WORKING PAPER 5/2022 (STATISTICS)

Modeling stock-oil co-dependence with Dynamic Stochastic MIDAS Copula models

Hoang Nguyen and Audrone Virbickaite

ISSN 1403-0586

Örebro University School of Business
SE-701 82 Örebro, Sweden

Modeling stock-oil co-dependence with Dynamic Stochastic MIDAS Copula models

Hoang Nguyen [*†]

Audronė Virbickaitė [‡]

Abstract

Stock and oil relationship is usually time-varying and depends on the current economic conditions. In this study, we propose a new Dynamic Stochastic Mixed data frequency sampling (DSM) copula model, that decomposes the stock-oil relationship into a short-run dynamic stochastic component and a long-run component, governed by related macro-finance variables. We find that inflation/interest rate, uncertainty and liquidity factors are the main drivers of the long-run co-dependence. We show that investment portfolios, based on the proposed DSM copula model, are more accurate and produce better economic outcomes as compared to other alternatives.

Keywords: Stock-Oil; Copula; MIDAS; SMC; Portfolio allocation; Hedging

JEL Classification: C32, C52, C58, G11, G12

[*]Corresponding author

[†]School of Business - Örebro University, Sweden, hoang.nguyen@oru.se

[‡]Department of Quantitative Methods, CUNEF Universidad, Madrid, Spain, audrone.virbickaite@cunef.edu

1 Introduction

Modeling and predicting the co-movements of financial returns has been one of major interests among the researchers and practitioners alike, especially in the presence of recent global crises (the 2008 financial crisis and the Covid pandemic). Understanding the equity return co-dependence is especially important in modern finance, in particular in derivative pricing, portfolio allocation, risk management and hedging. Commodities tend to have low or negative correlations with the conventional equity stocks, therefore, they have a high potential of providing diversification benefits (Sadorsky, 2014), especially in hedging the downside risk. With the financialization of the commodity markets, it is now straightforward to include certain commodities in one's investment portfolio.

The stock-oil relationship has received a great deal of attention not only because oil is often used to hedge the market risk, but also because oil shocks have direct effect on the financial markets, and vice-versa (Jones and Kaul, 1996; Sadorsky, 1999; Chao Wei, 2003; Kilian and Park, 2009). This co-dependence is not constant in two senses: first, the relationship fluctuates in the short-run and, secondly, it presents more substantial changes in levels in the long-run (Batten et al., 2017). Such changes in the co-dependence are affected by the changes in external political decisions, macroeconomic factors and the conditions in the financial markets (Huang et al., 1996; Amihud and Wohl, 2004; Kilian, 2008; Sukcharoen et al., 2014). For example, there has been a dramatic increase in the long-run co-dependence between stock-oil returns after the 2008-2009 crisis, documented by numerous studies (Aloui et al., 2012; Mollick and Assefa, 2013; Pan, 2014; Zhu et al., 2014). Kilian (2008) shows that oil demand shocks lead to a large and sustained increases in oil price. Mollick and Assefa (2013) consider the economic indicators as a signal for the future aggregate demand and aggregate uncertainty (inflation expectations, VIX).

With this work, we contribute to the existing literature investigating the co-movements of stock-oil returns in four major ways. First, we decompose the time-varying dependence into two parts: short- and long-run. We allow the short-run component to be parameter-driven, resulting into greater flexibility in the dynamics, as opposed to the observation-driven models. Secondly, we allow for the long-run component to be affected by some external macro-finance variables, sampled at lower frequencies. Following some of the most recent works (Asgharian et al., 2016; Nguyen and Javed, 2021), we group these macro-finance variables via principal components. Thirdly, we employ a novel Bayesian estimation strategy based on density-tempered sequential Monte Carlo sampler. The method is especially well suited for non-linear non-Gaussian state space models and produces marginal log likelihoods as a by-product, which allows for consistent model comparison even for non-nested models. And, finally, we present an extensive empirical illustration using stock-oil return data from more than 30 years (including the latest Covid period) and perform portfolio allocation exercises to quantify the economic gains of using our proposed approach.

The co-movement of financial returns is usually characterized via time-varying volatilities and correlations by employing multivariate GARCH or multivariate SV-type models (Bauwens et al., 2006; Asai et al., 2006; Chib et al., 2009). These models have been extensively applied to model the stock-oil co-dependence, see, for example, Choi and Hammoudeh (2010); Vo (2011); Sadorsky (2014); Basher and Sadorsky (2016); Pan et al. (2016), among others. The usual assumption is that the conditional joint distribution of the returns follows a multivariate Normal or multivariate Student-t distributions, which implies that it is the same for all marginals, and that the dependence structure must be linear. In order to relax the assumptions above, we model the stock-oil co-dependence via copulas (McNeil et al., 2015; Nelsen, 2006; Joe, 2015; Patton, 2009, 2012, 2013). The major advantage of the copula approach is that one can separate the modelling of the individual marginals from their dependence structure. Moreover, the marginals of the individual assets do not necessary have to be the same, as in the multivariate GARCH or SV setting; and the choice of the appropriate copula (symmetric or asymmetric) is completely independent from the marginals. Such approach is rather standard in financial econometrics literature, see for example, Dias and Embrechts (2004), Patton (2006), Jondeau and Rockinger (2006), Ausin and Lopes (2010), Virbickaite et al. (2022), among many others. Copulas have been used in modeling the stock-oil co-dependence in Aloui et al. (2013); Sukcharoen et al. (2014); Pan (2014); Mensi et al. (2017); Yu et al. (2020).

Copula function is characterized by a set of parameters, that can be static or time-varying. Hafner and Manner (2012), for example, have proposed to model this copula dependence parameter as a stochastic process via latent AR(1)-type equation. They show that such approach allows for greater flexibility as compared to the observation-driven models. Therefore, we build on the dynamic stochastic copula (DSC) model of Hafner and Manner (2012) and interpret the unconditional mean of the AR(1) process as the long-run stock-oil co-dependence, that can be affected by related macro-finance factors. Since most of the economic factors are observed at a lower frequency than the financial variables, in order to account for the frequency mismatch, we rely on a mixed data sampling method (MIDAS, Ghysels et al., 2004), where the effects of the lagged macro-finance variables are regularized by a polynomial weighting scheme. We refer to resulting model as the dynamic stochastic MIDAS (DSM) copula, in which the co-dependence of financial returns is decomposed into a short-run and long-run components. The DSM copula not only allows for a greater flexibility in the dynamic dependence of financial returns but also takes advantage of exogenous macro-finance variables to explain the long-run changes in the dependence structure. The co-dependence of stock-oil in the DSM copula models fluctuates around a long-run component rather than a mean reverting process, which is the case in the DSC model of Hafner and Manner (2012). Finally, modeling the co-dependence as a function of macroeconomic variables not only helps investors to build efficient investment strategies but also might serve as a warning signal for the changes of macroeconomic conditions, and vice-versa (Conrad and Loch, 2016).

There have been several studies that model the long-run component of the correlations of the financial returns based on the MIDAS framework. Conrad et al. (2014) employ the dynamic conditional correlation (DCC) MIDAS model (Engle, 2002a; Colacito et al., 2011) and find that the contraction in macroeconomic activity, measured using multiple current and forward-looking economic indicators, leads to an increase of the long-run correlation between crude oil and stock price returns, as vice-versa. Gong et al. (2018) employ a copula-MIDAS model to describe the asymmetric long-run return-liquidity dependence using the realized correlations. Gong et al. (2020) model the relationship between stock-oil returns via an observation-driven MIDAS copula model. The authors find that aggregate demand and stock-specific negative news have a significant impact on stock-oil co-dependence, however their proposed copula allows only for the positive dependence. Alternative to our proposal, Nguyen and Javed (2021) use a generalized autoregressive score mixed frequency data sampling (GAS MIDAS) copula model to analyze the time varying co-dependence between stock-bond returns. The authors find that the inflation and interest rate, the state of the economy and the illiquidity are significant in modeling the long-run dependence.

Most of the aforementioned studies are based on a class of observation-driven models and employ the maximum likelihood approach to estimate the parameters. The uncertainty of parameters is quantified using a numerical approximation of the information matrix that can lead to the divergence of the likelihood or cannot guarantee the positive semi-definiteness of the information matrix. Instead, we use the Bayesian approach for inference and prediction for our proposed parameter-driven DSM copula models. We apply the density-tempered sequential Monte Carlo (DTSMC) sampler of Tran et al. (2014) and Duan and Fulop (2015). This sampler combines the Anneal important sampling (Neal, 2001) and the Sequential Monte Carlo method (Del Moral et al., 2006). The DTSMC sampler provides an efficient alternative to the standard MCMC methods for estimating complex non-linear non-Gaussian state-space models; also it provides the estimate of the marginal log likelihood as a by-product, which is useful for model comparison using the Bayes factors, for example.

In an empirical illustration using stock-oil return data from more than 30 years we find that the stock-oil relationship is symmetric and fat-tailed. The inclusion of the macro-finance variables to model the stock-oil co-dependence increases the in-sample log marginal likelihoods. In particular, inflation/interest rate, uncertainty and liquidity factors are the main drivers of the long-run co-dependence. Moreover, the variance-covariance matrices, predicted using our proposed DSM copula models, are more accurate. Finally, the hedging and minimum variance portfolios, built using these predictions produce better economic outcomes out-of-sample.

The rest of the paper is organized as follows. Section 2 introduces the DSM copula model. Section 3 describes the Bayesian inference algorithm. Section 4 models the marginal distributions of the individual assets and then analyzes the fundamental factors that affect the co-dependence of the stock-oil returns. Section 5 presents two portfolio allocation exercises and conclusions are drawn in Section 6.

2 Model Specification

In this section, we describe the proposed DSM copula model. We start by briefly introducing the model for the marginals for the returns. We then proceed to define copulas and the DSC model of Hafner and Manner (2012). Finally, we extend the DSC model by incorporating the MIDAS effects into the long-run dynamics as a function of past macro-finance variables, sampled at lower frequencies.

2.1 Marginals for the returns

First, let $P_{i,t}$ be the daily price of a financial asset i (e.g. a stock index or a commodity, such as oil) for $t = 1, \ldots, T$ and define the daily log returns (in %) as $r_{i,t} = 100 \times (\ln P_{i,t} - \ln P_{i,t-1})$. Then, the dynamics of the individual asset log returns can be modeled as

$$r_{i,t} = \mu_i + \epsilon_{i,t}\sqrt{\sigma^2_{i,t}}. \tag{1}$$

Here μ_i is the unconditional mean, $\epsilon_{i,t}$ is an error term with a zero-mean and unit variance, which can be assumed to be Gaussian, Student-t, Generalized Error Distribution or any other parametric or non-parametric distribution. $\sigma^2_{i,t}$ is the volatility, which again, can follow either GARCH, or SV dynamics, or, alternatively, can be estimated using the realized volatility measure (given the high-frequency data necessary for estimation is available). Hafner and Manner (2012) found that the estimated volatilities using either GARCH or SV models are very similar hence we model the dynamics of the returns in Eq. (1) via GJR-GARCH model with skew-t errors. Nonetheless, any other specification for the distribution term and the dynamics of the volatility is in order (e.g. Liesenfeld and Jung, 2000; Engle, 2002b; Andersen et al., 2010; Loaiza-Maya et al., 2018; Conrad and Kleen, 2020), as long as the probability integral transform of the standardized returns is uniformly distributed. In other words, denote $F_i(\cdot)$ as the cumulative distribution function (CDF) of the $\epsilon_{i,t}$, then if the distribution term and the volatility dynamics are correctly specified, the $u_{i,t} = F_i((r_{i,t} - \mu_i)/\sqrt{\sigma^2_{i,t}})$ should be *iid* uniformly distributed. Once the marginals have been specified, one can model the co-dependence between the resulting $u_{i,t}$ using copula.

2.2 Copula

The construction of flexible multivariate distributions using copulas has started with the seminal work of Sklar (1959). For a formal introduction to copulas, we refer to Nelsen (2006); Joe (2015). Consider a collection of random variables $Y_1, \ldots, Y_d$ with corresponding marginal CDFs $F_i(y_i) = P[Y_i \leq y_i]$ for $i = 1 \ldots, d$ and a joint distribution function $H(y_1, \ldots, y_d) = P[Y_1 \leq y_1, \ldots, Y_d \leq y_d]$. According to Sklar (1959) there exists a copula C with parameter(s) θ such that $H(y_1, \ldots, y_d) = C_\theta(F_1(y_1), \ldots, F_d(y_d))$. Copulas are defined in the unit hypercube $[0,1]^d$, where d is the dimension of the data, and the univariate marginals $F_i(y_i) \forall i$ are uniformly distributed. Finally, let $h(y_1 \ldots, y_d)$ be the joint density of $Y_1, \ldots, Y_d$, and $f_i(y_i)$ the corresponding marginal probability density function (PDF) of Y_i. Then $h(\cdot)$ can be expressed as a product of the marginals and a copula density:

$$h(y_1, \ldots, y_d) = c_\theta(F_1(y_1), \ldots, F_d(y_d)) \cdot f_1(y_1) \cdots f_d(y_d).$$

For the sake of simplicity, in what follows we consider only bivariate case, i.e. $d = 2$. Joe (2015) summarizes several copula functions that allow for a flexible dependence in the bivariate context. Observing that elliptical copula and Archimedean copula families are most commonly used in finance literature due to the parsimonious specification and their ability to capture tail dependence, we employ the Gaussian copula, the Student-t copula (or simply t copula), the Clayton copula, the Gumbel copula, the Frank copula, and the Joe copula to model the dynamic dependence, see Online Appendix A. As the Archimedean copulas can model only positive or negative dependence, we propose modified Archimedean copulas based on a mixture of their own rotations so that,

$$c_{Archimedean}(u_1, u_2; \theta) = \begin{cases} c^+_{Archimedean}(u_1, u_2; \theta) = 0.5c(u_1, u_2; \theta) + 0.5c_{R180}(u_1, u_2; \theta), & \text{if } \theta > 0 \\ c^-_{Archimedean}(u_1, u_2; \theta) = 0.5c_{R90}(u_1, u_2; \theta) + 0.5c_{R270}(u_1, u_2; \theta), & \text{if } \theta < 0 \end{cases}$$

where c, c_{R90}, c_{R180}, c_{R270} are the Archimedean copula density and its 90-degrees, 180-degrees, 270-degrees rotations; (u_1, u_2) are uniformly distributed probability integral transforms $(F_1(y_1), F_2(y_2))$, and θ is a copula parameter. A key copula-related measure is Kendalls' tau $\tau_\kappa \in (-1, 1)$, also known as rank correlation coefficient. For most copulas there is a one-to-one relationship between the copula parameter θ and τ_κ. Kendall's tau provides a convenient way to compare the strength of the dependence across different copulas, because, unlike the copula parameter θ, τ_κ always lies in the same domain. As seen in the following section, we rely on Kendall's tau to map the copula-specific parameter θ to its corresponding domain. The copula CDFs, PDFs, and the $\theta \leftrightarrow \tau_\kappa$ transformations can be found in the Online Appendix A.

