline{\text{Rcov}}$ so that the long-term mean of the covariances is equal to the sample mean, $\overline{\text{Rcov}}$. Note that one could use either Wishart or inverse-Wishart models. However, Jin & Maheu (2016) have found that in their empirical application the inverse-Wishart specification outperformed the Wishart counterpart; also, inverse-Wishart coupled with Gaussian copula results into closed form marginal predictive distributions for the standardized returns.

Another closely related model is the CAW(p, q) model of Golosnoy et al. (2012):

$$\begin{aligned} \text{Rcov}_t &\sim \mathcal{W}(V_t/\nu, \nu), \\ V_t &= B_0 + \sum_{i=1}^{p} B_1^{(i)} \text{Rcov}_{t-1} B_1^{(i)'} + \sum_{j=1}^{q} B_2^{(j)} V_{t-1} B_2^{(j)'}, \end{aligned} \tag{4}$$

where $B_0 = CC'$ with C being a $d \times d$ lower triangular matrix and $B_1^{(i)}$, $B_2^{(j)}$ are $d \times d$ parameter matrices. To avoid curse-of-dimensionality, as noted in Golosnoy et al. (2012), a natural restriction is to impose a diagonal structure of the dynamics of V_t by assuming $B_1^{(i)}$ and $B_2^{(j)}$ are diagonal. In this manuscript, we assume the order-1 CAW model where $p = q = 1$ and employ covariance targeting such that $B_0 = (\iota\iota' - B_1 - B_2) \odot \overline{\text{Rcov}}$ with $B_1 = b_1^{(1)} b_1^{(1)'}$ and $B_2 = b_2^{(1)} b_2^{(1)'}$. We also replace the Wishart distribution with the inverse-Wishart, same as in the model of Jin & Maheu (2013, 2016) in Eq.(3) so that the marginal predictive density is of closed form.

2.3 Choosing the weights

In this article, we attempt to cover a large part of the types of linear pooling schemes by focusing on four different approaches: equally weighted, static (Geweke & Amisano 2011), naïve dynamic (Jore et al. 2010) and dynamic (Del Negro et al. 2016).

Geweke & Amisano (2011) have proposed to maximize the log predictive score function at each point in time:

$$\begin{aligned} \omega_{T+k+1}^{Gew} &= \arg\max_{\omega} f(\omega), \quad \text{such that} \\ f(\omega) &= \sum_{t=1}^{T+k} \log[\omega c(u_t|\mathcal{M}_{HF}) + (1-\omega) c(u_t|\mathcal{M}_{LF})], \end{aligned} \tag{5}$$

where $c(u_t|\mathcal{M}_{HF})$ and $c(u_t|\mathcal{M}_{LF})$ are predictive copula densities for u_t, and $k = 1, \ldots, K$ is the out of sample evaluation period. Even though the weights are recalculated at each time point, this weighting scheme is considered static because, for a large K, the weights reach a stable equilibrium (Del Negro et al. 2016). Another approach involves using the log-score rolling weights calculated at

each time t by using $\tilde{m}$ lags, as defined in Jore et al. (2010):

$$\omega^{Jore}_{T+k+1,\tilde{m}} = \frac{\exp[\sum_{\tau=T+k+1-\tilde{m}}^{T+k} \log c(u_\tau|\mathcal{M}_{HF})]}{\sum_{r=\{HF,LF\}} \exp[\sum_{\tau=T+k+1-\tilde{m}}^{T+k} \log c(u_\tau|\mathcal{M}_r)]}. \tag{6}$$

We call this a naïve time-varying weighting approach. The main difference between the weights in Eqs. (5) and (6) is that Geweke's approach considers the predictive densities from the entire sample, whereas Jore's weighting scheme places importance only on the last $\tilde{m}$ observations. Finally, as in Del Negro et al. (2016), we allow for persistence in weights by introducing a latent variable s_t, thus giving rise to a dynamic weighting scheme:

$$\begin{aligned} s_t &= \beta s_{t-1} + \sqrt{1-\beta^2}\xi_t, \quad \xi_t \sim \mathcal{N}(0,1) \\ \omega_t^{DN} &= \Phi(s_t). \end{aligned} \tag{7}$$

The unconditional mean of s_t is 0, and the unconditional variance is 1. Parameter β controls the persistence of the weight dynamics: when $\beta = 1$, the process reduces to a random walk; when $\beta = 0$, at each time t, the weights ω_t^{DN} will be Uniformly distributed *a priori*.

2.4 Competitor model

For comparison purposes, we also include a model that augments the low frequency DCC model with the high frequency information via the high frequency-based volatility (HEAVY) approach (Noureldin et al. 2012), thus resulting in a scalar DCC-HEAVY specification of Bauwens & Xu (2023). In this model, the lagged outer product of standardized returns in Eq.(2) is replaced by the lagged realized correlations. More specifically, we can model the correlation matrix directly, without using the variance-covariance matrix, as $\boldsymbol{R}_t = \bar{\boldsymbol{R}} + a(\text{Rcor}_{t-1} - \overline{\text{Rcor}}) + b(\boldsymbol{R}_{t-1} - \bar{\boldsymbol{R}})$ corresponding to the model in equations (11)-(12) in Bauwens & Xu (2023). Here, $a, b \geq 0$, $b = 0$ if $a = 0$, $b < 1$; $\bar{\boldsymbol{R}}$ is the $d \times d$ sample correlation matrix of z_t; and $\overline{\text{Rcor}}$ is the sample mean of the realized correlation matrices. We consider the DCC-HEAVY model with Gaussian and t copulas. For more details, readers are referred to Bauwens & Xu (2023).

3 Posterior Inference and Model Selection

3.1 Posterior inference

For posterior inference and forecasting, we rely on Bayesian computation, particularly Markov chain Monte Carlo (MCMC) methods. To estimate the density pool in Eq. (1), we first sample from the posterior of the individual models $\mathcal{M}_{HF}$ and $\mathcal{M}_{LF}$. Conditional on those samples, the density pooling weights in Eqs. (5)-(7) can be obtained. Next, we briefly describe the posterior sampling details for each of the models presented in Sections 2.1 and 2.2.

Static model. Consider an inverse-Wishart prior on the unconditional variance covariance matrix $\Omega \sim \mathcal{IW}(I_d(\nu_0 - d - 1), \nu_0)$, $\nu_0 \geq d + 1$, so that $E[\Omega] = I_d$ and I_d id the d-dimensional unit matrix. Given the observed standardized approximately Normally distributed data $z_t = (z_{1,t}, \ldots, z_{d,t})'$, where $z_{1:T} = (z_1', \ldots, z_T')'$, the parameter Ω can be sampled directly from the posterior $\Omega | z_{1:T} \sim \mathcal{IW}(z_{1:T} z_{1:T}' + I_d(\nu_0 - d - 1), \nu_0 + T)$; derivations are provided in Appendix A.1 in Online Supplementary Material. The correlation matrix used as a copula parameter is obtained as $\boldsymbol{R} = (\text{diag}\Omega)^{-1/2} \Omega (\text{diag}\Omega)^{-1/2}$.

DCC, DCC-t The parameters for the rank-1 DCC and DCC-t models (b_1', b_2', η) can be sampled via Random Walk Metropolis-Hastings (RWMH) within Gibbs. Call $\boldsymbol{b} = (b_1', b_2')$, then the priors for the model parameters are $\eta \sim \mathcal{E}_{\eta>4}(\xi_\eta)$, $\boldsymbol{b} \sim \mathcal{N}_{2d}(0, V_b \cdot I_{2d})$. Here $\mathcal{E}(\cdot)$ is an exponential distribution, $\mathcal{N}_{2d}(\cdot)$ is a $2d$-variate Normal distribution. We sample $(\boldsymbol{b}, \eta)$ iteratively from (multivariate) Normal proposals given some starting values $(\boldsymbol{b}, \eta)^{(0)}$. The first elements of the vectors b_1', b_2' are restricted to be positive for identification purposes; the draws that result in non-positive semidefinite variance-covariance matrices are rejected, as well as the draws for which $B_1 + B_2$ are larger than one in modulus. For the DCC model with Gaussian copula, we have only parameter vectors (b_1', b_2').

