The Journal of Economic Asymmetries 22 (2020) e00173

Contents lists available at ScienceDirect

The Journal of Economic Asymmetries

journal homepage: www.elsevier.com/locate/jeca

Bayesian sequential stock return prediction through copulas

Audronė Virbickaitė [a,*], Christoph Frey [b], Demian N. Macedo [a]

[a] Universitat de les Illes Balears, Cra. de Valldemossa, km 7.5, 07122, Palma de Mallorca, Spain
[b] Erasmus University Rotterdam, Burgemeester Oudlaan, 50, 3062, Rotterdam, Netherlands

ARTICLE INFO

JEL classification:
C58
C11
C53

Keywords:
Bayes factor
Sequential Monte Carlo
Particle filters

ABSTRACT

In this paper we perform density prediction for the equity returns in a non-linear manner by employing a copula-based approach. The use of asymmetric copulas allows to model asymmetric predictive densities and non-linear dependencies between equity returns and some predictor variable. In our proposed approach, the copula parameter and the marginals are estimated simultaneously by using Sequential Monte Carlo techniques. We apply proposed models to daily log returns of 20 assets traded at the NYSE. Among other findings, we show that in terms of predictive log Bayes Factors the asymmetric copula is preferred by more assets than the symmetric copula, advocating the use of non-linear models. Also, dividend yield is a better predictor variable than the lagged returns overall, but this result is reversed if we consider a volatile period only. These results have major implications for the investors when making portfolio decisions or measuring tail risk.

1. Introduction

This paper re-examines equity return predictability in a novel non-linear context. Even though the efficient market hypothesis states that equity returns are unpredictable, multiple empirical studies have demonstrated that some returns are, in fact, predictable (Lettau & Ludvigson, 2001). Nonetheless, the task is not trivial and requires careful consideration. Producing accurate predictive distributions of the returns has major implications for investors when making portfolio decisions or measuring tail risk, such as Value-at-Risk, for example. The vast majority of the literature consider regression-type models for the equity returns, imposing a linear dependence structure between the returns and the explanatory variable(s). However, there is plenty of empirical evidence that equity return response to some predictor variable does not necessary have to be linear, see Nam (2003), Kahra et al. (2018), for example. One way to incorporate such non-linearities is via structural breaks, regime-switches, etc., but then the number of model parameters increases dramatically and the model quickly loses its parsimonious representation.

Apart from the non-linear dependence structure, accounting for estimation uncertainty and volatility timing are essential for improving return forecasts, as demonstrated by Johannes et al. (2014). In their widely cited paper the authors compare a number of alternative models for equity returns and show that significant gains in return prediction are obtained only when the investor is Bayesian and time-varying volatility is included in the model. Therefore, in this paper we extend the work of Johannes et al. (2014) and propose to model the excess equity returns using copula functions, offering a parsimonious model that allows for non-linear dependencies between the returns and some predictor variable.

In particular, in this work we extend the models of Johannes et al. (2014) for return prediction in three directions. First, apart from the stochastic volatility (SV) models (Taylor, 1982), we also study realized volatility (RV) measures and their role in return prediction.

* Corresponding author.
E-mail address: audrone.virbickaite@uib.es (A. Virbickaitė).

https://doi.org/10.1016/j.jeca.2020.e00173
Received 10 January 2020; Received in revised form 29 May 2020; Accepted 1 June 2020

Secondly, next to the dividend yield as a predictor variable we also use lagged returns. Finally, we consider not only regression-type linear dependence structures, but we also propose a copula-based model that allows for asymmetric dependencies. We allow these dependence structures to be static, dynamic and hierarchical. By using historical data of 20 assets we show that an ensemble of these extra features improves equity return distribution forecasts, that consequently can be used by investor in constructing portfolios or calculating tail risk.

As noted in Johannes et al. (2014), volatility timing is crucial for return prediction. In their paper the authors rely on stochastic volatility models for modelling the volatility of the returns since these models provide more flexibility than other standard volatility specifications, see Kim et al. (1998), Yu (2002) and Broto and Ruiz (2004). Alternatively, Andersen and Bollerslev (1998) introduced the realized volatility (RV) measure that is obtained using the high frequency data. Ever since high frequency trading data became available to practitioners and researchers alike, there was a shift in paradigm in volatility modelling (Andersen et al., 2003; Barndorff-Nielsen & Shephard, 2004). Realized volatility is an actually observable volatility estimator and does not rely on any model assumptions (unlike SV), therefore, we also consider RV as an alternative specification for the return volatility.

Secondly, another important feature to consider in return prediction is the use of predictor variables. As noted in Lettau and Ludvigson (2001) " (…) it is now widely accepted that excess returns are predictable by variables such as dividend-price ratios, earnings-price ratios, dividend-earnings ratios, and an assortment of other financial indicators." Multiple empirical studies have demonstrated that the most powerful predictors are lagged returns, the dividend yield, the earnings-price ratio, to name a few (Bossaerts & Hillion, 1999; Fama, 1991; Xia, 2001). Stambaugh (1999), Johannes et al. (2014), for example, consider the dividend yield as a predictor variable. Cremers (2002) finds gains in prediction by using past returns and divided yield, among other variables. Therefore in this paper we employ lagged returns and lagged dividend yield as regressors for return prediction.

Finally, copulas have been applied in many fields in both social and natural sciences, especially in the context of financial time series, see Patton (2009), chap. 36 for an extensive review. Even though the majority of the copula-related literature focuses on modelling contemporaneous dependence between multiple time series (e.g.ben Brayek et al., 2015; Trabelsi, 2017, among many others), copulas also permit to model the temporal dependence of a univariate time series (Chen & Fan, 2006). The use of copulas in modelling temporal dependence of univariate time series relates to Markov processes and has been described in Darsow et al. (1992), Joe (2015), for example. By considering various possible marginal distributions with different copula specifications one can capture often observed features of univariate financial time series, such as skewness and fat tails. Moreover, depending on the copula family, it is possible to model non-linear temporal dependencies, as opposed to the standard linear regression-type models.

Important to note that in this work we do not pursue multivariate time series analysis, since it is out of the scope of the paper. However, the proposed framework could be extended to multivariate case by assuming some structure for joint modelling of the univariate processes, discussed in this paper. Neither we consider portfolio allocation exercise, especially given that Johannes et al. (2014) already demonstrated how a superior univariate density prediction for each asset separately translates into superior out-of-sample portfolio performance. Finally, same as in Johannes et al. (2014), our investor is fully Bayesian and model estimation is carried out in a simultaneous manner via Sequential Monte Carlo techniques, allowing for fast inference and consistent model comparison via Bayes Factors.

The paper is organized as follows. Section 2 introduces the proposed copula-based model and describes the estimation procedure. Section 3 summarizes the set-up of our empirical study and presents empirical findings. Section 4 concludes and gives an outlook on further generalizations.

2. Methodology

2.1. Marginals for the returns

Define r_t as the demeaned log-returns (in %) of some financial asset:

$$r_t = 100 \times (\log(P_t / P_{t-1}) - \mathrm{E}[\log(P_t / P_{t-1})]), \tag{1}$$

where P_{t-1} and P_t are the prices at the beginning and at the end of the period, respectively. Also, define $RV_t = \sum_{j=1}^{N} \tilde{r}_{j,t}^2$ as a realized *ex post* volatility measure, where $\tilde{r}_{j,t}$ is a m-minute intraday log-return for day t and N is the number of m-minute intervals in a trading day. It holds that $r_t = \sum_{j=1}^{N} \tilde{r}_{j,t}$. For introduction and review of realized volatility refer to Barndorff-Nielsen and Shephard (2002), Andersen et al. (2003), Barndorff-Nielsen and Shephard (2004), McAleer and Medeiros (2008), among others.

Then, the demeaned returns are standardized by the realized volatility measure. As seen in Andersen, Bollerslev, Diebold, and Labys (2000, 2001), it is safe to assume that the resulting standardized returns z_t are approximately Normally distributed:

$$z_t = r_t / \sqrt{RV_t}, \quad \text{such that} \quad z_t \overset{N}{\sim} \mathcal{N}(0,1). \tag{2}$$

As an illustration, Fig. 1 draws the histogram with the standard Normal density (solid line) and the QQ-plot for the daily standardized IBM returns. As seen from the plots, the returns are indeed approximately Normally distributed. We have also carried out the Kolmogorov-Smirnov and Jarque-Bera tests for Normality and Lagrange Multiplier test for ARCH effects for the RV-standardized return data considered in this paper, and the results show that the Normality assumption is, indeed, appropriate (see Table 2). Consequently, $\Phi(z_t) \equiv u_t^r \overset{N}{\sim} \mathcal{U}(0,1)$, where $\Phi(\cdot)$ is the CDF of the Normal distribution. In other words, the probability integral transforms of the returns,

u_t^r, are uniformly distributed. Since we are interested not only in the model fit, but prediction as well, we need to specify a dynamic model for the RV_t. $\log(RV_t)$ can be modelled in many manners, such as standard AR (1) or a more sophisticated Heterogeneous Autoregressive model (HAR) of Corsi (2009) or HARQ of Bollerslev et al. (2016) specifications. We have tried fitting AR (1) and HAR models for several data sets and evaluating their predictive performance at $t+n$ via Bayes Factors. HAR model performs better for $t+n$ horizons with $n>1$. For $t+1$ we did not find any substantial improvement as compared with a simple AR (1) specification, since we re-estimated the model after each new data point. Therefore, the model for the marginals of the returns is the following:

$$r_t = \varepsilon_t^r \sqrt{\sigma_t^2}, \tag{3}$$

$$\log(\sigma_t^2) = \mu^{(l)} + \varphi^{(l)} \log(\sigma_{t-1}^2) + \tau^{(l)} \varepsilon_t^{(l)}, \tag{4}$$

where $\sigma_t^2 \equiv RV_t$, $l = RV$, ε_t^r and ε_t^{RV} are independent $\overset{N}{\sim} \mathcal{N}(0,1)$. For the sake of comparison with benchmark models, we also consider a stochastic volatility process for the variance of the returns, first introduced by Taylor (1982). In this case, σ_t^2 is replaced with SV_t and $l = SV$ in (4). SV-based models for the distribution of the returns were considered in Johannes et al. (2014). Call $\Theta_V = (\mu^{(l)}, \varphi^{(l)}, \tau^{(l)})$ a set of volatility-related parameters, where $l = \{SV, RV\}$. If the marginal distributions are specified correctly, the probability integral transform should provide uniformly distributed variables, see Diebold et al. (1998). In this paper, for the sake of simplicity, we consider only AR (1) processes for the dynamics of the volatilities. However, more flexible specifications, such as asymmetric volatility models with leverage effect, could potentially improve the results (Aye et al., 2018; Fousekis, 2020; García-Centeno et al., 2010; Harvey & Shephard, 1996; Omori et al., 2007; Yu, 2005).

2.2. Copula

The construction of flexible multivariate distributions using copulas has started with the seminal work of Sklar (1959, pp. 229–231). It allows to combine a copula function with marginal distributions, which not necessary have to be the same and can be specified separately. Since then, copulas have been widely used in modelling temporal dependence between financial time series, because they can capture non-linear dependence, as opposed to the correlation coefficient. For a formal introduction and details on copulas the reader is referred to the books of Nelsen (2006) and Joe (2015), among others.

