

Forecasting Global Equity Indices Using Large Bayesian VARs

Huber, Florian; Krisztin, Tamas; Piribauer, Philipp

DOI:

[10.57938/ae02591a-21de-4bee-a8e1-bbf15ea86ec6](https://doi.org/10.57938/ae02591a-21de-4bee-a8e1-bbf15ea86ec6)

Published: 01/10/2014

Document Version

Publisher's PDF, also known as Version of record

[Link to publication](#)

Citation for published version (APA):

Huber, F., Krisztin, T., & Piribauer, P. (2014). *Forecasting Global Equity Indices Using Large Bayesian VARs*. Department of Economics Working Paper Series No. 184 <https://doi.org/10.57938/ae02591a-21de-4bee-a8e1-bbf15ea86ec6>

Department of Economics
Working Paper No. 184

Forecasting Global Equity Indices Using Large Bayesian VARs

Florian Huber
Tamás Krisztin
Philipp Piribauer

October 2014



Forecasting Global Equity Indices using Large Bayesian VARs

Florian Huber¹, Tamás Krisztin¹, and Philipp Piribauer ^{*1}

¹Vienna University of Economics and Business (WU)

September 5, 2014

Abstract

This paper proposes a large Bayesian Vector Autoregressive (BVAR) model with common stochastic volatility to forecast global equity indices. Using a dataset consisting of monthly data on global stock indices the BVAR model inherently incorporates co-movements in the stock markets. The time-varying specification of the covariance structure moreover 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 root mean squared errors as well as Bayesian log predictive scores. The BVAR model without stochastic volatility, on the other hand, underperforms relative to the random walk. In a portfolio allocation exercise we moreover show that it is possible to use the forecasts obtained from our BVAR 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.

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

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

*Correspondence to: Florian Huber, Vienna University of Economics and Business, Phone: +43 1 31336 5089 . E-mail: fhuber@wu.ac.at

1 Introduction

In recent decades, financial integration across economies led to increased co-movement between various asset prices. Especially for equities, this effect increased over the last ten years, where the global financial crisis 2008/2009 illustrated how correlations between different markets change over time. This regularity, which has received considerable attention in the academic literature on the dynamics of stock prices, has not yet been exploited in the forecasting literature.

A large number of contributions tried to forecast stock prices and volatilities using either atheoretical techniques, which tend to exploit information on the recent behaviour of stock prices or 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 which exclusively focus on the statistical characteristics of the time series involved have been slightly more successful, where the forecasters loss function have 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 \(1995\)](#). All these contributions conclude that empirical models with theoretical foundations are not able to exhibit more precise point forecasts as simple random walks. Thus again providing evidence for the unpredictability of stock prices.

Apart from focusing on the price or return of a given stock, a further important strand of the literature deals exclusively with the predictability of stock volatility. Prominent examples have been [French et al. \(1987\)](#); [Hamilton and Susmel \(1994\)](#) and [Bauer and Vorkink \(2011\)](#). These papers outline the great importance of accounting for non-linearities in the underlying data generating process.

Forecasting with Bayesian methods has a long standing tradition in macroeconomics. Recently, focus has shifted on high-dimensional models which explicitly allow for time-varying coefficients and stochastic volatility ([Cogley and Sargent, 2005](#); [Primiceri, 2005](#); [Clark, 2011](#); [Carriero et al., 2012](#)). Allowing for drifting parameters usually increases the precision of point- and density forecasts, but more importantly, improves the overall calibration of the model. Time-varying parameters, however, considerably increase the computational burden, rendering the usage of high-dimensional models effectively infeasible.

To alleviate such problems we focus on a relatively simple form of a constant parameter vector autoregressive model (VAR) with stochastic volatility. In fact, we follow [Carriero et al. \(2012\)](#) and allow the volatility of the system to be driven by a single latent stochastic process. 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 (especially at moderate- to high frequency time domains) ([Clark, 2011](#); [Carriero et al., 2012](#); [2014](#)). This so-called volatility clustering needs to be properly incorporated in the modeling framework to achieve proper calibration of the model involved.