2.3 Dynamic stochastic copula

After we have specified the appropriate marginal model for financial time series, the next step is to model the joint dependence through a dynamic copula model. Hafner and Manner (2012) propose to model the dynamics of the copula parameter as a transformation of a mean reverting process, giving rise to the DSC model:

$$
\begin{aligned}
(u_{1,t}, u_{2,t}) &\sim c(u_{1,t}, u_{2,t}; \theta_t), \\
\theta_t &= \Lambda(\lambda_t), \\
\lambda_t &= \lambda_0(1-\beta) + \beta\lambda_{t-1} + \sigma_e e_t, \quad e_t \sim N(0,1).
\end{aligned}
\tag{2}
$$

The function $\Lambda(\cdot)$ is a copula-specific transformation that ensures that a copula parameter θ_t is in the appropriate domain. Specifically, we use $\tau_\kappa = \Lambda(\lambda) = \frac{\exp(\lambda)-1}{\exp(\lambda)+1}$ to map λ to the Kendall-τ correlation, then find the parameter θ of the equivalent copula with τ_κ correlation, see the Online Appendix A. The dynamics of the transformed copula parameter λ_t can be modelled via AR(1) process with unconditional mean λ_0, persistence β and some variance σ_e^2. Note that we assume the persistence parameter $|\beta| < 1$ for stationary conditions and the variance $\sigma_e^2 > 0$. As a special case, Hafner and Manner (2012) show that the dynamic stochastic copula model becomes a bivariate Gaussian SV model with stochastic correlation when the copula and the marginals are all Gaussian. However, the use of other copula functions (potentially asymmetric and/or fat-tailed) allows for larger flexibility in modeling the tail dependence.

2.4 Dynamic stochastic MIDAS copula

The DSC model, defined in Eq. (2), assumes a fixed level of long-run dependence, controlled by parameter λ_0, which might be restrictive and not in-line with the actually observed empirical evidence. Same as in Nguyen and Javed (2021), we allow the long-run dependence parameter to be driven by some exogenous fundamental variables. Even observed at a lower frequency, the macro-finance variables often provide new information on the current economic situation and business outlook. Hence, taking advantage of these sources of information can help us to better understand the dynamic co-dependence among the financial returns. In particular, we allow the unconditional mean of the latent process to depend on some low-frequency explanatory macro-finance variables through a MIDAS regression, giving rise to a novel DSM copula model:

$$
\begin{aligned}
(u_{1,t}, u_{2,t}) &\sim c(u_{1,t}, u_{2,t}; \theta_t), \\
\theta_t &= \Lambda(\lambda_t), \\
\lambda_t &= \lambda_\tau(1-\beta) + \beta\lambda_{t-1} + \sigma_e e_t, \quad e_t \sim N(0,1), \\
\lambda_\tau &= \lambda_0 + \sum_{j=1}^{N} \delta_j \left[\sum_{k=1}^{K_j} \phi_k(\omega_j) X_{j,\tau-k} \right].
\end{aligned}
\tag{3}
$$

Here τ is an indicator for monthly time points which is related to the daily time points though $t = \tau L, \tau L+1, \ldots, (\tau+1)L$, where L is the number of trading days in a month; $X_\tau = (X_{1,\tau}, \ldots, X_{N,\tau})$ is N-dimensional vector of the low frequency explanatory variables at month τ, and $\phi_k(\omega_j)$ is the weighting function parameterized by ω_j of the variable j on its k^{th} lag, for $k = 1, \ldots, K_j$. Here, we use the restricted beta function to smooth the effect of the explanatory variables:

$$
\phi_k(w_j) = \frac{[1 - k/(K_j+1)]^{w_j - 1}}{\sum_{l=1}^{K_j} [1 - l/(K_j+1)]^{w_j - 1}}.
$$

This stochastic MIDAS extension allows for the flexible changes in the long-run dependence through a set of explanatory variables X_τ and their lags. The effect of the j^{th} variable can be seen through the size of the regression coefficient δ_j, where the weighting parameter ω_j regulates for how long the effect lasts. Higher value of ω_j means that the effect of the lagged values of the explanatory variable decays at a faster speed. The DSM copula model reduces to the DSC model when $\delta_j = 0 \forall j$. Some other alternative weighting functions could be considered, however, previous studies suggest that different weighting functions produce virtually identical estimation results (Nguyen and Javed, 2021). In the next section, we describe the Bayesian approach used for the parameter estimation.

3 Bayesian inference and model comparison

A standard way to estimate copula based models is via two-step approach: first, estimate the univariate marginals and obtain the probability integral transforms $(u_{1,t}, u_{2,t})$, and secondly, estimate the copula-related parameters, see Almeida and Czado (2012) for example. It is also possible to estimate the marginals and copula parameters in a one-step procedure, see Ausin and Lopes (2010), however, due to the latent nature of the evolution equation we opt for the two-step approach. The estimation of the univariate marginals can be performed using any standard estimation method, depending on the chosen specification for the dynamics of the individual volatilities and the distribution of the error term. Once the $(u_{1,t}, u_{2,t})$ data has been obtained, we then perform estimation of the copula related parameters.

The model in Eq. (3) is a non-linear and potentially non-Gaussian state space model. Estimation of such models entails estimating the set of fixed model parameters as well as the sequence of latent state variables λ_t. We rely on a version of a novel Bayesian estimation approach, known as annealed importance sampling with an estimated likelihood in Tran et al. (2014), and density-tempered marginalized sequential Monte Carlo sampler in Duan and Fulop (2015). We call this procedure the Density-tempered sequential Monte Carlo (DTSMC). Note that even though the name has a word "sequential" in it, the algorithm, same as the standard MCMC methods, performs batch inference, as opposed to "truly" sequential expanding-data approaches. However, differently than the standard MCMC methods, our approach is better suited for non-linear state space models, especially when the exact likelihood is not available. Finally, the estimate of the marginal likelihood, allowing for consistent model comparison via Bayes Factors, in our case is a by-product, meanwhile it is notoriously difficult to obtain in the classical MCMC setting.

3.1 Prior Distributions

Denote $\Theta = \{\beta, \sigma_e^2, \lambda_0, \delta_{1:N}, \omega_{1:N}\}$ as the set of fixed parameters of the DSM copula model. We use vague proper prior distributions: $\beta \sim \text{Beta}(20, 1.5)$ and $\sigma_e^2 \sim \text{Gamma}(\frac{1}{2}, \frac{1}{2B_\sigma})$ which is equivalent to $\pm\sqrt{\sigma_e^2} \sim \mathcal{N}(0, B_\sigma)$. Kastner and Frühwirth-Schnatter (2014) consider that the choice of the prior for σ_e is less influential when the true value is small, therefore, we set $B_\sigma = 1$. We consider the weakly informative priors of the parameters in the MIDAS regression: $\delta_j \sim N(0,1)$ and $\omega_j \sim U(1,30) \forall j$.

3.2 Density-tempered sequential Monte Carlo algorithm

Let $u_{1:T}$ be a collection of $(u_{1,t}, u_{2,t})$ for $t = 1, \dots, T$. We are interested in sampling from the posterior distribution:

$$
\begin{aligned}
p(\Theta|u_{1:T}) &\propto p(u_{1:T}|\Theta)p(\Theta), \text{ where} \\
p(u_{1:T}|\Theta) &= \int p(u_{1:T}|\lambda_{1:T}, \Theta)p(\lambda_{1:T}|\Theta)d\lambda_{1:T} \\
&= \int \prod_{t=1}^{T} c(u_{1t}, u_{2t}|\lambda_t, \Theta, u_{1:t-1})p(\lambda_t|\lambda_{t-1}, \Theta, u_{1:t-1})d\lambda_{1:T}.
\end{aligned}
$$

The likelihood $p(u_{1:T}|\Theta)$ is intractable due to the high dimensional integration over the latent process. Hafner and Manner (2012) use a particle filter to approximate the intractable likelihood and then use maximum likelihood to estimate Θ. Following this idea in the Bayesian context, we obtain an unbiased estimator $\widehat{p}(u_{1:T}|\Theta, e)$ of the likelihood $p(u_{1:T}|\Theta)$, where e is the set of pseudo random numbers used in the particle filter, and use the DTSMC algorithm to sample from the posterior distribution.

The DTSMC algorithm samples sequentially from the prior distribution $p(\Theta)$ to the posterior distribution $p(\Theta|u_{1:T})$ through a sequence of distributions $\pi_i(\Theta)$ that smoothly interpolate between the prior and the posterior distribution,

$$
\pi_i(\Theta) := p_i(\Theta|u_{1:T}) \propto \widehat{p}(u_{1:T}|\Theta, e)^{\gamma_i} p(\Theta),
$$

where γ_i is a level temperature for $i = 0, \dots, S$ such that $0 = \gamma_0 < \gamma_1 < \dots < \gamma_S = 1$. We choose the tempering sequence $\gamma_i = (i/S)^3$ for $i = 0, \dots, S$ so that the effective number of particles in each interpolation does not change dramatically.

In order to perform inference using the DTSMC algorithm, we first simulate a set of M weighted particles $\{W_0^j, \Theta_0^j\}_{j=1}^M$ from the prior distribution $p(\Theta)$ and a set of pseudo random e_0^j for the particle filter. For each level of temperature γ_i, the DTSMC algorithm performs reweighting, resampling and Markov moves. The reweighting step starts with a set of M weighted particles $\{W_{i-1}^j, \Theta_{i-1}^j\}_{j=1}^M$ obtained from the previous interpolation distribution $\pi_{i-1}(\Theta)$. The importance sampling is employed to re-weight the particles for approximating the target $\pi_i(\Theta)$,

$$w_i^j = W_{i-1}^j \widehat{p}(u_{1:T}|\Theta_{i-1}^j, e_{i-1}^j)^{\gamma_i - \gamma_{i-1}}, \quad j = 1, ..., M,$$
$$W_i^j = \frac{w_i^j}{\sum_{s=1}^M w_i^s}, \quad j = 1, ..., M.$$

If the efficiency of these weighted particles, measured by the effective sample size (ESS), is below some threshold, the particles are resampled. Then, in order to avoid particle degeneracy, the particles are refreshed by a Markov kernel whose invariant distribution is $\pi_i(\Theta)$. We also incorporate the Correlated Pseudo Marginal (CPM) approach by Deligiannidis et al. (2018) into the Markov move step to deviate from the unstable integrated likelihood estimated by the particle filter, and use a random walk proposal for the new value of Θ. The DTSMC algorithm can be run in parallel for the particle moves in the Markov move step, so we can take advantage of the parallel computation in a high-performance computing server. The DTSMC algorithm is described in detail in the Online Appendix B. Finally, the log marginal likelihood is computed as

$$\log \widehat{p}(u_{1:T}) = \sum_{i=1}^S \log \left(\sum_{j=1}^M w_i^j \right).$$

3.3 Simulation

The Online Appendix C contains a simulation study that compares the proposed DSM copula model with the alternative models such as the DCC and the DSC. We consider different stress scenarios for the correlation dynamics and compare the estimation accuracy. In general, the DSM copula is preferred in terms of MSE and MAE. The time of inference is quite fast when using a multi-core Ryzen 3950x CPU (16 cores at 3.5Hz). For a Gaussian bivariate copula with $T = 2000$ data points, it took on average 25 minutes and 28 minutes for DSC and DSM copula models respectively.

4 Empirical Illustration

This section contains the description of the data and the estimation results for the marginals, the DSC and DSM copula models for the stock-oil return data.

4.1 Data

We apply the proposed models on the stock and oil data returns. For stock data we use daily S&P 500 observations and for oil data we use the daily spot prices of West Texas Intermediate (WTI) from 01/01/1990 to 31/12/2021. The S&P 500 data is obtained from Bloomberg and WTI daily spot prices are obtained from the US Energy Information Administration website. The negative WTI spot price that happened on April 20th 2020 was removed. See Figure 1 for the daily prices, log returns and monthly correlations for both assets. Shaded areas are the National Bureau of Economic Research (NBER) based recession indicators: early 1990s and 2000s recessions, the global financial crisis of 2008 and the Covid pandemic. As seen from the top right and middle right plots, both asset returns present increased volatility during crisis episodes, which is in line with a well-documented stylized features of financial returns. The bottom plot draws the monthly stock-oil correlation, calculated as a 22-day sample correlation at the end of each period. Just by eye-balling the plot, it seems that the correlation presents a rather mixed picture: sharp decrease during the early 1990s recession, virtually unchanged level during the early 2000s, significant jump-type change in the level during the global financial crisis and rather inconclusive change in the co-dependence during the Covid pandemic. Contrary to the “conventional” asset returns, whose co-dependence increases during economic turmoils and decreases during the periods of calm economic conditions, the stock-oil relationship seems much more complex and possibly driven by a multitude of external factors.

Table 1 presents the descriptive statistics for S&P 500 and WTI returns. Both assets’ returns are negatively skewed, have large kurtosis, and present autocorrelation in squared returns, indicating the presence of ARCH effects. The average of daily stock returns is higher than oil returns while oil returns are more volatile and skewed. This could be due to the mismatch of oil supply and demand.