Nelsen (2006) defines copulas in the following manner. Consider a collection of random variables $Y_1, \ldots, Y_d$ with corresponding distribution functions $F_i(y_i) = P[Y_i \le y_i]$ for $i = 1\ldots, d$ and a joint distribution function $H(y_1, \ldots, y_d) = P[Y_1 \le y_1, \ldots, Y_d \le y_d]$. Then, according to a theorem by Sklar (1959, pp. 229–231), there exists a copula C with parameter θ such that

$$H(y_1, \ldots, y_d) = C(F_1(y_1), \ldots, F_d(y_d); \theta). \tag{5}$$

In other words, it is possible to model univariate marginals and the dependence structure separately. Copulas are defined in the unit hypercube $[0,1]^d$, where d is the dimension of the data, and all univariate marginals are uniformly distributed $u_1, \ldots, u_d \overset{N}{\sim} \mathcal{U}(0,1)$, where $F_i(y_i) = u_i \forall i = 1, \ldots, d$. Copulas are very flexible in the sense that (i) the marginal distributions $F(\cdot)$ can be modelled independently from the dependence structure $C(\cdot)$ and (ii) copulas are able to capture asymmetric dependencies, as opposed to the standard multivariate

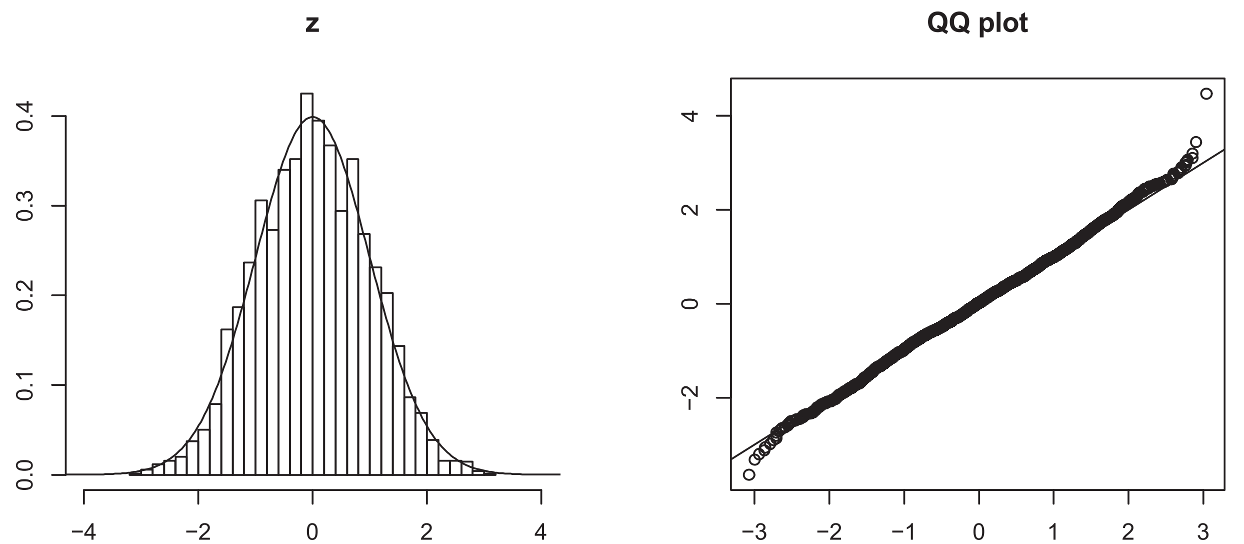

Fig. 1. Demeaned daily returns of IBM for the period of January 03, 2001–December 31, 2014 (3486 observations), standardized by the RV measure. On the left: historgam with the standard Normal density (solid line); on the right: QQ-plot against the standard Normal quantiles.

Table 1
Gaussian and Clayton copulas: CDF, PDF and Kendall's τs.

Gaussian	
CDF	$C(u,v;\theta) = \Phi_2(\Phi^{-1}(u), \Phi^{-1}(v);\theta),\ \theta \in [0,1]$
PDF	$c(u,v;\theta) = (1-\theta^2)^{-1/2}\exp\left\{\frac{2\theta xy - \theta^2(x^2+y^2)}{2(1-\theta^2)}\right\},\ x = \Phi^{-1}(u), y = \Phi^{-1}(v)$
$\theta \leftrightarrow \tau_\kappa$	$\tau_\kappa = 2\arcsin(\theta)/\pi,\ \theta = \sin(\pi\tau_\kappa/2)$
Clayton	
CDF	$C(u,v;\theta) = (u^{-\theta}+v^{-\theta}-1)^{-1/\theta},\ \theta \in (0,\infty)$
PDF	$c(u,v;\theta) = (\theta+1)(uv)^{-(\theta+1)}(u^{-\theta}+v^{-\theta}-1)^{-(2\theta+1)/\theta}$
$\theta \leftrightarrow \tau_\kappa$	$\tau_\kappa = \theta/(\theta+2),\ \theta = 2\tau_\kappa/(1-\tau_\kappa)$

distributions, such as Gaussian or Student's *t*. There is a vast selection of flexible bivariate one-parameter copulas, see Joe (2015) and Nelsen (2006) for example.

In this paper we consider some of the most popular one-parameter copulas (only bivariate cases), such as Gaussian and Clayton (referred as Mardia-Takahasi-Clayton-Cook-Johnson copula in Joe, 2015). An important notion associated with copulas is Kendall's τ_κ, a measure of dependence, which is given by $\tau_\kappa = 4\iint_{I^2} C(u,v)dC(u,v) - 1$ (Nelsen, 2006). See Table 1 for the CDFs, PDFs and Kendall's τs for Gaussian and Clayton copulas, and for detailed properties of these copulas refer to Joe (2015). The copula parameter θ for different copula families lies in different domains. Therefore, in order to be able to compare the dependence across different copulas, we rely on Kendall's τ_κ, where there is a one-to-one relationship between θ and τ_κ. But first we need to make sure it lies in the same domain for all copulas of interest. Note, that for Gaussian copula $\tau_\kappa \in [-1, 1]$, however, for standard Clayton copula $\tau_\kappa > 0$. Therefore, instead of considering standard Clayton copula (c_C), we couple this copula with its rotation and obtain the rotated Clayton copula (c_{RC}), that is defined as follows:

$$c_{RC}(u,v;\theta) = \left\{ \begin{array}{l} c_C(u,v;\theta) \text{ if } \theta \geq 0, \\ c_C(1-u,v;-\theta) \text{ if } \theta < 0. \end{array} \right\}$$

In order to be able to model the dynamics of τ_κ in an unconstrained manner, it is common to perform some deterministic transformations on this parameter and then model the behavior of the transformation *x*. In order to recover the copula parameter θ from the auxiliary latent variable *x*, we first convert *x* to Kendall's τ_κ thought the inverse of Fisher's z-transformation $\tau_\kappa = (\exp\{2x\}-1)/(\exp\{2x\}+1)$ and then the resulting τ_κ is converted to a copula parameter θ through one-to-one copula-specific function (see Table 1). Next, we consider two specifications for the copula parameter θ, or, to be more precise, for its transformation *x*: hierarchical and dynamic.

Hierarchical copula. For the hierarchical parameter case, θ (its auxiliary variable *x*) is fixed in time with a prior distribution $\mathscr{N}(m_x, V_x)$. In order to be completely uninformative we construct a hierarchical structure for *x* by putting a Normal-Inverse Gamma $\mathscr{NIG}$ hyperprior on m_x and V_x. In this manner θ (or *x*) is treated as a latent variable that needs to be filtered out during the estimation

Table 2
Descriptive statistics for the 20 assets for January 03, 2001–December 31, 2014 (3486 observations).

	mean	median	stdev.	skew.	kurt.	$\overline{\text{RV}}$	KS	JB	KS*	JB*	LM†	LM‡
AIG	−0.00	0.08	2.78	2.73	86.83	1.699	0.000	0.000	0.308	0.005	0.711	0.853
BA	0.00	−0.01	1.28	0.07	6.39	1.176	0.001	0.000	0.306	0.000	0.446	0.468
BAC	−0.00	0.02	1.95	0.79	25.46	1.393	0.000	0.000	0.106	0.000	0.818	0.062
C	−0.00	0.04	2.11	−0.53	28.63	1.558	0.000	0.000	0.572	0.010	0.147	0.147
F	0.00	0.01	1.90	−0.01	12.54	1.673	0.000	0.000	0.154	0.000	0.609	0.639
GE	−0.00	−0.03	1.35	0.28	10.09	1.162	0.000	0.000	0.247	0.003	0.592	0.539
GS	0.00	0.02	1.69	0.74	22.44	1.352	0.000	0.000	0.208	0.004	0.033	0.014
HD	0.00	0.00	1.39	0.49	9.03	1.223	0.000	0.000	0.274	0.011	0.307	0.564
HPQ	0.00	−0.04	1.57	0.25	11.43	1.380	0.000	0.000	0.334	0.035	0.432	0.105
IBM	0.00	−0.01	1.06	0.33	8.26	0.941	0.000	0.000	0.427	0.003	0.059	0.143
JNJ	0.00	−0.02	0.79	0.45	8.66	0.765	0.000	0.000	0.709	0.001	0.319	0.376
JPM	0.00	−0.00	1.71	0.66	13.73	1.411	0.000	0.000	0.448	0.001	0.145	0.110
KO	0.00	0.00	0.87	0.24	7.01	0.834	0.000	0.000	0.961	0.294	0.494	0.334
MO	0.00	0.01	1.04	−0.74	16.92	0.914	0.000	0.000	0.959	0.428	0.299	0.360
NKE	−0.00	−0.01	1.26	0.44	8.41	1.114	0.005	0.000	0.449	0.006	0.583	0.735
PFE	0.00	−0.01	1.08	0.16	6.28	1.035	0.000	0.000	0.150	0.000	0.367	0.325
PG	0.00	−0.01	0.79	−0.15	6.49	0.786	0.000	0.000	0.634	0.118	0.635	0.689
VZ	0.00	0.01	1.11	0.19	7.07	1.030	0.000	0.000	0.301	0.000	0.888	0.885
WMT	0.00	−0.01	0.97	0.57	10.84	0.926	0.000	0.000	0.783	0.067	0.351	0.438
XOM	−0.00	0.01	1.03	−0.24	10.24	0.993	0.000	0.000	0.551	0.001	0.395	0.541

Note: Here $\overline{\text{RV}}$ stands for the sample mean of the realized volatility; KS and JB report the *p*-values for the Kolmogorov-Smirnov and Jarque-Bera tests for Normality; KS* and JB* report the *p*-values for the tests for Normality for the RV-standardized returns; LM† reports the *p*-value for the Lagrange Multiplier test for ARCH effects in the RV-standardized returns with 30 lags; and LM‡ - with 40 lags.

procedure.

Dynamic copula. In dynamic case, the copula parameter θ_t (or its auxiliary process x_t) follows a random walk process: $x_t = x_{t-1} + V_x\eta_t, \quad \eta_t \overset{N}{\sim} \mathcal{N}(0,1)$. The unknown parameter V_x has an Inverse Gamma $\mathcal{IG}$ prior. In general, call Θ_C the set of parameters, associated with the evolution of x_t. Θ_C in hierarchical model contains parameters m_x and V_x, meanwhile in the dynamic case contains only the variance parameter V_x. Then, $\Omega = (\Theta_V, \Theta_C)$ is a complete set of model parameters.

Next, we describe the use of copulas in each specific case: when the predictor variable is the lagged return and when the predictor variable is the dividend yield.

2.3. Copulas for lagged returns as a predictor variable

As described in Chen and Fan (2006), let $\{R_t\}$ be a stationary first order Markov process whereas its probabilistic behavior is completely defined by joint distribution function $H(\cdot)$ between R_{t-1} and R_t. On the other hand, as seen in (5), using Sklar's theorem this joint can be expressed using a copula representation $H(r_t, r_{t-1}) = C(F(r_t), F(r_{t-1}); \theta)$, where $F(\cdot)$ is a marginal CDF of R_t and θ is a copula parameter. This allows to model a stationary Markov process using copula, where the transition kernel, determined by θ, is constant over time.

