

# FORECASTING GLOBAL EQUITY INDICES USING LARGE BAYESIAN VARs

*Florian Huber\**, *Tamás Krisztin†* and *Philipp Piribauer‡*

\*Vienna University of Economics and Business (WU), Vienna, Austria, †International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria, and ‡Austrian Institute of Economic Research (WIFO), Vienna, Austria

## ABSTRACT

This paper proposes a large Bayesian Vector Autoregressive (BVAR) model with common stochastic volatility to forecast global equity indices. Using a monthly dataset on global stock indices, the BVAR model controls for co-movement commonly observed in global stock markets. Moreover, the time-varying specification of the covariance structure accounts for sudden shifts in the level of volatility. In an out-of-sample forecasting application we show that the BVAR model with stochastic volatility significantly outperforms the random walk both in terms of point as well as density predictions. The BVAR model without stochastic volatility, on the other hand, shows some merits relative to the random walk for forecast horizons greater than six months ahead. In a portfolio allocation exercise we moreover provide evidence that it is possible to use the forecasts obtained from our model with common stochastic volatility to set up simple investment strategies. Our results indicate that these simple investment schemes outperform a naive buy-and-hold strategy.

*Keywords:* BVAR, equity indices, forecasting, log-scores, stochastic volatility

*JEL Classification:* C11, C22, C53, E17, G11

## I. INTRODUCTION

In recent decades, a marked increase in globalization led to stronger integration in the dynamics of various asset prices. The global financial crisis in 2008/09 illustrated how correlations between different global equity markets changed over time. In light of financial innovations and increased connectivity between international stock exchanges, accounting for the prevailing co-movement thus appears to be of predominant importance when modelling equity prices and returns. While the first moment of equity prices, i.e., the mean, displays similar movements across markets, the amount of co-movement in the volatilities also exhibits a large degree of similarity. Although this regularity has received considerable attention in the academic literature on the dynamics of stock prices, it has not yet been exploited in the forecasting literature. Practitioners in financial institutions like investment banks, pension funds or hedge funds naturally strive for models

Correspondence: Florian Huber, Institute for Macroeconomics, Vienna University of Economics and Business (WU), Welthandelsplatz 1, 1020, Vienna, Austria. E-mail: fhuber@wu.ac.at.

that are capable of producing precise out-of-sample predictions that might guide portfolio allocations. Moreover, central banks often manage large amounts of capital and tend to invest considerable amounts in equity markets. The sovereign wealth fund of Norway managed by the Norges Bank, for example, invests more than 60 percent of its available capital in global equity markets (Norges Bank Investment Management, 2014).

The vast majority of existing contributions on the prediction of stock prices focus either on atheoretical techniques, which tend to extrapolate recent behaviour of stock prices, or on theoretically motivated empirical models. Apart from purely statistical approaches, this strand of literature also includes techniques from machine learning and computational intelligence (see, for example, Chen *et al.*, 2003; Enke and Thawornwong, 2005). Contributions focusing exclusively on the statistical characteristics of the time series involved have been slightly more successful, in cases where the forecaster's loss function has been specified such that risk-return ratios of a given portfolio are optimized. Papers which follow a theoretically motivated empirical approach are, among others, Ou and Penman (1989); Holthausen and Larcker (1992) and Pesaran and Timmermann (1995). In addition, several studies exploit the relationship between dividend yields and excess returns for predictive purposes (Ang and Bekaert, 2007; Campbell and Thompson, 2008; Rapach *et al.*, 2010). All these studies rely on models that include a moderate to large number of predictors in the model. Another strand of the literature focuses on models that reduce the dimensionality of the problem. For instance, Kelly and Pruitt (2013) extract a latent factor from the cross-section of book-to-market ratios for the US stock market. Apart from focusing on the price or return of a given stock, an important strand of the literature exclusively deals with the predictability of stock market volatility. Prominent examples are French *et al.* (1987), Hamilton and Susmel (1994) and Bauer and Vorkink (2011).

The majority of the contributions mentioned so far conclude that empirical models with or without theoretical foundations are not able to produce more precise point forecasts than simple random walks, providing further evidence for the unpredictability of stock prices. The reason for this lack of out-of-sample performance mainly stems from three sources. First, most models used tend to be heavily parameterized, suggesting that the models overfit the data. This typically translates into weak out-of-sample predictive performance. The second reason could be due to the fact that financial markets are quite efficient, implying that if a model tends to generate robust predictions, traders would exploit this information until the advantage vanishes. Finally, most models employed to forecast equity prices are linear, implying that parameters are not allowed to change in response to different economic conditions.

In the present paper we alleviate such problems by focusing on a simple variant of a Bayesian vector autoregressive model with common stochastic volatility (BVAR-CSV) in the spirit of Carriero *et al.* (2015). Forecasting with Bayesian methods has a long standing tradition in macroeconomics. Recently, focus has shifted to high-dimensional models which explicitly allow for time-varying coefficients and stochastic volatility (Cogley and Sargent, 2005; Primiceri, 2005; Clark, 2011; Carriero *et al.*, 2015). We borrow strength from this literature by using a BVAR of moderate size that exploits two important characteristics of equity indices, namely the co-movement in the conditional mean and in the conditional volatility. More specifically, our model assumes that the volatility of the time series in the panel may be well described by a single latent factor that effectively scales the variance-covariance matrix of the VAR model. Several authors have emphasized the important role of stochastic volatility for producing accurate predictive densities, whereas research in finance mainly agrees on the importance to account for heteroskedasticity commonly observed in financial time series at moderate to high frequency time domains (Clark, 2011; Clark and Ravazzolo, 2015; Carriero *et al.*, 2014, 2015). This so-called volatility clustering needs to be properly incorporated in the modelling framework to obtain properly calibrated predictive densities.

This paper aims to contribute to the literature along several important dimensions. First, we provide evidence on the degree of co-movement between the volatility of equity indices, which supports the use of a single stochastic factor. Second, we benchmark our BVAR-CSV against the same BVAR without stochastic volatility and predict a panel of 18 well-known international equity indices. Albeit purely statistical in nature, our approach provides more precise point and density forecasts as compared to a simple random walk model. Finally, to assess whether the BVAR-CSV model performs well in terms of other loss functions, we use a simple portfolio allocation application. In this application we entertain the BVAR-CSV model to guide our investment decisions and show how investment strategies based on the BVAR-CSV produce larger investment returns relative to a naive investment scheme.

The remainder of the paper is structured as follows: Section II introduces the BVAR model with common stochastic volatility and discusses the prior setup employed. The design and evaluation of the forecasting application are outlined in Section III. Furthermore, we conduct a simple portfolio allocation exercise where different investment strategies based on the BVAR-CSV model are implemented and benchmarked against a naive investment strategy. Finally, the last section concludes the paper.

## II. THE ECONOMIC FRAMEWORK

This section outlines the econometric model. More specifically, after providing an overview of the statistical model we describe the prior and posterior distributions and give a brief overview of the Markov chain Monte Carlo (MCMC) algorithm.

### II.1 Bayesian vector autoregressive models

Let us consider the general VAR( $p$ ) model given by

$$Y_t = b_0 + B_1 Y_{t-1} + \dots + B_p Y_{t-p} + e_t, \quad (2.1)$$

where  $Y_t$  is a  $M \times 1$  vector of endogenous variables (equity indices) measured in time  $t$ ,  $b_0$  is a  $M \times 1$  intercept vector and  $B_1, \dots, B_p$  are conformable  $M \times M$  coefficient matrices. Finally,  $e_t$  denotes the usual vector of errors, where

$$e_t \sim \mathcal{N}(0, \Sigma_t). \quad (2.2)$$

Note that the  $M \times M$  matrix  $\Sigma_t$  is time-varying and depends on the  $M \times M$  variance-covariance matrix  $\Sigma$

$$\Sigma_t = \exp(h_t) \Sigma, \quad (2.3)$$

$$h_t = \gamma + \phi(h_{t-1} - \gamma) + \sigma u_t, \quad (2.4)$$

$$u_t \sim \mathcal{N}(0, 1). \quad (2.5)$$

Let  $\gamma \in \mathbb{R}$  and  $\phi \in (-1, 1)$  denote the level and the autoregressive parameter in (2.4), respectively. Finally,  $\sigma^2$  is the variance of the latent log-volatility process.

Note that we can decompose the time-invariant part of  $\Sigma_t$  as

$$\Sigma = \Upsilon S \Upsilon', \quad (2.6)$$

with  $\Upsilon$  being a lower triangular matrix with unit diagonal and typical free element denoted by  $\nu_{ij}$ . The matrix  $S = \text{diag}(s_1, \dots, s_M)$  stores the time-invariant idiosyncratic variances. For

identification purposes we set  $s_1 = 1$ , implying that the factor loading of the first element of  $e_t$  equals unity.