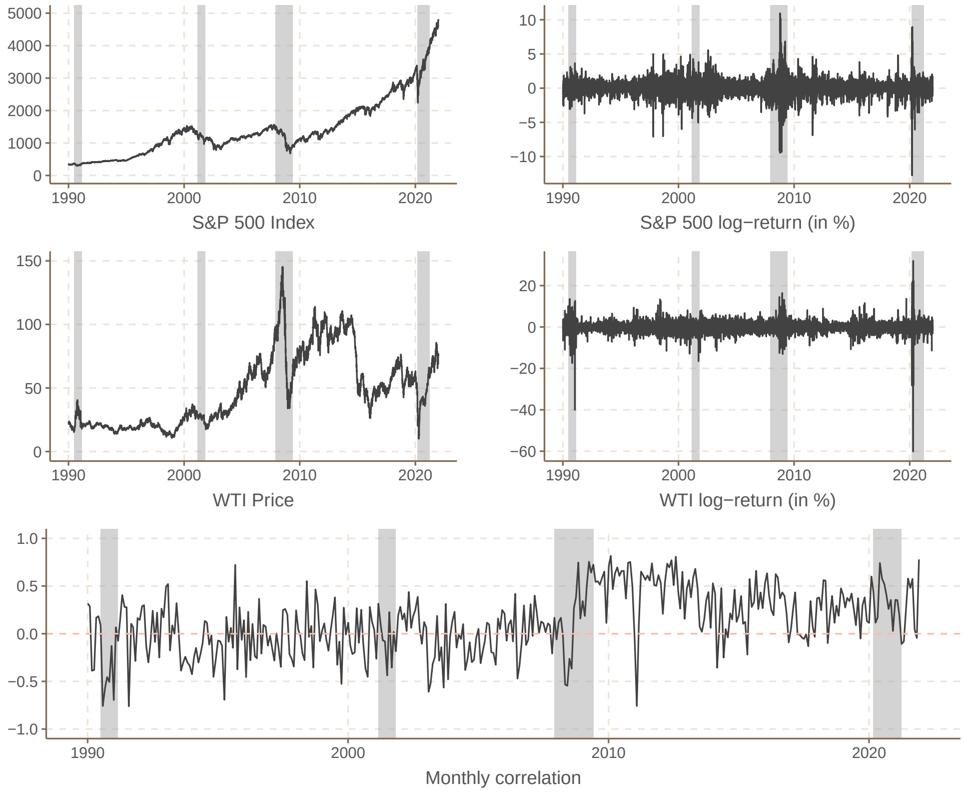

Figure 1: Daily prices, log returns and monthly correlations for S&P 500 and WTI.

The plots show the daily S&P 500 index and returns and the daily WTI oil prices and returns during the period from 01/01/1990 to 31/12/2021. The S&P 500 returns (WTI oil returns) are calculated as the log difference of the S&P 500 index (WTI oil price) multiplied by 100. Monthly correlation is calculated as a 22-day sample correlation at the end of each period. The shaded areas highlight the recession periods based on the NBER recession indicators.

Table 1: Descriptive statistics for S&P 500 and WTI log returns.

	mean	Q2	sd	skew	kurtosis	min	max	LB(10)	ARCH(10)	JB	n
S&P 500	0.03	0.06	1.14	-0.41	14.39	-12.77	10.96	0.00	0.00	0.00	8059
WTI	0.01	0.05	2.65	-1.87	54.26	-60.17	31.96	0.00	0.00	0.00	8059

Descriptive statistics for the log return data (in %) for 01/01/1990 to 31/12/2021. The LB(k) reports the p-value for the Ljung-Box test for autocorrelation of k lags, and ARCH(l) reports the p-value for the test for ARCH effects with l lags. The JB reports the p-value for the Jarque-Bera test for Normality.

4.2 Marginals

We first fit the GJR-GARCH model of Glosten et al. (1993) with skew-t errors on the daily log returns. As defined in Eq. (1), the log returns for asset $i = \{\text{S\&P 500, WTI}\}$ can be decomposed into time-invariant mean component and a heteroscedastic error term $r_{i,t} - \mu_i \equiv \varepsilon_{i,t} = \epsilon_{i,t}\sqrt{\sigma^2_{i,t}}$. Here $\epsilon_{i,t}$ follows a skew-t distribution with skew and shape parameters (ξ_i, ν_i) (Fernández and Steel, 1998; Ferreira and Steel, 2006). The GJR-GARCH model is defined as follows:

$$\sigma^2_{i,t} = \omega_i + (\alpha_i + \gamma_i I_{i,t-1})\varepsilon^2_{i,t-1} + \beta_i \sigma^2_{i,t-1},$$

where the indicator function $I_{i,t}$ takes value of 1 if $\varepsilon_{i,t} \leq 0$ and 0 otherwise; and $(\omega_i, \alpha_i, \beta_i, \gamma_i)$ are the parameters of the marginal returns where γ_i represents the leverage effect. The models were estimated using the `R` package `rugarch` (Ghalanos, 2022). Estimation results are in Table 2. The persistence parameter β_i for both assets takes a value close to one, a finding consistent with the stylized features of financial volatility. The leverage parameter γ_i is larger for S&P 500 than for WTI indicating a stronger asymmetric response of the volatility, but, nonetheless, both parameters are statistically significant. The degrees of freedom parameter ν_i is relatively low in both cases, indicating fat-tailed distribution. Table 3 presents the descriptive statistics for residuals of the GJR-GARCH skew-t model. According to Ljung-Box and ARCH tests, there is no leftover autocorrelation in the mean or squared residuals. The p-values of the Kolmogorov-Smirnov (KS) and Anderson-Darling (AD) tests for skew-t distribution with the estimated (ξ, ν) parameter values do not reject the null hypothesis, indicating the appropriateness of the distributional assumption. Finally, the Data Driven Smooth (DD) test for Uniformity, which checks whether the probability integral transforms of the residuals are uniformly distributed, does not reject the null either. In other words, the marginals have been modelled correctly and the resulting uniformly distributed data can be used in the copula estimation step.

4.3 Dynamic dependence via DSM copula

Finally, after the marginals have been fitted, we estimate the DSM copula to model the time-varying co-dependence between the stock-oil returns. Short-run co-dependence is modelled via latent AR-

Table 2: Estimation results for the marginals.

	μ_i	ω_i	α_i	β_i	γ_i	ξ_i	ν_i
S&P 500	0.029	0.017	0.002	0.894	0.18	0.893	7.052
	(0.008)	(0.003)	(0.008)	(0.012)	(0.021)	(0.013)	(0.568)
WTI	0.012	0.064	0.051	0.918	0.04	0.922	6.284
	(0.021)	(0.027)	(0.013)	(0.018)	(0.013)	(0.014)	(0.494)

Estimated parameters and standard errors (in parenthesis) for the GJR-GARCH model with skew-t errors for S&P 500 and WTI data.

Table 3: Descriptive statistics for GJR-GARCH model residuals.

	mean	Q2	sd	skew	kurtosis	min	max	LB(10)	ARCH(10)	KS	AD	DD
S&P 500	0.00	0.04	1.00	-0.57	5.42	-7.17	6.22	0.07	0.71	0.09	0.12	0.97
WTI	0.00	0.02	1.01	-0.42	6.22	-8.63	6.22	0.56	0.29	0.57	0.69	0.95

Descriptive statistics for GJR-GARCH skew-t model residuals. The LB(k) reports the p-value for the Ljung-Box test for autocorrelation of k lags, and ARCH(l) reports the p-value for the test for ARCH effects with l lags. KS and AD are the p-values for Kolmogorov-Smirnov and Anderson-Darling tests for skew-t distribution. Finally, DD is the p-value for the Data Driven Smooth test for Uniformity that checks whether the probability integral transforms of the residuals are uniformly distributed.

type process, where the long-run co-dependence is driven by several macro-finance factors. Kilian (2008), for example, explains the changes in oil price through the oil demand shocks while Mollick and Assefa (2013) use inflation expectation and VIX for the aggregate demand and aggregate uncertainty. Gong et al. (2020) utilize the supply and demand shocks to oil as the economic factors, however it is prone to model identification assumptions. Conrad et al. (2014), on the other hand, employs a number of indicators of current and future economic activity to model the stock-oil relationship. Such indicators include industrial production, unemployment, National Activity Index, among others. In the spirit of Asgharian et al. (2016); Nguyen and Javed (2021), we group the macro-finance variables into four major categories and use the first principal component (PC) as a representative factor for each group. Monthly data is obtained from FRED[1] and Bloomberg.

i. Inflation/interest rate variables (INR): Inflation (INF), calculated as 100Δ of the consumer price index; Effective Federal Funds Rate (IR EFFR); The spread between 10-Year Treasury Constant Maturity and 3-Month Treasury Constant Maturity (T10Y3M).

ii. Uncertainty variables (UNC): Economic policy uncertainty (EPU); Equity Market Volatility Tracker: Energy And Environmental Regulation (EMV); National Financial Conditions Index

[1] Federal Reserve Bank of St. Louis Economic Data

(NFCI).

iii. (Il)liquidity variables (LIQ): liquidity LIX and Amivest (Amihud et al., 1997; Marshall et al., 2012; Danyliv et al., 2014) and Amihud illiquidity (Amihud, 2002) measures are calculated daily and transformed to monthly by taking the average over the corresponding month:

$$LIX = L^{-1} \sum_{j=1}^{L} \log_{10} \left(\frac{Volume_j \times Close_j}{High_j - Low_j} \right),$$

$$\text{Amivest} = L^{-1} \sum_{j=1}^{L} \frac{Volume_j}{|r_j|},$$

$$\text{Amihud} = L^{-1} \sum_{j=1}^{L} \frac{|r_j|}{Volume_j \times Close_j}.$$

Here L is the number of trading days in a month, $Volume_j$, $Close_j$, $High_j$, Low_j and r_j are the daily trading volume, closing price, high price, low price and the daily log returns, respectively.

iv. State of the economy variables (SOE) [2]: changes (in %) of Industrial Production (IP) index; changes (in %) of Coincident Economic Activity Index (CEAI), as a proxy for nonfarm payrolls, the unemployment, average hours worked and wages and salaries; National Activity Index (NAI); changes (in %) of US Consumer Confidence Index (CCI); changes (in %) of US Business Confidence Index (BCI); changes (in %) of US Composite Leading Indicator (CLI).

Table 4 reports the cross-correlation matrix among the explanatory variables, where the monthly correlation (RCor) is calculated as a sample correlation between daily stock and oil returns during a month.

The INR factor is positively correlated with inflation and with the effective interest rate and negatively with the term spread. High interest rates encourage more people to save due to the flight-to-quality phenomenon. As investors receive more on their savings rate than in the market, they would be getting rid of risky stocks. Also, while the spread is positive, the long term rates

[2] In the state of the economy group IP and CEAI variables are yearly difference instead of monthly difference as during the pandemic, these variables dropped (increased) dramatically, and only a few months later they suffered an almost equal increase (decrease), returning to the (almost) pre-Covid level. Such noisy data might inadvertently affect the estimation results, therefore, for those variables we used lagged yearly data instead.

Table 4: Cross-correlation matrix for the macro-finance variables.

	RCor	PC INR	PC UNC	PC LIQ	PC SOE
INF	-0.192	0.556			
IR EFFR	-0.484	0.864			
T10Y3M	0.089	-0.768			
EPU	0.405		0.750		
EMV	0.217		0.675		
NFCI	0.179		0.724		
LIX wti	0.473			0.822	
Amih wti	-0.122			-0.572	
Amiv wti	0.265			0.625	
LIX sp500	0.110			0.817	
Amih sp500	-0.279			-0.833	
Amiv sp500	0.023			0.306	
IP	-0.329				0.865
CEAI	-0.261				0.782
NAI	-0.361				0.855
CCI	0.060				-0.291
BCI	0.068				-0.564
CLI	0.092				-0.458
PC INR	-0.363	1.000			
PC UNC	0.376	-0.295	1.000		
PC LIQ	0.335	-0.418	0.037	1.000	
PC SOE	-0.332	0.485	-0.218	-0.167	1.000

Cross-correlation matrix of individual macro-finance variables, the first principal component in each group and the monthly correlation (RCor) between oil-stock returns for 01/01/1990 till 31/12/2021 (n = 8059). Monthly RCor is calculated as a sample correlation using 22 days at the end of each period.

are higher than the short term rates: an indication of healthy economy, however, when the spread decreases (or becomes negative), it might be seen as a lead indicator of a recession, which, again encourages the selling-off of the stocks. On the other hand, when the inflation and interest rates are high, oil prices tend to increase. To sum up, increase in the INR factor signals the sell-off of the stock but not necessary oil, hence, it is negatively correlated with the realized stock-oil correlation. We expect that the sign of INR variable in the DSM copula model will be negative.

The UNC factor is positively correlated with the three variables: EPU, EMV and the NFCI. EPU measures policy-related economic uncertainty, so higher values mean more uncertainty. EMV moves with the VIX (CBOE Volatility Index), so increase in EMV means increase in market uncertainty. Finally, positive values of the NFCI indicate financial conditions that are tighter than average, therefore, more uncertainty. To sum up, increase in the UNC factor signals the increase in the overall uncertainty, which, in turn, makes both stock and oil risky investments. Investors would be getting rid of both risky assets and the stock-oil co-dependence would increase. Therefore, the uncertainty factor and the realized monthly correlation is positively correlated and we expect that the sign of UNC variable in the DSM copula model will be positive.

The LIQ factor presents the overall liquidity: it is positively correlated with LIX and Amivest liquidity measures, and negatively with Amihud illiquidity. Therefore, higher LIQ values mean higher liquidity overall. When the supply of oil and stock is abundant they can be easily bought and sold, therefore, the co-dependence increases and we expect positive sign of the LIQ variable in the DSM copula model.

The SOE factor is positively correlated with positive economic indicators, signaling favorable economic conditions. In general, when the economy is not in distress, the co-dependence between financial assets tends to become weaker. The opposite is also true: when the market is in turmoil the relationship becomes stronger. Therefore, the SOE factor is negatively correlated with the realized correlation and we expect the SOE variable to have a negative effect on the stock-oil co-dependence in the DSM copula model.

4.4 Results

In this section we first estimate the DSC and DSM copula models, with monthly lagged Rcor only as an explanatory variable, in order choose the most appropriate copula family based on the log marginal likelihood. Then, for the selected copula family, we identify the set of macro-finance factors that affect the dynamic co-dependence by estimating the DSM copula with other explanatory variables. The number of lags in each DSM copula model is fixed at 24 months so that the marginal likelihood is stable (increasing the lag beyond 24 months does not change the value of the log marginal likelihood.).