Let $h(\cdot)$ be the joint density of R_t and R_{t-1}, and $f(\cdot)$ the corresponding marginal PDF of R_t. Using copula representation in (5), $h(\cdot)$ can be expressed as a product of the marginals and a copula density, which defines the dependence structure:

$$h(r_t, r_{t-1}) = c(F(r_t), F(r_{t-1}); \theta) \cdot f(r_t) \cdot f(r_{t-1}). \tag{6}$$

Then the conditional distribution of r_t given r_{t-1} is

$$f(r_t | r_{t-1}) = \frac{h(r_t, r_{t-1})}{f(r_{t-1})} = c(F(r_t), F(r_{t-1}); \theta) \cdot f(r_t). \tag{7}$$

Parameter θ completely determines the dependence structure which is constant across time. Then the collection of $\{R_t\}$ follows a stationary first order Markov process with constant transition kernels. A natural extension is to relax the assumption of time-invariant dependence and consider dynamic copula approach by allowing θ to be time-varying, i.e. θ_t. This implies that $\{R_t\}$ is inhomogeneous first order Markov process with time-varying transition kernel.

Since in this work we consider Markov process of order one, bivariate copulas with a single parameter are sufficiently flexible. On the other hand, if we wish to consider a higher-order dependence structure, one-parameter copulas might be too restrictive, because it would imply the same strength of dependence between all lags of returns. Gaussian copula is symmetric and does not present any tail dependence, meanwhile Clayton copula can model lower tail dependence. Economic interpretation of strong lower tail dependence is the following: when the markets are in turmoil, the dependence between log returns and some predictor variable becomes stronger as compared to the calm periods.

2.4. Copulas for dividend yield as a predictor variable

Similarly to the case outlined above, for the predictor variable instead of lagged returns one can consider dividend yield, i.e. r_{t-1} is replaced by DY_{t-1} - the previous period's dividend yield. As seen in Lettau and Ludvigson (2001), Boudoukh et al. (2007), Johannes et al. (2014), dividend yield (or net pay-off) is a reasonable predictor variable for the log returns. In this case the CDF of dividend yield $F_{DY}(DY_{t-1})$ has to be estimated separately. In this work we consider a non-parametric estimator of the distribution function of the dividend yield. The observed dividend yield data is transformed to the unit interval via empirical cumulative distribution function (ECDF), adjusted by the $T/(T+1)$ factor, in order to avoid the unit at the end of the interval, as seen in Genest et al. (1995). Here T is the length of the sample size.

$$u_t^{DY} = T^{-1}\sum_{j=1}^{N} 1\{DY_t \le DY_j\}(T/(T+1)) = (T+1)^{-1}\sum_{j=1}^{N} 1\{DY_t \le DY_j\},$$

where $1\{\cdot\}$ is an indicator function. Then, $u_t^{DY} \overset{N}{\sim} \mathcal{U}(0,1)$ for $t = 1, \ldots, T$. Notice that there is no need to assume any dynamics for the dividend yield process, because at time t the dividend yield from $t-1$ is observed. We have decided to use the ECDF for the dividend yield in order not to impose any arbitrary distributional assumptions. As long as we have the uniformly distributed data (that is produced by applying the ECDF), the distribution of the dividend yield is irrelevant, see Equation (7). If one wished, one could choose any other parametric or non-parametric model for the dividend yield, given that the probability integral transforms of such model are *i.i.d.* Uniform.

Modelling the dependence between returns and the dividend yield via time varying copula parameter has economics justifications. The instability of parameters (non-constant regression coefficient) in return models that are based on dividend yield has been widely documented in the financial literature, see Goyal and Welch (2003), Paye and Timmermann (2006), Ang and Bekaert (2007), among others.

2.5. Conditional densities

The complete model for the univariate log returns r_t, written in a state-space representation, looks as follows:

$$\begin{aligned} & u_t = \Phi(r_t/\sigma_t), \\ & \left(u_t, u_{t-1}^{(j)}\right) \big| x_t \sim C\left(\left(u_t, u_{t-1}^{(j)}\right); \theta_t\right), \text{ where } \theta_t = f_\tau(x_t), \\ & x_t | \Theta_C, x_{t-1} \sim \mathscr{N}(x_t; m_x, V_x). \end{aligned} \tag{8}$$

Here σ_t^2 follows either RV or SV dynamics as in (4), x_t is either hierarchical with fixed m_x or follows a random walk process with $m_x = x_{t-1}$, $j \in \{DY, r\}$ and $f_\tau(\cdot)$ is a deterministic function that transforms the latent variable x to the copula parameter θ via Kendall's τ transformation. Note that this model is specified for the time series of univariate log returns, where the dependence between today's log return and some predictor variable, instead of relying on linear regression, is modelled via copula.

In many financial applications the estimation of parameters is not the ultimate goal. One is usually interested in estimating and forecasting conditional distributions and certain moments, such as the mean or variance for example. The predictive marginal density for one-step-ahead returns, given the all the information up till time t: $\mathscr{F}_{1:t} = (r_1, \ldots, r_t, DY_1, \ldots, DY_t, RV_1, \ldots, RV_t)$ is:

$$f(r_{t+1}|\mathscr{F}_{1:t}) = \iiint c\left(\Phi\left(\frac{r_{t+1}}{\sqrt{\sigma_{t+1}^2}}\right), u_t \middle| x_{t+1}, \sigma_{t+1}^2, \Omega\right) f\left(r_{t+1}|\sigma_{t+1}^2\right) f(x_{t+1}|\mathscr{F}_{1:t}, \Omega) f\left(\sigma_{t+1}^2|\mathscr{F}_{1:t}, \Omega\right) f\left(\Omega|\mathscr{F}_{1:t}\right) dx_{t+1} d\sigma_{t+1}^2 d\Omega$$

$$f\left(\sigma_{t+1}^2|\mathscr{F}_{1:t}, \Omega\right) f\left(\Omega|\mathscr{F}_{1:t}\right) dx_{t+1} d\sigma_{t+1}^2 d\Omega. \tag{9}$$

Moreover, the conditional k^{th} moment can be calculated as

$$\mathrm{E}\left[r_{t+1}^k \middle| \mathscr{F}_{1:t}\right] = \int r_{t+1}^k f(r_{t+1}|\mathscr{F}_{1:t}) dr_{t+1}.$$

2.6. Estimation

The usual Bayesian estimation approach relies on MCMC schemes. Ausín and Lopes (2010), for example, use a multivariate random walk Metropolis - Hastings in a one-step estimation procedure for the parameters of the marginals and the copula, where their time-varying copula parameter is observation driven. Meanwhile, Almeida and Czado (2012) employ a two-step estimation approach where the marginal series are estimated first. Then, conditioning on the estimated marginal parameters, use a similar method as Ausín and Lopes (2010) to model copula dynamics, and a coarse grid method for updating the unobserved states of the stochastic copula. Creal and Tsay (2015) also employ a MCMC estimation scheme for modelling large panels of financial assets using high dimensional dynamic stochastic copula models. In general, MCMC methods are inherently non-sequential and once a new data point is observed the algorithm has to be re-run all over again. Alternatively, Johannes et al. (2014) employ a sequential Monte Carlo (SMC) algorithm for their proposed return prediction models. Therefore, in the spirit of their paper, we also use an algorithm that performs sequential simultaneous estimation for the proposed model, where the marginals and the copula parameters are estimated simultaneously. In particular, we use a modified version of Particle Learning method (Carvalho et al., 2010a), that relies on the use of the sufficient statistics (Storvik, 2002) to allow for parameter learning. The use of sufficient statistics has been shown to increase the efficiency of the algorithm by reducing the variance of sampling weights, see Carvalho et al. (2010a). The proposed sequential scheme also filters the latent volatilities of the SV model (in RV model they are observed). However, if one wished to compare these filtered estimates with the ones produced by the MCMC-based Forward Filtering Backward Sampling (Carter & Kohn, 1994; Fruhwirth-Schnatter, 1994) algorithm, it is necessary to perform smoothing at the end of the estimation, which can be computationally demanding (Carvalho et al., 2010). For a detailed description of the algorithm see the Appendix at the end of the manuscript.

The priors for model parameters are chosen to be conditionally conjugate. For the set of copula parameters Θ_C:

$$V_x \sim \mathscr{I}\mathscr{G}(b_0/2, b_0 V_{x0}/2),$$

$$m_x | V_x \sim \mathscr{N}(m_m, V_m V_x).$$

And for the set of volatility parameters Θ_V:

$$\tau^{2(l)} \sim \mathscr{I}\mathscr{G}\left(b_0^{(l)}/2, b_0 \tau_0^{2(l)}/2\right),$$

$$\varphi^{(l)} \big| \tau^{2(l)} \sim \mathscr{T}\mathscr{N}_{(-1,1)}\left(m_\varphi^{(l)}, V_\varphi^{(l)} \tau^{2(l)}\right),$$

$$\mu^{(l)} \sim \mathscr{N}\left(m_\mu^{(l)}, V_\mu^{(l)}\right).$$

where $l = \{SV, RV\}$ and $\mathscr{T}\mathscr{N}_{(a,b)}$ stands for a truncated Normal distribution with truncation points at a and b. This restriction on the persistence parameter guarantees stationary process for the dependence parameter, however, it is not necessary (one could actually test if the process is stationary by removing the restriction). Initial states are $x_0 \sim \mathcal{N}(c_0^x, C_0^x)$ and $\log\sigma_0^2 \sim \mathcal{N}(c_0^h, C_0^h)$ (only for SV model). Here c_0^x, C_0^x, c_0^h, C_0^h, b_0, $b_0 V_{x0}$, m_m, V_m, $b_0^{(l)}$, $b_0 \tau_0^{2(l)}$, $m_\varphi^{(l)}$, $V_\varphi^{(l)}$, $m_\mu^{(l)}$, $V_\mu^{(l)}$ are the known hyper-parameters. The initial states at t_0 for all parameters and latent variables are simulated from their corresponding priors. Then, the SMC algorithm transports the set of N particles from time $t-1$ to time t, where at each step these particles are updated using the new information at time t. A set of particles is an approximate sample from the posterior distribution.

A well-known limitation of particle filters is called particle degeneracy: an ever-decreasing set of unique atoms in the particle approximation of the density of interest. It has been shown in numerous studies that increasing the number of observations will lead to degenerating paths, unless the number of particles is being increased simultaneously. Therefore, particle degeneracy has to be monitored carefully and this shortcoming can be seen as a trade-off between the sequential nature of the algorithm and stability of MCMC for very large samples. For example, in our analysis, we have plotted the posterior sequential paths for the median, 2.5% and 97.5% percentiles of the estimated parameters. These plots are not reported because of space constraints. Visual examination of these plots shows sufficient variation in particle approximation of the posterior densities, therefore, the selected number of particles is large enough. Another way to check for degeneracy is to monitor the number of distinct particles approximating the density of interest (Doucet & Johansen, 2009). For a general introduction to particle filters, comparison with MCMC, numerous empirical illustrations and discussion of the shortcomings of such estimation approach refer to Doucet and Johansen (2009), Carvalho et al. (2010a), Carvalho et al. (2010b), Chopin et al. (2011), Lopes and Tsay (2011), Virbickaite et al. (2019), among many others.

Finally, if the interest is not on-line type inference, MCMC is still a gold standard. Recently other approaches, such as Particle MCMC or SMC^2, have emerged, presenting an alternative to the proposed estimation scheme, see Andrieu et al. (2010), Pitt et al. (2012), Chopin et al. (2013), Fulop and Li (2013), among others.