This volatility specification – in contrast to the existing literature on stochastic volatility in multivariate dynamic systems – implies that the whole system is driven by a single volatility process, thus effectively imposing a factor structure on the volatilities<sup>1</sup>. As we will show in Subsection III.2, this specification is justified on the ground that the first principal component extracted from the volatilities of our dataset explains the majority of observed variation. In addition, as will be explained below, this specific volatility structure implies significant computational gains as compared to models where the volatilities are modeled as equation-specific (see Primiceri, 2005; Clark, 2011, for an application).

Our model thus combines two important empirical regularities commonly observed in financial markets. First, our framework permits us to account for volatility clustering among equity indices. Due to the time-varying specification of the variance-covariance matrix, the model effectively incorporates sudden shifts in the level of volatility. However, it is worth noting that the assumption of constant covariances  $\Sigma$  over time implies that the relationship between the equity indices included in our panel is assumed to be constant over time as well. Second, due to the large panel used, we are also able to exploit cross-sectional information from the equity indices in the sample. The BVAR model thus inherently accounts for interdependencies and co-movements between different equity indices.

The model in (2.1) can be rewritten more compactly as

$$Y_t = B'X_t + e_t, \quad (2.7)$$

where  $B = (B_1, \dots, B_p, b_0)'$ , which is a  $K \times M$  matrix, with  $K = Mp + 1$  and  $X_t = (Y'_{t-1}, \dots, Y'_{t-p}, 1)'$  is a  $K \times 1$  vector. Stacking the columns of (2.7) yields

$$Y = XB + e, \quad (2.8)$$

where  $Y$  and  $X$  are  $T \times M$  and  $T \times K$ , respectively. Additionally, it proves to be convenient to normalize the matrices  $X_t$  and  $Y_t$  by dividing through  $\exp(h_t/2)$ , that is

$$\tilde{X}_t = \exp(-h_t/2)X_t \text{ and } \tilde{Y}_t = \exp(-h_t/2)Y_t, \quad (2.9)$$

with the corresponding full-data matrices denoted as  $\tilde{X}$  and  $\tilde{Y}$ , respectively.

## II.2 Prior specification

The VAR described in the previous subsection typically suffers from the well-known curse of dimensionality. This implies that the apparent overparameterization of the model in equation (2.1) leads to in-sample overfitting, which typically translates into weak out-of-sample forecasting performance. To alleviate overfitting problems we introduce additional information in the model through Bayesian shrinkage priors in the spirit of Doan *et al.* (1984), Litterman (1986) and Sims and Zha (1998). This implies shrinking the model discussed above towards a prior model, which in our case is a random walk with drift. This is predicated by the fact that stock prices tend to follow a random walk process, which is typically a tough benchmark for more elaborate econometric models.

In general, a Bayesian framework requires the specification of prior distributions on all parameters in the model. Our prior setup is given by

$$\text{vec}(B) | \Sigma^{-1}, \theta \sim \mathcal{N}(\text{vec}(\underline{B}), \Sigma \otimes \underline{V}_B), \quad (2.10)$$

<sup>1</sup>See Huber (2016) for a systematic comparison between large-scale models that feature different stochastic volatility specifications.

where equation (2.10) represents the normally distributed prior on  $B$ , where  $\underline{B}$  and  $\underline{V}_B$  denote prior mean and variance, respectively. Note that we assume prior dependence between  $B$  and  $\Sigma$ , which leads to a conjugate prior specification. Conjugacy implies that the posterior distributions are available in closed-form and the Kronecker structure of the variance-covariance matrix for  $B$  leads to significant computational gains.<sup>2</sup>

Moreover, note that we also condition on a hyperparameter  $\theta$ , which controls the tightness of the prior. Following Giannone *et al.* (2015), we impose a Gamma prior with parameters  $a_0$  and  $b_0$  on  $\theta$ ,

$$\theta \sim \mathcal{G}(a_0, b_0). \tag{2.11}$$

$\theta$  is thus treated as an unknown quantity to be estimated jointly with the system described in (2.1)-(2.5). It is noteworthy that conditional on  $\theta$ , this prior setup can be implemented through suitable dummy observations. This captures the notion that the prior arises from a fictitious dataset. In general, let  $\underline{Y}$  and  $\underline{X}$  denote suitable dummy data matrices. Then, the prior variance on the coefficients and the prior mean are given by

$$\underline{V}_B = (\underline{X}'\underline{X})^{-1}, \tag{2.12}$$

$$\underline{B} = \underline{V}_B \underline{X}'\underline{Y}. \tag{2.13}$$

For  $s_j$  we elicit an inverted gamma prior with hyperparameters  $c_0$  and  $d_0$ ,

$$s_j \sim \mathcal{IG}(c_0, d_0), \tag{2.14}$$

where we set the prior hyperparameters such that the prior is effectively non-informative. Moreover, we impose a normally distributed prior on  $v_{ij}$ ,

$$v_{ij} \sim \mathcal{N}(\underline{\mu}_v, \underline{V}_v). \tag{2.15}$$

$\underline{\mu}_v$  and  $\underline{V}_v$  denote the prior mean and variance, respectively.

For the parameters of the log-volatility equation in (2.4), we use the prior setup proposed in Kastner and Frühwirth-Schnatter (2014). This implies specifying a Gaussian prior on  $\gamma$ ,

$$\gamma \sim \mathcal{N}(\underline{\mu}_\gamma, \underline{V}_\gamma), \tag{2.16}$$

with mean  $\underline{\mu}_\gamma$  and variance  $\underline{V}_\gamma$ . Moreover, on the autoregressive parameter we impose a beta distributed prior given by

$$\frac{\phi + 1}{2} \sim \mathcal{B}(a_1, b_1), \tag{2.17}$$

Note that the density for  $\phi$  is given by

$$p(\phi) = \frac{1}{2B(a_1, b_1)} \frac{(1 + \phi)^{(a_1-1)} (1 - \phi)^{(b_1-1)}}{2}, \tag{2.18}$$

where  $B(a_1, b_1)$  denotes the Beta function. A convenient feature of this prior setup is that it rules out explosive behavior of the log-volatility process because the support of this distribution is the unit interval  $(-1, 1)$ . Using the mean of the Beta distribution,  $\mathbb{E}(\phi)$  is given by

$$\mathbb{E}(\phi) = \frac{2a_1}{a_1 + b_1} - 1. \tag{2.19}$$

<sup>2</sup>This result holds true as long as we condition on  $h_t$  and  $\theta$ .

Hence, if  $a_1$  is greater than  $b_1$ , the prior mean is positive, whereas if  $b_1$  is greater than  $a_1$ , the prior mean would be negative.

Finally, we use a gamma prior on  $\sigma^2$ ,

$$\sigma^2 \sim \mathcal{G}(1/2, 1/2B_\sigma). \tag{2.20}$$

Here,  $B_\sigma$  is a scalar hyperparameter controlling the tightness of the prior. It is straightforward to show that this prior translates into a normally distributed prior on the signed square root of  $\sigma^2$  with mean zero and variance  $B_\sigma$ . This implies that, in contrast to the traditional inverted gamma prior, we do not bound  $\sigma^2$  away from zero a-priori and thus allow for more shrinkage. These choices are motivated in Frühwirth-Schnatter and Wagner (2010) and Kastner and Frühwirth-Schnatter (2014).

### II.3 Posterior distributions

Due to the specific form of the priors discussed in the previous subsection it is possible to derive well-known conditional posterior distributions for  $B$  and  $\Sigma^{-1}$ , which facilitate a simple Gibbs sampling scheme.

Under the prior assumptions (2.10) - (2.20), the conditional posterior for  $B$  is given by

$$\text{vec}(B) | \Sigma^{-1}, \theta, h, \mathcal{D} \sim \mathcal{N}(\text{vec}(\bar{B}), \Sigma \otimes \bar{V}_B), \tag{2.21}$$

with

$$\bar{V}_B = (\bar{X}'\bar{X})^{-1}, \tag{2.22}$$

$$\bar{B} = \bar{V}_B \bar{X}' \bar{Y}, \tag{2.23}$$

where  $\bar{X} = (\tilde{X}', \underline{X}')$ ,  $\bar{Y} = (\tilde{Y}', \underline{Y}')$ ,  $h = (h_1, \dots, h_T)$  and  $\mathcal{D}$  denotes the available data. Note that conditional on  $h$ , posterior quantities are standard results found in many sources (see, for example, Kadiyala and Karlsson, 1997; Koop and Korobilis, 2010; Karlsson, 2013).