Table 5 reports the estimation results for the DSC models (without MIDAS effects) for a number of different copulas. Overall, the elliptical copula family is preferred over the Archimedean copula family according the the LML. The persistence parameter β is close to unity in all cases which allows the dependence parameter to follow an almost random walk process and quickly move into different regimes rather than into a mean reverting process. This is in-line with the empirically observed co-dependence: before the global financial crisis of 2008, the realized stock-oil correlations fluctuated around zero, and during/after the crisis turned positive. In order to capture such jump-type change with an AR(1)-type process, the persistence has to be very close to unity. The best log marginal likelihood is obtained with the t copula with estimated 19 degrees of freedom. This indicates that the stock-oil co-dependence is moderately fat-tailed and tail-symmetric, in accordance with the results in Avdulaj and Barunik (2015). Next, Table 6 reports the estimation results for the DSM copula models with RCor as an explanatory variable. In all cases, the marginal likelihoods are improved over the corresponding DSC alternatives. The DSM with t copula is the most preferred, however there is an increase in the parameter of the degrees of freedom and the persistence parameter is smaller than the previous cases. The decrease of β is not surprising: now the jump component is partially captured by the exogenous RCor variable. Hence, the RCor is effectively guiding the long-run stock-oil co-dependence, and the short-run component becomes a mean reverting process. The effect of lagged monthly RCor δ is positive and statistically significant for all copulas. In the DSM t copula, the weighting parameter of the restricted beta function implies that the effect of RCor decays after approximately 12 months.

Next, we identify the set of macro-finance factors that affect the dynamic co-dependence by

Table 5: Estimation results for the DSC model.

	λ_0	β	σ_e	ν	LML
DSC Gaussian	0.232	0.995	0.037		242.691
	(0.090;0.364)	(0.992;0.998)	(0.027;0.051)		
DSC Student	0.189	0.996	0.033	18.977	253.255
	(0.074;0.301)	(0.994;0.998)	(0.025;0.042)	(14.191;25.576)	
DSC Clayton	0.280	0.999	0.015		172.000
	(0.121;0.477)	(0.998;1.000)	(0.012;0.020)		
DSC Gumbel	0.271	0.997	0.023		214.275
	(0.095;0.412)	(0.995;0.999)	(0.012;0.032)		
DSC Frank	0.295	0.998	0.025		230.173
	(0.146;0.448)	(0.996;0.999)	(0.017;0.034)		
DSC Joe	0.167	0.998	0.011		144.187
	(0.088;0.245)	(0.998;0.999)	(0.008;0.013)		

The table reports the estimation results for the DSC model with different copulas. The numbers in the brackets show the [10% - 90%] credible intervals.

Table 6: Estimation results for the DSM copula model using RCor.

	λ_0	β	σ_e	δ	ω_2	ν	LML
DSM Gaussian	0.067	0.903	0.141	1.323	4.848		250.758
	(0.030;0.103)	(0.847;0.950)	(0.096;0.190)	(1.167;1.475)	(2.862;6.968)		
DSM Student	0.084	0.960	0.078	1.144	6.230	23.838	256.180
	(0.033;0.134)	(0.934;0.991)	(0.039;0.109)	(0.830;1.417)	(2.547;11.766)	(16.066;32.743)	
DSM Clayton	0.062	0.975	0.035	0.899	6.831		170.227
	(0.019;0.104)	(0.955;0.993)	(0.022;0.050)	(0.717;1.062)	(2.349;14.443)		
DSM Gumbel	0.078	0.979	0.044	1.045	5.456		214.449
	(0.032;0.134)	(0.964;0.993)	(0.028;0.061)	(0.820;1.267)	(2.031;10.155)		
DSM Frank	0.071	0.953	0.083	1.240	5.131		234.804
	(0.027;0.113)	(0.918;0.980)	(0.053;0.120)	(1.065;1.421)	(2.820;7.593)		
DSM Joe	0.069	0.991	0.017	0.638	16.264		146.258
	(0.029;0.112)	(0.985;0.997)	(0.012;0.023)	(0.416;0.833)	(4.075;27.951)		

The table reports the estimation results of the DSM copula model with different copulas. The long term dependence component is modelled using the RCor as an explanatory variable with the restricted beta weighting function. The selected lag length is such that the maximum likelihood becomes insensitive to the its choice ($K = 24$). The numbers in the brackets show the [10% - 90%] credible intervals.

estimating the DSM copula with one of the macro-finance factors as explanatory variables. Note that we use the t copula only, because it provided the largest log marginal likelihood in previous estimations. Table 7 reports the estimation results for the DSM copula models with one macro-finance explanatory variable. Only the model with UNC factor is comparable to the model with RCor in terms of log marginal likelihood. Among the four factors, only SOE gives an insignificant effect, while the signs of INR, UNC, and LIQ factors are in line with the economic reasoning and empirical findings. When the inflation and interest rates are high, oil prices may increase while the stock return is negative due to the flight-to-quality phenomenon. On the other hand, increase in the Uncertainty and Liquidity factors drive the stock-oil return co-dependence up. Uncertainty makes both stock and oil risky investments, thus, in uncertain times they are both sold-off in exchange for safer investments. When the supply of stock and oil is abundant (high liquidity), both assets are bought and sold at higher rates, therefore, the co-dependence moves higher. Finally, when the state of the economy is good, the co-dependence decreases, a well-known stylized factor of the financial returns (e.g. Conrad et al., 2014).

Table 7: Estimation results for the DSM copula model with one macro-finance factor.

	λ_0	β	σ_e^2	δ	ω_2	ν	LML
DSM RCor	0.084	0.960	0.078	1.144	6.230	23.838	256.180
	(0.033;0.134)	(0.934;0.991)	(0.039;0.109)	(0.830;1.417)	(2.547;11.766)	(16.066;32.743)	
DSM INR	0.234	0.997	0.027	-0.241	17.056	16.676	253.087
	(0.092;0.371)	(0.995;0.999)	(0.019;0.037)	(-0.462;-0.069)	(6.066;28.237)	(12.810;20.943)	
DSM UNC	0.193	0.987	0.043	0.237	5.226	18.841	256.466
	(0.135;0.249)	(0.979;0.994)	(0.027;0.059)	(0.172;0.298)	(1.348;15.596)	(13.862;24.322)	
DSM LIQ	0.197	0.996	0.030	0.101	14.432	18.022	246.742
	(0.059;0.340)	(0.994;0.999)	(0.021;0.038)	(0.005;0.182)	(4.233;25.855)	(13.056;24.166)	
DSM SOE	0.239	0.996	0.030	-0.074	13.451	17.803	250.586
	(0.134;0.341)	(0.993;0.998)	(0.022;0.039)	(-0.151;0.001)	(2.703;27.551)	(13.565;23.078)	

The table reports the estimation results of the DSM copula model (with t copula). The long term dependence component is modelled using one macro-finance factor as an explanatory variable with the restricted beta weighting function. The selected lag length is such that the maximum likelihood becomes insensitive to the its choice ($K = 24$). The numbers in the brackets show the [10% - 90%] credible intervals.

Figure 2 plots the equivalent long-run Pearson correlation between S&P 500 and WTI returns using DSM copula model with one macro-finance explanatory variable in comparison to the rolling correlation (RC1Y), calculated at the end of each month using previous 252 daily observations. The RCor model, as expected, tracks closely the long-run correlation. The models with INR and UNC (panels (b) and (c)) factors have a similar pattern with the RC1Y except during 2002-2008, which shows that one factor model might not be sufficient to capture the shifts of co-dependence in a long run. The DSM copula models with LIQ and SOE factors estimate a flat long run co-dependence.

Finally, we estimate the DSM copula model with two explanatory variables, see Table 8. The top panel of the table contains results for RCor plus one of the macro-finance factors. In general, the log marginal likelihood in these two factor models is better than the model with one macro-finance factor, i.e., including RCor improves model fit. The signs of the coefficients of the macro-finance variables are consistent, although, the size of the effect is smaller. This is expected, as RCor variable accounts for some long-run variation in the co-dependence, thus reducing the effect of the other explanatory variables.

Table 8: Estimation results for the DSM copula model with two macro-finance factors.

	λ_0	β	σ_e	δ_1	ω_1	δ_2	ω_2	ν	LML
RCor-INR	0.087	0.953	0.080	1.100	5.112	-0.047	14.848	24.002	252.059
	(0.043;0.132)	(0.928;0.977)	(0.053;0.109)	(0.902;1.337)	(2.663;7.780)	(-0.083;-0.010)	(3.183;27.522)	(16.067;34.531)	
RCor-UNC	0.110	0.945	0.089	0.921	5.087	0.125	8.808	23.023	258.572
	(0.063;0.161)	(0.915;0.969)	(0.064;0.118)	(0.623;1.183)	(2.436;8.559)	(0.071;0.189)	(1.506;22.692)	(15.892;31.533)	
RCor-LIQ	0.095	0.976	0.059	1.096	4.375	0.021	13.054	19.305	254.595
	(0.045;0.149)	(0.961;0.990)	(0.041;0.077)	(0.786;1.385)	(2.132;7.059)	(-0.023;0.063)	(3.081;25.013)	(14.496;25.320)	
RCor-SOE	0.114	0.984	0.045	0.699	9.714	-0.036	10.676	19.485	254.344
	(0.038;0.204)	(0.966;0.996)	(0.027;0.071)	(-0.049;1.267)	(2.673;23.675)	(-0.096;0.016)	(1.458;23.549)	(14.190;25.638)	
INR-UNC	0.250	0.993	0.034	-0.087	10.249	0.248	13.734	16.603	253.026
	(0.183;0.318)	(0.987;0.997)	(0.023;0.047)	(-0.155;-0.022)	(1.310;24.184)	(0.171;0.328)	(3.549;24.873)	(12.751;21.209)	
INR-LIQ	0.194	0.990	0.044	-0.145	17.783	0.085	11.229	19.987	250.519
	(0.127;0.266)	(0.986;0.995)	(0.033;0.056)	(-0.208;-0.081)	(6.542;27.645)	(0.038;0.131)	(1.910;23.505)	(14.302;26.383)	
INR-SOE	-0.150	0.999	0.023	-0.077	18.213	0.152	15.069	15.273	247.804
	(-0.737;0.296)	(0.998;1.000)	(0.018;0.029)	(-0.495;0.418)	(7.546;27.683)	(-0.229;0.588)	(4.423;25.758)	(11.966;18.847)	
UNC-LIQ	0.203	0.985	0.045	0.287	2.598	0.082	14.546	19.390	253.390
	(0.158;0.248)	(0.977;0.991)	(0.032;0.059)	(0.231;0.342)	(1.365;3.632)	(0.044;0.119)	(4.188;26.207)	(14.599;25.382)	
UNC-SOE	0.241	0.984	0.050	0.388	1.740	0.093	18.043	20.099	256.640
	(0.194;0.289)	(0.976;0.991)	(0.037;0.064)	(0.295;0.478)	(1.111;2.099)	(0.037;0.147)	(7.261;27.591)	(14.691;27.401)	
LIQ-SOE	0.204	0.989	0.052	0.109	12.980	-0.049	17.586	19.680	246.371
	(0.135;0.272)	(0.984;0.994)	(0.039;0.065)	(0.059;0.162)	(3.845;23.584)	(-0.100;-0.001)	(7.028;26.365)	(14.186;26.223)	

The table reports the estimation results of the DSM copula model (with t copula). The long term dependence component is modelled using two explanatory variables with the restricted beta weighting function. Top panel contains results for the RCor plus one other macro-finance factor, and bottom panel contains results for the two macro-finance factors as explanatory variables. The selected lag length in the weighting function is such that the maximum likelihood becomes insensitive to the its choice ($K = 24$). The numbers in the brackets show the [10% - 90%] credible intervals.

Figure 3 shows the equivalent long-run Pearson correlation between S&P 500 and WTI returns using DSM copula model with Rcor plus one of the macro-finance factors as explanatory variable. The long-run components are quite similar among two factor models but there are a few exceptions. The RCor-UNC (panel (b)) predicts that the long-run co-dependence is higher during the two most recent recessions while the RCor-SOE (panel (d)) estimates lower long-run co-dependence overall.

The bottom part of Table 8 reports the DSM copula model with two macro-finance factors. The signs of the effects are consistent with the previous results, and the three factors, INR, UNC and LIQ, remain statistically significant. However, we do not observe a significant over the one factor model in terms of log marginal likelihood. Some of the macro-finance factors are correlated,

Figure 2: Long-run co-dependence between S&P 500 and WTI returns: one explanatory variable.

The figure shows the long-run co-dependence (the equivalent Person's correlation of λ_τ) between S&P 500 and WTI returns using DSM copula models with one explanatory variable in comparison to the rolling correlation (RC1Y), calculated at the end of each month using previous 252 daily observations. The PC factor is shown as a bar chart. The shaded areas highlight the recession periods based on the NBER indicators.

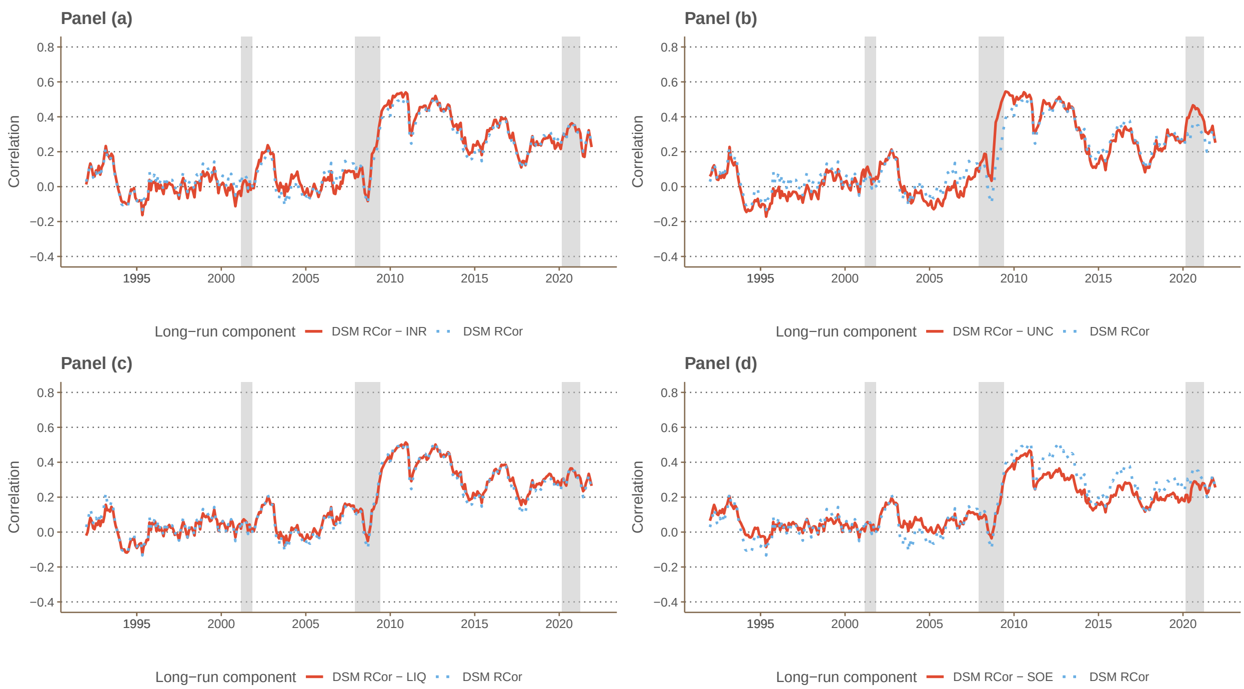

Figure 3: Long-run co-dependence between S&P 500 and WTI returns: RCor and one other macro-finance explanatory variable.