DCC-HEAVY and DCC-HEAVY-t. The parameters for the DCC-HEAVY and DCC-HEAVY-t models (a, b, η) can be sampled via RWMH. The priors for the parameters (a, b) are assumed Beta so that $0 < a, b < 1$, and the prior for the degrees of freedom of the Student t distribution, η, is exponential. We sample (a, b) and η iteratively in two steps: (a, b) from a bivariate Normal proposal distribution

given some starting values $(a, b)^{(0)}$ and η from a truncated Normal given some starting value $\eta^{(0)}$. For the DCC-HEAVY model with Gaussian copula, we have only parameters (a, b).

AIW and CAW. For estimation of the AIW and CAW models, as in Jin & Maheu (2013, 2016), we use RWMH within Gibbs. We assume $L = 2$ (for AIW) and $p = q = 1$ (for CAW) and call $\boldsymbol{b} = (b_1', b_2')$. The priors for the model parameters are $\nu \sim \mathcal{E}_{\nu>d+1}(\xi_\nu)$, $\boldsymbol{b} \sim \mathcal{N}_{2d}(0, V_b \cdot I_{2d})$, $l_2 \sim \mathcal{U}_Z(a_l, b_l)$. Here $\mathcal{U}_Z(\cdot)$ is a discrete Uniform distribution. Given some starting values $(l_2, \nu, \boldsymbol{b})^{(0)}$, the algorithm iterates through the following for $m = 1, \ldots, M$:

1. Sample ν via RWMH from the conditional posterior:
   $p(\nu|l_2, \boldsymbol{b}, \text{Rcov}_{1:T}) \propto \pi(\nu) \prod_t g_{IW}(\text{Rcov}_t|l_2, \nu, \boldsymbol{b})$, where g_{IW} is the density function of the inverse-Wishart distribution and $\pi(\cdot)$ denotes a prior density.

2. Sample $\boldsymbol{b} = (b_1', b_2')$ via RWMH jointly from the $2d$-variate Normal proposal, where the first elements of each vector are truncated to be positive, for identification purposes. As in Jin & Maheu (2013, 2016) we reject such draws of $\boldsymbol{b}$ where B_0 is not positive definite, or the absolute value of any element of $\sum_{i=1}^{2} B_i$ is not smaller than one.

3. Sample l_2 via RWMH by using Poisson increments that can be either positive or negative with equal probability. This step is relevant only for the AIW model.

Pooling weights. Estimation of the static and naïve time-varying weights is straightforward and can be performed by applying the formulas in Eqs. (5) and (6) on the log predictive scores at each MCMC iteration after the estimation is performed for all models individually. For the time-varying persistent weights ω_t^{DN}, we use a variant of particle MCMC called particle marginal Metropolis-Hastings sampler (Andrieu et al. 2010). In particular, we use a bootstrap filter of Gordon et al. (1993) for the latent state s_t filtering and a standard MH step with Normal prior truncated at (-1,1) $\beta \sim \mathcal{TN}_{(-1,1)}(m_\beta, V_\beta)$ with a random walk proposal for the persistence parameter β.

3.2 Model selection

For illustration purposes, in this article we focus one-step-ahead density forecasts. One-step-ahead horizon has also been considered by Billio et al. (2013), Jin & Maheu (2013), for example. The fore-

casting procedure includes combining the one-step-ahead forecasts for the marginal models (see Appendix C.2 in the Online Supplementary Material for details regarding the specification and estimation of the HAR model of Corsi 2009) and the one-step-ahead predictive copula density. The one-step-ahead joint density for the returns is then given by $h(r_{t+1}|r_{1:t}) = c(F_i(r_{1,t+1}),\ldots,F_d(r_{d,t+1})|r_{1:t}) \cdot \prod_{i=1}^{d} f_i(r_{i,t+1}|r_{1:t})$, where $F_i(\cdot)$ and $f_i(\cdot)$ are the marginal distribution and density functions of $r_{i,t}$, and the copula density can be compactly re-written as $c(u_{t+1}|z_{1:t})$.

For the Static model, the marginal predictive is available analytically:

$$c^{\text{static}}(u_{t+1}|z_{1:t}) = x^{-1} t_{d,\nu_0+t-d+1}\left(z_{t+1}\left|\frac{I_d(\nu_0-d-1)+z'_{1:t}z_{1:t}}{\nu_0+t-d+1}\right.\right),$$

where $z_{t+1} = (\Phi^{-1}(u_{1,t+1}),\ldots,\Phi^{-1}(u_{d,t+1}))'$ and $x = \prod_{i=1}^{d}\phi_1(z_{i,t+1})$. Of note, the marginal predictive for the Static model is the Student t density, which is a result of the Normal-inverse Wishart conjugacy. For the DCC and DCC-HEAVY models, with Gaussian and t copulas, the posterior predictive distributions are given by

$$\begin{aligned} c^{\text{DCC}}(u_{t+1}|z_{1:t},\theta_{DCC}) &= x^{-1}\phi_d(z_{t+1}|\boldsymbol{R}_{t+1}(\theta_{DCC})), \\ c^{\text{DCCt}}(u_{t+1}|z_{1:t},\theta_{DCCt}) &= \left(\prod_{i=1}^{d} t_{1,\eta}(T_{1,\eta}^{-1}(u_{i,t+1}))\right)^{-1} t_{d,\eta}(u_{t+1}|\boldsymbol{R}_{t+1}(\theta_{DCCt})). \end{aligned}$$

Here, θ_{DCC} and θ_{DCCt} are the estimated parameters for the DCC/DCC-HEAVY and DCC/DCC-HEAVY-t models. Finally, the posterior predictive density for the AIW and CAW models is:

$$c^{\text{AIW/CAW}}(u_{t+1}|z_{1:t},\theta_{AIW/CAW}) = x^{-1} t_{d,v-d+1}\left(z_{t+1}\left|\frac{v-d-1}{v-d+1}V_{t+1}\right.\right),$$

where $\theta_{AIW/CAW}$ is a vector of the parameters in either AIW or CAW model. The marginal predictive densities $p(u_{t+1}|z_{1:t})$ that account for parameter uncertainty for the DCC, DCC-t, DCC-HEAVY, DCC-HEAVY-t, and AIW models can be obtained by using the MCMC output:

$$c(u_{t+1}|z_{1:t}) = \int c(u_{t+1}|z_{1:t},\theta)p(\theta|z_{1:t})d\theta \approx \frac{1}{M}\sum_{m=1}^{M} c(u_{t+1}|z_{1:t},\theta^{(m)}),$$

where $(\theta^{(1)},\ldots,\theta^{(M)})$ are the M posterior samples obtained from the MCMC.

The model comparison is carried out via predictive Bayes factors (BF) given K out of sample observations, where T is the sample size used for estimation. The BF between model 0 ($\mathcal{M}_0$) and model 1 ($\mathcal{M}_1$) is defined as (West 1986, Kass & Raftery 1995):

$$BF_{T:T+K} = \frac{c(u_{T:T+K}|z_{1:T}, \mathcal{M}_0)}{c(u_{T:T+K}|z_{1:T}, \mathcal{M}_1)},$$

where $c(u_{T:T+K}|z_{1:T}, \mathcal{M}_r) = \prod_{k=1}^{K} c(u_{T+k}|z_{1:T+k-1}, \mathcal{M}_r)$. The exact calculation of $c(u_{T:T+K}|z_{1:T}, \mathcal{M}_r)$ is time-consuming because of an expanding time horizon, i.e., the model must be re-estimated K times. For notational convenience, we omit conditioning on the model $\mathcal{M}_r$; and instead of conditioning on $z_{1:T}$ we condition on $u_{1:T}$, because $u_{i,t} = \Phi(z_{i,t})$. Then we can write:

$$\begin{aligned}
c(u_{T:T+K}|u_{1:T}) &= \prod_{k=1}^{K} c(u_{T+k}|u_{1:T+k-1}) \\
&= \prod_{k=1}^{K} \int c(u_{T+k}|\theta) p(\theta|u_{1:T+k-1}) d\theta \\
&\overset{\text{T large}}{\approx} \prod_{k=1}^{K} \int c(u_{T+k}|\theta) \hat{p}(\theta) d\theta, \quad \text{where} \quad \theta^{(1)}, \ldots, \theta^{(M)} \sim \hat{p}(\theta), \\
&\approx \prod_{k=1}^{K} \frac{1}{M} \sum_{m=1}^{M} c(u_{T+k}|\theta^{(m)}).
\end{aligned}$$

The marginal predictive distribution of $u_{T:T+K}$ can be approximated by using a posterior sample of estimated model parameters until time T (instead of re-estimating the model K times) with density $\hat{p}(\theta)$.