2.7. Evaluation

The model comparison is carried out via sequential predictive log Bayes Factors (BF). As pointed out in Koop (2003), Bayes Factors permit consistent model comparison even for non-nested models. Also, it contains rewards for model fit, accounts for coherency between the prior and the information arising from the data, as well as rewards parsimony. Bayes Factor between two competing models $\mathcal{M}_k, k \in \{1,2\}$ is defined as (Kass & Raftery, 1995):

$$BF_{12} = \frac{p(D|\mathcal{M}_1)}{p(D|\mathcal{M}_2)},$$

where $p(D|\mathcal{M}_k)$ is the marginal likelihood for data D given a model $\mathcal{M}_k$. Then the log predictive Bayes Factor at time t is defined as

$$\log BF_{12,t} = \sum_{s=1}^{N} \log p\left(r_s|\mathcal{M}_1, \mathcal{F}_{1:s-1}) - \sum_{s=1}^{N} \log p\,(r_s|\mathcal{M}_2, \mathcal{F}_{1:s-1}\right),$$

and the marginal predictive $p(r_t|\mathcal{M}_k, \mathcal{F}_{1:t-1})$ for model k, given all the information available till time $t-1$, is obtained as in (2.5). The integral in (2.5) is rarely analytically tractable and can be approximated using costly MCMC-based procedures. On the other hand, sequential Monte Carlo approach produces marginal likelihoods as by-products of the estimation procedure, therefore, sequential model comparison via log predictive Bayes Factors can be carried out without any additional computational cost.

Alternatively, we also employ a log predictive score (LPS_k) measure for model k, which is defined as follows:

$$LPS_k = T^{-1} \sum_{t=1}^{N} \log p(r_t|\mathcal{M}_k, \mathcal{F}_{1:t-1}),$$

that is calculated once at the end of the sample period (time T). Notice, that the model is estimated using the data available up to time $t-1$, therefore, LPS is a predictive, not an in-sample score. It is easy to show that the difference between two competing log predictive scores coincides with the log predictive Bayes Factor at time T. One could also consider other comparison metrics, such as Value-at-Risk measure or mean squared error for example. However, (marginal) predictive likelihoods are more informative in the sense that they focus on the entire distribution of the returns, not only on the tails or certain moments.

3. Empirical analysis

3.1. Data and set-up

We investigate the forecast performance of the proposed copula model using daily log returns of 20 US stocks traded at the NYSE, that are coded as: "AIG" - American International Group, Inc.; "BA" - The Boeing Company; "BAC" - Bank of America Corporation; "C" - Citigroup Inc.; "F" - Ford Motor Company; "GE" - General Electric Company; "GS" - The Goldman Sachs Group, Inc.; "HD" - The Home Depot, Inc.; "HPQ" - HP Inc.; "IBM" - International Business Machines Corporation; "JNJ" - Johnson & Johnson; "JPM" - JPMorgan Chase & Co.; "KO" - The Coca-Cola Company; "MO" - Altria Group, Inc.; "NKE" - NIKE, Inc.; "PFE" - Pfizer Inc.; "PG" - The Procter & Gamble

Company; "VZ" - Verizon Communications Inc.; "WMT" - Wal-Mart Stores, Inc.; "XOM" - Exxon Mobil Corporation.

We consider daily observations from January 03, 2001–December 31, 2014 (3486 data points) and at each point in time perform one-step-ahead marginal density prediction. The realized volatility measures are based on 10-min intraday price intervals. We have also tried 2, 5 and 15-min intraday returns for the RV_t measure and from the signature plot we concluded that 10-min intervals are big enough not to be affected by the market micro-structure noise and small enough to obtain efficient estimates. The main descriptive statistics can be found in Table 2. As seen from the table, apart from the AIG asset, the rest of the assets do not display extreme skewness, and only 5 out of 20 are negatively skewed. All assets present excess kurtosis in varying degrees: from as low as 6.28 (PFE) to as high as 86.83 (AIG). The table also reports the Kolmogorov-Smirnov (KS) and Jarque-Bera (JB) tests for Normality for the returns and the RV-standardized returns. As expected, the unconditional distribution is not Normal for any of the returns, which is consistent with the findings in the related financial econometrics literature. However, once the returns are standardized by their corresponding RV measure, they are approximately Normally distributed. The *p*-values for the KS test do not reject Normality for all 20 assets, meanwhile JB test rejects Normality for 13 out of 20 assets (at 1% level), but only marginally. Finally, the last two columns report the *p*-values for the Lagrange Multiplier test for ARCH effects (Engle, 1982). The null hypothesis of this LM test is that a series exhibits no ARCH effects. If rejected, ARCH effects with a given number of lags are present. Number of lags is usually specified by the user; we chose 30 and 40 lags, as seen in Ausín and Lopes (2010). All *p*-values are larger than 1%, therefore, there are no ARCH effects remaining in the RV-standardized returns.

We also consider two sub-periods in order to get a better understanding how the models perform during calm (Jan 2002–Dec 2006) and nervous (Jan 2007–Dec 2012) periods. Separating data into several subperiods in order to understand how the dependence structure changes is a rather common practice (e.g.ben Brayek et al., 2015, among others). The competing models are all possible combinations of the following, resulting into 16 models:

1. Realized volatility (RV) or stochastic volatility (SV) based models.
2. Normal Copula (NC) or Clayton Copula (CC) models.
3. Lagged returns (r) or dividend yield (DY) as a predictor variable.
4. Hierarchical (h) and dynamic (d) dependence structures.

Note that Normal copula based model allows for linear (regression-type) dependence structure. Finally, for comparison purposes we include four static linear models oftenly used as benchmarks, similar to the ones seen in Johannes et al. (2014):

$$r_t = \theta r_{t-1} + \varepsilon_t^r \sigma_t^{(l)}, \tag{10}$$

$$r_t = \theta DY_{t-1} + \varepsilon_t^r \sigma_t^{(l)}. \tag{11}$$

The four models are all static (s) in θ, RV or SV based such that $l = \{SV, RV\}$, using either returns (10) or dividend yield (11) as a predictor variable. These static models are also estimated using the SMC algorithm. The dependence parameter θ between the returns and the predictor variable is assumed to have a Normal prior distribution with some known hyper-parameter values. Our proposed

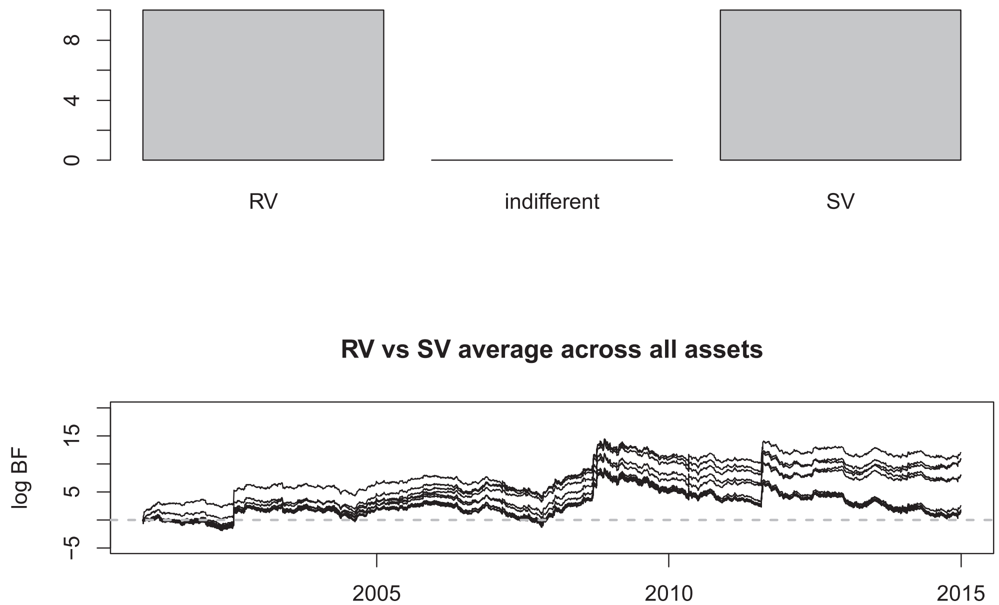

Fig. 2. RV vs. SV-based models. Top plot: number of assets (out of 20) that prefer each model; bottom plot: the average strength of preference (over 20 assets) measured by the log predictive Bayes Factors, SV as benchmark.

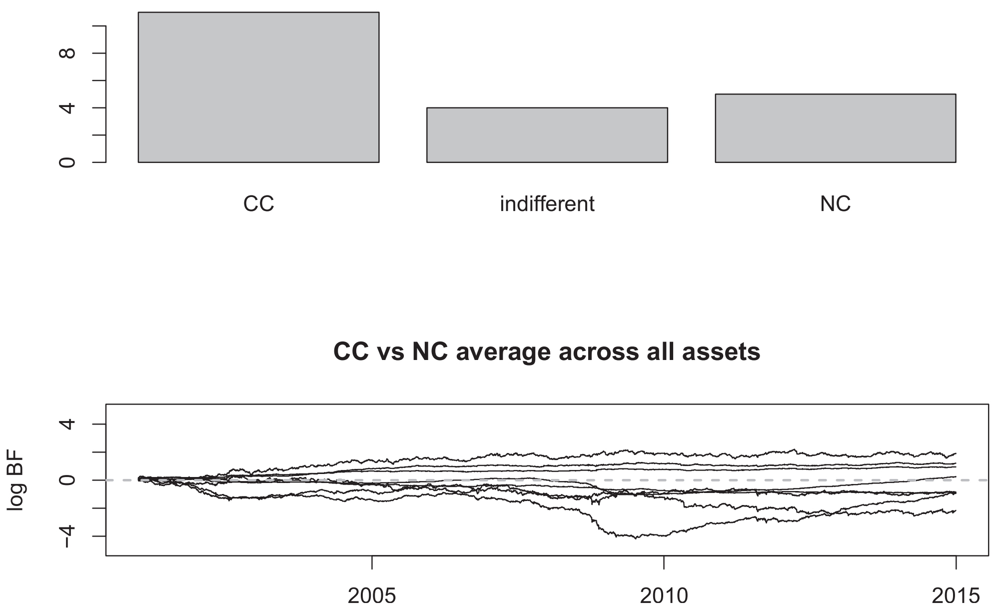

Fig. 3. Clayton copula (CC) vs. Normal copula (NC)-based models. Top plot: number of assets (out of 20) that prefer each model; bottom plot: the average strength of preference (over 20 assets) measured by log predictive Bayes Factors, NC as benchmark.

hierarchical structure is more flexible in the sense that the hyper-parameters have their own priors and are estimated, also, θ is a latent variable that is filtered out. For a complete list of models refer to Table 3 in the Appendix. The hyper-parameter values are as follows: $c_0^x = 0$, $C_0^x = 0.1$, $c_0^h = RV_1$, $C_0^h = 0.36$, $b_0 = 6$, $b_0 V_{x0} = 0.004$, $m_m = 0$, $V_m = 0.1$, $b_0^{(l)} = 3$, $b_0 \tau_0^{2(l)} = 0.36$, $m_\varphi = 0.85$, $V_\varphi = 1$, $m_\mu = 0$, $V_\mu = 0.01$. The values were chosen either to represent uninformative priors, or to match the unconditional sample moments (for example, the variance of RV), or by employing previous knowledge from numerous empirical studies (for example, it is known that persistence parameter in RV or SV models is close to 1). The number of particles is set to be very large, $N = 500$k, in order to ensure that there are no particle degeneracy problems. Model monitoring can be performed via sequential log predictive Bayes Factors.