The figure shows the long-run co-dependence (the equivalent Person's correlation of λ_τ) between S&P 500 and WTI returns using DSM copula model with Rcor plus one of the macro-finance factors as explanatory variable. The shaded areas highlight the recession periods based on the NBER indicators.

see Table 4, inducing possible multicollinearity of the explanatory variables, hence the choice of the two macro-finance factors in the DSM copula model should be performed with care. Figure 4 plots the equivalent long-run Pearson correlation of the two macro-finance factor models. The DSM copula models with INR-UNC and UNC-LIQ as explanatory variables track the long run correlation reasonably well (panels (a) and (d)), meanwhile the long-run trends produced by the combination of INR-LIQ and INR-SOE factors deviate from the true correlation (panels (b) and (c)). In these cases the persistence coefficient of the short-run component is very close to unity (0.990 and 0.999, see bottom part of Table 8) in order to compensate the lack of proper fit from the long-run component.

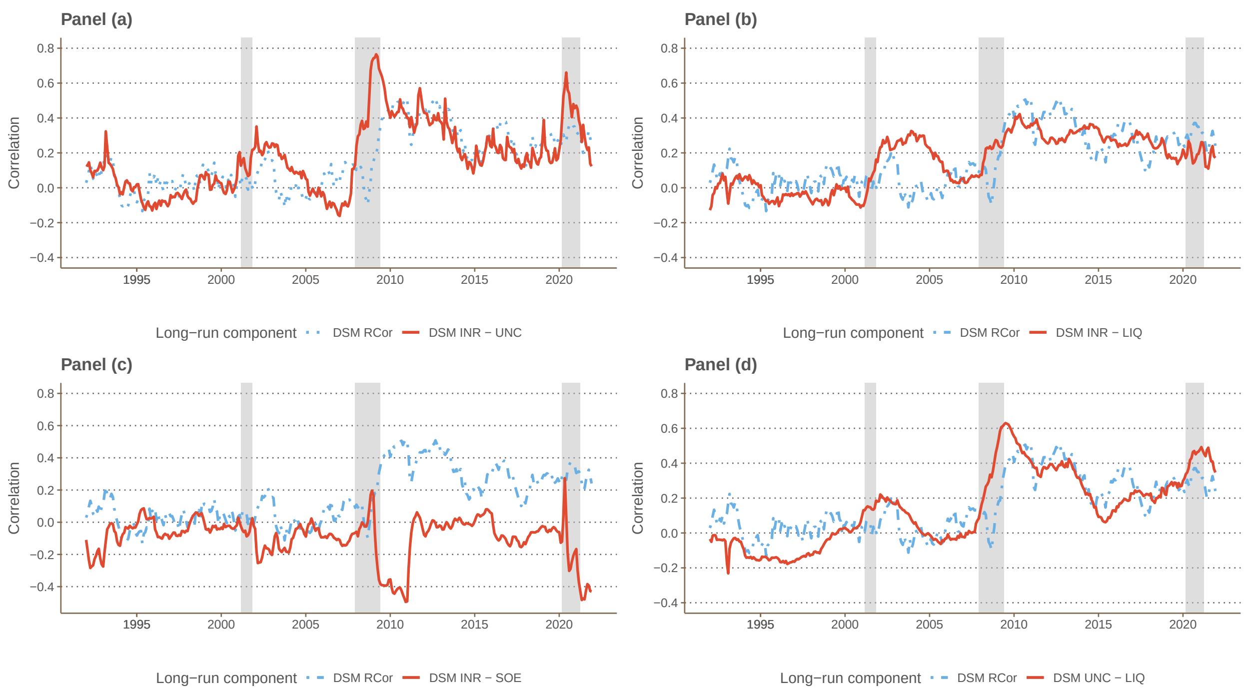

Figure 4: Long-run co-dependence between S&P 500 and WTI returns: two macro-finance explanatory variables.

The figure shows the long-run co-dependence (the equivalent Person's correlation of λ_τ) between S&P 500 and WTI returns using DSM copula models with two macro-finance explanatory variables in comparison to the rolling correlation (RC1Y), calculated at the end of each month using previous 252 daily observations. The shaded areas highlight the recession periods based on the NBER indicators.

Figure 5 shows the equivalent total correlation (the equivalent Person's correlation of λ_t) between S&P 500 and WTI returns using DSC model (without MIDAS effects), DSM copula with RCor, and DSM copula with the RCor-UNC factor, all with t copula. The total correlation is quite similar among all models since the total time varying co-dependence is mainly driven by the short-run component. However, the long-run component plays an important role in long-run forecasts, which can have an effect on the optimal portfolio allocation. Hence in the next section we perform k-step ahead prediction for the variance-covariance matrix and portfolio allocation exercises in order to quantify the economic gains from using the information from the macro-finance variables in the proposed DSM copula models.

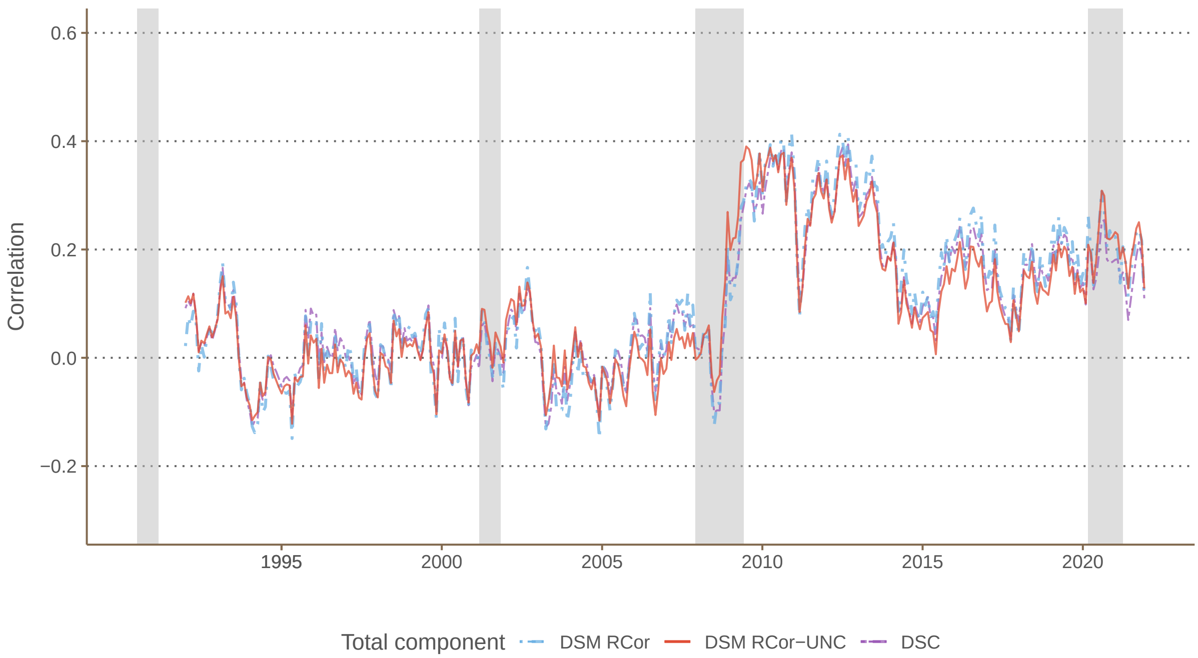

Figure 5: Total co-dependence between S&P 500 and WTI returns: DSC, DSM copula RCor and DSM copula with RCor-UNC factor models.

The figure shows the time varying co-dependence between S&P 500 and WTI returns using DSC model and DSM RCor, DSM RCor-UNC. The shaded areas highlight the recession periods based on the NBER indicators.

5 Economic evaluation

In this section, we evaluate the potential economic gains of using the DSM copula models. We split the sample into an in-sample period (from January 1990 to January 2018, 7059 observations) and an out-of-sample period (from January 2018 to December 2021 using the last 1000 observations). We re-estimate the DSM copula models using the in-sample data and use the out-of-sample observations to evaluate the performance of DSM copula model in portfolio allocation decisions.

5.1 Predictive variance and covariance

We estimate the GJR-GARCH model for daily log returns $r_{i,t}$, where $i = \{\text{S\&P 500}, \text{WTI}\}$, and obtain the set of marginal parameters $\widehat{\Theta}_i = (\widehat{\mu}_i, \widehat{\omega}_i, \widehat{\alpha}_i, \widehat{\beta}_i, \widehat{\gamma}_i, \widehat{\xi}_i, \widehat{\nu}_i)$. Define the realized cumulative returns and the realized cumulative covariance of stock and oil during the k periods at time t:

$$r_{i,t+1:t+k} = \sum_{j=1}^{k} r_{i,t+j},$$
$$\text{RCov}_{t+1:t+k} = \sum_{j=1}^{k} r_{1,t+j} r_{2,t+j}.$$

Following Engle (2009), the k-step-ahead forecast of the cumulative covariance can be approximated by

$$\text{Cov}(r_{1,t+1:t+k}, r_{2,t+1:t+k}) \approx \sum_{j=1}^{k} \rho_{t+j|t} \sqrt{\sigma^2_{1,t+j|t} \sigma^2_{2,t+j|t}}.$$

The Online Appendix D derives the k-step-ahead forecasts for the variance and cumulative variance of the GJR-GARCH model. For the k-step-ahead forecast of the correlation, we obtain $\rho_{t+j|t}$ as the equivalent Pearson correlation implied by copula models. At first, we obtain the k-step-ahead forecast of the stochastic process λ_t and use the transformation function $\Lambda_\rho = \sin(\frac{\pi}{2}\tau_\kappa) = \sin(\frac{\pi}{2}\frac{\exp(\lambda)-1}{\exp(\lambda)+1})$ to map the process into a restricted domain of the correlation:

$$\lambda_{t+k|t} = \lambda_\tau + (\lambda_t - \lambda_\tau)^k,$$
$$\rho_{t+k|t} = \Lambda_\rho(\lambda_{t+k|t}).$$

We compare the accuracy of covariance forecasts using the Mincer-Zarnowitz regression (Mincer and Zarnowitz, 1969) for the in-sample data. We regress the *ex-post* monthly realized covariance on a constant and the 22-days ahead cumulative covariance forecast. Table 9 reports the regression R^2 and the p-value of the F-test that the intercept is zero and the slope coefficient is one. The test does not reject the null at 5% confidence level for none of the models, and the R^2 is higher in DSM copula models with two factors than the simple DSC model. In other words, the DSM copula models produce 22-days ahead cumulative covariance forecasts that are closest to the actually observed realized covariance.

Table 9: Mincer-Zarnowitz Regressions for the in-sample data.

Model	R^2	p-value
DSC	0.292	0.077
DSM-RCor	0.262	0.374
DSM-UNC	0.398	0.433
DSM-RCor-UNC	0.384	0.656
DSM-INR-UNC	0.384	0.254
DSM-UNC-LIQ	0.397	0.956

The table reports the results for the Mincer-Zarnowitz regressions of monthly realized covariance on a constant and the 22-days ahead cumulative covariance forecast. The p-value is for the F-test that the intercept is zero and the slope coefficient is one.

5.2 Portfolio allocation

In this section we quantify the economic gains by performing two portfolio allocation exercises: minimum variance portfolio and hedging portfolio. In particular, we compare portfolios formed using the proposed models with MIDAS effects against a benchmark model (BM) without MIDAS. The benchmark model is the DSC, and the models for comparison are DSM copula models with either one or two explanatory factors: monthly realized correlation RCor, Uncertainty factor UNC, Inflation/interest rates factor INR and liquidity factor LIQ. We select the DSM copula models with statistically significant effects of the explanatory variables and a reasonable economic interpretation. Same as before, all models are with t copula.

Minimum variance portfolio

For the first portfolio allocation exercise we consider the classical dynamic minimum variance portfolio. Such portfolio minimizes the overall risk independently of the portfolio returns. The investor can obtain optimal weights by solving the following minimization problem:

$$\arg\min_{\mathbf{w}_t} \mathbf{w}_t^{'} \mathbf{H}_{t,k} \mathbf{w}_t, \qquad s.t. \mathbf{w}_t^{'} \mathbf{1} = 1,$$

where $\mathbf{w}_t = [w_{1,t}, w_{2,t}]'$ is the vector of optimal weights and the k-step-ahead forecast of the conditional variance-covariance is as follows: $\mathbf{H}_{t,k} = \begin{bmatrix} \operatorname{Var}(r_{1,t+1:t+k}) & \operatorname{Cov}(r_{1,t+1:t+k}, r_{2,t+1:t+k}) \\ \operatorname{Cov}(r_{1,t+1:t+k}, r_{2,t+1:t+k}) & \operatorname{Var}(r_{2,t,k}) \end{bmatrix}$. If there is no short-selling constraint, then the the minimum variance portfolio weights can be obtained analytically: $\mathbf{w}_t = (\mathbf{H}_{t,k}^{-1}\mathbf{1})/(\mathbf{1}'\mathbf{H}_{t,k}^{-1}\mathbf{1})$. We assume that the investor holds this portfolio for k-days and readjusts the portfolio weights after that.