The conditional posterior of  $\Sigma^{-1}$  is of Wishart form

$$\Sigma^{-1} | B, \theta, h, \mathcal{D} \sim \mathcal{W}(\bar{v}, \bar{S}), \tag{2.24}$$

with  $\bar{v} = \underline{v} + T$  and  $\bar{S} = (\bar{Y} - \bar{X}B)'(\bar{Y} - \bar{X}B)$ .

Unfortunately, the conditional posterior distributions for the remaining parameters are of no well-known form. This implies that  $p(h|B, \Sigma^{-1}, \theta, \mathcal{D})$  and  $p(\theta|B, \Sigma^{-1}, h, \mathcal{D})$  are not readily available, which prevents the usage of simple Gibbs steps for the aforementioned parameters.

### II.4 Prior implementation and posterior simulation

To estimate the BVAR model we have to specify the hyperparameters for the priors discussed above. Starting with the prior on  $B$ , we follow Bańbura *et al.* (2010) and Koop (2013), and construct the following dummy observations to implement a variant of the Minnesota prior (Litterman, 1986). This implies choosing  $\underline{V}_B$  and  $\underline{B}$  such that the prior model equals the naive random walk with drift and the prior variance is set so that coefficients on higher lag orders are shrunk aggressively towards zero. More specifically, the following dummy observations are used to match the Minnesota moments

$$\underline{Y} = \begin{pmatrix} \text{diag}(\underline{b}_1 s_1, \dots, \underline{b}_M s_M) / \theta \\ \mathbf{0}_{M(p-1) \times M} \\ \text{diag}(s_1, \dots, s_M) \\ \mathbf{0}_{1 \times M} \end{pmatrix}, \underline{X} = \begin{pmatrix} J_p \otimes \text{diag}(s_1, \dots, s_M) / \theta & \mathbf{0}_{M_p \times 1} \\ \mathbf{0}_{M \times M_p} & \mathbf{0}_{M \times 1} \\ \mathbf{0}_{1 \times M_p} & \kappa \end{pmatrix}, \tag{2.25}$$

where  $J_p = \text{diag}(1, \dots, p)$  and  $b_1, \dots, b_M$  denote the diagonal elements of the first  $M$  rows and columns of  $\underline{B}$ , which just equals the identity matrix under the traditional Minnesota prior. Note, that the first block in (2.25) implements the prior on the first lag of the endogenous variables, whereas the second block implements the prior on the variance-covariance matrix. The last block implements the prior on the intercept, where  $\kappa = 1/1000$  is set such that the prior on the constant is effectively non-informative. Following Litterman (1986) and Sims and Zha (1998),  $s_1, \dots, s_M$  denote standard deviations obtained by estimating univariate autoregressive models of order  $p$ . Usually, the tightness parameter is assumed to be constant and known a priori. However, following Giannone *et al.* (2015) we treat  $\theta$  as an unknown quantity to be estimated. For the gamma prior associated with  $\theta$ , we set the hyperparameters equal to  $a_0 = 1, b_0 = 1$ . For the free elements in  $\Upsilon$ , we set the prior mean  $\underline{\mu}_v$  equal to zero and the variance  $\underline{V}_v$  equal to ten. Given the scale of our data this choice proves to be rather uninformative. Moreover, for the prior on  $s_j$  we set  $c_0 = d_0 = 0.01$ . For the log-volatility equation, we use the following set of hyperparameters for the priors. First, for the Beta prior on  $\phi$  we set  $a_1 = 5$  and  $b_1 = 1.5$ , resulting in a prior which puts considerable mass on high-persistence regions of  $\phi$ . Second, the prior mean on the level of the log-volatility is set equal to zero, with variance set to 100. This translates into a diffuse prior on  $\gamma$ . Finally, for  $\sigma$  we set  $B_\sigma = 1$ .

This leads us directly to the specific design of our MCMC algorithm. The following MCMC algorithm is employed to perform posterior inference:

1. Initialize the parameters of the model using maximum likelihood estimates or draws from the prior.
2. Sample  $B$  from  $\mathcal{N}(\text{vec}(\bar{B}), \Sigma \otimes \bar{V}_B)$ .
3. Conditional on  $B$ , draw  $\Sigma^{-1}$  from  $\mathcal{W}(\bar{v}, \bar{S})$ .
4. Obtain a draw from  $p(h|B, \Sigma^{-1}, \theta, \mathcal{D})$  (and the parameters of the log-volatility equation) using the algorithm outlined in Kastner and Frühwirth-Schnatter (2014).
5. Finally, sample  $\theta$  using a random walk Metropolis step with acceptance probability

$$\alpha(\theta^*|\theta) = \frac{p(\mathcal{D}|\theta^*)p(\theta^*)}{p(\mathcal{D}|\theta)p(\theta)}. \tag{2.26}$$

Let  $\theta^* \sim \mathcal{N}(\theta, \bar{\delta})$  denote the proposed value of  $\theta$ ,  $\bar{\delta}$  be a scaling parameter<sup>3</sup> and  $p(\mathcal{D}|\theta)$  denotes the marginal likelihood.

Steps 2. and 3. can be implemented using simple Gibbs steps. For the components of the log-volatility equation and consequently the history of log-volatilities, we use the so-called ancillarity sufficiency interweaving strategy put forward by Kastner and Frühwirth-Schnatter (2014).<sup>4</sup> Step 5. is implemented using a random walk Metropolis step where the marginal likelihood is available in closed form due to (conditional) conjugacy. More specifically, the marginal likelihood is given by

$$p(\mathcal{D}|\theta) \propto (|\underline{L}||\bar{V}^{-1}|)^{\frac{M}{2}} |\bar{S}|^{-\frac{T+M+p-1}{2}}. \tag{2.27}$$

This completes the description of the MCMC algorithm employed.

<sup>3</sup>This parameter is specified such that the acceptance probability lies between 20 and 40 percent, see Karlsson (2013).

<sup>4</sup>This step is implemented by means of the R package *stochvol* (Kastner, 2016).

TABLE 1  
*Stock indices used in the BVAR model*

<i>Region</i>	<i>Name</i>	<i>Description</i>
Northern America	DJIA	Dow Jones Industrial Average
	SPX	S&P 500 Index
	COMPX	NASDAQ Composite Index
	OSPTX	S&P/Toronto Stock Exchange Composite Index
Latin America	MEXBOL	Mexican Stock Exchange Mexican Bolsa IPC Index
	IBOV	Ibovespa Brasil Sao Paulo Stock Exchange Index
Europe	SX5E	EURO STOXX 50 Index
	FTSE	Financial Times Stock Exchange 100 Index
	CAC	Cotation Assistée en Continu 40 Index
	DAX	Deutsche Boerse AG German Stock Index
	IBEX	Iberia 35 Index
	XMB	Financial Times Stock Exchange/Milano Italia Borsa Index
	AEX	Amsterdam Exchange Index
	OMX	OMX Stockholm 30 Index
	SMI	Swiss Market Index
Asia-Pacific	NKY	Nikkei 225 Index
	HSI	Hong Kong Hang Seng Index
	SPASX	S&P/Australian Securities Exchange 200 Index

### III. FORECASTING RETURNS OF EQUITY INDICES

In this section we provide a brief overview on the data and the model specification adopted in this study, assess whether our proposed modelling approach is at odds with the data, and present the main findings of our forecasting exercise. Moreover, the last subsection illustrates the advocated approach within a simple portfolio exercise.

#### *III.1 Data overview and model specification*

In order to account for the co-movement between the indices under scrutiny, we use a large sample of stock indices. The dynamic behaviour of a stock index is thus not only explained by its own movement, but also by the dynamics of the other indices. Our dataset comprises the most important equity indices (in terms of market capitalization) across North America, Europe, Asia and Latin America. More specifically, the dataset includes major equity indices of Northern America (4 indices), Europe (9 indices), Latin America (2 indices) as well as Asia and Pacific (3 indices). Further details on the included stock indices are depicted in Table 1. Data on stock prices are obtained on a daily basis from the *Yahoo! Finance* database for the time period ranging from 1 January 1998 to 31 July 2014. We then constructed monthly averages of all equity indices under consideration.

All equity indices enter the model in log levels and we include a lag length of  $p = 4$ . For the MCMC algorithm we use 10,000 iterations with the first 5,000 being discarded as burn-ins.<sup>5</sup>

<sup>5</sup>Estimation of the model has been carried out in R, where all estimation files are available upon request.