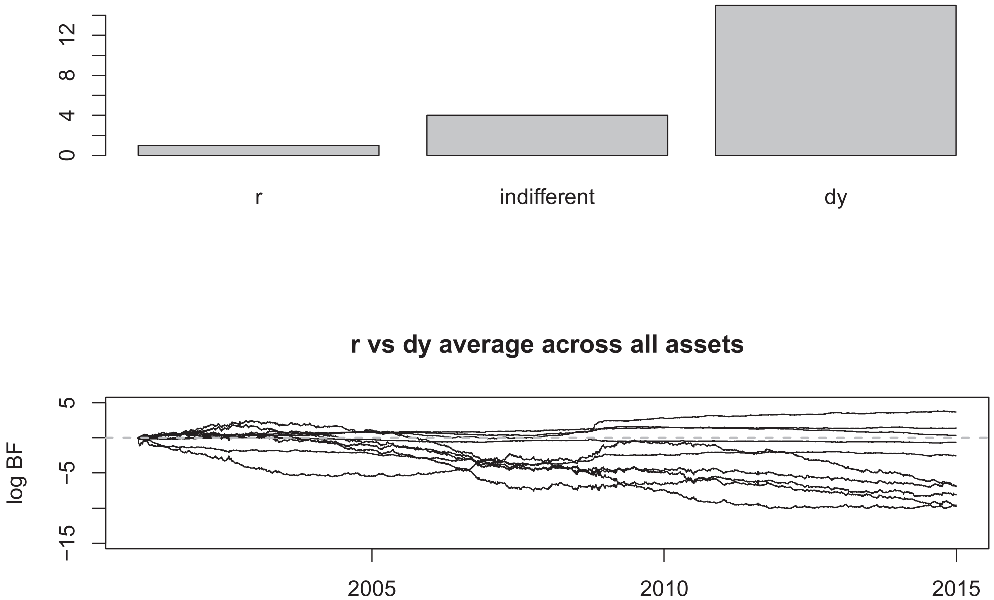

Fig. 4. Lagged returns (r) vs. dividend yield (DY)-based models. Top plot: number of assets (out of 20) that prefer each model; bottom plot: the average strength of preference (over 20 assets) measured by the log predictive Bayes Factors, DY as benchmark.

3.2. Results

Since there are 20 models and 20 assets, there are 20 × 20 estimation results (only for *LPS*) to be reported. Tables A1, A2 and A3 in the Appendix include the complete model list and report the *LPS* for all 20 models and 20 assets for the entire sample and two subsamples: calm (Jan 2002–Dec 2006) and volatile (Jan 2007–Dec 2012). The numbers in bold indicate the highest LPS for each asset. For now we will not discuss the magnitude of these differences (this is done in the later paragraphs), but rather see how many times one or another model is preferred.

Firstly, return based models and dividend yield based models appear as best models for equal number of assets (10÷ 10, such that $r \div DY$) during the entire period. An interesting finding though is that for the calm period, see Table A2 dividend yield based models are the best for 16 out of 20 assets (4÷ 16). However, for the crisis period, see Table A.3, this ratio almost reverses: now return based models are better for a bigger number of assets as compared to the dividend yield based models (13 ÷ 7). This has some important implications: even though for the entire sample dividend yield is a better predictor variable for the majority of the assets, this result might not hold depending if we are considering calm or crisis period. For calm period dividend yield seems to provide better predictive power, meanwhile during crisis periods lagged return is a better predictor variable for the majority of the assets.

As for the dependence structure, hierarchical model (h), as compared to dynamic (d) or static (s) models, is always preferred: in full sample and both subsamples. In the full sample and during the crisis period, hierarchical models are the best for 18 and 17 assets (out of 20), respectively. Only for the calm period hierarchical model is the best for 10 assets, meanwhile dynamic and static models share the rest, 3 and 7 assets respectively. This result implies that for calm period it is reasonable to consider less flexible dependence structure, however, as expected, for the crisis period a more flexible hierarchical dependence structure is preferred.

Next, we present four main figures, that can summarize the principal findings of the paper for the entire period taking into account the magnitude of the preferences via sequential predictive log Bayes Factors. In the following figures we report the number of assets that prefer one or another model type. Preference is measured in terms of differences between *LPS*, and if $LPS_{A,i} - LPS_{B,i} > 0$, it is said that model *A* is preferred to model *B* by the asset *i*. This classification loses information about the strength of the preference. Therefore, we also report the *average* sequential predictive log Bayes Factor. Averaging across assets is not intuitively appealing, however, it can convey important information on the average strength of the preference.

Fig. 2 compares how many assets in total prefer RV versus SV-based models. Out of 20 models in total, half of them (ten) are RV-based, and the other half are SV-based (for a complete list of models refer to Table A1 in the Appendix). For example, model (1) is compared to its 'counterpart' model (5), where both are the same in almost every aspect, except for volatility modelling approach, i.e. RV vs. SV. Then, for asset *i*, we perform ten comparisons and if RV-based model is preferred > 5 times, we say that asset *i* prefers RV-based models. If asset *i* prefers RV-based models half of the time, then we say that asset *i* is indifferent. As seen in the top part of Fig. 2, exactly for half of the assets RV-based models perform better, and the other half prefer the SV-based models. The bottom plot of the figure draws sequential predictive log Bayes Factor, averaged across all assets, for 10 RV vs. SV-based model combinations. Here we can see that even though half of the assets prefer one or another volatility model, the preference for RV-based models is much stronger on average. A log Bayes Factor > 5 indicates a strong preference for one model over another (Kass & Raftery, 1995). This is an important result in the sense that availability of high frequency data can improve predictive performance of the one step ahead log returns. Finally,

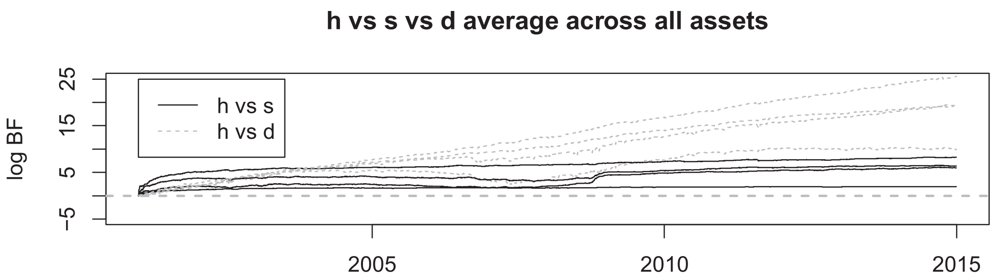

Fig. 5. Hierarchical (h) vs. static (s) vs. dynamic (d) models. Top plot: number of assets (out of 20) that prefer each model; bottom plot: the average strength of preference (over 20 assets) measured by the log predictive Bayes Factors, static and dynamic models as benchmark.

one can see an increase in sequential predictive log Bayes Factors around the year 2008, meaning that the predictive power of RV-based models increases during financial turmoils.

- Assets that prefer the RV-based models are: “AIG” “BA” “BAC” “C” “F” “GS” “HPQ” “JPM” “NKE” “XOM”.
- Assets that prefer the SV-based models are: “GE” “HD” “IBM” “JNJ” “KO” “MO” “PFE” “PG” “VZ” “WMT”.

Fig. 3 compares how many assets prefer Clayton copula (CC) based models vs. Normal copula (NC) based models. Out of 20 models in total, 8 are CC-based and 8 are NC-based. In the top plot we can observe that more assets prefer CC-based models. Also, there are 4 assets in the ‘indifferent’ category, meaning that half of the time these assets preferred CC, another half NC-based models. The bottom plot of the figure indicates that the preference on average is not conclusive, meaning that it is asset-specific. In other words, some assets exhibit symmetric and others - asymmetric dependence structures. The first draft of the manuscript also included Gumbel copula based models, however, it always performed the worst and due to space restrictions we are not reporting estimation results.

- Assets that prefer the CC-based models are: “AIG” “BA” “F” “GS” “HD” “JNJ” “KO” “PFE” “PG” “WMT” “XOM”.
- Assets that are indifferent to CC vs. NC-based models are: “BAC” “HPQ” “IBM” “NKE”.
- Assets that prefer the NC-based models are: “C” “GE” “JPM” “MO” “VZ”.

Next, Fig. 4 presents which is a more powerful predictor for the daily log returns: lagged log returns or lagged dividend yield. In total there are 10 lagged return-based and 10 dividend yield-based models. As seen from the top plot, for the majority of the assets dividend yield is a stronger predictor variable than the lagged returns. The strength of the preference is moderate, in the sense that the sequential predictive log Bayes Factors in the bottom plot seem to favour dividend yield based models.

- Assets that prefer the lagged returns-based models are: “BA”.
- Assets that are indifferent to lagged returns vs. dividend yield as predictor variable are: “AIG” “HPQ” “JNJ” “VZ”.
- Assets that prefer the dividend yield-based models are: “BAC” “C” “F” “GE” “GS” “HD” “IBM” “JPM” “KO” “MO” “NKE” “PFE” “PG” “WMT” “XOM”.

Fig. 5 compares only Normal copula based (i.e. linear dependence structure) models across different ways of modelling the dependence parameter: hierarchical (h) and dynamic (d). Also, we include the four linear static (s) models, that are defined in (10) and (11). Hierarchical dependence structure allows more flexibility, since the hyper-parameters (mean and variance) are estimated and the copula parameter θ, governing dependence structure is treated as a latent random variable. Finally, dynamic model follows a random walk structure, where apart from filtering the latent variable we also estimate the variance parameter. The top plot of the figure indicates that all assets prefer hierarchical rather than static dependence structures. We have not included the dynamic models in the top figure, because they performed even worse than the static models. As seen from the bottom plot of Fig. 5, the hierarchical model (h) on average is strongly preferred to the dynamic model (d) and preferred to the static (s) dependence structure. In general, this figure summarizes two important results. Firstly, contrary to Johannes et al. (2014), dynamic dependence structure produces lower Bayes Factors than the static dependence structure. Although important to notice that Johannes et al. (2014) investigate monthly data, whereas our results are based on daily data. Secondly, hierarchical dependence structure seems to be the most flexible, at least for daily data.

4. Discussion and conclusion

In this paper, we consider static, hierarchical and time-varying stochastic copulas to model dependence between a single financial asset and two alternative predictor variables: its lagged value and dividend yield. We have designed a fast one-step estimation procedure based on the SMC techniques, that allow for consistent model comparison via log predictive Bayes Factors. We apply the proposed models to daily log returns of 20 assets traded at the NYSE and we find a number of important results.

Firstly, on average, RV-based models outperform the SV-based models in terms of sequential predictive log Bayes Factors. Moreover, more assets exhibit asymmetric dependence structure preferring Clayton copula to Normal copula models. Also, for majority of the assets dividend yield is a better predictor variable than the lagged returns. Finally, flexible hierarchical dependence structure is preferred by all assets versus dynamic random walk or static dependence structures.

Considering two sub-periods separately (calm and volatile), we have found that for calm period dividend yield is a better predictor variable, meanwhile during the crisis periods lagged return is a better predictor variable for the majority of the assets. As for the dependence structure, we find that for calm period it is reasonable to consider less flexible dynamics for the dependence parameter, however, for the crisis period a more flexible hierarchical dependence structure is preferred.

These findings show that the use of an ensemble of extra features, such as non-linearity, realized measures and flexible parameter structure, can improve equity return distribution forecasts. From the investor’s perfective, this translates into better performing portfolios.

CRediT authorship contribution statement

Audronė Virbickaitė: Conceptualization, Methodology, Software, Validation, Formal analysis, Data curation, Writing - original draft, Supervision, Project administration. **Christoph Frey:** Conceptualization, Methodology, Investigation, Resources, Visualization. **Demian N. Macedo:** Conceptualization, Investigation, Writing - review & editing, Project administration, Funding acquisition.

Declaration of competing interest

There is no conflict of interest.

Acknowledgements

We would like to thank Winfried Pohlmeier, anonymous Referee and the Editor for the helpful comments. We are also grateful to Roxana Halbleib for providing us the data. Financial support by the German Science Foundation (DFG - Projektnummer 321011636) and the Spanish Ministry of Science, Innovation and Universities (grant ECO2017-83255-C3-2-P) is gratefully acknowledged. All remaining errors are ours.

Appendix

For a state-space model described in (2.5) particle filter solves hidden state filtering problem for a set of fixed parameters Θ_C and Θ_V. A set of sufficient statistics S_t contains all updated hyper-parameters, necessary for the parameter simulation, as well as filtered state variables x_t, all approximated by a set of N particles. For $t = 1\ldots, T$ iterate through the following steps:

1. (Blind) Propagating

Sample new hidden states x_{t+1} from $x_{t+1}\tilde{}p(x_{t+1}|x_t,\Theta_C)$ and obtain θ_{t+1} deterministically. Call $\log\sigma_t^2 \equiv h_t$ and sample h_{t+1} from $p(h_{t+1}|h_t,\Theta_V)$ (only for the SV model).