This paper aims to contribute to the literature along several important dimensions. First, we use a large Bayesian vector autoregressive model (BVAR) in the spirit of [Bańbura et al. \(2010\)](#) to forecast a panel of well-known equity indices. This approach is purely statistical, only exploiting the past dynamics and co-movements between the included variables. Second, we augment the BVAR to allow for a simplified version of stochastic volatility in the errors. Especially for financial data, allowing for stochastic volatility might lead to large gains in terms of forecasting accuracy. This directly leads to the final contribution, where in addition to the usual analysis based exclusively on point forecasts, we also investigate the predictive densities by means of log predictive scores. In a simple portfolio allocation application we moreover utilize the BVAR model with common stochastic volatility as an investment strategy. We show that our BVAR model yields significantly larger returns as compared to naive trading schemes.

The remainder of the paper is structured as follows: Section 2 introduces the BVAR model with common stochastic volatility and discusses the prior setup employed. The design and evaluation of the forecasting application is outlined in Section 3. Furthermore, Section 3 presents a simple portfolio allocation exercise to illustrate the performance of the BVAR model as a trading scheme. The final section concludes.

2 The Econometric Framework

The following 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 on the Markov chain Monte Carlo (MCMC) algorithm.

2.1 Bayesian VARs

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 Σ_t is time-varying, and depends on the $M \times M$ variance-covariance matrix Σ

$$\Sigma_t = \exp(h_t/2)\Sigma \quad (2.3)$$

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

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

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

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. This specification is justified on the ground that the first principal component of our dataset explains the majority of variation observed. 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 Σ 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

$$Y_t = B'X_t + u_t \tag{2.6}$$

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 (2.6) and transposing leads to

$$Y = XB + u \tag{2.7}$$

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

$$\tilde{X}_t = \exp(-h_t/2)X_t \text{ and } \tilde{Y}_t = \exp(-h_t/2)Y_t \tag{2.8}$$

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

2.2 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. This is predicated by the fact that stock prices tend to follow random walks, which is typically a tough benchmark.

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

$$\text{vec}(B)|\Sigma^{-1}, \theta \sim \mathcal{N}(\text{vec}(\underline{B}), \Sigma \otimes \underline{V}_B) \quad (2.9)$$

$$\theta \sim \mathcal{G}(a_0, b_0) \quad (2.10)$$

$$\Sigma^{-1} \sim \mathcal{W}(\underline{S}, \underline{v}) \quad (2.11)$$

$$\gamma \sim \mathcal{N}(\underline{\mu}_\gamma, \underline{V}_\gamma) \quad (2.12)$$

$$\frac{\phi + 1}{2} \sim \mathcal{B}(a_1, b_1) \quad (2.13)$$

$$\sigma \sim \mathcal{G}(1/2, 1/2B_\sigma) \quad (2.14)$$

where equation (2.9) 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 Σ , 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.¹ Moreover, note that we also condition on a hyperparameter θ , which controls the tightness of the prior. Following [Giannone et al. \(2012\)](#), we impose a Gamma prior with parameters a_0 and b_0 on θ . Thus, θ is treated as an unknown quantity to be estimated jointly with the system described in (2.1). Note that a natural conjugate prior can always 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 equal

$$\underline{V}_B = (\underline{X}'\underline{X})^{-1} \quad (2.15)$$

$$\underline{B} = \underline{V}_B \underline{X}'\underline{Y} \quad (2.16)$$

respectively. For the prior on Σ^{-1} , the prior scale matrix \underline{S} is then simply

$$\underline{S} = (\underline{Y} - \underline{X}\underline{B})'(\underline{Y} - \underline{X}\underline{B}). \quad (2.17)$$

¹This result holds true as long as we condition on h_t and θ .

Moreover, \underline{v} denotes the prior degrees of freedom. Following [Bańbura et al. \(2010\)](#) and [Koop \(2013\)](#), the general form of \underline{Y} and \underline{X} are, respectively

$$\underline{Y} = \begin{pmatrix} V_B^{-\frac{1}{2}} B \\ \underline{S}^{\frac{1}{2}} \end{pmatrix}, \underline{X} = \begin{pmatrix} V_B^{-\frac{1}{2}} \\ 0_{M \times MK} \end{pmatrix}, \quad (2.18)$$

where $\underline{V} = (V^{\frac{1}{2}})' V^{\frac{1}{2}}$ and $\underline{S} = (S^{\frac{1}{2}})' S^{\frac{1}{2}}$.