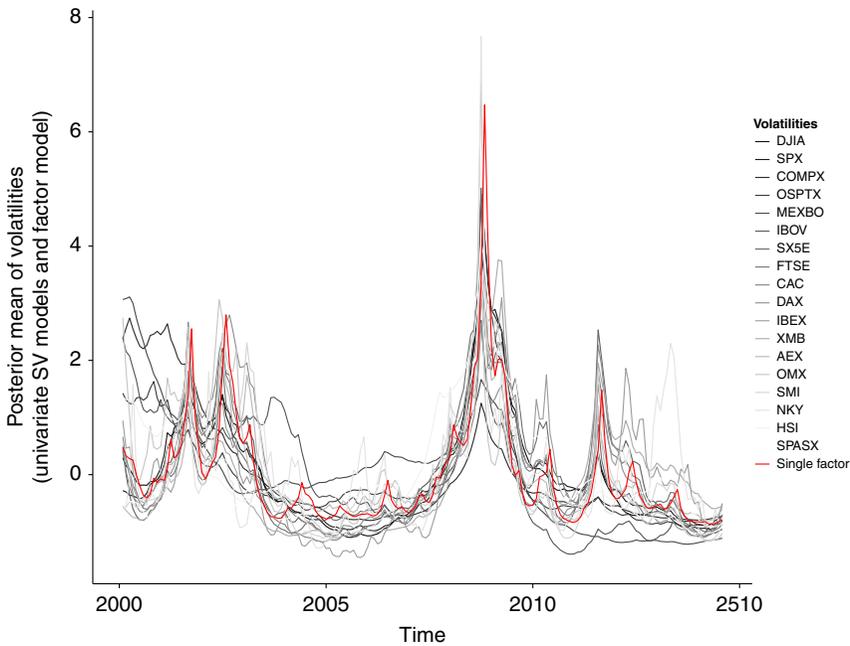


Fig. 1. Volatilities of univariate stochastic volatility models and the posterior mean of  $\exp(h_t)$ .

### III.2 Is there a common factor in the volatilities of equity price returns?

Before we proceed with the actual forecasting exercise it is worth noting that our modelling approach seems to be overly simplistic at a first glance. However, equity markets tend to be strongly integrated across the globe, implying that shocks that hit a particular country (most notably the USA) influence equity prices internationally.

Figure 1 presents the volatilities obtained by running a set of univariate stochastic volatility models of the form

$$\Delta y_{it} = \exp(h_{it}/2)\eta_{it}, \quad (3.1)$$

where  $\Delta y_{it}$  is the first difference of the  $i$ th element of  $Y_t$  and  $\eta_{it} \sim \mathcal{N}(0, 1)$  is a standard normally distributed random variate. In addition, Figure 1 also presents the common volatility factor that is obtained by estimating the BVAR-CSV over the full sample.

A few interesting findings emerge. Note that all volatilities tend to follow a similar pattern, implying that global equity markets display a strong tendency towards co-movement in volatility. Equity price volatility increased across the globe, especially in the early 2000s. This can be traced back to the burst of the ‘dot-com’ bubble and the 11 September terrorist attacks. Both shocks could be viewed as being purely US based. However, the pronounced increase in volatility for most indices considered suggests that the aforementioned country-specific shocks quickly turned into global shocks, leading to sharp declines in equity prices. The financial crisis of 2008/2009, that originated from the US housing market, also led to severe drops in equity prices and a sudden spike in the volatility of equity returns. For this period, we find again a general increase in volatility for most indices considered.

Apart from small idiosyncratic deviations from this general pattern, visual inspection suggests that the variance structure of a broad set of equity indices can be well captured by our simplistic

TABLE 2  
Statistical evidence for a common volatility structure

<i>Proportion of variance explained by the first four factors</i>				
	<i>Component 1</i>	<i>Component 2</i>	<i>Component 3</i>	<i>Component 4</i>
Prop. of Variance	0.76	0.09	0.07	0.02
Cumulative Prop.	0.76	0.85	0.92	0.94
<i>Classical Selection Criteria</i>				
<i>Kaiser</i>	<i>Opt. Coord.</i>	<i>Parallel. Analysis</i>	<i>Opt. Acceleration</i>	
2	2	2	1	

The upper panel depicts the proportion of variance explained by the first four principal components (first row) and the cumulative proportion of variance (second row). The lower panel presents the results of several classical information criteria.

single factor model for the variance-covariance matrix. To provide some statistical information on the appropriateness of our approach, Table 2 presents evidence on the amount of variation explained due to the first four principal components and a set of selected classical information criteria.

The upper panel of Table 2 suggests that over the estimation period, the first principal component explains around 76 percent of total variation of the panel, capturing the majority of variation in volatilities across all 18 indices considered. The lower panel of Table 2 presents the results obtained by estimating several classical information criteria to select the optimal number of factors. The criteria we use are the well-known Kaiser and parallel analysis criteria, that both rely on the inspection of the sample correlation matrix. The optimal coordinates (Opt. Coord.) criterion is obtained by linearly extrapolating the coordinates of the preceding eigenvalues through a linear regression. Finally, the optimal acceleration criterion inspects the inflection point of the scree plot. All measures point towards one or two factors. Inspection of the corresponding scree plot (not shown) also points towards two factors (albeit the eigenvalue associated with the second principal component only marginally exceeds unity). Our analysis thus suggests that the optimal number of factors lies between one and two and supports our assumption of one stochastic factor.

### III.3 Design of the forecasting exercise

We propose the following recursive forecasting exercise.<sup>6</sup> In the first step, we start with an initial estimation period ranging from 1998:M01 to 2011:M07. The remaining observations are used as a verification sample to assess the predictive accuracy of our modeling approach. Then, we use the BVAR to simulate  $k$ -step ahead predictive densities. After obtaining the predictions, we expand the initial estimation window by  $k$  steps. This procedure is repeated until we reach the end of our data sample (2014:M07).

In our framework the  $k$ -step ahead predictive density is given by

$$p(Y_{\tau+k}|\mathcal{D}_{\tau}) = \int_{\Xi} p(Y_{\tau+k}|\mathcal{D}_{\tau}, \Xi)p(\Xi|Y_{1:\tau})d\Xi, \quad (3.2)$$

<sup>6</sup>The results based on a rolling estimation window yields similar results.

where  $\mathcal{D}_\tau$  denotes the information set up to time  $\tau$  and  $\Xi$  collects all available parameters of the model. Equation (3.2) can be easily approximated using Monte Carlo integration. As a point estimator, we utilize the mean of the predictive density, denoted by  $\bar{Y}_{\tau+k}$ .

We base our forecasting comparison upon two measures, the root mean square error (RMSE), which is a well-known measure for the accuracy of point forecasts. The RMSE is defined as

$$\text{RMSE} = \sqrt{\frac{\sum_{\tau=t_0}^T (Y_\tau^O - \bar{Y}_\tau)^2}{T}}, \tag{3.3}$$

where  $t_0$  and  $T$  denotes the first and last period of the verification sample, respectively. The actual outcome at time  $\tau$  is denoted by  $Y_\tau^O$ .

To assess how well our model performs in terms of density predictions, we moreover adopt the log predictive score (LPS). The LPS is a well-known Bayesian evaluation criterion, motivated recently in Geweke and Amisano (2010). In general, the log predictive score is the logarithm of the predictive density  $\log p(Y_{\tau+k}|\mathcal{D}_\tau)$  evaluated at  $Y_{\tau+k}^O$ . This implies that the log predictive score is given by

$$\text{LPS} = \sum_{\tau=t_0}^{T-k} \log p(Y_{\tau+k} = Y_{\tau+k}^O | \mathcal{D}_\tau). \tag{3.4}$$

Conjugacy of the model described above implies that the one-step ahead predictive density is available in closed form. However, for  $k > 1$  we have to perform posterior simulation. Evaluation of the predictive density is then done using the quadratic approximation put forward in Adolfson *et al.* (2007), which is given by

$$\text{LPS}(Y_{\tau+k}^O) = -0.5(M \log(2\pi) + \log |\bar{V}_{\tau+k|\tau}| + (Y_{\tau+k}^O - \bar{Y}_{\tau+k})' \bar{V}_{\tau+k|\tau}^{-1} (Y_{\tau+k}^O - \bar{Y}_{\tau+k})), \tag{3.5}$$

where  $\bar{V}_{\tau+k|\tau}$  and  $\bar{Y}_{\tau+k}$  denote the posterior variance and mean of the predictive density, respectively. Another convenient property of this approximation is that we can easily obtain the corresponding marginal predictive densities (i.e., the predictive density for a given element of  $Y_{\tau+k}$  after integrating out the remaining elements of  $Y_{\tau+k}$ ) by simply dropping the corresponding rows and columns of  $\bar{V}_{\tau+k|\tau}$  and  $\bar{Y}_{\tau+k}$ .

### III.4 Forecasting results

This section provides details on the evaluation of out-of-sample forecasts of the equity indices in our dataset. Table 3 summarizes the results of the forecasting exercise for different forecasting horizons and model settings. BVAR-CSV and BVAR denote the BVAR models with and without common stochastic volatility, respectively. In both cases out-of-sample forecasts ranging from one to twelve months ahead are reported.