(a.1) For the hierarchical model, where $x = x_{t+1}$, the dependence on the hyperparamaters m_x and V_x can be integrated out analytically:

$$
\begin{aligned}
p(x_{t+1}) &= \iint p(x_{t+1}|m_x, V_x)p(m_x, V_x)dm_x dV_x \\
&= \iint f_N(x_{t+1}|m_x, V_x) f_N(m_x|V_x) f_{IG}(V_x)dm_x dV_x \\
&= \frac{\Gamma\left(\frac{b_0+1}{2}\right)}{\Gamma\left(\frac{b_0}{2}\right)}(\pi(b_0V_{x0})(V_m+1))^{-1/2}\left(1+\frac{(x_{t+1}-m_m)^2}{(b_0V_{x0})(V_m+1)}\right)^{-\frac{b_0+1}{2}},
\end{aligned}
$$

which implies that x_{t+1} has a Student-t distribution with $b_0 > 0$ degrees of freedom, location parameter m_m, and scale parameter $(b_0V_{x0})(V_m + 1)/b_0 > 0$.

(a.2) For the dynamic model for x_{t+1} the dependence on the hyperparameter V_x, same as before, can be integrated out analytically:

$$
\begin{aligned}
p(x_{t+1}|x_t) &= \int p(x_{t+1}|x_t, V_x)p(V_x|x_t)dV_x \\
&= \int f_N(x_{t+1}|x_t, V_x) f_{IG}(V_x)dV_x \\
&= \frac{\Gamma\left(\frac{b_0+1}{2}\right)}{\Gamma\left(\frac{b_0}{2}\right)}(\pi(b_0V_{x0}))^{-1/2}\left(1+\frac{(x_{t+1}-x_t)^2}{(b_0V_{x0})}\right)^{-\frac{b_0+1}{2}},
\end{aligned}
$$

which implies that x_{t+1} has a Student-t distribution with $b_0 > 0$ degrees of freedom, location parameter x_t, and scale parameter $(b_0V_{x0})/b_0 > 0$.

2. Resampling

Resample old particles (parameters and the set of sufficient statistics, including states) with weights $w\propto p(u_{t+1}, u_t^{(j)};\theta_{t+1})$, that are proportional to the predictive density of the $(u_{t+1},u_t^{(j)})$, where $u_{t+1} = \Phi(r_{t+1}/\sigma_{t+1}^{(l)})$, such that $l = \{SV,RV\}$. $u_t^{(j)}$ is either the transformed lagged return, or the transformed dividend yield, i.e. $j \in \{r, DY\}$. The components of Θ_C and Θ_V have been simulated at the end of the previous period. The resampled particles are denoted by a tilde above the particle, as in $\tilde{\Theta}$, for example. Note, that $\sigma_{t+1}^{(l)}$ is either propagated in the previous step (for the SV model), or is observed (for the RV model).

3. Propagating sufficient statistics and learning Θ_C

For the hierarchical model:

(c.1) Sample V_x from $\mathcal{IG}(V_x; b_0^\star/2, b_0^\star V_{x0}^\star/2)$, where

$$b_0^\star = \tilde{b}_0 + 1 \quad \text{and} \quad b_0^\star V_{x0}^\star = \tilde{b}_0 \tilde{V}_{x0} + \frac{(\tilde{x}_{t+1} - \tilde{m}_m)^2}{1 + \tilde{V}_m}.$$

(c.2) Sample m_x from $\mathcal{N}(m_x; m_m^\star, V_m^\star V_x)$, where

$$m_m^\star = \frac{\tilde{m}_m + \tilde{V}_m \tilde{x}_{t+1}}{1 + \tilde{V}_m} \quad \text{and} \quad V_m^\star = \frac{\tilde{V}_m}{1 + \tilde{V}_m}.$$

For the dynamic model:

(c.3) Sample V_x from $\mathcal{IG}(V_x; b_0^\star/2, b_0^\star V_{x0}^\star/2)$, where

$$b_0^\star = \tilde{b}_0 + 1 \quad \text{and} \quad b_0^\star V_{x0}^\star = \tilde{b}_0 \tilde{V}_{x0} + (\tilde{x}_{t+1} - \tilde{x}_t)^2.$$

4. ***Propagating sufficient statistics and learning*** Θ_V*, where* h_{t+1} *is either observed (RV) or filtered (SV) log-volatility, depending on the model.*

(d.1) Sample $\tau^{2(i)}$ from $\mathcal{IG}(\tau^{2(i)}; b_0^{(i)\star}/2, b_0^\star \tau_0^{2\star}/2)$, where

$$b_0^\star = \tilde{b}_0 + 1 \quad \text{and} \quad b_0^\star \tau_0^{2\star} = \tilde{b}_0 \tilde{\tau}_0^2 + \frac{(\tilde{m}_\varphi \tilde{h}_t - (\tilde{h}_{t+1} - \tilde{\mu}))^2}{1 + \tilde{V}_\varphi \tilde{h}_t^2}.$$

(d.2) Sample φ from $\mathcal{N}(\varphi; m_\varphi^\star, V_\varphi^\star \tau^2)$, where

$$m_\varphi^\star = \frac{\tilde{m}_\varphi + \tilde{V}_\varphi \tilde{h}_t (\tilde{h}_{t+1} - \tilde{\mu})}{1 + \tilde{V}_\varphi \tilde{h}_t^2} \quad \text{and} \quad V_\varphi^\star = \frac{\tilde{V}_\varphi}{1 + \tilde{V}_\varphi \tilde{h}_t^2}.$$

(d.3) Sample μ from $\mathcal{N}(\mu; m_\mu^\star, V_\mu^\star)$, where

$$m_\mu^\star = \frac{\tilde{m}_\mu \tau^2 + \tilde{V}_\mu (\tilde{h}_{t+1} - \varphi \tilde{h}_t)}{\tau^2 + \tilde{V}_\mu} \quad \text{and} \quad V_\mu^\star = \frac{\tau^2 \tilde{V}_\mu}{\tau^2 + \tilde{V}_\mu}.$$

Table A.1
Average *LPS* for 20 assets for 20 models (entire sample: January 03, 2001–December 31, 2014)

	AIG	BA	BAC	C	F	GE	GS	HD	HPQ	IBM
(1) NC-RV-d-r	−1.761	−1.546	−1.619	−1.693	−1.837	−1.459	−1.679	−1.555	−1.699	−1.285
(2) NC-RV-d-dy	−1.754	−1.547	−1.617	−1.686	−1.834	−1.455	−1.678	−1.550	−1.697	−1.282
(3) NC-RV-h-r	−1.756	−1.542	−1.613	−1.687	−1.831	−1.453	−1.673	−1.549	−1.692	−1.278
(4) NC-RV-h-dy	−1.757	−1.542	−1.613	−1.686	−1.831	−1.452	−1.673	−1.548	−1.692	−1.278
(5) NC-SV-d-r	−1.772	−1.551	−1.622	−1.705	−1.844	−1.451	−1.684	−1.556	−1.709	−1.284
(6) NC-SV-d-dy	−1.763	−1.552	−1.621	−1.698	−1.840	−1.452	−1.685	−1.551	−1.709	−1.282
(7) NC-SV-h-r	−1.766	−1.546	−1.616	−1.698	−1.836	−1.445	−1.678	−1.546	−1.703	−1.277
(8) NC-SV-h-dy	−1.766	−1.546	−1.615	−1.696	−1.835	−1.445	−1.677	−1.546	−1.702	−1.277
(9) CC-RV-d-r	−1.759	−1.546	−1.619	−1.692	−1.839	−1.459	−1.678	−1.554	−1.698	−1.284
(10) CC-RV-d-dy	**−1.754**	−1.548	−1.618	−1.688	−1.834	−1.456	−1.677	−1.550	−1.700	−1.283
(11) CC-RV-h-r	−1.756	−1.542	**−1.613**	**−1.686**	−1.831	−1.453	−1.674	−1.548	**−1.692**	−1.278
(12) CC-RV-h-dy	−1.756	**−1.542**	−1.613	−1.687	**−1.831**	−1.451	**−1.673**	−1.548	−1.693	−1.278
(13) CC-SV-d-r	−1.772	−1.552	−1.624	−1.706	−1.844	−1.456	−1.684	−1.553	−1.709	−1.287
(14) CC-SV-d-dy	−1.760	−1.554	−1.622	−1.701	−1.838	−1.452	−1.684	−1.550	−1.709	−1.283
(15) CC-SV-h-r	−1.765	−1.545	−1.615	−1.695	−1.836	−1.446	−1.676	**−1.546**	−1.702	**−1.277**
(16) CC-SV-h-dy	−1.765	−1.546	−1.616	−1.698	−1.835	**−1.444**	−1.676	−1.546	−1.703	−1.277
(17) L-RV-s-r	−1.756	−1.543	−1.613	−1.687	−1.832	−1.453	−1.674	−1.549	−1.693	−1.279

(*continued on next page*)

Table A.1 (*continued*)

	AIG	BA	BAC	C	F	GE	GS	HD	HPQ	IBM
(18) L-RV-s-dy	−1.764	−1.544	−1.614	−1.687	−1.832	−1.453	−1.674	−1.549	−1.694	−1.280
(19) L-SV-s-r	−1.767	−1.547	−1.620	−1.698	−1.836	−1.445	−1.683	−1.556	−1.704	−1.279
(20) L-SV-s-dy	−1.767	−1.548	−1.617	−1.697	−1.835	−1.449	−1.678	−1.547	−1.704	−1.278
	JNJ	**JPM**	**KO**	**MO**	**NKE**	**PFE**	**PG**	**VZ**	**WMT**	**XOM**
(1) NC-RV-d-r	−1.016	−1.643	−1.138	−1.266	−1.489	−1.390	−1.070	−1.362	−1.224	−1.316
(2) NC-RV-d-dy	−1.011	−1.641	−1.133	−1.263	−1.487	−1.387	−1.069	−1.361	−1.223	−1.315
(3) NC-RV-h-r	−1.011	−1.637	−1.132	−1.262	−1.484	−1.385	−1.065	−1.357	−1.219	−1.311
(4) NC-RV-h-dy	−1.011	−1.637	−1.131	−1.262	**−1.484**	−1.385	−1.064	−1.358	−1.219	−1.311
(5) NC-SV-d-r	−1.010	−1.654	−1.135	−1.267	−1.496	−1.388	−1.064	−1.359	−1.223	−1.320
(6) NC-SV-d-dy	−1.009	−1.652	−1.134	−1.263	−1.494	−1.387	−1.064	−1.358	−1.221	−1.318
(7) NC-SV-h-r	−1.003	−1.646	−1.129	−1.258	−1.488	**−1.381**	−1.058	**−1.350**	−1.215	**−1.310**
(8) NC-SV-h-dy	−1.003	−1.645	−1.128	**−1.258**	−1.488	−1.382	**−1.058**	−1.351	−1.215	−1.310
(9) CC-RV-d-r	−1.014	−1.644	−1.137	−1.266	−1.489	−1.389	−1.069	−1.363	−1.223	−1.315
(10) CC-RV-d-dy	−1.012	−1.643	−1.132	−1.265	−1.489	−1.386	−1.068	−1.361	−1.221	−1.314
(11) CC-RV-h-r	−1.011	−1.637	−1.131	−1.262	−1.484	−1.385	−1.065	−1.357	−1.219	−1.310
(12) CC-RV-h-dy	−1.012	−1.638	−1.131	−1.263	−1.484	−1.385	−1.065	−1.357	−1.219	−1.310
(13) CC-SV-d-r	−1.010	−1.654	−1.139	−1.265	−1.496	−1.389	−1.066	−1.359	−1.224	−1.319
(14) CC-SV-d-dy	−1.009	−1.655	−1.134	−1.266	−1.494	−1.387	−1.064	−1.359	−1.222	−1.317
(15) CC-SV-h-r	**−1.003**	−1.644	−1.129	−1.259	−1.488	−1.382	−1.058	−1.351	−1.215	−1.310
(16) CC-SV-h-dy	−1.003	−1.646	**−1.128**	−1.258	−1.488	−1.382	−1.058	−1.352	**−1.215**	−1.311
(17) L-RV-s-r	−1.012	**−1.637**	−1.132	−1.263	−1.484	−1.386	−1.066	−1.358	−1.220	−1.312
(18) L-RV-s-dy	−1.014	−1.639	−1.132	−1.265	−1.485	−1.385	−1.068	−1.357	−1.220	−1.313
(19) L-SV-s-r	−1.004	−1.650	−1.129	−1.260	−1.491	−1.385	−1.061	−1.351	−1.217	−1.312
(20) L-SV-s-dy	−1.005	−1.649	−1.131	−1.261	−1.489	−1.384	−1.061	−1.353	−1.216	−1.311