Finally, equations (2.12) - (2.14) constitute the prior distributions for the log-volatility equation in (2.4). We follow [Kastner and Frühwirth-Schnatter \(2014\)](#) and impose a Normal prior with mean $\underline{\mu}_\gamma$ and variance \underline{V}_γ on the level of the log-volatility process. For the autoregressive parameter we impose a Beta prior and for σ we use a non-conjugate Gamma prior. These choices are motivated in [Kastner and Frühwirth-Schnatter \(2014\)](#) and [Frühwirth-Schnatter and Wagner \(2010\)](#). Note that the density for ϕ is given by

$$p(\phi) = \frac{1}{2B(a_1, b_1)} \frac{(1 + \phi)^{(a_1-1)}}{2} \frac{(1 - \phi)^{(b_1-1)}}{2} \quad (2.19)$$

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 $(-1, 1)$. Note that the mean of a Beta distribution is given by

$$\mathbb{E}(\phi) = \frac{2a_1}{a_1 + b_1} - 1. \quad (2.20)$$

Thus, 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.

2.3 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 Σ^{-1} , which facilitate a simple Gibbs sampling scheme.

Under the prior assumptions (2.9) - (2.14), the conditional posterior for B is given by

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

with

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

$$\bar{B} = \bar{V}_B\bar{X}'\bar{Y} \quad (2.23)$$

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, 2012](#)).

The conditional posterior of Σ^{-1} is of Wishart form, implying that

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

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

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.

2.4 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 such that coefficients on higher lag orders are shrunk aggressively towards zero. More specifically, the following dummy observations are used to match the Minnesota moments

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

where $J_p = \text{diag}(1, \dots, p)$ and $\underline{b}_1, \dots, \underline{b}_M$ 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. (2012) we treat θ as an unknown quantity to be estimated. For the gamma prior associated with θ , we set the hyperparameters equal to $a_0 = 1, b_0 = 1$. For the log-volatility equation, we use the following set of hyperparameters for the priors. First, for the Beta prior on ϕ we set $a_1 = 5$ and $b_1 = 1.5$, resulting in a prior which puts considerable mass on high-persistence regions of ϕ . 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 γ . Finally, for σ 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 $\mathcal{N}(\text{vec}(\overline{B}), \Sigma \otimes \overline{V}_B)$
3. Conditional on B , draw Σ^{-1} from $\mathcal{W}(\overline{v}, \overline{S})$
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 θ using a random walk Metropolis step with the following proposal density $p(\theta|\mathcal{D}) \propto p(\mathcal{D}|\theta)p(\theta)$ where $p(\mathcal{D}|\theta)$ denotes the marginal likelihood

Steps 2. to 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). 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

$$p(\mathcal{D}|\theta) \propto (|\underline{V}||\overline{V}^{-1}|)^{\frac{M}{2}} |\overline{S}|^{-\frac{T+M+g-1}{2}}. \quad (2.26)$$

This completes the description of the MCMC algorithm employed.

3 Forecasting Returns of Equity Indices

3.1 Data Overview

Our dataset comprises the most important equity indices (in market capitalization) across North America, Europe, Asia and Latin America. More specifically, the dataset includes major stock 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. Including a wide sample of stock indices in the BVAR model aims to account for the prevailing co-movement between the indices under scrutiny. The development of a stock index is thus not only explained by its own historical behaviour, but also by the movements of the other indices. Data on stock movements was obtained on a daily basis from the *Yahoo! Finance* database for the time period ranging from 1st of January 1998 to 31st of July 2014. We then constructed monthly averages of the stock indices under scrutiny.

[Table 1 about here.]

3.2 Design of the forecasting exercise

We propose the following recursive forecasting exercise. In the first step, we start with an initial estimation period ranging from 1998:M01 to 2011:M07. 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

$$p(Y_{\tau+k}|\mathcal{D}_\tau) = \int_{\Xi} p(Y_{\tau+k}|\mathcal{D}_\tau, \Xi) p(\Xi|Y_{1:T}) d\Xi \quad (3.1)$$

where \mathcal{D}_τ denotes the information set up to time τ and Ξ collects all available parameters of the model. Equation (3.1) can be approximated easily 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 and the log predictive score (LPS). The RMSE is defined as

$$\text{RMSE} = \sqrt{\frac{\sum_{\tau=t_0}^T (Y_\tau^O - \bar{Y}_\tau)^2}{T}} \quad (3.2)$$

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

The LPS is a well-known Bayesian evaluation criterion, motivated recently in [Geweke and Amisano \(2010\)](#). In general, the log predictive score is just the predictive density $p(Y_{\tau+k}^O | \mathcal{D}_\tau)$ evaluated at $Y_{\tau+k}^O$. This implies that the log predictive score is given by

$$\text{LPS} = \sum_{\tau=t_0}^{T-k} p(Y_{\tau+k} = Y_{\tau+k}^O | \mathcal{D}_\tau). \quad (3.3)$$