Hedging portfolio

For the second portfolio allocation exercise, similarly as in Conrad and Stürmer (2017), we consider a dynamic hedging portfolio. In such portfolio the investor holds stock and then adds oil to reduce risk. The portfolio weights can be obtained by solving the following:

$$\arg\min_{\mathbf{w}_t} \mathbf{w}_t' \mathbf{H}_{t,k} \mathbf{w}_t, \qquad s.t. \mathbf{w}_t' \boldsymbol{\mu} = \mu_0,$$

where $\boldsymbol{\mu} = [\mu_0, 0]'$ is the vector of the expected excess returns on stocks and oil. Here μ_0 is the required excess return, which is the same as the expected excess returns of stock. Then the optimal hedging portfolio weights are given by:

$$\mathbf{w}_t = (1, -\operatorname{Cov}(r_{1,t+1:t+k}, r_{2,t+1:t+k}) / \operatorname{Var}(r_{2,t+1:t+k}))'.$$

Since $\mathbf{w}_t = [w_{1,t}, w_{2,t}]' = [1, w_{2,t}]'$, then, $1 - w_{1,t} - w_{2,t} = -w_{2,t}$ is held in cash. Apart from holding stocks, investor will either hold or short-sell oil. Because the sign of the weight of oil $w_{2,t}$ is opposite to the sign of the covariance, it means that investor will hold oil if covariance is negative and short-sell oil if covariance is positive.

Comparison in terms of portfolio variance

First, we are interested in whether the DSM copula models deliver smaller portfolio variance than the DSC model. Engle and Colacito (2006) show that the portfolio variance is minimized when using the true conditional covariance. Engle and Colacito (2006) propose a test for the equality of the variances of two portfolios based on the predictive covariance $\mathbf{H}_{t,k}$. The difference between

k-step-ahead portfolio variances at time t is defined as follows:

$$d_{t,k} = \left(\mathbf{w}_t^{'}(\mathbf{r}_{t,t+1:t+k} - k\bar{\mathbf{r}})\right)^2 - \left(\mathbf{w}_{BM,t}^{'}(\mathbf{r}_{t,t+1:t+k} - k\bar{\mathbf{r}})\right)^2,$$

where $\mathbf{w}_{BM,t}$ are the benchmark portfolio weights and $\bar{\mathbf{r}}$ is the sample mean of $\mathbf{r}_t$. The null hypothesis is that the mean of $d_{t,k}$ is 0. The testing can be done via Diebold-Mariano type test.

Similarly as in Conrad and Stürmer (2017), we estimate the gain/loss (G/L) of using the DSM copula model as compared to the DSC model (the benchmark):

$$G/L = 100\%(\sigma_{BM} - \sigma_P)/\sigma_P, \tag{4}$$

where σ_{BM}^2 and σ_P^2 are the overall variances of the portfolios that are formed based on the predicted variance-covariance matrices of the DSC model and some DSM copula model, respectively. The G/L ratio shows the percentage of the expected return gain if we use a DSM copula model over the benchmark. Portfolio with positive and higher G/L ratio is preferred to the benchmark and other portfolios.

Comparison in terms of other portfolio metrics

A commonly used metric for portfolio comparison is the total portfolio turnover from period $t-1$ to period t, defined as

$$TO_t = \sum_{i=1}^{2} \left| w_{i,t} - w_{i,t-1} \frac{1 + r_{i,t-1}}{1 + r_{P,t-1}} \right|.$$

Larger turnover implies larger transaction costs, therefore, the portfolio return needs to be adjusted by such costs, resulting into portfolio excess return $r_{E,t} = r_{P,t} - cTO_t$. Here c is some small constant, which we set to 1%. Finally, we also calculate the portfolio Sharpe ratio $SR = r_E/\sigma_P$. Note that for simplicity in all portfolio comparison metrics we assume that the risk-free interest rate is equal to zero.

5.3 Portfolio results

We consider dynamic portfolio allocation for the in-sample and out-of-sample periods, with latter also split in pre-Covid and post-Covid. The in-sample evaluation period is from January 1992 to January 2018 which contains 7059 daily and 337 monthly observations. The out-of-sample evaluation period is up to December 2021 which contains 1000 daily and 48 monthly observations split at the end of January 2020 into pre- and post-Covid (518 and 482 daily observations, respectively). We consider the daily and monthly portfolios, where monthly portfolio is re-balanced every month. Monthly re-balancing means forming the portfolio, keeping it for 22 days and then observing the realized returns.

Tables 10 and 11 report the gain/loss criteria, defined in (4), for in- and out-of-sample periods. The G/L ratio is calculated with respect to the benchmark portfolio, which is based on the DSC model (without MIDAS effects). As seen in Table 10, the in-sample G/L ratio is almost always positive, indicating that all MIDAS specifications produce portfolios with smaller variances, meaning that the covariance matrices are closer to the "true" conditional covariance matrix than the one produced by the benchmark model (Engle and Colacito, 2006). The asterisk next to the G/L ratio indicates that the difference in variances is statistically significant at 5% confidence level. Out-of-sample results are not as consistent, see Table 11. Top panel contains results for the minimum variance portfolio. Here the daily G/L is always positive, but the difference in portfolio variances is not statistically significant. Monthly portfolio results are not so overwhelmingly favorable to the DSM copula models, especially if we split the sample into pre- and post-Covid. However, overall, the DSM copula specifications produce portfolios with either the same or statistically smaller variances than the benchmark model, except for one instance (pre-Covid period, the DSM-INR-UNR model). The bottom panel displays results for the hedge portfolio. Here during the pre-Covid period daily and monthly portfolios are always better, albeit, statistically insignificant. For the post-Covid and the total in-sample period the results are mixed, but two models consistently provide positive G/L values over the benchmark: DSM copula with RCor and DSM copula with RCor and UNC factors. Incidentally, these two models are the ones with the highest overall log marginal likelihoods. However, the differences in variances is statistically insignificant. To sum up, results from Tables 10- 11 suggest that the portfolios formed using the DSM copula models have statistically equal or smaller variances than the benchmark model, both in-sample and out-of-sample, for daily and monthly portfolios, with only one exception.

Table 10: In-sample portfolio evaluation: G/L.

Model	Minimum variance portfolio		Hedge portfolio	
	Daily	Monthly	Daily	Monthly
DSM-RCor	0.520*	0.338	1.543*	-0.232
DSM-UNC	0.299	0.603	1.099*	4.055
DSM-RCor-UNC	0.931*	0.798	3.051*	3.816
DSM-INR-UNC	0.255	0.406	0.763	3.145
DSM-UNC-LIQ	0.538*	0.969	1.654*	3.985

The table presents the G/L in percentage for the in-sample period (01/01/1992 - 10/01/2018) for daily and monthly minimum variance and hedge portfolios. The G/L ratio shows the percentage of the expected return gain if we use a DSM copula model over the benchmark. We test for the equality of the portfolio variances using the Diebold Mariano test. The asterisk next to the G/L ratio indicates that the test is significant at 5% level.

Table 11: Out-of-sample portfolio evaluation: G/L.

Model	Daily	Monthly	Daily	Monthly	Daily	Monthly
	Pre-Covid		Post-Covid		Total	
Minimum variance portfolio						
DSM-RCor	0.309	-1.001	0.115	-0.042	0.157	-0.218
DSM-UNC	0.128	0.143	0.893	1.130*	0.725	0.971*
DSM-RCor-UNC	0.050	-2.010*	0.566	0.365	0.452	-0.064
DSM-INR-UNC	0.122	-0.048	0.778	1.221*	0.634	1.016*
DSM-UNC-LIQ	0.186	1.031*	0.685	1.259*	0.575	1.238*
Hedge portfolio						
DSM-RCor	0.243	0.769	0.626	0.017	0.527	0.265
DSM-UNC	0.023	0.723	-0.768	2.451	-0.566	1.418
DSM-RCor-UNC	0.093	0.091	0.527	2.679	0.415	1.227
DSM-INR-UNC	0.155	0.975	-0.863	4.670	-0.603	2.846
DSM-UNC-LIQ	0.348	1.406	-0.289	-4.005	-0.126	-1.877

The table presents the G/L in percentage for the out-of-sample period (11/01/2018 - 31/12/2021) for daily and monthly minimum variance and hedge portfolios. The out-of-sample is split on 31/01/2020 for pre- and post-Covid. The G/L ratio shows the percentage of the expected return gain if we use a DSM copula model over the benchmark. We test for the equality of the portfolio variances using the Diebold Mariano test. The asterisk next to the G/L ratio indicates that the test is significant at 5% level.

Finally, we investigate the performance of the minimum variance and hedge portfolios in economic terms. We are especially interested in the portfolio standard deviation (σ_P) for the global minimum variance portfolio and the portfolio Sharpe ratio for the hedge portfolio. The Sharpe ratio is calculated as the ratio of portfolio excess return over the standard standard deviation; where the excess return is obtained as the overall portfolio return minus $0.01\times$ the turnover (also reported).

Left panel of Table 12 contains the results for the in-sample period for the minimum variance portfolio. The first three smallest overall variances are produced by the DSM copula specifications with RCor-UNC, UNC-LIQ and Rcor factors as explanatory variables, in that order, for daily; and UNC-LIQ, Rcor-UNC and UNC factors as explanatory variables, in that order, for monthly portfolios. The results remain virtually the same for the out-of-sample period as well, see left panel of Table 13. The first three smallest overall variances are produced by the DSM copula specifications: with UNC, INR-UNC and UNC-LIQ factors as explanatory variables, in that order, for daily; and UNC-LIQ, Rcor-UNC and UNC factors as explanatory variables - identical to the in-sample ordering - for monthly portfolios. To sum up, incorporating information from low frequency macro-finance variables into modeling the co-dependence of stock-oil returns results into more precise 1- and 22-step ahead predicted variance covariance matrices. This translates into daily and monthly global minimum variance portfolios with the lowest variance overall, in- and out-of-sample.

As for hedging portfolio, we are primarily interested in the Sharpe ratios. In our case the calculated Sharpe ratio is a rather comprehensive metric for portfolio comparison because it also penalizes the large turnover. For the in-sample period (right panel of the Table 12), the highest Sharpe ratios are produced by the DSM copula models with essentially the same explanatory variables as in the best minimum variance portfolios: Uncertainty, Liquidity and RCor factors, either alone or in combinations. The result is the same for daily and monthly horizons, with a slight change in model ordering. Out-of-sample results change only very slightly (right panel of the Table 13), where the DSC specification (without MIDAS effects) is among the top three models, and the remaining best models are again the ones using RCor, UNC and LIQ as explanatory variables (either alone or in combinations), for daily and monthly hedging portfolios.

6 Conclusions

The paper has proposed a new approach to model the time varying stock-oil co-dependence via dynamic stochastic mixed data frequency sampling (MIDAS) copula model, DSM copula in short. Such model decomposes the stock-oil relationship into a short-run dynamic stochastic component

Table 12: In-sample portfolio evaluation: economic gains.

Model	σ_P	r_E	Sharpe	TO	σ_P	r_E	Sharpe	TO
Daily	Minimum variance portfolio				Hedge portfolio			
DSC	15.654	9.281	0.593	7.724	16.734	7.578	0.453	2.062
DSM-RCor	15.613	9.403	0.602	7.996	16.606	7.781	0.469	2.567
DSM-UNC	15.630	9.223	0.590	7.774	16.643	7.839	0.471	2.068
DSM-RCor-UNC	15.581	9.345	0.600	8.061	16.484	8.111	0.492	2.662
DSM-INR-UNC	15.634	9.228	0.590	7.765	16.670	7.799	0.468	2.056
DSM-UNC-LIQ	15.612	9.290	0.595	7.852	16.597	7.888	0.475	2.231
Monthly	Minimum variance portfolio				Hedge portfolio			
DSC	15.295	6.139	0.401	0.413	14.199	7.845	0.553	0.180
DSM-RCor	15.271	6.216	0.407	0.418	14.216	7.783	0.548	0.201
DSM-UNC	15.248	6.022	0.395	0.412	13.920	8.109	0.583	0.197
DSM-RCor-UNC	15.234	6.085	0.399	0.416	13.937	8.081	0.580	0.206
DSM-INR-UNC	15.264	6.043	0.396	0.412	13.982	8.075	0.578	0.200
DSM-UNC-LIQ	15.222	6.079	0.399	0.411	13.927	8.138	0.584	0.202

The table presents various portfolio metrics, all annualized, for the in-sample period (01/01/1992 - 10/01/2018) for daily and monthly minimum variance and hedge portfolios. σ_P is portfolio standard deviation, r_E is portfolio excess return calculated as the overall portfolio return minus $0.01 \times TO$, where TO is the turnover.

Table 13: Out-of-sample portfolio evaluation: economic gains.

Model	σ_P	r_E	Sharpe	TO	σ_P	r_E	Sharpe	TO
Daily	Minimum variance portfolio				Hedge portfolio			
DSC	22.881	2.464	0.108	8.107	20.290	18.869	0.930	2.573
DSM-RCor	22.863	2.127	0.093	8.347	20.237	18.836	0.931	2.880
DSM-UNC	22.799	2.365	0.104	7.991	20.349	18.587	0.913	2.636
DSM-RCor-UNC	22.830	1.874	0.082	8.401	20.249	18.430	0.910	2.970
DSM-INR-UNC	22.809	2.324	0.102	8.012	20.353	18.456	0.907	2.629
DSM-UNC-LIQ	22.816	2.235	0.098	8.243	20.303	18.864	0.929	2.826
Monthly	Minimum variance portfolio				Hedge portfolio			
DSC	27.365	5.958	0.218	0.461	13.812	18.369	1.330	0.359
DSM-RCor	27.393	5.709	0.208	0.455	13.811	18.096	1.310	0.344
DSM-UNC	27.243	5.870	0.215	0.448	13.758	17.741	1.289	0.387
DSM-RCor-UNC	27.376	5.561	0.203	0.448	13.774	17.654	1.282	0.340
DSM-INR-UNC	27.235	5.783	0.212	0.450	13.664	17.536	1.283	0.386
DSM-UNC-LIQ	27.205	5.882	0.216	0.450	13.969	18.215	1.304	0.401

The table presents various portfolio metrics, all annualized, for the out-of-sample period (11/01/2018 - 31/12/2021) for daily and monthly minimum variance and hedge portfolios. σ_P is portfolio standard deviation, r_E is portfolio excess return calculated as the overall portfolio return minus $0.01 \times TO$, where TO is the turnover.

and a long-run component, driven by a set macro-finance factors, assembled into four major groups via principal components: inflation/interest rate, uncertainty, liquidity and state of the economy factors.

The proposed approach leads to a better understanding of the economic drivers of the stock-oil co-dependence. In particular, by using data from more that 30 years, we found that inflation/interest rate factor has a negative and statistically significant effect on the co-movement of the stock-oil returns. Moreover, uncertainty and liquidity factors have a positive and statistically significant effect on the co-dependence. We have shown that incorporating low frequency macro-finance factors leads to better in-sample fit in terms of marginal log likelihoods as compared to the benchmark specification of Hafner and Manner (2012).

