



Universidad
Carlos III de Madrid



This is a postprint version of the following published document:

Fresoli, D., Ruiz, E. y Pascual, L. (2015). Bootstrap Multi-step Forecasts of Non-Gaussian VAR Models. *International Journal of Forecasting*, v. 31, n. 3, pp. 834-84.
Available in: <http://dx.doi.org/10.1016/j.ijforecast.2014.04.001>

© Elsevier



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

Bootstrap multi-step forecasts of non-Gaussian VAR models

Diego Fresoli ^a, Esther Ruiz ^{a,b,*}, Lorenzo Pascual ^c

^a Dpt. de Estadística, Universidad Carlos III de Madrid, Spain

^b Instituto Flores de Lemus, Universidad Carlos III de Madrid, Spain

^c EDP-Energias de Portugal, S.A., Unidade de Negócio de Gestão da Energia, Portugal

ARTICLE INFO

Keywords:

Bias correction
DCC model
Forecast density
Forecast regions
High density regions
Lag order uncertainty
Multivariate forecast
Resampling methods

ABSTRACT

In this paper, we establish the asymptotic validity and analyse the finite sample performance of a simple bootstrap procedure for constructing multi-step multivariate forecast densities in the context of non-Gaussian unrestricted VAR models. This bootstrap procedure avoids the backward representation, and, as a consequence, can be used to obtain multivariate forecast densities in, for example, VARMA or VAR-GARCH models. In the context of bivariate stationary VAR(p) models, we show that its finite sample properties are comparable to those of alternatives based on the backward representation. The bootstrap procedure is also implemented in a VAR-DCC model which lacks a backward representation. Finally, joint forecast densities of US quarterly inflation, unemployment and GDP growth are obtained.

1. Introduction

Since Sims (1980), Vector Autoregressive (VAR) models have been an essential tool for policy making and forecasting in the context of macroeconomic multivariate time series; see Stock and Watson (2001) for the advantages and limitations of VAR models. In this paper, we focus on the forecasting ability of VAR models. It is well known that, in practice, VAR forecasts of large macroeconomic systems may be very imprecise because of the large number of parameters to be estimated relative to the available sample sizes. However, VARs are still very popular when forecasting small or moderate systems in which the parameters can be estimated with an acceptable level of precision. For some selected examples of useful VAR forecasts, see D'Agostino, Gambetti, and Giannone (2013)

and Marcellino, Stock, and Watson (2003) for unemployment, Batchelor, Alizadeh, and Visvikis (2007) for international freight prices, Gupta, Kabundi, and Miller (2011) for US house prices, Baumeister and Kilian (2012) for oil prices, Polito and Wickens (2012) for fiscal forecasts, and Kilian and Vigfusson (2013) for US growth.

Most of the literature dealing with VAR forecasts focuses on marginal point forecasts of each of the variables in the system. However, policy makers and forecasters are increasingly interested in metrics that require joint multivariate forecasts. Komunjer and Owyang (2012) point out the importance of recognizing the multivariate nature of most forecasting problems, which has fundamental implications for the prospects of rational expectations in macroeconomic models. Furthermore, joint multivariate forecasts are also important when forecasting future values of one variable conditional on particular values of other variables in the system; see Baumeister and Kilian (2012), Doan, Litterman, and Sims (1984), and Waggoner and Zha (1999). In order to define spillover measures, Diebold and Yilmaz (2009) also consider multivariate

* Correspondence to: Dpt. Statistics, Universidad Carlos III de Madrid, C/ Madrid 126, 28903 Getafe (Madrid), Spain. Tel.: +34 91 6249851; fax: +34 91 6249849.

E-mail address: ortega@est-econ.uc3m.es (E. Ruiz).

multi-period-ahead forecasts; see also [Klobner and Wagner \(2014\)](#). On the other hand, the focus of the forecasting literature is moving from point forecasts to density forecasts that incorporate the uncertainty about the future evolution of the variables of interest; see [Bache, Jore, Mitchell, and Vahey \(2011\)](#), [Clark \(2011\)](#), [Diebold, Hanh, and Tay \(1999\)](#), and [Jore, Mitchell, and Vahey \(2010\)](#), among others. Traditionally, a multivariate forecast density for a given horizon can be obtained by assuming Gaussian forecast errors and a known lag order and model parameters. However, it has long been recognized that the parameter uncertainty may be an important issue when dealing with VAR forecasts in practice; see [Fair and Shiller \(1990\)](#) and [Lewis and Reinsel \(1985\)](#) for early references. Furthermore, [Kilian \(1998a\)](#) points out the problems associated with assuming a known lag order when it needs to be estimated. Finally, the empirical evidence suggests that departures from Gaussianity are quite plausible when dealing with economic time series; see for example [Harvey and Newbold \(2003\)](#) and [Kilian \(1998b\)](#). These departures are a serious concern when forecasting with VAR models, calling into question traditional techniques for constructing joint multivariate forecast densities.