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

$$\text{LPS}(Y_{\tau+k}^O) = -0.5 \left(M \log(2\pi) + \log |\bar{V}_{\tau+k|\tau}| + (Y_{\tau+k}^O - \bar{Y}_{\tau+k})' \bar{V}_{\tau+k|\tau}^{-1} (Y_{\tau+k}^O - \bar{Y}_{\tau+k}) \right) \quad (3.4)$$

where $\bar{V}_{\tau+k|\tau}$ and $\bar{Y}_{\tau+k}$ denote the posterior variance and mean of the predictive density, respectively.

3.3 Evaluation of point forecasts

This section provides details on the evaluation of out-of-sample point forecasts of the equity indices in our dataset. [Table 2](#) summarizes the results of the forecasting exercise for different forecasting horizons and model settings. 'BVAR-CSV' and 'BVAR' denote the BVAR settings with and without common stochastic volatility, respectively. In both cases out-of-sample forecast horizons ranging from one to twelve months ahead are re-

ported. The RMSEs presented in the table are moreover reported relative to the RMSEs of random walk forecasts. This means that values below unity indicate outperformance relative to the benchmark, whereas values above unity indicate underperformance.

Table 2 reveals that in the BVAR setting without common stochastic volatility, the random walk proves to be a rather tough benchmark. The table shows that 'BVAR' barely outperforms the random walk in any of the stock indices in the sample. This holds true for all the forecasting horizons considered. Relative to the random walk, average out-of-sample performance seems to work best in short-term (one month ahead) and long-term (twelve months ahead).

The forecasting performance of the BVAR model with common stochastic volatility ('BVAR-CSV'), however, shows significant improvements relative to the random walk. Irrespectively 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 striking 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, are particularly striking with relative RMSEs ranging from 0.76 to 0.88 and 0.65 and 0.90, respectively.

[Table 2 about here.]

3.4 Evaluation of density forecasts

Due to the fact that RMSEs fully disregard the uncertainty surrounding the point forecasts, we also focus on the log predictive score.

[Table 3 about here.]

Table 3 presents the results for the BVAR-CSV and the homoskedastic BVAR relative to the random walks LPS. Numbers greater than zero indicate outperformance of the respective model whereas numbers small than zero 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 worth noting. First, the BVAR-CSV outperforms the random walk benchmark at all time horizons considered. Note that the average difference in log scores is substantially greater than zero for all time horizons considered, as can be

seen in the last row of [Table 3](#). Furthermore, we find that especially for short-term forecasts, the premium in terms of log scores provided by the CSV specification is large, falling with the forecasting horizon. This corroborates the findings in [Carriero et al. \(2014\)](#), where the fall in the accuracy premium is explained by the fact that for higher forecasting horizons, the predicted volatilities converge towards their long-run mean. This implies that the differences in conditional volatilities between the two BVAR specifications vanishes.

Note that across indices, the only difference in log scores which is always negative is the one associated with the OMX index. Other indices like COMPX display negative differences for three to nine month time horizons. Thus as can be seen from [Table 3](#), even though on average the BVAR-CSV performs quite well in this density forecasting exercise, there seem to be some interesting differences between the individual stock indices. These differences could be attributed to the fact that whereas the factor structure of our log-volatility might be appropriate for the majority of indices included, some markets like Norway tend to experience a broadly different pattern of estimated volatilities.

Comparing the differences between the homoskedastic BVAR and the CSV specification reveals that on average, allowing for stochastic volatility improves the accuracy of the density forecasts by large margins. The last subsection outlined the premium in terms of point forecasts relative to random walks and the standard BVAR, however, those differences were rather small. However, financial time series usually exhibit significant changes 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 is much larger. This can be seen in [Table 3](#), where especially for the three- to six steps ahead density forecasts the outperformance is particularly large, leading to the conclusion that for the dataset employed, stochastic volatility exhibits significant positive effects in terms of density predictions.

Furthermore, allowing for stochastic volatility also robustifies the analysis with respect to large economic shocks. Inspection of the dynamics of the log scores over time reveals that especially in downturns, the LPS of the CSV stays relatively robust as compared to the log scores of the random walk and the BVAR. 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.

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 10%. This is, again, in line with the results described above, where the inclusion of stochastic volatility leads to more reliable density forecasts.