Another necessary measure used for calculating the pooling weights is the log predictive score (LPS):

$$LPS = \sum_{t=T}^{T+K-1} \log c(u_{t+1}|z_{1:t}). \tag{8}$$

Finally, we also compare the predictive model performance for the lower q^* percentile. Similar metrics have also been considered by Delatola & Griffin (2011) and Opschoor et al. (2021), among others. We define the log predictive tail score ($LPTS$) measure as follows:

$$LPTS_{q^*} = \sum_{t=T}^{T+K-1} I[u_{t+1} < q] \times \log c(u_{t+1}|z_{1:t}),$$

where q is a $d \times 1$ vector, and $I[u_{t+1} < q] = \prod_{i=1}^{d} I[u_{i,t+1} < q_i]$ with $q_i \in [0,1]$. Here $I[a]$ denotes the indicator function, which equals 1 if condition a is fulfilled and 0 otherwise. We select $q = [q_1, \ldots, q_d]$ such that $K^{-1} \sum_{t=T}^{T+K-1} I[u_{t+1} < q] = q^*$, for $q^* = 0.05, 0.10, 0.25$ (Opschoor et al. 2021). That is, we examine the LPS from Eq.(8) only when the d-variate data are jointly in the lower region $[0, q_1] \times \ldots \times [0, q_d]$.

4 Empirical Application

4.1 Data description

The intraday exchange rate data are from the `histdata.com` website. The data are from 2008/01/03 to 2023/01/31, and, after treatment, contain 3881 daily data points in total. Similar to Noureldin et al. (2012) we remove the first and last 15 minutes trading to avoid potential overnight effects, which should be minimal given continuous trading information. Overall, the overnight effect does not pose a problem as the forex market trading takes place 24 hours on weekdays and our high-frequency data is recorded without any breaks [3]. High frequency returns and the realized covariance measures are extracted by using the `highfrequency` package in `R` (Boudt et al. 2022) and employing the 10-minute sampling with 2-minute subsampling estimator of Zhang et al. (2005), similar to Noureldin et al. (2012). For detailed high frequency data treatment, refer to Appendix B.1 in the Online Supplementary Material. The dataset contains 14 exchange rates: EUR/USD, EUR/GBP, EUR/JPY, EUR/AUD, USD/CAD, USD/JPY, GBP/JPY, GBP/USD, AUD/USD, EUR/CAD, AUD/CAD, CAD/JPY, GBP/AUD, GBP/CAD. Table 1 presents the descriptive statistics for the crude returns (open-to-close). All of the assets present excess kurtosis, indicating fat-tailed unconditional distribution. The Kolmogorov-Smirnov and Jarque-Bera tests for Normality reject the Normally distributed returns for all assets and all confidence levels.

[3] However, when dealing for instance with stock market data, one should consider that stock markets are open only for a certain number of hours per day. In this case, the overnight effect might become an important factor in the forecasting ability of the model, and should be carefully accounted for, see for example Ahoniemi & Lanne (2013).

Ljung-Box Q-test for autocorrelation rejects the absence of autocorrelation in the mean for more than half of the exchange rates and the ARCH test rejects the absence of heteroscedasticity for all exchange rates.

Table 1: Descriptive statistics for the log return data and p-values for Kolmogorov-Smirnov (KS), and Jarque-Bera (JB) tests for Normality, Ljung-Box Q-test for autocorrelation and ARCH test for heteroscedasticity for lag 10.

	mean	median	sd	skew.	kurt.	KS	JB	LB(10)	ARCH(10)
EURUSD	-0.01	-0.00	0.58	0.14	5.36	0.0000	0.0000	0.5850	0.0000
EURGBP	0.01	0.01	0.53	0.11	5.45	0.0000	0.0000	0.2874	0.0000
EURJPY	0.01	0.01	0.73	-0.05	11.93	0.0000	0.0000	0.1004	0.0000
EURAUD	0.00	-0.02	0.68	0.79	12.14	0.0000	0.0000	0.1673	0.0000
USDCAD	0.01	0.01	0.56	0.09	6.14	0.0000	0.0000	0.0010	0.0000
USDJPY	0.01	0.01	0.62	-0.03	8.59	0.0000	0.0000	0.3006	0.0000
GBPJPY	0.01	0.02	0.79	-0.31	11.06	0.0000	0.0000	0.0157	0.0000
GBPUSD	-0.00	-0.00	0.60	-0.21	6.17	0.0000	0.0000	0.0070	0.0000
AUDUSD	0.01	0.03	0.81	-0.58	12.74	0.0000	0.0000	0.0016	0.0000
EURCAD	0.01	0.01	0.57	0.13	5.44	0.0000	0.0000	0.8227	0.0000
AUDCAD	0.02	0.04	0.58	-0.24	13.03	0.0000	0.0000	0.0008	0.0000
CADJPY	0.01	0.02	0.85	-0.39	9.55	0.0000	0.0000	0.0147	0.0000
GBPAUD	0.00	0.00	0.69	0.47	10.47	0.0000	0.0000	0.0040	0.0000
GBPCAD	0.02	0.02	0.58	-0.09	5.38	0.0000	0.0000	0.0000	0.0000

To preserve space, the first set of descriptive plots is for EUR/USD and EUR/GBP returns. Figure 2 draws the log returns together with the realized standard deviations, QQ-plots and histograms for the standardized returns $(z_{1,t}, z_{2,t})$ against the Normal distribution, and the probability integral transforms against the Uniform distribution. As described in Section 2, standardized returns $z_{i,t}$ are obtained by dividing the log returns $r_{i,t}$ by the corresponding realized volatilities and then de-meaning and scaling by σ_i. For all exchange rates, the scaling factor σ_i is less than one, meaning that the estimated realized volatilities have to be scaled down. As seen from the plots, the time series data includes calm and volatile episodes. The QQ-plots indicate that the data, standardized by the RV measure, are approximately Normally distributed, as shown by Andersen et al. (2000, 2001). This is also confirmed by the probability integral transforms of the standardized returns $u_{i,t} = \Phi(z_{i,t})$, which are Uniformly distributed. The corresponding plots for all 14 exchange rates, as well as the main descriptive statistics for the standardized log returns, can be found in Online Supplementary Material Appendix B.2.