Table A.2

Average *LPS* for 20 assets for 20 models (calm subsample: January 02, 2002–December 29, 2006)

	AIG	BA	BAC	C	F	GE	GS	HD	HPQ	IBM
(1) NC-RV-d-r	−1.473	−1.564	−1.195	−1.366	−1.802	−1.362	−1.581	−1.570	−1.756	−1.333
(2) NC-RV-d-dy	−1.463	−1.567	−1.188	−1.355	−1.799	−1.353	−1.577	−1.566	−1.750	−1.325
(3) NC-RV-h-r	−1.469	−1.560	−1.188	−1.359	−1.795	−1.355	−1.576	−1.564	−1.748	−1.326
(4) NC-RV-h-dy	−1.470	−1.560	−1.188	−1.358	−1.795	−1.354	−1.575	−1.563	−1.748	−1.326
(5) NC-SV-d-r	−1.485	−1.559	−1.207	−1.388	−1.810	−1.356	−1.583	−1.568	−1.774	−1.330
(6) NC-SV-d-dy	−1.475	−1.562	−1.202	−1.377	−1.808	−1.352	−1.582	−1.563	−1.770	−1.328
(7) NC-SV-h-r	−1.479	−1.555	−1.201	−1.382	−1.803	−1.348	−1.580	−1.558	−1.767	−1.325
(8) NC-SV-h-dy	−1.480	−1.555	−1.199	−1.379	−1.803	−1.350	−1.578	**−1.558**	−1.767	−1.324
(9) CC-RV-d-r	−1.468	−1.562	−1.192	−1.363	−1.801	−1.361	−1.579	−1.566	−1.755	−1.334
(10) CC-RV-d-dy	**−1.459**	−1.566	−1.186	**−1.352**	−1.799	−1.356	−1.578	−1.564	−1.751	−1.327
(11) CC-RV-h-r	−1.468	−1.559	−1.187	−1.358	−1.795	−1.354	−1.575	−1.563	−1.748	−1.326
(12) CC-RV-h-dy	−1.468	−1.560	−1.186	−1.355	**−1.795**	−1.354	**−1.575**	−1.564	**−1.748**	−1.326
(13) CC-SV-d-r	−1.481	−1.558	−1.206	−1.390	−1.812	−1.360	−1.583	−1.564	−1.774	−1.336
(14) CC-SV-d-dy	−1.471	−1.563	−1.201	−1.381	−1.807	−1.352	−1.585	−1.559	−1.769	−1.330
(15) CC-SV-h-r	−1.478	−1.555	−1.200	−1.378	−1.803	−1.350	−1.578	−1.558	−1.766	−1.324
(16) CC-SV-h-dy	−1.478	−1.555	−1.200	−1.384	−1.803	**−1.345**	−1.580	−1.558	−1.767	−1.325
(17) L-RV-s-r	−1.469	−1.559	−1.188	−1.359	−1.796	−1.355	−1.576	−1.564	−1.749	−1.326
(18) L-RV-s-dy	−1.464	−1.561	**−1.182**	−1.355	−1.796	−1.356	−1.576	−1.565	−1.752	−1.325
(19) L-SV-s-r	−1.479	**−1.554**	−1.204	−1.380	−1.804	−1.347	−1.585	−1.565	−1.768	−1.325
(20) L-SV-s-dy	−1.468	−1.556	−1.193	−1.375	−1.804	−1.355	−1.581	−1.559	−1.768	**−1.324**
	JNJ	**JPM**	**KO**	**MO**	**NKE**	**PFE**	**PG**	**VZ**	**WMT**	**XOM**
(1) NC-RV-d-r	−1.111	−1.513	−1.197	−1.365	−1.430	−1.453	−1.122	−1.463	−1.342	−1.398
(2) NC-RV-d-dy	−1.100	−1.506	−1.189	−1.366	−1.428	−1.450	−1.121	−1.457	−1.335	−1.395
(3) NC-RV-h-r	−1.106	−1.505	−1.190	−1.364	−1.424	−1.448	−1.114	−1.457	−1.335	−1.392
(4) NC-RV-h-dy	−1.106	−1.504	−1.190	−1.364	−1.425	−1.448	−1.114	−1.458	−1.335	−1.392
(5) NC-SV-d-r	−1.098	−1.520	−1.199	−1.369	−1.438	−1.466	−1.115	−1.457	−1.344	−1.390
(6) NC-SV-d-dy	−1.094	−1.517	−1.197	−1.369	−1.437	−1.462	−1.114	−1.456	−1.337	−1.385
(7) NC-SV-h-r	−1.089	−1.512	−1.192	−1.366	−1.430	−1.459	−1.107	−1.452	−1.334	**−1.379**
(8) NC-SV-h-dy	−1.089	−1.510	−1.190	−1.363	−1.431	−1.459	**−1.107**	−1.452	−1.334	−1.380
(9) CC-RV-d-r	−1.111	−1.510	−1.197	−1.366	−1.427	−1.452	−1.119	−1.465	−1.340	−1.396
(10) CC-RV-d-dy	−1.103	−1.508	**−1.188**	−1.369	−1.430	−1.451	−1.124	−1.458	−1.336	−1.392
(11) CC-RV-h-r	−1.106	**−1.504**	−1.190	−1.363	**−1.424**	−1.448	−1.114	−1.458	−1.334	−1.391
(12) CC-RV-h-dy	−1.108	−1.504	−1.190	−1.365	−1.425	−1.448	−1.117	−1.457	−1.336	−1.390
(13) CC-SV-d-r	−1.098	−1.515	−1.203	−1.374	−1.437	−1.467	−1.113	−1.459	−1.343	−1.389
(14) CC-SV-d-dy	−1.090	−1.521	−1.195	−1.368	−1.437	−1.464	−1.115	−1.456	−1.337	−1.384
(15) CC-SV-h-r	−1.089	−1.507	−1.191	−1.363	−1.431	−1.459	−1.107	−1.452	−1.333	−1.380
(16) CC-SV-h-dy	**−1.089**	−1.512	−1.191	−1.366	−1.431	−1.459	−1.107	−1.454	−1.333	−1.381
(17) L-RV-s-r	−1.106	−1.505	−1.191	−1.364	−1.425	−1.449	−1.115	−1.457	−1.335	−1.392

(*continued on next page*)

Table A.2 (*continued*)

	AIG	BA	BAC	C	F	GE	GS	HD	HPQ	IBM
(18) L-RV-s-dy	−1.105	−1.506	−1.190	−1.367	−1.425	**−1.447**	−1.115	−1.455	**−1.332**	−1.392
(19) L-SV-s-r	−1.090	−1.514	−1.193	−1.363	−1.432	−1.461	−1.109	−1.450	−1.335	−1.381
(20) L-SV-s-dy	−1.089	−1.513	−1.194	**−1.362**	−1.433	−1.459	−1.108	**−1.450**	−1.333	−1.381

Table A.3
Average LPS for 20 assets for 20 models (volatile subsample: January 03, 2007–December 31, 2012)

	AIG	BA	BAC	C	F	GE	GS	HD	HPQ	IBM
(1) NC-RV-d-r	−2.180	−1.580	−2.037	−2.062	−2.017	−1.600	−1.857	−1.607	−1.610	−1.258
(2) NC-RV-d-dy	−2.177	−1.582	−2.037	−2.060	−2.014	−1.599	−1.856	−1.603	−1.614	−1.254
(3) NC-RV-h-r	−2.174	−1.576	−2.031	−2.056	−2.010	−1.594	−1.850	−1.600	−1.604	−1.252
(4) NC-RV-h-dy	−2.175	**−1.576**	−2.031	−2.056	−2.010	−1.593	−1.850	−1.600	−1.604	−1.252
(5) NC-SV-d-r	−2.188	−1.596	−2.033	−2.069	−2.027	−1.593	−1.867	−1.614	−1.612	−1.262
(6) NC-SV-d-dy	−2.180	−1.595	−2.035	−2.067	−2.020	−1.595	−1.866	−1.611	−1.615	−1.256
(7) NC-SV-h-r	−2.181	−1.590	−2.029	−2.061	−2.019	−1.587	−1.859	−1.605	−1.605	−1.254
(8) NC-SV-h-dy	−2.182	−1.590	−2.028	−2.061	−2.018	**−1.587**	−1.858	−1.605	−1.605	−1.254
(9) CC-RV-d-r	−2.180	−1.581	−2.038	−2.063	−2.019	−1.600	−1.858	−1.606	−1.609	−1.257
(10) CC-RV-d-dy	−2.179	−1.584	−2.039	−2.066	−2.014	−1.602	−1.855	−1.605	−1.618	−1.257
(11) CC-RV-h-r	−2.174	−1.576	−2.031	**−2.056**	−2.010	−1.595	−1.851	**−1.600**	**−1.604**	−1.252
(12) CC-RV-h-dy	−2.176	−1.576	−2.032	−2.060	**−2.010**	−1.594	−1.851	−1.600	−1.606	**−1.251**
(13) CC-SV-d-r	−2.190	−1.597	−2.038	−2.071	−2.027	−1.600	−1.867	−1.613	−1.612	−1.264
(14) CC-SV-d-dy	−2.175	−1.598	−2.037	−2.069	−2.018	−1.596	−1.862	−1.611	−1.616	−1.257
(15) CC-SV-h-r	−2.181	−1.589	**−2.027**	−2.061	−2.019	−1.589	−1.858	−1.605	−1.605	−1.254
(16) CC-SV-h-dy	−2.181	−1.590	−2.029	−2.062	−2.018	−1.590	−1.857	−1.606	−1.606	−1.254
(17) L-RV-s-r	**−2.173**	−1.577	−2.031	−2.056	−2.011	−1.594	**−1.850**	−1.600	−1.604	−1.252
(18) L-RV-s-dy	−2.198	−1.576	−2.036	−2.060	−2.010	−1.595	−1.851	−1.600	−1.605	−1.253
(19) L-SV-s-r	−2.183	−1.590	−2.029	−2.061	−2.018	−1.588	−1.858	−1.612	−1.605	−1.255
(20) L-SV-s-dy	−2.195	−1.590	−2.033	−2.065	−2.017	−1.588	−1.858	−1.605	−1.606	−1.256