3.5 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 predict the future direction of that variable. This is of key interest to practitioners in financial institutions or central banks which 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 index goes up or down), we propose a simple portfolio management exercise. Recently, several studies emphasized the importance of judging a models' predictive capabilities by using economic measures. [Carriero et al. \(2009\)](#) benchmark their BVAR using a simple trading strategy and evaluate the corresponding Sharpe ratios. They find that, as compared to simple autoregressive models, using the forecasts obtained from the BVAR generally improves Sharpe ratios. [Costantini et al. \(2014\)](#) show in another contribution that it is possible to use forecasts obtained from several macroeconomic models to guide investment decisions. They conclude that it is possible to improve upon several benchmark strategies using combinations of forecasts from different models.

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

1. We assume that the investor is allowed to only enter long positions
2. Furthermore, our investor is not allowed to borrow money, i.e. to leverage positions
3. The investor starts with an equally weighted portfolio in t_0

4. We assume that there are no transaction costs involved
5. Investors are only allowed to change their positions once per time period considered (i.e. month)
6. Finally, our investor is risk-neutral and maximizes expected profits

Under assumptions 1. to 3. 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, we exclude the index for that 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

$$w_{i,t} = \frac{g_{i,t+1}}{\sum_{j=1}^M g_{j,t+1}}. \quad (3.5)$$

Note that if $g_{i,t+1} < 0$ we set $w_{i,t} = 0$. Equation (3.5) 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. This strategy is labeled the *Active* investment style.

As a second 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 portfolio consists of a single equity index. We call this strategy the *Max* investment strategy.

For the third strategy, we have to relax assumption (1) above. Hitherto we 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 size, 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 labeled the *Long/Short* investment strategy.

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

equally distributed across all equity indices and those shares are held constant over time.

[Figure 1 about here.]

Figure 1 presents the evolution of our portfolio over the time period ranging from September 2008 to July 2014. All portfolios start initially with 18 USD worth of capital, spread equally across indices. Especially in the last quarter of 2008 and the first quarter of 2009, all portfolios considered experience heavy losses. This is partly due to the fact that correlation between equity indices increased in that given time period, rendering portfolio diversification ineffective. In the second quarter of 2009, however, most actively managed portfolio appreciated sharply, whereas the passive investment strategy reacts slower overall. Especially the *Max* investment style exhibits a strong performance during the financial downturn in 2008/09, recovering quite fast. However, evaluation of the portfolio value in July 2014 reveals that the *Max* strategy is the only investment style considered which failed to fully profit from the sharp increase in equity prices observed recently. Another interesting result is the dismal performance of the *Long/Short* investing style. Here it can be seen that, even though the portfolio recovers the losses suffered in 2008 marginally faster than the naive portfolio, the overall value of the portfolio under the *Long/Short*-strategy is the lowest among all competing strategies. All other strategies considered managed to improve upon the simple equal weighting strategy.

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.

4 Conclusion

This paper puts forth a large dimensional BVAR model to forecast equity indices. This approach improves the precision of point and density forecasts, by allowing for stochastic volatility and taking into account the recent dynamics and co-movements of the included equity indices.

The performance of our approach is evaluated on an out-of-sample forecasting exercise. To effectively capture the interdependencies of the global market, the forecast

is 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. The BVAR without stochastic volatility hardly outperforms the random walk. In contrast, the BVAR with common stochastic volatility produces significantly better out-of-sample forecasts – especially for horizons greater than nine months – as compared to the no-change forecast. These results are further validated by the comparison of log predictive scores. Examining the time-related changes of log scores reveals that not only does the BVAR with CSV significantly improve the density of the forecasts (as compared to the homoskedastic BVAR and the random walk), but also provides more robust forecasts with respect to large economic shocks.

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 three 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.