Finally, even though the in-sample returns are modeled non-parametrically, in order to perform individual marginal density forecasts we specify a log-HAR(1,5,22) model of Corsi (2009) for the RV dynamics. According to the six criteria outlined in Hansen & Lunde (2006), as well Mincer Zarnowitz

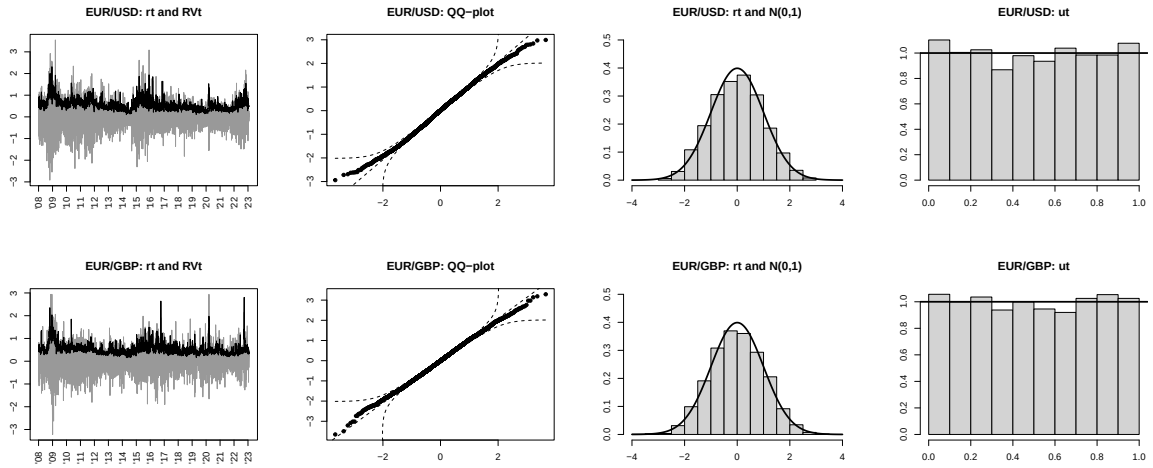

Figure 2: First column: log returns (open-to-close, in gray) and the square root of realized volatility (in black). Second column: QQ-plots of the standardized returns against the Normal distribution with 95% confidence intervals (dashed line). Third column: histograms of the standardized returns against the standard Normal distribution. Last column: histograms of the probability integral transforms against the Uniform density for EUR/USD (top row) and EUR/GBP (bottom row) exchange rates.

regressions, the log-HAR(1,5,22) produces the most precise one-step-ahead point forecasts. Therefore, we use this model in the portfolio allocation exercise for forecasting the realized volatility of each individual return series. All details are in Online Supplementary Material Appendix C.2.

4.2 Prior specification and estimation

The prior hyperparameters for the variance-covariance matrix in the static model are set to $\Omega \sim \mathcal{IW}(I_d, 10)$; for the DCC-HEAVY-t, the priors are $a \sim \mathcal{B}(3,10)$, $b \sim \mathcal{B}(10,3)$, $\eta \sim \mathcal{E}(0.01)$, where $\mathcal{B}(a,b)$ is the Beta distribution with shape parameters a and b. The priors for the rank-1 DCC, DCC-t, AIW and CAIW models are $\eta \sim \mathcal{E}_{\eta>4}(0.01)$, $\nu \sim \mathcal{E}_{\nu>d+1}(0.01)$, $\boldsymbol{b} \sim \mathcal{N}_{2d}(0, 10 \cdot I_{2d})$ and $l_2 \sim \mathcal{U}_Z(2,100)$. In general, all priors are somewhat uninformative but proper. The size of the MCMC chain is $M = 100k$ for all models; the first half is retained as burn-in, and thinning is performed every 50th observation from the second half, thus resulting in posterior samples of 1000 observations. For the RWMH steps, the proposal variances are adjusted such that the acceptance rate is approximately 0.5 for univariate parameter vectors and 0.10 to 0.30 for multivariate parameter vectors. For sampling l_2 Poisson increments have a rate parameter equal to either 1.5 or 2, depending on the acceptance probability. All MCMC chains have converged after $100k$ iterations. Online Supplementary Material

Appendix C.1 contains parameter estimation results.

4.3 Full model forecasting results

For estimation, we use 12 years of data, plus three more years for the out-of-sample performance evaluation. In particular, the data used for estimation are from 2008/01/03 to 2019/12/31 (3084 data points), and the out-of-sample evaluation period is from 2020/01/02 to 2023/01/31 (797 data points). Table 2 presents the average LPS for the $K = 797$ out of sample observations for three low frequency data-based models, two high frequency models, and the natural competitor DCC-HEAVY model with t copula (results for the Gaussian copula are not reported, because of the considerably poorer performance). According to the LPS, the DCC-t model performs best among the low frequency data-based models, and CAW performs best among the high frequency data-based models, and a competitor DCC-HEAVY-t performs best overall. Figure 3 draws expanding-window predictive log BFs for each of the models, wherein the Static model is the benchmark. Positive BFs indicate that the model outperforms the Static specification. Not surprisingly, all time-varying models provide superior out-of-sample density forecasts.

Table 2: 1-step-ahead LPS for all individual models: Static, DCC, DCC-t, AIW, CAW, and DCC-HEAVY model with t copula for 2020/01/02 - 2023/01/31 out-of-sample period (K = 797 observations).

Static	DCC	DCC-t	AIW	CAW	DCC-HEAVY-t
-5163.32	-3500.67	-3155.91	-2841.37	-2774.60	-2702.95

Next, we perform the density forecast density combination exercise, as described in Section 2.3. As seen in Figure 3, model preference is non-constant, therefore, instead of choosing a single model for density forecasting, we combine predictive densities by using several alternative weighting schemes. We combine the DCC-t and CAW models, which are the best models in the LF and HF model classes. In fact, the results reported in Online Supplementary Material Appendix D.1 show that combining any other LF model with the HF CAW model still yields superior forecasts as compared to the best individual model, the CAW. This is not the case for model pools within the same frequency, which perform systematically worse. According to those results, we conclude that the predictive gains are primarily originating from the incorporation of low- and high frequency information and not from the

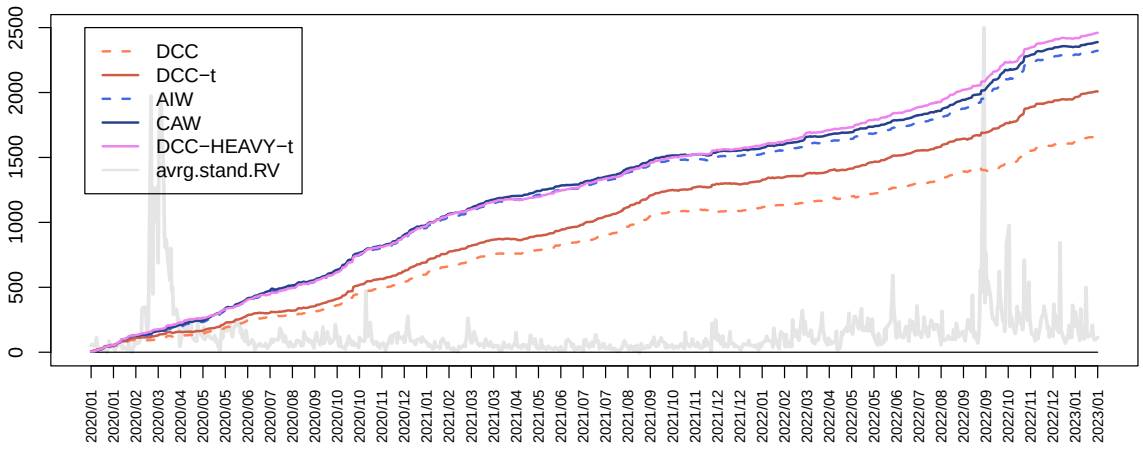

Figure 3: Expanding-window predictive log Bayes factors with a Static model as a benchmark for all individual models: DCC, DCC-t, AIW, CAW, and DCC-HEAVY model with t copula for 2020/01/02 to 2023/01/31 out-of-sample period ($K = 797$ observations). Average standardized realized volatility (in gray) in the background.

model pooling itself.

Top plot of Figure 4 shows the posterior average of the weights for the HF component (CAW model) for the four weighting schemes: Geweke's as in Eq. (5), Jore's with $\tilde{m} = \{1, 10\}$ as in Eq. (6) and Del Negro (DN) as in Eq. (7). Jore's weights are more volatile because they take into consideration only the last $\tilde{m}$ observations, whereas Geweke's weight takes into consideration the entire out of sample period until the time when the weights are calculated, and reaches a seemingly stable level of approximately 0.7. Drawing the 95% credible intervals around Geweke's weight indicates that the HF component weight is different from 0.5. The weight dynamics in Del Negro's pool follow an AR(1) representation, which, in a sense, gives them a longer memory compared to Jore1's weights, making them smoother; however, both follow similar patterns.