The RMSEs presented in the upper part of the table are moreover reported relative to the RMSEs of random walk forecasts, implying that values below unity signal outperformance, while values exceeding unity indicate underperformance, relative to the benchmark. To assess the significance of our findings, we also report significance levels obtained by running the Diebold-Mariano test (see Diebold and Mariano 1984) and the Amisano-Giacomini test<sup>7</sup> (Amisano and

<sup>7</sup>Note, that the specific design of our forecasting exercises implies that the size of the estimation and verification sample expand over time, suggesting that a key condition for the validity of the Amisano-Giacomini test is violated. Nevertheless, we still include confidence levels to provide a rough gauge on the significance of the differences in LPS.

TABLE 3  
 Out-of-sample performance relative to the random walk model in terms of the sum of the root mean square error (RMSE) and log predictive scores (LPS): 2011:M08 to 2014:M07

Index	Root mean square errors relative to the random walk model					BVAR months				
	1	3	6	9	12	1	3	6	9	12
DJIA	0.774**	0.885	0.713	0.794**	0.833**	0.905	1.145	0.746	1.249**	1.005
SPX	0.912	0.744*	0.787**	0.778**	0.820**	1.023	0.982	0.976**	0.962**	0.983**
COMPX	0.950	0.880	0.905*	0.844**	0.849**	0.990	1.013	0.991	0.981	0.965**
OSPTX	1.113	1.135*	1.131**	1.048	1.063**	1.090**	1.076**	1.029	1.055**	1.055**
MEXBO	0.990	0.991	1.004	0.956**	0.882**	0.981	1.022	0.995	0.967**	0.993
IBOV	1.007	0.959	1.019	0.913**	1.008	1.018	0.978	1.014	0.915*	1.010
SX5E	0.983	0.960	0.974	0.899**	0.857**	1.015	1.019	1.036	1.043**	1.025*
FTSE	0.975	0.948	0.986	0.748**	0.808**	0.981*	0.982	1.017	0.958**	1.010*
CAC	1.001	0.998	0.984	0.916**	0.896**	1.020*	1.032	0.980	1.052**	1.044**
DAX	0.988	0.945	0.932	0.810**	0.821**	0.985	0.966	0.954**	0.943**	0.955**
IBEX	1.005	1.002	0.997	1.013*	1.023	1.029	1.045	1.028	1.068**	1.052**
XMB	1.002	0.966	0.973	0.946**	0.949**	1.027	1.046	1.012	1.090**	1.065**
AEX	0.923	0.818	0.827	0.723**	0.629**	0.929	0.872	0.857*	0.884**	0.820**
OMX	0.931	0.835	0.944	0.653**	0.718**	0.903	0.859	0.919	0.760**	0.765**
SMI	0.929	0.917**	0.845**	0.902**	0.901**	0.994	1.001	0.959*	0.993	0.985*
NKY	0.965	0.956	0.899**	0.915**	0.949**	1.004	1.005	0.986**	0.996*	0.999
HSI	1.036	1.053	1.068**	1.051	1.031	1.032	1.033	1.045*	0.990	1.068**
SPASX	1.006	1.035	0.982	1.029	0.992	1.032	1.058	1.069	1.072	1.076**
Average	0.972	0.946*	0.943**	0.885**	0.890**	0.998	1.007	0.978	0.999*	0.993*

(Continued)

TABLE 3  
Continued

Index	BVAR-CSV months						Log predictive scores relative to the random walk model						BVAR months		
	1	3	6	9	12	1	3	6	9	12	1	3	6	9	12
DJIA	1.623*	1.204	1.663	-0.881	2.225*	0.331	-0.167	-0.242	-1.138*	1.594*					
SPX	-3.484*	-0.016	6.009	19.306**	25.666**	-5.294*	-4.002	0.562	10.477*	13.659**					
COMPX	-1.244	-1.743	-0.026	5.958**	8.998**	-2.382	-3.586	-2.394	1.267	3.342*					
OSPTX	-5.412*	-5.459*	-2.542	1.535	7.795	-6.237**	-6.484**	-1.924	0.226	7.217*					
MEXBO	-2.863	-3.422*	-3.282*	-0.152	-2.100*	-3.208*	-4.065**	-2.920	0.621	-3.218*					
IBOV	-1.082	-0.432	-1.595*	3.150**	-2.694*	-1.589	-0.675	-1.328	4.569**	-2.144*					
SX5E	-1.332	-0.931	0.715	7.263*	9.011**	-1.986	-2.220*	-1.145	0.208	0.919					
FTSE	-4.841**	-4.502*	-5.743*	-0.782	2.862	-5.131	-4.790	-6.788**	-2.963*	-0.969					
CAC	-1.989	-1.736	-0.542	5.779**	3.648*	-2.540	-2.528	-2.277*	-1.199	-2.188*					
DAX	-1.598	-1.379	1.021	8.009**	14.825**	-1.595	-1.885	-0.325	2.826	7.885*					
IBEX	-0.226	2.146	12.511**	10.403**	16.613**	-0.144	-0.458	10.189**	2.212	6.427*					
XMB	-0.496	1.173	6.169**	11.313	15.719**	-0.698	-1.462	4.111*	0.094	4.584					
AEX	-1.412	-0.976	-0.045	2.331*	2.606*	-1.888	-1.416	-0.937	0.069	1.213					
OMX	0.321	0.774	0.051	1.590	2.070*	-0.853	-0.759	-1.829	-0.675	0.361					
SMI	-2.240*	-1.019	2.434	12.767**	18.582**	-3.407*	-2.135	-0.935	7.159*	11.435**					
NKY	5.428**	11.334**	24.105**	29.714**	32.118**	3.223	5.662*	16.335**	17.922**	5.372*					
HSI	-2.269	-2.723*	-2.775	-2.880*	-2.219	-3.292	-3.612	-2.961	-1.432	-4.162					
SPASX	-3.887	-3.343*	-4.117*	-4.257	7.481**	-4.598	-3.745	-5.944**	-5.149	3.999					
Overall	-4.821*	22.572**	84.447**	162.171**	259.540**	-13.460*	8.514*	59.845**	123.523**	204.358**					

CSV stands for common stochastic volatility. Mean squared errors are measured relative to random walk forecasts. 'Average' refers to the root mean square error averaged over all equity indices considered. Overall refers to the LPS based on the full predictive density approximated through the quadratic approximation put forward in Adolfson *et al.* (2007). Values below unity indicate outperformance relative to the random walk. The figures in the lower panel refer to the difference in log predictive scores relative to the random walk LPS. Values above zero indicate outperformance relative to the random walk. Significance levels for the root mean square errors and the log predictive scores are based on the Diebold-Mariano test and the Amisano-Giacomini test, respectively.

\* indicates significance at the 10% level

\*\* indicates significance at the 5% level.

Giacomini, 2007) with two asterisks marking the five percent significance level and one asterisk signaling significance at the ten percent level.

Table 3 reveals that in the BVAR setting without common stochastic volatility, the random walk proves to be a rather tough benchmark. The table shows that the BVAR barely outperforms the random walk for any of the stock indices in the sample. This holds true for all the forecasting horizons considered. Relative to the random walk, the average performance across all indices considered reveals that the BVAR only outperforms the random walk in terms of nine- and twelve-steps-ahead forecasts in a statistically significant fashion.

The forecasting performance of the BVAR model with common stochastic volatility (BVAR-CSV), however, shows significant improvements relative to the random walk. Irrespective of the forecasting horizon, BVAR-CSV produces more accurate out-of-sample predictions for most of the equity indices under scrutiny as compared to the benchmark. On average, the table reveals that the outperformance is especially pronounced for longer horizons (nine and twelve months ahead). Forecasting European and North American equity indices produces the smallest RMSEs relative to the random walk. The results for the Dow Jones Industrial Average (DJIA) and the S&P 500 Index (SPX), for example, particularly confirm this pattern with relative RMSEs ranging from 0.71 to 0.89 and 0.74 to 0.91, respectively. While improvements for short-term predictions (one and three months ahead) appear to be modest, being insignificant for the vast majority of indices considered, longer term forecasts (greater than six months ahead) tend to be significant for most indices. Note that the overall differences in terms of RMSE are significant even at shorter forecast horizons, outperforming the random walk significantly at the ten percent levels at the three-months-ahead horizon and at the critical five percent levels for the remaining time horizons (except at the one-month-ahead horizon). This result is due to the fact that the evaluation of a the full matrix of forecast errors across all indices considered raises the power of the test.