References

- Adolfson, Malin, Jesper Lindé, and Mattias Villani**, “Forecasting performance of an open economy DSGE model,” *Econometric Reviews*, 2007, *26* (2-4), 289–328.
- Bañbura, Marta, Domenico Giannone, and Lucrezia Reichlin**, “Large Bayesian vector auto regressions,” *Journal of Applied Econometrics*, 2010, *25* (1), 71–92.
- Bauer, Gregory H and Keith Vorkink**, “Forecasting multivariate realized stock market volatility,” *Journal of Econometrics*, 2011, *160* (1), 93–101.
- Carriero, Andrea, George Kapetanios, and Massimiliano Marcellino**, “Forecasting exchange rates with a large Bayesian VAR,” *International Journal of Forecasting*, 2009, *25* (2), 400–417.
- , **Todd Clark, and Massimiliano Marcellino**, “Common drifting volatility in large Bayesian VARs,” 2012.
- , **Todd E Clark, and Massimiliano Giuseppe Marcellino**, “No Arbitrage Priors, Drifting Volatilities, and the Term Structure of Interest Rates,” 2014.
- Chen, An-Sing, Mark T Leung, and Hazem Daouk**, “Application of neural networks to an emerging financial market: forecasting and trading the Taiwan Stock Index,” *Computers & Operations Research*, 2003, *30* (6), 901–923.
- Clark, Todd E**, “Real-time density forecasts from Bayesian vector autoregressions with stochastic volatility,” *Journal of Business & Economic Statistics*, 2011, *29* (3).
- Cogley, Timothy and Thomas J Sargent**, “Drifts and volatilities: monetary policies and outcomes in the post WWII US,” *Review of Economic dynamics*, 2005, *8* (2), 262–302.
- Costantini, Mauro, Jesus Crespo Cuaresma, and Jaroslava Hlouskova**, “Can Macroeconomists Get Rich Forecasting Exchange Rates?,” 2014.
- Doan, Thomas, Robert Litterman, and Christopher Sims**, “Forecasting and conditional projection using realistic prior distributions,” *Econometric reviews*, 1984, *3* (1), 1–100.
- Enke, David and Suraphan Thawornwong**, “The use of data mining and neural networks for forecasting stock market returns,” *Expert Systems with applications*, 2005, *29* (4), 927–940.
- French, Kenneth R, G William Schwert, and Robert F Stambaugh**, “Expected stock returns and volatility,” *Journal of financial Economics*, 1987, *19* (1), 3–29.
- Frühwirth-Schnatter, Sylvia and Helga Wagner**, “Stochastic model specification search for Gaussian and partial non-Gaussian state space models,” *Journal of Econometrics*, 2010, *154* (1), 85–100.
- Geweke, John and Gianni Amisano**, “Comparing and evaluating Bayesian predictive distributions of asset returns,” *International Journal of Forecasting*, 2010, *26* (2), 216–230.
- Giannone, Domenico, Michele Lenza, and Giorgio E Primiceri**, “Prior selection for vector autoregressions,” Technical Report, National Bureau of Economic Research 2012.
- Hamilton, James D and Raul Susmel**, “Autoregressive conditional heteroskedas-

- ticity and changes in regime,” *Journal of Econometrics*, 1994, *64* (1), 307–333.
- Holthausen, Robert W and David F Larcker**, “The prediction of stock returns using financial statement information,” *Journal of Accounting and Economics*, 1992, *15* (2), 373–411.
- Kadiyala, K and Sune Karlsson**, “Numerical Methods for Estimation and Inference in Bayesian VAR-Models,” *Journal of Applied Econometrics*, 1997, *12* (2), 99–132.
- Karlsson, Sune**, “Forecasting with Bayesian vector autoregressions,” *prepared for the Handbook of Economic Forecasting*, 2012, *2*.
- Kastner, Gregor and Sylvia Frühwirth-Schnatter**, “Ancillarity-sufficiency interweaving strategy (ASIS) for boosting MCMC estimation of stochastic volatility models,” *Computational Statistics & Data Analysis*, 2014, *76*, 408–423.
- Koop, Gary and Dimitris Korobilis**, *Bayesian multivariate time series methods for empirical macroeconomics*, Now Publishers Inc, 2010.
- Koop, Gary M**, “Forecasting with medium and large Bayesian VARs,” *Journal of Applied Econometrics*, 2013, *28* (2), 177–203.
- Litterman, Robert B**, “Forecasting with Bayesian vector autoregressions five years of experience,” *Journal of Business & Economic Statistics*, 1986, *4* (1), 25–38.
- Ou, Jane A and Stephen H Penman**, “Financial statement analysis and the prediction of stock returns,” *Journal of accounting and economics*, 1989, *11* (4), 295–329.
- Pesaran, M Hashem and Allan Timmermann**, “Predictability of stock returns: Robustness and economic significance,” *The Journal of Finance*, 1995, *50* (4), 1201–1228.
- Primiceri, Giorgio E**, “Time varying structural vector autoregressions and monetary policy,” *The Review of Economic Studies*, 2005, *72* (3), 821–852.
- Sims, Christopher A and Tao Zha**, “Bayesian methods for dynamic multivariate models,” *International Economic Review*, 1998, pp. 949–968.