The bottom plot of Figure 4 draws expanding-window predictive log BFs for density combinations and individual models, with the CAW model as the benchmark. Geweke's and equal weights show the poorest performance, mainly because the weights are not re-balanced to adjust to a rapidly changing environment. Del Negro's scheme performs better than Geweke's and equal weights, but not as well as Jore1. Even though DN and Jore1 move in very similar patterns, the more extreme movements of Jore1's weights appear to be the source of the superior forecasting performance. However, independently on the weighting scheme, all pooled models produce significant improvements in one-

step-ahead density forecasts over individual LF and HF models, and over the natural mixed-frequency competitor - DCC-HEAVY-t model.

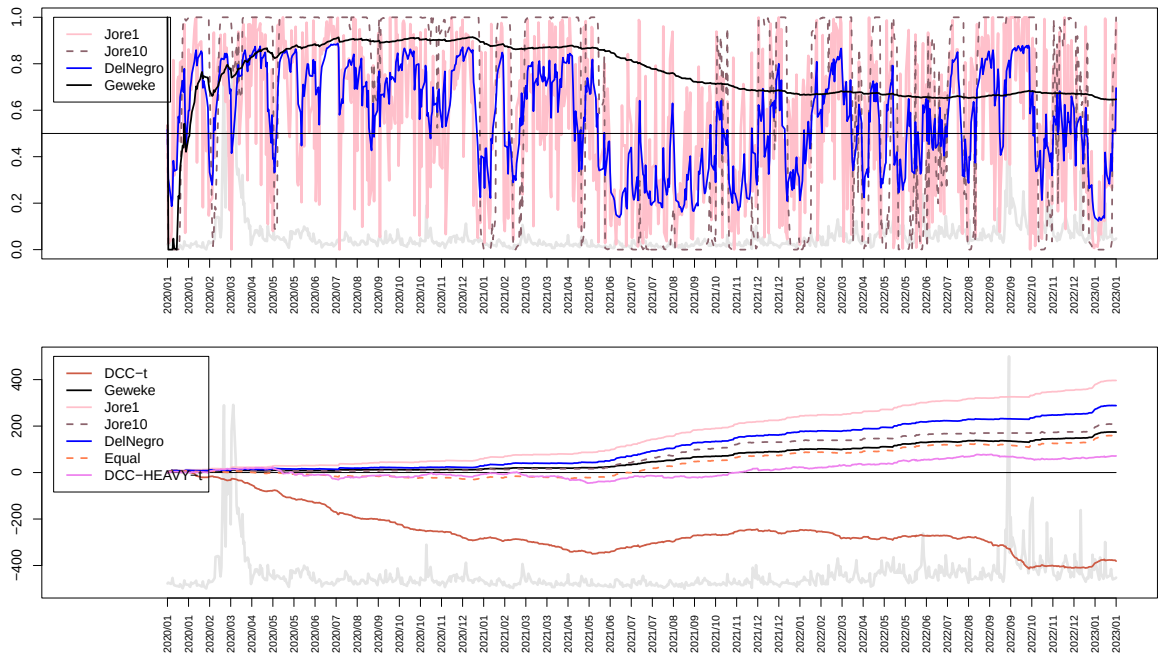

Figure 4: Results for pooling the high frequency CAW-Gaussian and low frequency DCC-t copulas. Top plot: posterior means of the high frequency component weight for different weighting schemes. Bottom plot: expanding-window predictive log Bayes factor for density combinations and individual models, with CAW model as the benchmark, for 2020/01/02 to 2023/01/31 out-of-sample period ($K = 797$ observations). Average standardized realized volatility (in gray) in the background.

Next, Figure 5 draws the posterior densities for the *average per observation* out-of-sample LPS and $LPTS_{q^*}$ for lower 25%, 10% and 5% quantiles, only for some individual and some pooled models (for ease of readability of the graph). The posterior densities also indicate whether the differences in these average LPS and $LPTS$ are statistically significant. The top left plot indicates that the difference in average overall LPS is statistically significant, pooled models provide the best predictive out-of-sample performance, and the DCC-t model has the poorest performance. The results change somewhat within the 25% lower quantile (top right plot). Here, the model ordering changes, but Jore1's scheme still continues to perform best. Finally, in the 10% and 5% lower quantiles (bottom row), Jore1's pooling scheme continues to perform the best, but with overlapping intervals from the competitor DCC-HEAVY-t model. These results show that different models perform differently depending on the metric used (whole distribution vs the tail of the distribution) and confirms the general consensus

that finding a universally best model is conceptually impossible. In this article, the preferred model is characterized as that providing the best forecast as measured by the log predictive score, because we are interested in the entire predictive distribution of the returns. Nonetheless, if one is interested exclusively in the tails, for example, the log predictive tail score would be a more appropriate metric for calculating the pooling weights. For example, Kapetanios et al. (2015) have proposed to model weights dependent on some variable of interest, which could be some measure related to the lower region of a predictive density.

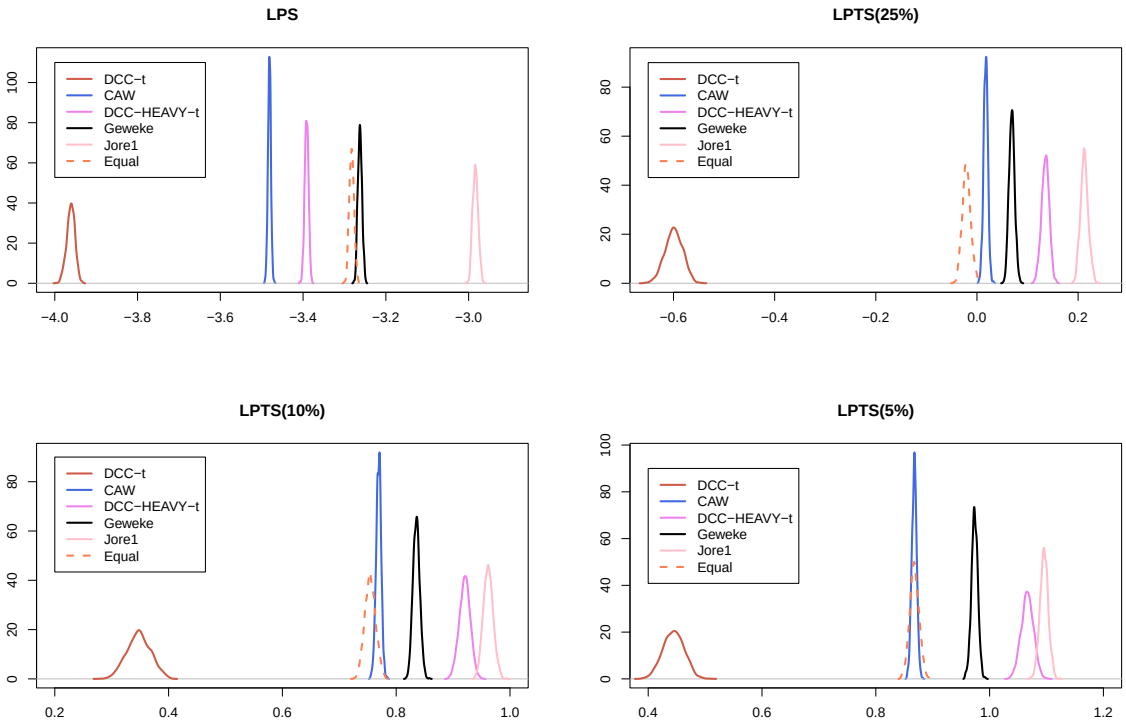

Figure 5: Posterior densities for the average per observation one-step-ahead LPS and $LPTS$ for the lower 25%, 10% and 5% quantiles for 2020/01/02 to 2023/01/31 out-of-sample period ($K = 797$ observations). CAW and DCC-t are the high- and low frequency models, and the pooled models are according to Geweke's, Jore's and equally weighted schemes.