Due to the fact that RMSEs neglect the uncertainty surrounding the point forecasts, we also focus on the log predictive score. The lower panel of Table 3 presents the results for the BVAR-CSV and the homoskedastic BVAR relative to the random walk's LPS. Numbers greater than zero indicate outperformance of the respective model whereas negative values indicate outperformance of the random walk. Note that we simulate the predictive density from a random walk model by exactly imposing the prior in the standard BVAR.

Several things are noteworthy. First, the final row of Table 3 suggests that the BVAR-CSV outperforms the random walk benchmark for time horizons greater than one-month-ahead in terms of relative overall LPS. The homoskedastic BVAR model also improves upon the random walk model, although the accuracy gains are somewhat more muted. The overall LPS figures are substantially greater than zero for forecast horizons ranging from three- to twelve-months-ahead, being statistically significant at all horizons considered. This suggests that allowing for more flexibility of the error variance-covariance matrix yields pronounced accuracy improvements in terms of the first and second moment of the corresponding predictive density.

Across equity indices and by considering the marginal predictive density provides no clear-cut evidence in favour of the CSV specification. In the short-run, random walk predictions tend to be more precise as compared to the BVAR and even the BVAR-CSV. In the case of the BVAR-CSV, this finding could reflect the fact that the volatility structure closely tracks equity market volatility in the USA and Europe. This treatment might be appropriate for the majority of indices, however, it might be inappropriate for other markets. In addition, note that marginal predictive densities ignore the rich covariance structure contained in  $\Sigma_t$ , neglecting one important advantage of a large-scale multivariate model.<sup>8</sup>

<sup>8</sup>Technically, this stems from the fact that we evaluate a set of  $M$  univariate predictive densities, implying that the relationship between different equity markets has been integrated out.

Comparing the differences between the homoskedastic BVAR and the CSV specification reveals that when the predictive density of the full model is evaluated, allowing for stochastic volatility improves the accuracy of the density forecasts by large margins. Financial time series usually exhibit significant shifts in volatility, which translates into situations where the predictive density is expected to become more dispersed in times of crisis and more concentrated in 'normal' times. Even though such behaviour also influences point predictions, the effects on density predictions are much larger. This can be seen in Table 3, where especially for the three- to six-steps-ahead density forecasts the outperformance is particularly pronounced. This leads to the conclusion that stochastic volatility exerts significant positive effects in terms of density predictions.

We have made several attempts to ensure the robustness of our findings. First, increasing the length of the verification period to include the great crisis of 2008/2009 leads to similar results. However, in terms of RMSEs the differences between both BVAR specifications and the random walk tend to disappear. This does not carry over to the log scores, where the differences tend to increase by margins up to ten percent. This is again in line with the results described above, where the inclusion of stochastic volatility leads to more reliable density forecasts.

### Performance over time

Table 3 presents the average RMSEs and the overall LPS across the whole verification sample. To gain some insights on the degree of time variation of our findings and to what extent the changes in the LPS are driven by different moments of the corresponding predictive density, Figure 2a to Figure 2e display the evolution of the cumulative log predictive score (vis-à-vis the random walk) and the cumulative (average) squared forecast errors over the verification sample.

For the one-month-ahead time horizon, LPS of both models considered tend to be smaller than the random walks' LPS. However, the LPS of both models improve upon the random walks' LPS in the first half of 2013. This roughly coincides with a general increase in volatility as shown in Figure 1. Afterwards the LPS tend to fall relative to the random walk, with the BVAR being outperformed by the BVAR-CSV and the random walk. Note that both models display a similar performance in terms of point forecasts, with the CSV specification providing slightly more precise point forecasts.

The finding that the relative LPS increases sharply when volatility increases carries over to all other forecast horizons considered. The BVAR-CSV generally outperforms all competing models for horizons greater than one-month-ahead in terms of point and density predictions. Note that up to the second half of 2012, both BVARs tend to be inferior to the random walk in terms of LPS. However, the period from the beginning of 2013 onwards marks a serious regime shift with a sharp increase in volatility. Both models that aim to estimate a full covariance matrix exploit cross-sectional information, ultimately providing superior density predictions. In addition, the single scaling factor used to adjust the variance-covariance matrix of the system leads to appropriate forecast intervals. This is also due to the fact that the CSV specification reacts to changes in volatility, making the prediction intervals wider when necessary, thus also covering observations which would be highly unlikely under the BVAR specification.<sup>9</sup>

### III.5 A simple portfolio exercise

Even though LPS allow us to unveil the ability of the BVAR to properly predict the density associated with some variable of interest, it is not possible to directly judge the ability to

<sup>9</sup>All results and codes are available upon request.

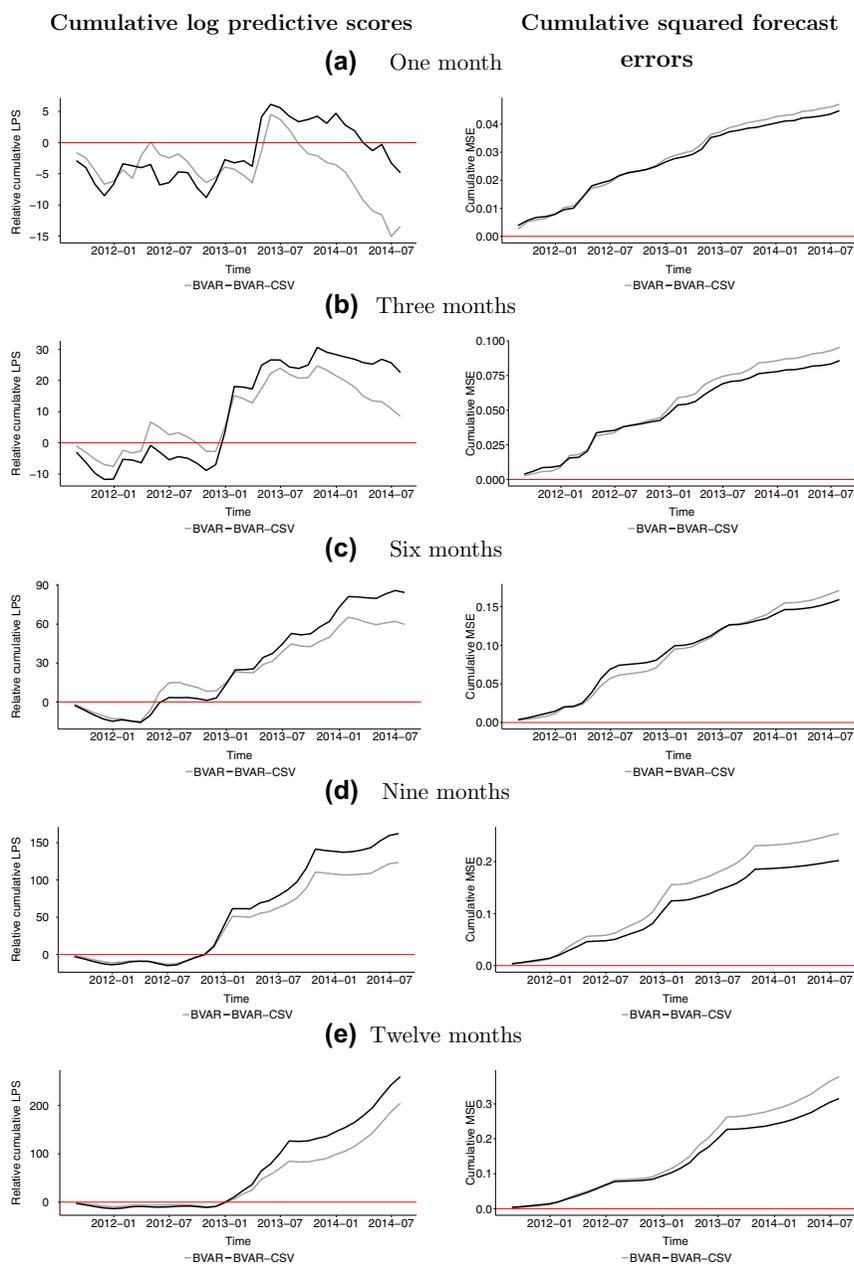


Fig. 2. Evolution of the cumulated LPS and the cumulated squared forecast errors over the verification sample.

predict the future direction of that variable. This is of key interest to practitioners in financial institutions or central banks who base their decisions on the most likely path of some financial or macroeconomic quantity. Since our goal is to show that the BVAR produces reliable directional forecasts (i.e., whether some indices go up or down), we demonstrate the performance of our approach in a simple portfolio management exercise.

Recently, several studies have emphasized the importance of judging a models' predictive capabilities by using economic measures. Carriero *et al.* (2009), for example, benchmark their BVAR using a simple trading strategy and evaluate the corresponding Sharpe ratios. They find that forecasts obtained from the BVAR generally improves Sharpe ratios as compared to simple autoregressive models.