Table 1: Stock indices used in the BVAR model

Region	Name	Description
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

Table 2: Evaluation of point forecasts

Index	BVAR-CSV												BVAR											
	months						months						months						months					
	1	3	6	9	12		1	3	6	9	12		1	3	6	9	12		1	3	6	9	12	
DJIA	0.8777	0.7765	0.7558	0.8195	0.8617		0.9836	1.1995	0.9055	1.0505	1.0475		0.9836	1.1995	0.9055	1.0505	1.0475		0.9836	1.1995	0.9055	1.0505	1.0475	
SPX	0.8506	0.6478	0.8251	0.8626	0.8952		1.0417	1.1013	1.0847	1.1451	1.0535		1.0417	1.1013	1.0847	1.1451	1.0535		1.0417	1.1013	1.0847	1.1451	1.0535	
COMPX	0.9113	0.7218	0.8433	0.8507	0.8648		0.9303	1.0762	0.8439	0.9426	0.8261		0.9303	1.0762	0.8439	0.9426	0.8261		0.9303	1.0762	0.8439	0.9426	0.8261	
OSPTX	1.0254	1.0487	1.0755	0.9687	0.9908		1.1709	1.2979	1.1554	1.2442	1.0620		1.1709	1.2979	1.1554	1.2442	1.0620		1.1709	1.2979	1.1554	1.2442	1.0620	
MEXBO	1.1517	1.2522	1.0406	0.9877	0.7956		1.2205	1.5330	1.5413	1.4329	1.4756		1.2205	1.5330	1.5413	1.4329	1.4756		1.2205	1.5330	1.5413	1.4329	1.4756	
IBOV	1.1065	1.2324	1.1192	1.0815	0.9977		1.1199	1.2788	1.2350	1.4183	1.2623		1.1199	1.2788	1.2350	1.4183	1.2623		1.1199	1.2788	1.2350	1.4183	1.2623	
SX5E	0.9764	0.8952	0.9727	0.8642	0.8750		0.9994	1.2273	1.0495	1.2493	1.0302		0.9994	1.2273	1.0495	1.2493	1.0302		0.9994	1.2273	1.0495	1.2493	1.0302	
FTSE	0.9593	0.9382	0.9398	0.8367	0.7832		1.0432	1.2307	1.0795	1.3648	1.0592		1.0432	1.2307	1.0795	1.3648	1.0592		1.0432	1.2307	1.0795	1.3648	1.0592	
CAC	0.9959	1.0008	1.0086	0.9391	0.9105		1.0488	1.3654	0.9270	1.2484	0.9558		1.0488	1.3654	0.9270	1.2484	0.9558		1.0488	1.3654	0.9270	1.2484	0.9558	
DAX	0.9547	0.9180	0.8791	0.8395	0.8250		0.9357	1.1236	0.9786	1.1927	1.0192		0.9357	1.1236	0.9786	1.1927	1.0192		0.9357	1.1236	0.9786	1.1927	1.0192	
IBEX	0.9577	0.9793	1.0240	0.9155	0.9949		0.9895	1.0824	1.0451	1.1771	0.9745		0.9895	1.0824	1.0451	1.1771	0.9745		0.9895	1.0824	1.0451	1.1771	0.9745	
XMB	0.9828	0.9564	0.9472	0.8832	0.9398		0.9724	1.2147	0.9204	1.2358	0.9099		0.9724	1.2147	0.9204	1.2358	0.9099		0.9724	1.2147	0.9204	1.2358	0.9099	
AEX	0.9560	0.8956	0.8710	0.8261	0.8889		1.1410	1.6058	1.2145	1.3631	1.2696		1.1410	1.6058	1.2145	1.3631	1.2696		1.1410	1.6058	1.2145	1.3631	1.2696	
OMX	0.8882	0.7734	1.0985	0.7095	0.8209		1.1120	1.2712	1.6504	0.7557	0.7889		1.1120	1.2712	1.6504	0.7557	0.7889		1.1120	1.2712	1.6504	0.7557	0.7889	
SMI	0.9167	0.8645	0.9317	0.9028	0.9173		1.2447	1.4964	1.5048	1.4368	1.2329		1.2447	1.4964	1.5048	1.4368	1.2329		1.2447	1.4964	1.5048	1.4368	1.2329	
NKY	0.9700	0.9255	0.9240	0.9153	0.9679		1.0571	1.1224	1.0381	1.1826	1.0467		1.0571	1.1224	1.0381	1.1826	1.0467		1.0571	1.1224	1.0381	1.1826	1.0467	
HSI	1.0104	1.1777	1.1381	1.1047	1.0999		1.1176	1.3923	1.1186	1.2276	1.1147		1.1176	1.3923	1.1186	1.2276	1.1147		1.1176	1.3923	1.1186	1.2276	1.1147	
SPASX	0.9172	1.1187	0.9011	0.9820	0.8951		1.0085	1.3715	1.0279	1.5693	1.0556		1.0085	1.3715	1.0279	1.5693	1.0556		1.0085	1.3715	1.0279	1.5693	1.0556	
Average	0.9671	0.9513	0.9609	0.9050	0.9069		1.0632	1.2772	1.1289	1.2354	1.0658		1.0632	1.2772	1.1289	1.2354	1.0658		1.0632	1.2772	1.1289	1.2354	1.0658	