Finally, we sought to determine whether the preference for the high frequency model might correlate with overall market conditions, proxied by the market volatility. The preference for the high frequency model is measured as the pooling weight of the high frequency component in various pooling schemes. We also consider the difference between the predictive log likelihoods between the two best models: CAW and DCC-t. Positive values would indicate that the CAW model is preferred, whereas

negative values would indicate that the CAW model is outperformed by the DCC-t. As a proxy for the market volatility, we take the average standardized realized volatility over the 14 exchange rates. As alternative proxies, we also consider the equally weighted market portfolio realized volatility, and daily VIX index[4] for the corresponding period. The top part of Table 3 shows that all three volatility proxies are highly correlated, as expected. The bottom part of Table 3 reports the posterior medians of the sample correlation coefficients between the preference for the high frequency weight and the market volatility proxies. Except for Geweke's weights[5], the rest are positively correlated, meaning that the preference for the HF model is positively correlated with the market volatility. The correlations are of varying magnitudes, reaching as high as 0.37 in VIX-Del Negro pair, and all of them are statistically significant in the sense that the 95% posterior credible intervals exclude zeros (posterior percentile tables are available in the Appendix C.3 in the Online Supplementary Material). We argue that, as the market volatility increases, the high frequency data based models are more agile in responding to the fast-changing market conditions. On the other hand, when the markets are calm the low frequency data-based models provide more stable estimates of the co-dependence structure. Such results have important implications from the investor's perspective: investors anticipating bull (or bear) market conditions, might choose to rely on low frequency (or high frequency) models to produce their density forecasts.

Table 3: Posterior medians of sample correlations between three proxies for the market volatility and the preference for high-frequency model for 2020/01/02 to 2023/01/31 out-of-sample period ($K = 797$ observations). The proxies for the market volatility are: average standardized realized volatility (avrg RV), equally weighted market portfolio realized volatility (Mkt:eql) and VIX index. The preference for the high-frequency model is measured as a high-frequency component weight in various pooling schemes as well as the difference between the daily log likelihood (logLik diff) between the CAW and DCC-t models.

	avrg RV	Mkt:eql	VIX	Geweke	Jore1	DelNegro	logLik diff
avrg RV	1.000						
Mkt:eql	0.914	1.000					
VIX	0.646	0.538	1.000				
Geweke	-0.094	-0.161	0.165	1.000			
Jore1	0.094	0.098	0.219	0.159	1.000		
DelNegro	0.149	0.159	0.365	0.317	0.702	1.000	
logLik diff	0.090	0.117	0.179	0.126	0.903	0.635	1.000

[4]VIX is the Chicago Board Options Exchange's volatility index, based on S&P500 index options.

[5]Geweke's weights are bound to converge almost surely to some stable equilibrium (Geweke & Amisano 2011), therefore, the correlation between those weights and time-varying market volatility should be interpreted with caution.

4.4 Portfolio allocation exercise

Next, we are interested in quantifying how the pooling of LF and HF models translates to a better performing portfolios in terms of economic gains. To this end, we consider two portfolios: the Global Minimum Variance (GMV) and the minimum Conditional Value-at-Risk (CVaR) portfolios, both with short-sale constraints. GMV portfolio takes into account only the second moment of the return distribution as it minimizes the overall variance of the portfolio. On the other hand, the CVaR portfolio is tail-oriented, as it aims to minimize the expected losses in the lower tail of the return distribution, usually targeting losses in the lower 1%, 5%, or 10th percentile (Rockafellar & Uryasev 2002). We note that, even though we have the explicit form of the K one-step-ahead forecast densities for the 14-variate exchange rate series at each MCMC iteration, given the complex objective function (in the CVaR case) and short-sale restrictions, we need to rely on numerical methods for portfolio optimization. Therefore, we use a similar approach to those in Ausín & Lopes (2010) and Opschoor et al. (2021), wherein at each MCMC iteration, and for each out-of-sample point, we draw N replications from the 14-variate predictive distribution, where N is a large number.[6] Given this simulated data, we then can calculate the one-step-ahead GMV and CVaR portfolio weights. The procedure can be summarized as follows. For each $m = 1, \dots, M$ and for each $k = 1, \dots, K$:

1. Simulate N replications of $u_{T+k}^{(m)}$ from $c(u_{T+k}|z_{1:T+k-1}, \theta^{(m)})$, where $\theta^{(m)}$ is the collection of model parameters. Then transform the Uniformly distributed data to predictive returns $r_{T+k}^{*(m)}$ via the corresponding quantile function. Because in the copula setting, the modeling of the marginals is performed separately from the dependence structure, all predictive returns have the same marginals (across models, not across assets). The realized volatility forecasts are obtained using the log-HAR(1,5,22) model of Corsi (2009), as described in the Online Supplementary Material Appendix C.2.

2. Given N simulated scenarios of predictive returns $r_{T+k}^{*(m)}$ perform the portfolio optimization numerically to obtain the portfolio weights $w_{GMV,T+k}^{(m)}$ and $w_{CVaR,T+k}^{(m)}$. For the CVaR portfolio we consider two levels of risk: lower 5% and 10% quantiles. Optimization was performed using `R` package `NMOF` (Gilli et al. 2019, Schumann 2023).

[6] $N = 10,000$ in our case.

3. Given the estimated predictive optimal portfolio weights $w^{(m)}_{GMV,T+k}$, $w^{(m)}_{CVaR,T+k}$ and the actual *ex post* observed returns r_{T+k}, calculate *ex post* portfolio returns and related metrics.

Ideally, for each MCMC iteration, one would calculate the $K = 797$ sequence of the realized portfolio returns. However, due to extreme computational cost associated with return simulation and numerical optimization, we have performed the above procedure *at the posterior medians* of the estimated parameter values $(\theta^{(1)}, \ldots, \theta^{(M)})$. Finally, given the *ex post* observed returns r_{T+k}, we can obtain the average portfolio return, overall portfolio standard deviation, adjusted[7] Sharpe ratio (the ratio between the expected return and the standard deviation), the average return in the empirical 5% and 10% quantiles (rCVaR05 and rCVaR10), average turnover (TO) and average concentration (CO), see Opschoor et al. (2021) for details.

Table 4 reports the portfolio metrics at the posterior medians of the estimated parameters. The best-performing portfolio is in gray and the second best - underlined. For the GMV portfolio (first eight rows) Jore1's pooling scheme provides the lowest annualized portfolio standard deviation as well as the highest Sharpe ratio. The annualized portfolio standard deviations are rather low, however, exchange rate data usually exhibits low volatility in absolute terms as compared to the conventional assets, see for example Patton (2006b), Gong et al. (2022) among many others. For the CVAR05 and CVaR10 portfolios we are more concerned with the behavior in the lower tail, i.e., the *ex post* realized CVaRs. In both portfolios, Jore1's pool is always the best.

Following Opschoor et al. (2021) we asses the differences of the standard deviation values by means of the Model Confidence Set (MCS) procedure of Hansen et al. (2011) and provide the p-values in Table 4. Here the best performance corresponds to the lowest standard deviation. The hypothesis of equal performance is rejected for all confidence levels for the Jore1 weighting scheme under all portfolio objectives: GMV, CVaR05 and CVaR10. However, other models such as DCC-HEAVY-t, CAW and Geweke perform reasonably well and are often among the superior ones at 20% level. This result is not surprising. As pointed out by Opschoor et al. (2021) "To reconcile the findings in terms of economic performance with those of the density forecast evaluation (...), it is important to note that

[7]For calculating the Adjusted Sharpe ratio we use the 10-year U.S. Treasury yield as a proxy for the risk-free interest rate, data from Federal Reserve Bank of St. Louis (FRED).

Table 4: Portfolio allocation results based on 1-step-ahead predictions for 2020/01/02 to 2023/01/31 out-of-sample period (K = 797 observations). The three portfolios are: Global Minimum Variance (GMV), and minimum Conditional Value at Risk for lower 5 and 10 percentiles (CVaR05 and CVaR10), all with short-sale constraints. The table reports average portfolio return (avrg), overall standard deviation in % (sd), the p-value corresponding to the model confidence set of Hansen et al. (2011) using a significance level of 5%, adjusted Sharpe ratio (SR), realized Conditional Value at Risk for lower 5 and 10 percentiles (rCVaR05 and rCVaR10), turnover (TO), and concentration (CO), all quantities annualized, for the pooled models (Geweke's, Jore's and equally weighted), two best individual models (CAW and DCC-t) and a competitor model (DCC-HEAVY-t). In gray is highlighted the best performing portfolio and the second best is underlined.