In the spirit of the aforementioned study we use the predictions from our BVAR model to guide the investment process of a portfolio manager. First, we have to make several assumptions characterizing the investors' behaviour, which in turn allow us to formulate three simple investment strategies.

1. We assume that the investor is only allowed to enter long positions.
2. Furthermore, our investor is not allowed to borrow money, i.e., to leverage positions.
3. We assume that there are no transaction costs involved.
4. Investors are only allowed to change their positions once per time period considered (i.e., per month).
5. The investor starts with an equally weighted portfolio in  $t_0$ .
6. Finally, our investor maximizes the risk/reward ratio by picking the strategy that yields the highest expected return for a given level of volatility.

Before turning to the trading strategy, a few remarks on the employed assumptions appear necessary. While the first assumption will be relaxed in an alternative investment strategy, 2. to 4. are assumptions in order to keep the employed investment strategies simple. The equally weighted portfolio assumption in  $t_0$  reflects an agnostic stance by preventing an asymmetric portfolio allocation a priori. Conditional on the level of uncertainty (volatility), the last assumption ensures that the investor maximizes the risk/reward ratio.

Under these assumptions, we propose the following simple trading strategy. At time  $t$ , use the point forecast for the  $i$ th index,  $\bar{Y}_{i,t+1}$ , and compare it with the current outcome,  $Y_{i,t}^O$ . Compute the percentage difference denoted by  $g_{i,t+1}$ . If this difference is greater than zero (i.e., the index is expected to increase in value), we include it in our portfolio. Otherwise, if the expected change is negative (or equal to zero), we exclude the index for the given time period. Computing  $g_{i,t+1}$  for all  $i = 1, \dots, M$  allows us to calculate portfolio weights at time  $t$ ,  $w_t$ . The  $i$ th element of  $w_t$  is given by

$$w_{i,t} = \frac{g_{i,t+1}}{\sum_{j=1}^M g_{j,t+1}}. \quad (3.6)$$

Note that if  $g_{i,t+1} < 0$  we set  $w_{i,t} = 0$ . Equation (3.6) implies that if the expected percentage increase is high, we overweight that index/market in our portfolio, whereas for low/negative increases, the respective index is underweighted/excluded in our portfolio. If all expected returns happen to be negative the investor is not investing in equities, effectively selling all positions for that given point in time. Afterwards, if the expected returns of some indices turns positive again, the investor consequently re-enters the market. This strategy is labelled the *Active* investment style.

For the second strategy, we relax the first assumption above. Hitherto we have assumed that the investor is not allowed to bet on falling markets by short-selling a given equity index. We relax this assumption by assuming that the investor is also allowed to invest in markets where the forecast for  $t + 1$  is smaller than the current value at time  $t$ . Furthermore, this strategy assumes that all positions in the portfolio are equal in value both in relative as well as in absolute terms (i.e., we equally weight all indices included). This implies that the  $M$  positions included in our portfolio only differ whether they are long or short positions. This strategy is labelled the *Long/Short* investment strategy.

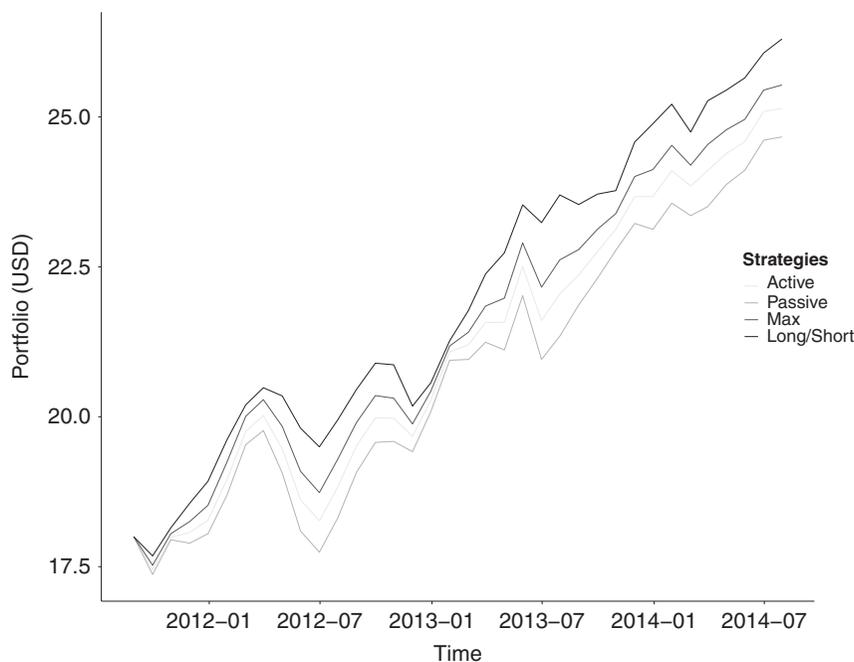


Fig. 3. Comparison of alternative investment strategies over the verification sample.

As a third benchmark strategy, we assume that the investor invests all available capital in the index, which has the highest expected profit from period  $t$  to period  $t + 1$ . This implies that the entire portfolio consists of only a single equity index. We call this strategy the *Max* investment strategy. However, it is worth noting that in the case of the highest expected profit being negative, the strategy involves short-selling the equity index with the lowest expected profit.

As the natural competitor to the aforementioned strategies, we also investigate the effects of a *Passive* investment style. This corresponds to the case where money is kept equally distributed across all equity indices over the entire investment horizon.

Figure 3 presents the posterior distribution of the evolution of our portfolio over the time period ranging from August 2011 to July 2014. All portfolios start initially with 18 USD worth of capital, spread equally across indices (i.e., at  $t_0$  one USD is invested in each equity index under scrutiny). For simplicity we moreover assume that exchange rates are mutually constant relative to the US dollar. This rules out effects related to movements in exchange rates.<sup>10</sup>

Apparently, the *Passive* investment strategy is outperformed by all other strategies considered. It is worth noting that the *Long/Short* displays the strongest performance among the strategies that exploit the BVAR-CSV model to generate trading signals, with the *Max* trading scheme ranking second. Note that while all strategies suffer severe losses in the first half of 2012, active strategies tend to profit from the fact that the weightings are adjusted rather rapidly, effectively diminishing the adverse effect of a global downturn in equity prices on the investor's portfolio. In the first half of 2013, all strategies except the *Long/Short* strategy face sharp losses, recovering quite fast afterwards.

<sup>10</sup>Relaxing this assumption would complicate the analysis, since we would also have to forecast a broad panel of exchange rates relative to the US dollar.

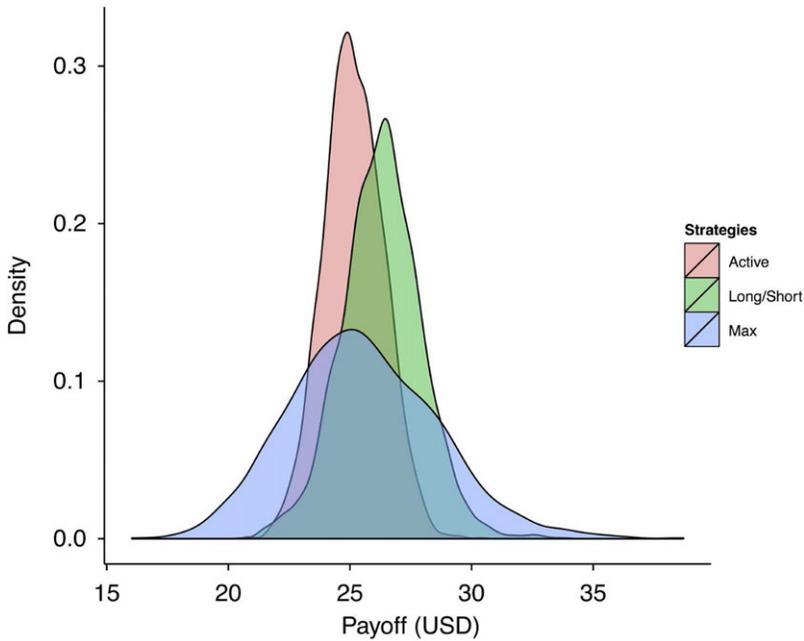


Fig. 4. Posterior distribution of the portfolio value at time  $t = 2014/07$ .

To gain some information on the dispersion of the portfolio, Figure 4 depicts the posterior distribution of the portfolio value in July 2014. Both the *Active* and *Long/Short* strategy display a similar standard deviation. However, in concordance with the finding described above, the mean of the distribution of the latter investment scheme is shifted to the right, implying a stronger return in the end. Note that the standard deviation of the *Max* strategy is remarkably high. This is simply a result of the lack of portfolio diversification. Since this strategy invests the available capital in the index that is expected to show the highest absolute change in the future, gains from diversification do not materialize, leading to a much higher variance of the portfolio.