CSV stands for common stochastic volatility. Mean squared errors are measured relative to random walk forecasts. Values below unity indicate outperformance relative to the random walk.

Table 3: Evaluation of density forecasts

Index	BVAR-CSV												BVAR																																																																																																																																																																																	
	months												months																																																																																																																																																																																	
	1	3	6	9	12	1	3	6	9	12	1	3	6	9	12	1	3	6	9	12																																																																																																																																																																										
DJIA	0.4366	0.3025	0.2964	0.2216	0.1922	-0.0575	-0.1933	-0.1419	-0.1740	-0.1382	0.4098	0.2707	0.2215	0.2073	0.1258	-0.0838	-0.2244	-0.2214	-0.2074	-0.1324	0.1508	-0.0273	-0.0474	-0.0064	0.0281	-0.0017	-0.2221	-0.1853	-0.1408	-0.0276	0.4053	0.2272	0.1882	0.1813	0.1072	-0.0389	-0.2122	-0.1695	-0.2092	-0.1574	0.1294	-0.0286	-0.0712	-0.0641	-0.0201	-0.0412	-0.1990	-0.2645	-0.2630	-0.2362	0.0085	-0.1488	-0.1607	-0.1859	-0.1557	-0.0262	-0.1843	-0.1969	-0.2563	-0.1969	0.2094	0.0792	0.0468	0.0760	0.0135	-0.0190	-0.2123	-0.1749	-0.2235	-0.1556	0.4234	0.2561	0.2454	0.2556	0.2258	-0.0615	-0.2610	-0.2524	-0.2600	-0.2307	0.2303	0.0907	0.0827	0.0643	0.0480	-0.0405	-0.2264	-0.1579	-0.2186	-0.1634	0.1792	0.0381	0.0430	0.0869	0.0989	-0.0199	-0.1339	-0.1383	-0.1557	-0.0589	0.1545	0.0333	0.0098	0.0769	0.0233	-0.0253	-0.1713	-0.1006	-0.1564	0.0301	0.0830	-0.0278	-0.0165	0.0141	-0.0114	-0.0207	-0.1733	-0.0717	-0.1860	0.0040	0.0169	-0.1252	-0.1573	-0.1112	-0.1746	-0.0304	-0.1746	-0.1448	-0.1970	-0.1819	-0.1405	-0.2979	-0.3533	-0.2887	-0.2922	0.0083	-0.1007	-0.1312	-0.0433	-0.0440	0.3479	0.1986	0.1734	0.1576	0.1177	-0.0517	-0.2669	-0.2661	-0.3024	-0.2181	0.1211	0.0446	0.0540	0.0588	0.1073	-0.0656	-0.1799	-0.1018	-0.1877	0.0529	0.1391	-0.0248	-0.0491	-0.0604	-0.0681	-0.0517	-0.2290	-0.1912	-0.2219	-0.1783	0.4636	0.2709	0.2958	0.2527	0.1781	-0.0585	-0.1994	-0.1887	-0.2270	-0.1943	0.2094	0.0629	0.0445	0.0520	0.0302	-0.0381	-0.1980	-0.1722	-0.2017	-0.1237

CSV stands for common stochastic volatility. Log predictive scores are measured relative to random walk forecasts. Positive values indicate outperformance relative to the random walk.



Figure 1: Comparison between Active and Passive Investment strategies