We have performed several portfolio allocation exercises and have obtained three major results. Firstly, the portfolios formed using the DSM copula models have statistically equal or smaller variances than the benchmark model, both in-sample and out-of-sample, for daily and monthly portfolios, with one exception only. According the Engle and Colacito (2006) this implies that the estimated variance-covariance matrix is closer to the true one. Secondly, DSM copula models produce daily and monthly global minimum variance portfolios with the lowest variance overall, in- and out-of-sample. Past realized correlation, liquidity and uncertainty are the most useful explanatory factors. Finally, the same realized correlation, liquidity and uncertainty factors remain essential for forming hedge portfolios with the highest Sharpe ratios: in-sample best daily and monthly portfolios are produced using the DSM copula model with either one or combination of two factors, and for out-of-sample the models using these factors are still in the top three best models, with a benchmark DSC model sometimes being better.

Finally, even though the estimation using the DTSMC approach is time-consuming, the multi-core server can reduce the burden of the inference significantly. It is also straightforward to extend the DSM model in higher dimensions, as a bivariate copula is a building block for vine copulas and factor copulas (Czado, 2019; Joe, 2015; Nguyen et al., 2020). One can also introduce hierarchical priors to account for a model with many regressors or propose a recurrent neutral network to track down the dynamic co-dependence similar to Nguyen et al. (2022) proposal of an univariate stochastic volatility model.

References

Almeida, C. and Czado, C. (2012). "Efficient Bayesian Inference for Stochastic Time-Varying Copula Models." *Computational Statistics & Data Analysis*, 56(6): 1511–1527.

Aloui, C., Nguyen, D. K., and Njeh, H. (2012). "Assessing the impacts of oil price fluctuations on stock returns in emerging markets." *Economic Modelling*, 29(6): 2686–2695.

Aloui, R., Hammoudeh, S., and Nguyen, D. K. (2013). "A time-varying copula approach to oil and stock market dependence: The case of transition economies." *Energy Economics*, 39: 208–221.

Amihud, Y. (2002). "Illiquidity and stock returns: Cross-section and time-series effects." *Journal of Financial Markets*, 5(1): 31–56.

Amihud, Y., Mendelson, H., and Lauterbach, B. (1997). "Market microstructure and securities values: Evidence from the Tel Aviv stock exchange." *Journal of Financial Economics*, 45(3): 365–390.

Amihud, Y. and Wohl, A. (2004). "Political news and stock prices: The case of Saddam Hussein contracts." *Journal of Banking and Finance*, 28: 1185–1200.

Andersen, T. G., Bollerslev, T., Frederiksen, P., and Ørregaard Nielsen, M. (2010). "Continuous-time models, realized volatilities, and testable distributional implications for daily stock returns." *Journal of Applied Econometrics*, 25: 233–261.

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

Asgharian, H., Christiansen, C., and Hou, A. J. (2016). "Macro-finance determinants of the long-run stock–bond correlation: The DCC-MIDAS specification." *Journal of Financial Econometrics*, 14(3): 617–642.

Ausin, M. C. and Lopes, H. F. (2010). "Time-varying joint distribution through copulas." *Computational Statistics & Data Analysis*, 54(11): 2383–2399.

Avdulaj, K. and Barunik, J. (2015). "Are benefits from oil–stocks diversification gone? New evidence from a dynamic copula and high frequency data." *Energy Economics*, 51: 31–44.

Basher, S. A. and Sadorsky, P. (2016). "Hedging emerging market stock prices with oil, gold, VIX, and bonds: A comparison between DCC, ADCC and GO-GARCH." *Energy Economics*, 54(Feb): 235–247.

Batten, J. A., Kinateder, H., Szilagyi, P. G., and Wagner, N. F. (2017). "Can stock market investors hedge energy risk? Evidence from Asia." *Energy Economics*, 66(Aug): 559–570.

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

Chao Wei (2003). "Energy , the Stock Market , and the Putty-Clay Investment Model." *The American Economic Review*, 93(1): 311–323.

Chib, S., Omori, Y., and Asai, M. (2009). "Multivariate Stochastic Volatility." In Mikosch, T., Kreiß, J.-P., Davis, R. A., and Andersen, T. G. (eds.), *Handbook of Financial Time Series*, chapter 16, 365–400. Berlin, Heidelberg: Springer Berlin Heidelberg.

Choi, K. and Hammoudeh, S. (2010). "Volatility behavior of oil, industrial commodity and stock markets in a regime-switching environment." *Energy Policy*, 38(8): 4388–4399.

Colacito, R., Engle, R. F., and Ghysels, E. (2011). "A component model for dynamic correlations." *Journal of Econometrics*, 164(1): 45–59.

Conrad, C. and Kleen, O. (2020). "Two are better than one: Volatility forecasting using multiplicative component GARCH-MIDAS models." *Journal of Applied Econometrics*, 35(1): 19–45.

Conrad, C. and Loch, K. (2016). "Macroeconomic expectations and the time-varying stock-bond correlation: international evidence." Vfs annual conference 2016 (augsburg): Demographic change, Verein für Socialpolitik / German Economic Association.

Conrad, C., Loch, K., and Rittler, D. (2014). "On the macroeconomic determinants of long-term volatilities and correlations in US stock and crude oil markets." *Journal of Empirical Finance*, 29(Dec): 26–40.

Conrad, C. and Stürmer, K. (2017). "On the Economic Determinants of Optimal Stock-Bond Portfolios: International Evidence." *University of Heidelberg, Department of Economics, Discussion Paper Series*, 636: 1–54.

Czado, C. (2019). *Analyzing dependent data with vine copulas*. Springer.

Danyliv, O., Bland, B., and Nicholass, D. (2014). "Convenient liquidity measure for Financial markets." *https://arxiv.org/abs/1412.5072*.

Del Moral, P., Doucet, A., and Jasra, A. (2006). "Sequential monte carlo samplers." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68(3): 411–436.

Deligiannidis, G., Doucet, A., and Pitt, M. K. (2018). "The correlated pseudomarginal method." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 80(5): 839–870.

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

Duan, J. C. and Fulop, A. (2015). "Density-Tempered Marginalized Sequential Monte Carlo Samplers." *Journal of Business and Economic Statistics*, 33(2): 192–202.

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.

— (2002b). "New frontiers for ARCH models." *Journal of Applied Econometrics*, 17: 425–446.

— (2009). *Anticipating correlations*. Princeton University Press.

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

Fernández, C. and Steel, M. F. (1998). "On Bayesian modeling of fat tails and skewness." *Journal of the American Statistical Association*, 93(441): 359–371.

Ferreira, J. T. and Steel, M. F. (2006). "A constructive representation of univariate skewed distributions." *Journal of the American Statistical Association*, 101(474): 823–829.

Ghalanos, A. (2022). "rugarch: Univariate GARCH models." R package version 1.4-7.

Ghysels, E., Santa-Clara, P., and Valkanov, R. (2004). "The MIDAS Touch: Mixed Data Sampling Regression Models." Cirano working papers, CIRANO.

Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993). "On the relation between the expected value and the volatility of the nominal excess return on stocks." *Journal of Finance*, 48(5): 1779–1801.

Gong, Y., Bu, R., and Chen, Q. (2020). "What Affects the Relationship Between Oil Prices and the US Stock Market? A Mixed-Data Sampling Copula Approach." *Journal of Financial Econometrics*, 20(2): 253–277.

Gong, Y., Chen, Q., and Liang, J. (2018). "A mixed data sampling copula model for the return-liquidity dependence in stock index futures markets." *Economic Modelling*, 68(Jan): 586–598.

Hafner, C. M. and Manner, H. (2012). "Dynamic stochastic copula models: Estimation, inference and applications." *Journal of Applied Econometrics*, 27(2): 269–295.

Huang, R. D., Masulis, R. W., and Stoll, H. R. (1996). "Energy Shocks and Financial Markets." *Journal of Futures Markets*, 16(1): 1–27.

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

Jondeau, E. and Rockinger, M. (2006). "The Copula-GARCH Model of Conditional Dependencies: An International Stock Market Application." *Journal of International Money and Finance*, 25(5): 827–853.

Jones, C. M. and Kaul, G. (1996). "Oil and the Stock Markets." *Journal of Finance*, 51(2): 463–491.

Kastner, G. and Frühwirth-Schnatter, S. (2014). "Ancillarity-sufficiency interweaving strategy (ASIS) for boosting MCMC estimation of stochastic volatility models." *Computational Statistics & Data Analysis*, 76(Aug): 408–423.

Kilian, L. (2008). "The economic effects of energy price shocks." *Journal of Economic Literature*, 46(4): 871–909.

Kilian, L. and Park, C. (2009). "The impact of oil price shocks on the U.S. stock market." *International Economic Review*, 50(4): 1267–1287.

Liesenfeld, R. and Jung, R. C. (2000). "STOCHASTIC VOLATILITY MODELS: CONDITIONAL NORMALITY VERSUS HEAVY-TAILED DISTRIBUTIONS." *Journal of Applied Econometrics*, 15: 137–160.

Loaiza-Maya, R., Smith, M. S., and Maneesoonthorn, W. (2018). "Time series copulas for heteroskedastic data." *Journal of Applied Econometrics*, 33: 332–354.

Marshall, B. R., Nguyen, N. H., and Visaltanachoti, N. (2012). "Commodity liquidity measurement and transaction costs." *Review of Financial Studies*, 25(2): 599–638.

McNeil, A. J., Frey, R., and Embrechts, P. (2015). *Quantitative risk management: concepts, techniques and tools-revised edition*. Princeton university press.

Mensi, W., Hammoudeh, S., Shahzad, S. J. H., and Shahbaz, M. (2017). "Modeling systemic risk and dependence structure between oil and stock markets using a variational mode decomposition-based copula method." *Journal of Banking and Finance*, 75(Feb): 258–279.

Mincer, J. A. and Zarnowitz, V. (1969). "The evaluation of economic forecasts." In *Economic forecasts and expectations: Analysis of forecasting behavior and performance*, 3–46. NBER.

Mollick, A. V. and Assefa, T. A. (2013). "U.S. stock returns and oil prices: The tale from daily data and the 2008-2009 financial crisis." *Energy Economics*, 36(Mar): 1–18.

Neal, R. M. (2001). "Annealed importance sampling." *Statistics and computing*, 11(2): 125–139.

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

Nguyen, H., Ausín, M. C., and Galeano, P. (2020). "Variational inference for high dimensional structured factor copulas." *Computational Statistics & Data Analysis*, 151: 107012.

Nguyen, H. and Javed, F. (2021). "Dynamic relationship between Stock market and Bond market: A GAS MIDAS copula approach." *Working Paper (Statistics)*, 15: 1–45.

Nguyen, T.-N., Tran, M.-N., Gunawan, D., and Kohn, R. (2022). "A Statistical Recurrent Stochastic Volatility Model for Stock Markets." *Journal of Business & Economic Statistics*, 0(just-accepted): 1–15.

Pan, Z. (2014). "Modelling tail dependence between energy market and stock markets in the BRIC countries." *Applied Economics Letters*, 21(11): 789–794.

Pan, Z., Wang, Y., and Liu, L. (2016). "The relationships between petroleum and stock returns: An asymmetric dynamic equi-correlation approach." *Energy Economics*, 56(May): 453–463.

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

— (2009). "Copula - Based Models for Financial Time Series." In Andersen, T. G., Mikosch, T., Kreiß, J.-P., and Davis, R. A. (eds.), *Handbook of Financial Time Series*, chapter 36, 767–785. Berlin, Heidelberg: Springer Berlin Heidelberg.

— (2012). "A review of copula models for economic time series." *Journal of Multivariate Analysis*, 110(Sep): 4–18.

— (2013). "Copula methods for forecasting multivariate time series." *Handbook of Economic Forecasting*, 2(May): 899–960.

Sadorsky, P. (1999). "Oil price shocks and stock market activity." *Energy Economics*, 21(5): 449–469.

— (2014). "Modeling volatility and correlations between emerging market stock prices and the prices of copper, oil and wheat." *Energy Economics*, 43(May): 72–81.

Sklar, M. (1959). *Fonctions de répartition à n dimensions et leurs marges*. Université Paris 8.

Sukcharoen, K., Zohrabyan, T., Leatham, D., and Wu, X. (2014). "Interdependence of oil prices and stock market indices: A copula approach." *Energy Economics*, 44(Jul): 331–339.

Tran, M.-N., Scharth, M., Pitt, M. K., and Kohn, R. (2014). "Annealed Important Sampling for Models with Latent Variables." *Technical Report, University of New South Wales, Australia*, arXiv:1402.6035(5): 1–27.

Tse, Y. K. and Tsui, A. K. C. (2002). "A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations." *Journal of Business & Economic Statistics*, 20(3): 351–362.

Virbickaite, A., Lopes, H. F., and Zaharieva, M. D. (2022). "Dynamic mixed frequency density pooling." *Working Paper*, 0(0): 1–35.

Vo, M. (2011). "Oil and stock market volatility: A multivariate stochastic volatility perspective." *Energy Economics*, 33(5): 956–965.

Yu, L., Zha, R., Stafylas, D., He, K., and Liu, J. (2020). "Dependences and volatility spillovers between the oil and stock markets: New evidence from the copula and VAR-BEKK-GARCH models." *International Review of Financial Analysis*, 68(Mar): 101280.

Zhu, H. M., Li, R., and Li, S. (2014). "Modelling dynamic dependence between crude oil prices and Asia-Pacific stock market returns." *International Review of Economics and Finance*, 29: 208–223.

Online Appendix

A Copulas

Here we summarize the CDFs, PDFs and the relationship between the copula-specific parameter θ and the Kendall's tau τ_κ for the bivariate Gaussian, Student-t, Clayton, Gumbel, Frank and Joe copulas (Joe, 2015).