		Jore1	Geweke	Equal	CAW	DCC-t	DCC-HEAVY-t
GMV	avrg	6.766	6.734	6.790	6.770	6.752	6.687
	sd	1.334	1.356	1.357	1.345	1.368	1.345
	(p-value)	(1.000)	(0.265)	(0.111)	(0.560)	(0.178)	(0.905)
	SR	3.708	3.627	3.665	3.682	3.606	3.621
	rCVaR05	-40.488	-40.758	-40.508	-40.760	-41.947	-41.194
	rCVaR10	-31.260	-31.422	-31.240	-31.504	-31.763	-31.527
	CO	0.466	0.465	0.466	0.463	0.470	0.469
	TO	0.419	0.417	0.420	0.409	0.416	0.411
CVaR05	avrg	6.856	6.782	6.867	6.835	6.781	6.718
	sd	1.337	1.361	1.364	1.351	1.373	1.351
	(p-value)	(1.000)	(0.199)	(0.043)	(0.381)	(0.133)	(0.660)
	SR	3.768	3.648	3.701	3.713	3.615	3.628
	rCVaR05	-40.430	-40.878	-40.644	-40.882	-42.110	-41.494
	rCVaR10	-31.061	-31.385	-31.314	-31.458	-31.812	-31.667
	CO	0.467	0.465	0.467	0.463	0.471	0.470
	TO	0.419	0.418	0.420	0.410	0.417	0.412
CVaR10	avrg	6.839	6.787	6.850	6.827	6.773	6.730
	sd	1.340	1.359	1.365	1.349	1.374	1.353
	(p-value)	(1.000)	(0.322)	(0.040)	(0.552)	(0.154)	(0.654)
	SR	3.746	3.657	3.686	3.712	3.607	3.631
	rCVaR05	-40.641	-40.985	-40.847	-40.953	-42.219	-41.567
	rCVaR10	-31.230	-31.412	-31.421	-31.541	-31.879	-31.675
	CO	0.467	0.465	0.467	0.463	0.471	0.470
	TO	0.419	0.417	0.420	0.410	0.417	0.412

the GMVP evaluation takes a very specific perspective. The GMVP focuses on an area of the forecast distribution where differences are more concentrated by design: all models focus on a portfolio with ex-ante minimum variance. If the different models are any good, differences in this concentrated performance measure are harder to obtain."

The table also reports the annualized average return, average turnover and average concentration. In all portfolios the equal weighting scheme produces the highest average portfolio return. In terms of turnover and concentration the high frequency data based model is always the best. However, it is important to note, that these measures are not directly targeted by the investor, and performing model ranking based on these metrics should be done with caution. Notably, the low frequency DCC-t model

not only has the lowest log predictive score and tail score, but it is always the worst for all portfolios and all metrics.

To sum up, the portfolio allocation results show how superior model performance in terms of log predictive score translates into superior portfolios; also, it confirms favorable economic outcomes for density pooling, as compared with the individual models.

5 Discussion and Conclusion

In this article, we propose a mixed frequency copula-based approach for forecasting the co-dependence between financial returns by using information arising from data sampled at different frequencies. In particular, we pool two copula densities, wherein the parameters are obtained from low- and high frequency data. For the high frequency copula parameter, we use a realized correlation measure. We model the dynamics of the realized variance-covariance matrices via two alternative high frequency data based specifications - AIW and CAW - with a Gaussian copula; meanwhile, for the low frequency dependence structure, we consider three standard models: Static, DCC with Gaussian copula, and DCC with t copula.

In the empirical application using more than 15 years of exchange rate data we show that even though the overall log predictive scores favor the CAW model, incorporating information arising from the low frequency data improves the forecasting outcomes. In addition, the density pool shows an improvement in predictive performance over those of other mixed frequency models, such as the natural competitor the DCC-HEAVY model. We also show that the gains arise not from density pooling itself, but from pooling different frequencies. Finally, the portfolio allocation exercise illustrates how the use of pooled models directly translates into favorable economic outcomes.

For future research, focusing on a more flexible pooling scheme, such as Bayesian predictive synthesis (McAlinn 2021) or infinite pools (Jin et al. 2022), is worth examining further. Ideally, one would also consider pooling marginal volatilities from high and low frequency data based models, however, this goes beyond the scope of the paper.

Acknowledgments

References

Ahoniemi, K. & Lanne, M. (2013), 'Overnight stock returns and realized volatility', *International Journal of Forecasting* **29**(4), 592–604.

Andersen, T. G., Bollerslev, T., Diebold, F. X. & Ebens, H. (2001), 'The distribution of realized stock return volatility', *Journal of Financial Economics* **61**(1), 43–67.

Andersen, T. G., Bollerslev, T., Diebold, F. X. & Labys, P. (2000), 'Exchange rate returns standardized by Realized Volatility are (Nearly) Gaussian', *NBER Working Paper Series* **7488**, 1–21.

Andersen, T. G., Bollerslev, T., Diebold, F. X. & Labys, P. (2003), 'Modeling and forecasting realized volatility', *Econometrica* **71**(2), 579–625.

Andrieu, C., Doucet, A. & Holenstein, R. (2010), 'Particle Markov Chain Monte Carlo methods', *Journal of the Royal Statistical Society Series B: Statistical Methodology* **72**(3), 269–342.

Asai, M., McAleer, M. & Yu, J. (2006), 'Multivariate Stochastic Volatility: A Review', *Econometric Reviews* **25**(2-3), 145–175.

Ausín, M. C. & Lopes, H. F. (2010), 'Time-Varying Joint Distribution Through Copulas', *Computational Statistics and Data Analysis* **54**(11), 2383–2399.

Barndorff-Nielsen, O. E. & 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* **66**(2), 253–280.

Barndorff-Nielsen, O. E. & Shephard, N. (2004), 'Econometric analysis of realized covariation: high frequency covariance, regression and correlation in financial economics', *Econometrica* **72**(3), 885–925.

Bassetti, F., Casarin, R. & Ravazzolo, F. (2018), 'Bayesian nonparametric calibration and combination of predictive distributions', *Journal of the American Statistical Association* **113**(522), 675–685.

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

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

Billio, M., Casarin, R., Ravazzolo, F. & Van Dijk, H. K. (2013), 'Time-varying combinations of predictive densities using nonlinear filtering', *Journal of Econometrics* **177**(2), 213–232.

Boudt, K., Kleen, O. & Sjørup, E. (2022), 'Analyzing intraday financial data in R: The highfrequency package', *Journal of Statistical Software* **104**(8), 1–36.

Chiriac, R. & Voev, V. (2011), 'Modelling and Forecasting Multivariate Realized Volatility', *Journal of Applied Econometrics* **26**(6), 922–947.

Clemen, R. T. & Winkler, R. L. (2007), Aggregating probability distributions, *in* E. Ward, R. Miles & D. von Winterfeldt, eds, 'Advances in Decision Analysis: From Foundations to Applications', Cambridge University Press, pp. 154–176.

Corsi, F. (2009), 'A Simple Approximate Long-Memory Model of Realized Volatility', *Journal of Financial Econometrics* **7**(2), 1–23.

Del Negro, M., Hasegawa, R. B. & Schorfheide, F. (2016), 'Dynamic prediction pools: An investigation of financial frictions and forecasting performance', *Journal of Econometrics* **192**(2), 391–405.

Delatola, E.-I. & Griffin, J. E. (2011), 'Bayesian Nonparametric Modelling of the Return Distribution with Stochastic Volatility', *Bayesian Analysis* **6**(4), 901–926.

Dias, A. & Embrechts, P. (2004), 'Dynamic copula models for multivariate high-frequency data in finance', *Manuscript, ETH Zurich* **81**(November), 1–42.

Ding, Z. & Engle, R. F. (2001), 'Large Scale Conditional Covariance Matrix Modeling, Estimation and Testing', *NYU Working Paper No. FIN-01-029* .

Engle, R. (2002*a*), 'Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models', *Journal of Business and Economic Statistics* **20**(3), 339–350.