Note, that this overly simplistic example can also be extended to allow for using leverage, i.e., use debt-financed investing. This could lead to further improvements in terms of expected returns. In addition, using a shorter trading time frame would lead both strategies to converge in terms of average returns, due to the optimistic sentiment in the stock markets.

#### IV. CONCLUSION

This paper puts forth a large dimensional BVAR model to forecast equity indices. This approach is empirically justified by the finding that a few latent factors explain the majority of the time-varying volatility of equity indices. Our model thus provides a parsimonious representation of the data, capable of incorporating several stylized facts commonly observed in the analysis of equity markets.

The performance of our approach is evaluated in an out-of-sample forecasting exercise. To effectively capture the interdependencies of the global market, forecasts are carried out on a sample of eighteen major equity indices. We compare the performance of the BVAR model, with and without stochastic volatility, to that of a naive random walk forecast. Both models tend

to perform well when the forecasting horizon is increased, outperforming the benchmark model both in terms of point and density predictions. Inspection of the evolution of cumulative log Bayes factors over time reveals that especially in times of economic stress and uncertainty, our model with stochastic volatility proves to be reliable and relatively precise.

In addition to the analysis above, the paper also presents a simple trading exercise. The BVAR-CSV model is used to efficiently allocate available capital across a portfolio of stock indices using a set of different investment styles. This exercise aims to demonstrate the ability of the BVAR-CSV to properly predict possible directions of the underlying equity indices. Most strategies considered clearly outperform a simple buy-and-hold strategy with fixed and equal capital allocation. This provides some evidence that institutional investors should incorporate models in their investment decisions that capture correlation in the conditional mean (i.e., the VAR part of our model) and in the conditional variance (captured through the stochastic volatility factor).

#### REFERENCES

- Adolfson, M., Lindé, J. and Villani, M. (2007). 'Forecasting performance of an open economy DSGE model', *Econometric Reviews*, 26, pp. 289–328.
- Amisano, G. and Giacomini, R. (2007). 'Comparing density forecasts via weighted likelihood ratio tests', *Journal of Business and Economic Statistics*, 2007, 25, pp. 177–90.
- Ang, A. and Bekaert, G. (2007). 'Stock return predictability: is it there?', *Review of Financial Studies*, 20, pp. 651–707.
- Bañbura, M., Giannone, D. and Reichlin, L. (2010). 'Large Bayesian vector auto regressions', *Journal of Applied Econometrics*, 25, pp. 71–92.
- Bauer, G. H. and Vorkink, K. (2011). 'Forecasting multivariate realized stock market volatility', *Journal of Econometrics*, 160, pp. 93–101.
- Campbell, J. Y. and Thompson, S. B. (2008). 'Predicting excess stock returns out of sample: can anything beat the historical average?', *Review of Financial Studies*, 21, pp. 1509–31.
- Carriero, A., Kapetanios, G. and Marcellino, M. (2009). 'Forecasting exchange rates with a large Bayesian VAR', *International Journal of Forecasting*, 25, pp. 400–17.
- Clark, T. E. and Marcellino, M. (2014). 'No arbitrage priors, drifting volatilities, and the term structure of interest rates', CEPR Discussion Paper No. DP9848.
- Clark, T. E. and Marcellino, M. (2015). 'Common drifting volatility in large Bayesian VARs', *Journal of Business and Economic Statistics*, 10.1080/07350015.2015.1040116.
- Chen, A.-S., Leung, M. T. and Daouk, H. (2003). 'Application of neural networks to an emerging financial market: forecasting and trading the Taiwan Stock Index', *Computers and Operations Research*, 30, pp. 901–23.
- Clark, T. E. (2011). 'Real-time density forecasts from Bayesian vector autoregressions with stochastic volatility', *Journal of Business and Economic Statistics*, 29, pp. 327–41.
- Clark, T. E. and Ravazzolo, F. (2015). 'Macroeconomic forecasting performance under alternative specifications of time-varying volatility', *Journal of Applied Econometrics*, 30, pp. 551–75.
- Cogley, T. and Sargent, T. J. (2005). 'Drifts and volatilities: Monetary policies and outcomes in the post WWII US', *Review of Economic Dynamics*, 8, pp. 262–302.
- Diebold, F. X. and Mariano, R. S. (1984). 'Comparing predictive accuracy', *Journal of Business and Economic Statistics*, 13, pp. 253–63.
- Doan, T., Litterman, R. and Sims, C. A. (1984). 'Forecasting and conditional projection using realistic prior distributions', *Econometric Reviews*, 3, pp. 1–100.
- Enke, D. and Thawornwong, S. (2005). 'The use of data mining and neural networks for forecasting stock market returns', *Expert Systems with Applications*, 29, pp. 927–40.

- French, K. R., Schwert, G. W. and Stambaugh, R. F. (1987). 'Expected stock returns and volatility', *Journal of Financial Economics*, 19, pp. 3–29.
- Frühwirth-Schnatter, S. and Wagner, H. (2010). 'Stochastic model specification search for Gaussian and partial non-Gaussian state space models', *Journal of Econometrics*, 154, pp. 85–100.
- Geweke, J. and Amisano, G. (2010). 'Comparing and evaluating Bayesian predictive distributions of asset returns', *International Journal of Forecasting*, 26, pp. 216–30.
- Giannone, D., Lenza, M. and Primiceri, G. E. (2015). 'Prior selection for vector autoregressions', *Review of Economics and Statistics*, 97, pp. 436–51.
- Hamilton, J. D. and Susmel, R. (1994). 'Autoregressive conditional heteroscedasticity and changes in regime', *Journal of Econometrics*, 64, pp. 307–33.
- Holthausen, R. W. and Larcker, D. F. (1992). 'The prediction of stock returns using financial statement information', *Journal of Accounting and Economics*, 15, pp. 373–411.
- Huber, F. (2016). 'Density forecasting using Bayesian global vector autoregressions with stochastic volatility', *International Journal of Forecasting*, 32, pp. 818–37.
- Kadiyala, K. and Karlsson, S. (1997). 'Numerical methods for estimation and inference in Bayesian VAR-models', *Journal of Applied Econometrics*, 12, pp. 99–132.
- Karlsson, S. (2013). 'Forecasting with Bayesian vector autoregressions', in Elliott, G. and Timmermann, A. (eds), *Handbook of Economic Forecasting*, Vol. 2, Amsterdam, Oxford: North-Holland, pp. 791–897.
- Kastner, G. (2016). 'Dealing with stochastic volatility in time series using the R package stochvol', *Journal of Statistical Software*, 69, pp. 1–30.
- Kastner, G. and Frühwirth-Schnatter, S. (2014). 'Ancillarity-sufficiency interweaving strategy (ASIS) for boosting MCMC estimation of stochastic volatility models', *Computational Statistics and Data Analysis*, 76, pp. 408–23.
- Kelly, B. and Pruitt, S. (2013). 'Market expectations in the cross-section of present values', *The Journal of Finance*, 68, pp. 1721–56.
- Koop, G. (2013). 'Forecasting with medium and large Bayesian VARs', *Journal of Applied Econometrics*, 28, pp. 177–203.
- Koop, G. and Korobilis, D. (2010). *Bayesian Multivariate Time Series Methods for Empirical Macroeconomics*. Boston, MA, Delft: Now Publishers Inc.
- Litterman, R. B. (1986). 'Forecasting with Bayesian vector autoregressions—Five years of experience', *Journal of Business and Economic Statistics*, 4, pp. 25–38.
- Norges Bank Investment Management (2014). 'Annual report: government pension fund global', Technical Report, Norges Bank Investment Management, Oslo. Available at: <http://www.nbim.no/contentassets/0ff9cd1d5c8e4737a7b7262d3ec167d4/norges-24bank-investment-management-annual-report-2014.pdf> (accessed 22 November 2015).
- Ou, J. A. and Penman, S. H. (1989). 'Financial statement analysis and the prediction of stock returns', *Journal of Accounting and Economics*, 11, pp. 295–329.
- Pesaran, M. H. and Timmermann, A. (1995). 'Predictability of stock returns: robustness and economic significance', *The Journal of Finance*, 50, pp. 1201–28.
- Primiceri, G. E. (2005). 'Time varying structural vector autoregressions and monetary policy', *The Review of Economic Studies*, 72, pp. 821–52.
- Rapach, D. E., Strauss, J. K. and Zhou, G. (2010). 'Out-of-sample equity premium prediction: combination forecasts and links to the real economy', *Review of Financial Studies*, 23, pp. 821–62.
- Sims, C. A. and Zha, T. (1998). 'Bayesian methods for dynamic multivariate Models', *International Economic Review*, 39, pp. 949–68.