Bivariate Gaussian copula:

Let $x = \Phi^{-1}(u)$ and $y = \Phi^{-1}(v)$, where Φ^{-1} is the inverse of the univariate standard normal distribution function. The bivariate Gaussian copula CDF and PDF are given by:

$$\begin{aligned} C(u,v;\theta) &= \Phi_2\left(x,y;\theta\right), \quad 0<u,v<1, \quad \theta \in [-1,1], \\ c(u,v;\theta) &= \frac{\phi_2\left(\Phi^{-1}(u),\Phi^{-1}(v);\theta\right)}{\phi\left(\Phi^{-1}(u)\right)\cdot\phi\left(\Phi^{-1}(v)\right)} = (1-\theta^2)^{-1/2}\exp\left\{\frac{2\theta xy-\theta^2(x^2+y^2)}{2(1-\theta^2)}\right\}, \end{aligned}$$

where $\Phi_2(\cdot;\theta)$ is the standard bivariate normal distribution function with correlation coefficient θ. The relationship between Kendall's tau and copula parameter is given by:

$$\tau_\kappa = 2\arcsin(\theta)/\pi, \qquad \theta = \sin(\pi\tau_\kappa/2).$$

Bivariate t copula:

Let $x = T_\nu^{-1}(u)$ and $y = T_\nu^{-1}(v)$, where T_ν^{-1} is the inverse of the univariate t-distribution function with ν degrees of freedom. The bivariate t copula CDF and PDF are given by:

$$\begin{aligned} C(u,v;\theta,\nu) &= T_{2,\nu}(x,y;\theta), \quad 0<u,v<1, \quad \theta\in[-1,1], \\ c(u,v;\theta,\nu) &= \frac{t_{2,\nu}(x,y;\theta)}{t_\nu(x)\cdot t_\nu(y)} = \frac{1}{\sqrt{1-\theta^2}}\frac{\Gamma((\nu+2)/2)\Gamma(\nu/2)}{\Gamma^2((\nu+1)/2)}\frac{\left(1+\frac{x^2+y^2-2\theta xy}{\nu(1-\theta^2)}\right)^{-(\nu+2)/2}}{\left(1+\frac{x^2}{2}\right)^{-(\nu+1)/2}\left(1+\frac{y^2}{2}\right)^{-(\nu+1)/2}}, \end{aligned}$$

where $T_{2,\nu}(\cdot;\theta)$ is the bivariate t distribution function with correlation coefficient θ and degrees of freedom parameter ν. The relationship between Kendall's tau and copula parameter is the same as in Gaussian copula and is given by:

$$\tau_\kappa = 2\arcsin(\theta)/\pi, \qquad \theta = \sin(\pi\tau_\kappa/2).$$

Bivariate Clayton copula:

The CDF and PDF for Clayton copula are given by:

$$\begin{aligned} C(u,v;\theta) &= \left(u^{-\theta}+v^{-\theta}-1\right)^{-1/\theta}, \quad 0 \le u,v \le 1, \quad 0 \le \theta < \infty, \\ c(u,v;\theta) &= (1+\theta)[uv]^{-\theta-1}\left(u^{-\theta}+v^{-\theta}-1\right)^{-2-1/\theta}. \end{aligned}$$

The relationship between Kendall's tau and copula parameter is given by:

$$\tau_\kappa = \frac{\theta}{\theta+2}, \qquad \theta = \frac{2\tau_\kappa}{1-\tau_\kappa}.$$

Bivariate Gumbel copula:

Let $x = -\log u$ and $y = -\log v$. The CDF and PDF are given by:

$$\begin{aligned} C(u,v;\theta) &= \exp\left\{-\left(x^\theta+y^\theta\right)^{1/\theta}\right\}, \quad 0 \le u,v \le 1, \quad 1 \le \theta < \infty, \\ c(u,v;\theta) &= (uv)^{-1}\cdot\exp\left\{-[x^\theta+y^\theta]^{1/\theta}\right\}\left[(x^\theta+y^\theta)^{1/\theta}+\theta-1\right]\left[x^\theta+y^\theta\right]^{1/\theta-2}(xy)^{\theta-1}. \end{aligned}$$

The relationship between Kendall's tau and copula parameter is given by:

$$\tau_\kappa = \frac{\theta-1}{\theta}, \qquad \theta = \frac{1}{1-\tau_\kappa}.$$

Bivariate Frank copula:

The CDF and PDF are given by:

$$C(u,v;\theta) = -\theta^{-1}\log\left(\frac{1-e^{-\theta}-(1-e^{-\theta u})(1-e^{-\theta v})}{1-e^{-\theta}}\right), \quad 0 \le u,v \le 1, \quad -\infty < \theta < \infty,$$
$$c(u,v;\theta) = \frac{\theta(1-e^{-\theta})e^{-\theta(u+v)}}{[1-e^{-\theta}-(1-e^{-\theta u})(1-e^{-\theta v})]^2}.$$

The relationship between Kendall's tau and copula parameter is given by:

$$\tau_\kappa = 1 + 4\theta^{-1}\left[\theta^{-1}\int_0^\theta \frac{t}{(e^t-1)}dt - 1\right].$$

Bivariate Joe copula:

Let $x = 1-u$ and $y = 1-v$. The CDF is given by:

$$C(u,v;\theta) = 1 - \left(x^\theta + y^\theta - x^\theta y^\theta\right)^\theta, \quad 0 \le u,v \le 1, \quad 1 \le \theta < \infty,$$
$$c(u,v;\theta) = (x^\theta + y^\theta - x^\theta y^\theta)^{-2+1/\theta} x^{\theta-1} y^{\theta-1}[\theta - 1 + x^\theta + y^\theta - x^\theta y^\theta].$$

The relationship between Kendall's tau and copula parameter is given by:

$$\tau_\kappa = 1 + 2(2-\theta)^{-1}[\text{digamma}(2) - \text{digamma}(2/\theta + 1)]$$

There is no closed form expression between θ and τ_κ (numerical inversion).

Rotation of Bivariate Archimedian copulas

The CDF, PDF of the 90-degree, 180-degree, 270 degree rotation of Clayton, Gumbel, and Joe copula model can be derived from its CDF and PDF,

$$C_{90}(u,v;\theta) = v - C(1-u,v;-\theta),$$
$$C_{180}(u,v;\theta) = u + v - 1 + C(1-u,1-v;\theta),$$
$$C_{270}(u,v;\theta) = u - C(u,1-v;-\theta),$$

$$c_{90}(u, v; \theta) = c(1 - u, v; -\theta),$$

$$c_{180}(u, v; \theta) = c(1 - u, 1 - v; \theta),$$

$$c_{270}(u, v; \theta) = c(u, 1 - v; -\theta).$$

B DTSMC algorithm

1. Sample $\Theta_0^j \sim p(\Theta)$, $e_0^j \sim p(e)$ and set $W_0^j = 1/M$ for $j = 1...M$ with $M = 1000$ particles
2. For $i = 1, ..., S$, with $S = 10000$ level temperatures,

Step 1: Reweighting: Calculate the unnormalized weights and the normalized weights

$$w_i^j = W_{i-1}^j \frac{\widehat{p}(u_{1:T}|\Theta_{i-1}^j, e_{i-1}^j)^{\gamma_i} p(\Theta_{i-1}^j)}{\widehat{p}(u_{1:T}|\Theta_{i-1}^j, e_{i-1}^j)^{\gamma_{i-1}} p(\Theta_{i-1}^j)} = W_{i-1}^j \widehat{p}(u_{1:T}|\Theta_{i-1}^j, e_{i-1}^j)^{\gamma_i - \gamma_{i-1}}, \ \ j = 1, ..., M,$$

$$W_i^j = \frac{w_i^j}{\sum_{s=1}^M w_i^s}, \ \ j = 1, ..., M.$$

Step 2: Calculate the effective sample size (ESS): ESS $= \frac{1}{\sum_{j=1}^M \left(W_i^j\right)^2}$.

if ESS $< cM$ for $c = 0.8$

(i) Resampling: Resampling from $\{\Theta_{i-1}^j, e_{i-1}^j\}_{j=1}^M$ using the weights $\{W_i^j\}_{j=1}^M$, and then set $W_i^j = 1/M$ for $j = 1...M$, to obtain the new equally-weighted particles $\{\Theta_i^j, e_i^j, W_i^j\}_{j=1}^M$.

(ii) Markov move: Parallel for each $j = 1, ..., M$, move the samples Θ_i^j, e_i^j for $Q = 10$ CPM steps (Deligiannidis et al., 2018):

(a) Sample Θ_i^{j*} from the random walk proposal density $q(\Theta_i^{j*}|\Theta_i^j)$.

(b) Sample $\epsilon^j \sim \mathbf{N}(\mathbf{0}, \mathbf{I})$ and set $e_i^{j*} = \rho e_i^j + \sqrt{1-\rho^2}\epsilon^j$ with $\rho = 0.999$ is a correlation factor.

(c) Compute the estimated likelihood $\widehat{p}(u_{1:T}|\Theta_i^{j*}, e_i^{j*})$ using a bootstrap particle filter

(d) Set $\Theta_i^j = \Theta_i^{j*}$ and $e_i^j = e_i^{j*}$ with the probability

$$\min\left(1, \frac{\widehat{p}(u_{1:T}|\Theta_i^{j*}, e_i^{j*})^{\gamma_i} p(\Theta_i^{j*})}{\widehat{p}(u_{1:T}|\Theta_i^j, e_i^j)^{\gamma_i} p(\Theta_i^j)} \frac{q(\Theta_i^j|\Theta_i^{j*})}{q(\Theta_i^{j*}|\theta_i^j)}\right),$$

otherwise keep Θ_i^j, e_i^j unchanged.

3. The log of marginal likelihood estimate is

$$\log\widehat{p}(u_{1:T}) = \sum_{i=1}^K \log\left(\sum_{j=1}^M w_i^j\right).$$

Algorithm 1: The DTSMC algorithm

C Simulation study

In this section, we compare the proposed DSM copula models with the Exponentially Weighted Moving Average (EWMA), the Dynamic Conditional Correlation (DCC) model (Engle, 2002a; Tse and Tsui, 2002), and the DSC when the true correlation structure is known, in different stress scenarios based on the proposal of Engle (2002a) and (Hafner and Manner, 2012). We simulate $T = 2000$ observations from a bivariate Gaussian copula with time-varying correlation parameter ρ_t. We consider five models for the behavior of ρ_t such that,

1. Constant: $\rho_t = 0.8$.
2. Sine: $\rho_t = 0.5cos(2\pi t/250)$.
3. Fast Sine: $\rho_t = 0.5cos(2\pi t/25)$.
4. Step: $\rho_t = 0.5 - I(t > 1000)$.
5. Ramp: $\rho_t = ((t \mod 200) - 100)/101$.

Engle (2002a) considers that these stress tests mimic different realistic contexts that the correlation can be constant, gradual changes, rapid changes, and abrupt changes. We generate 100 datasets for each stress test and obtain the estimate of the correlation process $\hat{\rho}_t$. We calculate the 22-day realized correlation (RCor) as a low-frequency explanatory variable for the long-run change in the correlation. We measure the accuracy of each model based on the mean absolute error (MAE) and the mean-squared error (MSE),

$$MAE = \frac{1}{T}\sum_{t=1}^{T} |\hat{\rho}_t - \rho_t|,$$
$$MSE = \frac{1}{T}\sum_{t=1}^{T} (\hat{\rho}_t - \rho_t)^2.$$

Table 14 compares the relative MAE and MSE of the estimation of the correlation using the EWMA, the DSC, the DSM copula over the benchmark DCC model. With the exception of constant correlation scenarios, the DSM copula has a smallest MAE and MSE due to the ability of capturing the long-run dependence. We also observe that when the correlation changes quickly which makes it hard to extract the long run signal, the DSM copula model is still on par with the DSC model. The choice of lag number of the explanatory variable is also robust to the estimation results.

Table 14: MAE and MSE results: a simulation study

	Constant	Sine	Fast sine	Step	Ramp
			(a) MAE		
EWMA	5.316	1.184	1.191	1.000	1.364
DCC	1.000	1.000	1.000	1.000	1.000
DSC	0.796	0.970	0.924	0.980	0.951
DSM	0.902	0.829	0.924	0.875	0.925
			(b) MSE		
EWMA	27.679	1.345	1.422	0.864	1.915
DCC	1.000	1.000	1.000	1.000	1.000
DSC	0.618	0.942	0.902	1.006	0.944
DSM	0.798	0.691	0.906	0.855	0.880

The table shows the relative MAE and MSE of the estimation of the correlation using the EWMA, the DSC, the DSM copula over the benchmark DCC model. We use the restricted beta weighting function and $K = 12$ lags of monthly RCor as a low-frequency explanatory variable for the long-run change in the correlation. We generate 100 pseudo datasets for each stress test and calculate the average of MAE and MSE. The entries less than 1 indicate that the given model is better.

D Volatility forecast

The k-step-ahead forecast of the volatility using the GJR-GARCH model at time t is $\sigma^2_{i,t+k|t}$:

$$\begin{aligned}
\sigma^2_{i,t+1|t} &= E_t(\sigma^2_{i,t+1}) = \omega + (\alpha + \gamma \mathbf{1}\{\varepsilon_{i,t} < 0\})\varepsilon^2_{i,t} + \beta\sigma^2_{i,t}, \\
\sigma^2_{i,t+2|t} &= E_t(\sigma^2_{i,t+2}) = \omega + (\alpha + 0.5\gamma + \beta)\sigma^2_{i,t+1|t} \\
&= \frac{\omega}{1-\delta} + \delta(\sigma^2_{i,t+1|t} - \frac{\omega}{1-\delta}) \text{ where } \delta = \alpha + 0.5\gamma + \beta, \\
\sigma^2_{i,t+k|t} &= E_t(\sigma^2_{i,t+k}) = \frac{\omega}{1-\delta} + \delta^{k-1}(\sigma^2_{i,t+1|t} - \frac{\omega}{1-\delta}).
\end{aligned}$$

The k-step-ahead forecast of the cumulative volatility forecast using the GJR-GARCH model at time t is $\sigma^2_{i,t+1:t+k|t}$:

$$\sigma^2_{i,t+1|t} = E_t(\sigma^2_{i,t+1}) = \omega + (\alpha + \gamma \mathbf{1}\{\varepsilon_{i,t} < 0\})\varepsilon^2_{i,t} + \beta\sigma^2_{i,t},$$

$$\sigma^2_{i,t+1:t+2|t} = E_t(\sum_{j=1}^{2} \sigma^2_{i,t+j}) = \frac{2\omega}{1-\delta} + \frac{1-\delta^2}{1-\delta}(\sigma^2_{i,t+1|t} - \frac{\omega}{1-\delta}),$$

$$\sigma^2_{i,t+1:t+k|t} = E_t(\sum_{j=1}^{k} \sigma^2_{i,t+j}) = \frac{k\omega}{1-\delta} + \frac{1-\delta^k}{1-\delta}(\sigma^2_{i,t+1|t} - \frac{\omega}{1-\delta}).$$