	JNJ	JPM	KO	MO	NKE	PFE	PG	VZ	WMT	XOM
(1) NC-RV-d-r	−0.934	−1.842	−1.101	−1.207	−1.570	−1.372	−1.051	−1.322	−1.167	−1.347
(2) NC-RV-d-dy	−0.934	−1.843	−1.095	−1.204	−1.566	−1.372	−1.055	−1.323	−1.168	−1.344
(3) NC-RV-h-r	−0.929	−1.838	−1.094	−1.201	−1.562	−1.368	−1.048	−1.315	−1.164	−1.342
(4) NC-RV-h-dy	−0.929	−1.838	−1.093	−1.201	−1.562	−1.368	−1.047	−1.316	−1.164	−1.342
(5) NC-SV-d-r	−0.934	−1.856	−1.098	−1.206	−1.575	−1.360	−1.045	−1.316	−1.167	−1.364
(6) NC-SV-d-dy	−0.934	−1.855	−1.094	−1.203	−1.573	−1.363	−1.048	−1.315	−1.165	−1.361
(7) NC-SV-h-r	−0.926	−1.849	−1.090	**−1.194**	−1.566	−1.355	−1.042	**−1.304**	−1.159	−1.353
(8) NC-SV-h-dy	−0.926	−1.849	−1.089	−1.195	−1.566	−1.356	**−1.041**	−1.306	**−1.159**	−1.354
(9) CC-RV-d-r	−0.932	−1.845	−1.100	−1.205	−1.569	−1.372	−1.051	−1.321	−1.167	−1.346
(10) CC-RV-d-dy	−0.929	−1.847	−1.094	−1.204	−1.568	−1.367	−1.051	−1.320	−1.164	−1.347
(11) CC-RV-h-r	−0.929	−1.839	−1.094	−1.201	**−1.562**	−1.368	−1.048	−1.315	−1.164	**−1.341**
(12) CC-RV-h-dy	−0.928	−1.839	−1.093	−1.202	−1.562	−1.367	−1.047	−1.315	−1.162	−1.342
(13) CC-SV-d-r	−0.933	−1.859	−1.102	−1.200	−1.575	−1.362	−1.051	−1.313	−1.168	−1.363
(14) CC-SV-d-dy	−0.935	−1.860	−1.094	−1.206	−1.573	−1.361	−1.047	−1.315	−1.168	−1.361
(15) CC-SV-h-r	**−0.926**	−1.849	−1.090	−1.198	−1.565	**−1.355**	−1.042	−1.305	−1.159	−1.353
(16) CC-SV-h-dy	−0.926	−1.850	**−1.089**	−1.195	−1.566	−1.356	−1.041	−1.306	−1.159	−1.355
(17) L-RV-s-r	−0.929	**−1.837**	−1.095	−1.201	−1.562	−1.369	−1.048	−1.316	−1.164	−1.342
(18) L-RV-s-dy	−0.933	−1.839	−1.094	−1.201	−1.563	−1.369	−1.051	−1.317	−1.166	−1.343
(19) L-SV-s-r	−0.926	−1.851	−1.089	−1.195	−1.568	−1.358	−1.042	−1.306	−1.159	−1.354
(20) L-SV-s-dy	−0.927	−1.851	−1.090	−1.198	−1.566	−1.358	−1.043	−1.307	−1.162	−1.353

References

Almeida, C., & Czado, C. (2012). Efficient Bayesian inference for stochastic time-varying copula models. *Computational Statistics & Data Analysis, 56*(6), 1511–1527.

Andersen, T. G., & Bollerslev, T. (1998). Answering the Skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review, 39*(4), 885–905.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Ebens, H. (2001). The distribution of realized stock return volatility. *Journal of Financial Economics, 61*(1), 43–67.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2000). Exchange rate returns standardized by Realized Volatility are (Nearly) Gaussian. *NBER Working Paper Series, 7488*, 1–21.

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

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

Ang, A., & Bekaert, G. (2007). Stock return predictability: Is it there? *Review of Financial Studies, 20*(3), 651–707.

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

Aye, G. C., Balcilar, M., Demirer, R., & Gupta, R. (2018). Firm-level political risk and asymmetric volatility. *The Journal of Economic Asymmetries, 18*, Article e00110. November.

Barndorff-Nielsen, O. E., & Shephard, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B, 66*(2), 253–280.

Barndorff-Nielsen, O. E., & Shephard, N. (2004). Econometric analysis of realised covariation: High frequency covariance, regression and correlation in financial economics. *Econometrica, 72*(3), 885–925.
Bollerslev, T., Patton, A. J., & Quaedvlieg, R. (2016). Exploiting the errors: A simple approach for improved volatility forecasting. *Journal of Econometrics, 192*(1), 1–49.
Bossaerts, P., & Hillion, P. (1999). Implementing statistical criteria to select return forecasting models: What do we learn? *Review of Financial Studies, 12*(2), 405–428.
Boudoukh, J., Michaely, R., Richardson, M., & Roberts, M. (2007). On the importance of measuring payout yield: Implications for empirical asset pricing. *The Journal of Finance, 62*(2), 877–915.
ben Brayek, A., Sebai, S., & Naoui, K. (2015). A study of the interactive relationship between oil price and exchange rate: A copula approach and a DCC-MGARCH model. *The Journal of Economic Asymmetries, 12*(2), 173–189.
Broto, C., & Ruiz, E. (2004). Estimation methods for stochastic volatility models: A survey. *Journal of Economic Surveys, 18*(5), 613–650.
Carter, A. C. K., & Kohn, R. (1994). On gibbs sampling for state space models. *Biometrika, 81*(3), 541–553.
Carvalho, C. M., Johannes, M. S., Lopes, H. F., & Polson, N. G. (2010a). Particle learning and smoothing. *Statistical Science, 25*(1), 88–106.
Carvalho, C. M., Lopes, H. F., Polson, N. G., & Taddy, M. A. (2010b). Particle learning for general mixtures. *Bayesian Analysis, 5*(4), 709–740.
Chen, X., & Fan, Y. (2006). Estimation of copula-based semiparametric time series models. *Journal of Econometrics, 130*(2), 307–335.
Chopin, N., Iacobucci, A., Marin, J.-M., Mengersen, K. L., Robert, C. P., Ryder, R., & Schäfer, C. (2011). *"On particle learning." comments on "Particle learning for sequential Bayesian computation" by Lopes, Carvalho, Johannes, and polson*. In J. M. Bernardo, M. J. Bayarri, J. O. Berger, A. P. Dawid, D. Heckerman, A. F. M. Smith, & M. West (Eds.) (pp. 1–35). arXiv:1006.
Chopin, N., Jacob, P. E., & Papaspiliopoulos, O. (2013). SMC2: An efficient algorithm for sequential analysis of state space models. *Journal of the Royal Statistical Society: Series B, 75*(3), 397–426.
Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics, 7*(2), 1–23.
Creal, D. D., & Tsay, R. S. (2015). High dimensional dynamic stochastic copula models. *Journal of Econometrics, 189*(2), 335–345.
Cremers, K. M. (2002). Stock return predictability: A Bayesian model selection perspective. *Review of Financial Studies, 15*(4), 1223–1249.
Darsow, W. F., Nguyen, B., & Olsen, E. T. (1992). Copulas and Markov processes. *Illinois Journal of Mathematics, 36*(4), 600–642.
Diebold, F. X., Gunther, T. A., & Tay, A. S. (1998). Evaluating density forecasts with applications to financial risk management. *International Economic Review, 39*(4), 863–883.
Doucet, A., & Johansen, A. M. (2009). A tutorial on particle filtering and smoothing: Fifteen years later. *In Handbook of nonlinear filtering, 12*, 656–704.
Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica, 50*(4), 987–1008.
Fama, E. F. (1991). Efficient capital markets: II. *The Journal of Finance, 46*(5), 1575–1617.
Fousekis, P. (2020). Sign and size asymmetry in the stock returns-implied volatility relationship. *The Journal of Economic Asymmetries, 21*, Article e00162. March.
Fruhwirth-Schnatter, S. (1994). Data augmentation and dynamic linear models. *Journal of Time Series Analysis, 15*(2), 183–202.
Fulop, A., & Li, J. (2013). Efficient learning via simulation: A marginalized resample-move approach. *Journal of Econometrics, 176*(2), 146–161.
García-Centeno, M.d. C., Fernández-Avilés, G., & Montero, J. M. (2010). Asymmetries in the volatility of precious metals returns: The TA-ARSV modelling strategy. *The Journal of Economic Asymmetries, 7*(1), 23–41.
Genest, C., Ghoudi, K., & Rivest, L.-P. (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika, 82*(3), 543–552.
Goyal, A., & Welch, I. (2003). Predicting the equity premium with dividend ratios. *Management Science, 49*(5), 639–654.
Harvey, A. C., & Shephard, N. (1996). Estimation of an asymmetric stochastic volatility model for asset returns. *Journal of Business & Economic Statistics, 14*(4), 429–434.
Joe, H. (2015). *Dependence modeling with copulas*. Boca Raton, FL: CRC Press.
Johannes, M. S., Korteweg, A., & Polson, N. G. (2014). Sequential learning, predictability, and optimal portfolio returns. *The Journal of Finance, 69*(2), 611–644.
Kahra, H., Martin, V. L., & Sarkar, S. (2018). A nonlinear model of asset returns with multiple shocks. *Studies in Nonlinear Dynamics & Econometrics, 23*(1), 1–45.
Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association, 90*(430), 773–795.
Kim, S., Shephard, N., & Chib, S. (1998). Stochastic volatility: Likelihood inference and comparison with ARCH models. *The Review of Economic Studies, 65*(3), 361–393.
Koop, G. (2003). *Bayesian econometrics*. John Wiley & Sons.
Lettau, M., & Ludvigson, S. (2001). Consumption, aggregate wealth, and expected stock returns. *The Journal of Finance, 56*(3), 815–849.
Lopes, H. F., & Tsay, R. S. (2011). Particle filters and Bayesian inference in financial econometrics. *Journal of Forecasting, 30*(1), 168–209.
McAleer, M., & Medeiros, M. C. (2008). Realized volatility: A review. *Econometric Reviews, 27*(1–3), 10–45.
Nam, K. (2003). The asymmetric reverting propertyof stock returns. *Studies in Nonlinear Dynamics & Econometrics, 6*(4), 1–18.
Nelsen, R. B. (2006). *An introduction to copulas* (2nd ed.). Springer.
Omori, Y., Chib, S., Shephard, N., & Nakajima, J. (2007). Stochastic volatility with leverage: Fast and efficient likelihood inference. *Journal of Econometrics, 140*(2), 425–449.
Patton, A. J. (2009). Copula - based models for financial time series. In T. G. Andersen, T. Mikosch, J.-P. Kreiß, & R. A. Davis (Eds.), *Handbook of financial time series* (pp. 767–785). Berlin, Heidelberg: Springer Berlin Heidelberg (chapter 36).
Paye, B. S., & Timmermann, A. (2006). Instability of return prediction models. *Journal of Empirical Finance, 13*(3), 274–315.
Pitt, M. K., dos Santos Silva, R., Giordani, P., & Kohn, R. (2012). On some properties of Markov chain Monte Carlo simulation methods based on the particle filter. *Journal of Econometrics, 171*(2), 134–151.
Sklar, A. (1959). *Fonctions de répartition à n dimensions et leurs marges*. *Publications Inst. Statis., Univ. Paris-VIII*.
Stambaugh, R. F. (1999). Predictive regressions. *Journal of Financial Economics, 54*(3), 375–421.
Storvik, G. (2002). Particle filters for state-space models with the presence of unknown static parameters. *IEEE Transactions on Signal Processing, 50*(2), 281–289.
Taylor, S. J. (1982). Financial returns modeled by the product of two stochastic processes - a study of daily sugar prices. In O. D. Anderson (Ed.), *Time series analysis: Theory and practice 1* (pp. 203–226). Amsterdam: North-Holland.
Trabelsi, N. (2017). Asymmetric tail dependence between oil price shocks and sectors of Saudi Arabia System. *The Journal of Economic Asymmetries, 16*, 26–41. January.
Virbickaite, A., Lopes, H. F., Ausín, M. C., & Galeano, P. (2019). Particle learning for Bayesian semi-parametric stochastic volatility model. *Econometric Reviews, 38*(9), 1007–1023.
Xia, Y. (2001). Learning about predictability: The effects of parameter uncertainty on dynamic asst allocation. *The Journal of Finance, 56*(1), 205–246.
Yu, J. (2002). Forecasting volatility in the New Zealand stock market. *Applied Financial Economics, 12*(3), 193–202.
Yu, J. (2005). On leverage in a stochastic volatility model. *Journal of Econometrics, 127*(2), 165–178.