Engle, R. F. (2002*b*), 'New Frontiers for ARCH Models', *Journal of Applied Econometrics* **17**(5), 425–446.

Fan, Y. & Patton, A. J. (2014), 'Copulas in econometrics', *Annual Review of Economics* **6**, 179–200.

Fengler, M. R. & Okhrin, O. (2016), 'Managing risk with a realized copula parameter', *Computational Statistics and Data Analysis* **100**, 131–152.

Geweke, J. & Amisano, G. (2011), 'Optimal prediction pools', *Journal of Econometrics* **164**(1), 130–141.

Ghysels, E., Santa-Clara, P. & Valkanov, R. (2004), 'The MIDAS Touch: Mixed Data Sampling Regression Models', *UCLA: Finance* pp. 1–33.

Ghysels, E., Santa-Clara, P. & Valkanov, R. (2005), 'There is a risk-return trade-off after all', *Journal of Financial Economics* **76**(3), 509–548.

Gilli, M., Maringer, D. & Schumann, E. (2019), *Numerical Methods and Optimization in Finance*, second edn, Elsevier/Academic Press, Waltham, MA, USA. ISBN 978-0128150658.

Golosnoy, V., Gribisch, B. & Liesenfeld, R. (2012), 'The conditional autoregressive Wishart model for multivariate stock market volatility', *Journal of Econometrics* **167**(1), 211–223.

Gong, Y., Ma, C. & Chen, Q. (2022), 'Exchange rate dependence and economic fundamentals: A copula-midas approach', *Journal of International Money and Finance* **123**.

Gordon, N., Salmond, D. & Smith, A. (1993), 'Novel approach to nonlinear/non-Gaussian Bayesian state estimation', *IEE Proceedings F (Radar and Signal Processing)* **140**(2), 107–113.

Gourieroux, C., Jasiak, J. & Sufana, R. (2009), 'The Wishart Autoregressive process of multivariate stochastic volatility', *Journal of Econometrics* **150**(2), 167–181.

Hansen, P. R. & Lunde, A. (2006), 'Realized variance and market microstructure noise', *Journal of Business and Economic Statistics* **24**(2), 127–161.

Hansen, P. R., Lunde, A. & Nason, J. M. (2011), 'The model confidence set', *Econometrica* **79**, 453–497.

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

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

Jin, X., Maheu, J. M. & Yang, Q. (2022), 'Infinite markov pooling of predictive distributions', *Journal of Econometrics* **228**(2), 302–321.

Joe, H. (2015), *Dependence Modeling with Copulas*, CRC Press, Boca Raton, FL.

Jore, A. S., Mitchell, J. & Vahey, S. P. (2010), 'Combining forecast densities from VARs with uncertain instabilities', *Journal of Applied Econometrics* **25**(4), 621–634.

Kapetanios, G., Mitchell, J., Price, S. & Fawcett, N. (2015), 'Generalised density forecast combinations', *Journal of Econometrics* **188**(1), 150–165.

Kass, R. E. & Raftery, A. E. (1995), 'Bayes Factors', *Journal of the American Statistical Association* **90**(430), 773–795.

Koopman, S. J., Lit, R., Lucas, A. & Opschoor, A. (2018), 'Dynamic discrete copula models for high-frequency stock price changes', *Journal of Applied Econometrics* **33**(7), 966–985.

Lyócsa, Š., Molnár, P. & Výrost, T. (2021), 'Stock market volatility forecasting: Do we need high-frequency data?', *International Journal of Forecasting* **37**(3), 1092–1110.

McAleer, M. & Medeiros, M. C. (2008), 'Realized volatility: A review', *Econometric Reviews* **27**(1-3), 10–45.

McAlinn, K. (2021), 'Mixed-frequency Bayesian predictive synthesis for economic nowcasting', *Journal of the Royal Statistical Society Series C: Applied Statistics* **70**(5), 1143–1163.

McNeil, A. J., Frey, R. & Embrechts, P. (2005), *Quantitative risk management: Concepts, techniques, and tools*, Princeton University Press.

Nelsen, R. B. (2006), *An Introduction to Copulas, 2nd Edition*, Springer.

Nguyen, H. & Javed, F. (2021), 'Dynamic relationship between stock market and bond market: A GAS MIDAS copula approach', *Working Paper, Örebro University School of Business, Örebro* **15**, 1–45.

Nguyen, H., Virbickaitė, A., Ausín, M. C. & Galeano, P. (2024), 'Structured factor copulas for modeling the systemic risk of european and united states banks', *arXiv:2401.03443* pp. 1–30.

Noureldin, D., Shephard, N. & Sheppard, K. (2012), 'Multivariate High-Frequency-Based Volatility (HEAVY) Models', *Journal of Applied Econometrics* **27**(6), 907–933.

Oh, D. H. & Patton, A. J. (2017), 'Modelling Dependence in High Dimensions with Factor Copulas', *Journal of Business and Economic Statistics* **35**(1), 139–154.

Okhrin, O. & Tetereva, A. (2017), 'The Realized Hierarchical Archimedean Copula in Risk Modelling', *Econometrics* **5**(2), 26.

Opschoor, A., Lucas, A., Barra, I. & van Dijk, D. (2021), 'Closed-Form Multi-Factor Copula Models With Observation-Driven Dynamic Factor Loadings', *Journal of Business and Economic Statistics* **39**(4), 1066–1079.

Opschoor, A., Van Dijk, D. & van der Wel, M. (2017), 'Combining density forecasts using focused scoring rules', *Journal of Applied Econometrics* **32**(7), 1298–1313.

O'Doherty, M., Savin, N. E. & Tiwari, A. (2012), 'Modeling the cross section of stock returns: A model pooling approach', *Journal of Financial and Quantitative Analysis* **47**(6), 1331–1360.

Patton, A. J. (2006a), 'Estimation of multivariate models for time series of possibly different lengths', *Journal of Applied Econometrics* **21**(2), 147–173.

Patton, A. J. (2006b), 'Modelling Asymmetric Exchange Rate Dependence', *International Economic Review* **47**(2), 527–556.

Patton, A. J. (2009), Copula - Based Models for Financial Time Series, *in* T. G. Andersen, T. Mikosch, J.-P. Kreiß & R. A. Davis, eds, 'Handbook of Financial Time Series', Springer Berlin Heidelberg, Berlin, Heidelberg, chapter 36, pp. 767–785.

Patton, A. J. (2012), 'A review of copula models for economic time series', *Journal of Multivariate Analysis* **110**, 4–18.

Rockafellar, R. T. & Uryasev, S. (2002), 'Conditional value-at-risk for general loss distributions', *Journal of banking & finance* **26**(7), 1443–1471.

Salvatierra, I. D. L. & Patton, A. J. (2015), 'Dynamic copula models and high frequency data', *Journal of Empirical Finance* **30**, 120–135.

Schumann, E. (2023), *Numerical Methods and Optimization in Finance (NMOF) Manual. Package version 2.8-0)*.

Shephard, N. & Sheppard, K. (2010), 'Realising the Future: Forecasting with High-Frequency-Based Voality (HEAVY) Models', *Journal of Applied Econometrics* **25**(2), 197–231.

Silvennoinen, A. & Teräsvirta, T. (2009), Multivariate GARCH models, *in* T. G. Andersen, R. A. Davis, J. Kreiss & T. Mikosch, eds, 'Handbook of Financial Time Series', Springer-Verlag, pp. 201–226.

Sklar, A. (1959), 'Fonctions de répartition à n dimensions et leurs marges', *Publications Inst. Statis., Univ. Paris-VIII* pp. 229–231.

Stone, M. (1961), 'The opion pool.', *Annals of Mathematical Statistics* **32**, 1339–1342.

Timmermann, A. (2018), 'Forecasting methods in finance', *Annual Review of Financial Economics* **10**(11), 449–479.

West, M. (1986), 'Bayesian Model Monitoring', *Journal of the Royal Statistical Society: Series B (Methodological)* **48**(1), 70–78.

Zhang, L., Mykland, P. A. & 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**, 1394–1411.