Forecast densities that incorporate the parameter and lag order uncertainties without relying on particular assumptions on the error distribution can be obtained using bootstrap procedures; see [Holmes, Morris, Tibshirani, and Efron \(2003\)](#) for an interview with Bradley Efron about the advantages of bootstrap procedures. In the context of forecasting stationary VAR(p) models, bootstrap methods are introduced by [Kim \(1999\)](#), who extends the original procedure proposed by [Thombs and Schucany \(1990\)](#) for univariate AR(p) processes; see [Berkowitz and Kilian \(2000\)](#) for a review of bootstrap procedures for time series. Because of the biases associated with the Least Squares (LS) estimator of the VAR parameters, [Kim \(2001, 2004\)](#) considers bias-corrected forecast regions. The bootstrap procedure proposed by [Kim \(1999\)](#) has been implemented for dealing with various different issues in the context of forecasting using multivariate VAR and periodic state-space models; see, for instance, [Grigoletto \(2005, 2012\)](#), [Guerbyenne and Hamdi \(in press\)](#) and [Staszewska-Bystrova \(2011\)](#). It uses the backward representation (BR) of the VAR model to generate the bootstrap samples used to obtain replicates of the estimated parameters. As a consequence, its asymptotic validity relies on the assumption of Gaussian errors; see [Kim \(2001\)](#). Given that one of the main attractions of bootstrap procedures is their ability to make predictions in the context of non-Gaussian VAR models, this is an important drawback. Furthermore, bootstrap procedures based on the BR can only be implemented in models with such representations, which excludes, for example, multivariate models with Moving Average (MA) components or GARCH disturbances; see [Athanasopoulos and Vahid \(2008\)](#) and [Lütkepohl \(2006\)](#) for forecasting using VARMA models and [Kavussanos and Visvikis \(2004\)](#) for an empirical example of forecasting with a cointegrated VAR-GARCH model. Alternatively, [Eklund \(2007\)](#) implements a very simple bootstrap procedure for obtaining multivariate bootstrap forecasts of several variables of the Icelandic economy that do not require the BR.

However, the bootstrap procedure implemented by [Eklund \(2007\)](#) does not incorporate the parameter uncertainty. Finally, using arguments put forward by [Pascual, Romo, and Ruiz \(2004a\)](#) in the context of univariate ARIMA models, one can implement simple bootstrap procedures that incorporate the parameter uncertainty without requiring the BR. For example, [Lütkepohl, Staszewska-Bystrova, and Winker \(2015\)](#), [Staszewska-Bystrova and Winker \(2013\)](#) and [Wolf and Wunderli \(2012\)](#) implement the bootstrap procedure originally described by [Pascual, Ruiz, and Fresoli \(2011\)](#)¹ for constructing bands for forecast paths. It is important to note that the forward bootstrap procedure implemented in these papers is closely related to the bootstrap procedure proposed by [Kilian \(1998a,b,c\)](#) for the construction of confidence bands in the context of impulse response functions.

In this paper, we provide a theoretical justification of the forward bootstrap procedure. We establish its asymptotic out-of-sample validity in the context of VAR(p) models, without relying on particular distributions of the forecast errors. Furthermore, Monte Carlo experiments are carried out to analyse its finite sample performance when it is used to construct joint forecast regions. The forward bootstrap regions are compared with traditional and backward bootstrap regions. We show that, regardless of the error distribution, if the VAR(p) model is persistent and the sample size is not very large relative to the number of parameters, the finite sample properties of the bootstrap regions are clearly better than those based on Gaussian densities. Furthermore, we show that, when the VAR(p) model is far from having a unit root and the forecast errors are truly Gaussian, the loss incurred by using bootstrapping is not large, while the improvement in coverage is moderate if the errors are non-Gaussian. In any case, the bootstrap procedures provide similar coverage accuracy levels, regardless of whether they are based on the backward representation or not.

We also illustrate the good performance of the forward bootstrap procedure by constructing forecast densities in the context of a Dynamic Conditional Correlation (DCC) model which does not have a BR. The importance of constructing forecast regions which take the non-Gaussianity of the variables into account is illustrated by using the forward bootstrap to obtain joint forecast densities of US quarterly inflation, unemployment and growth rates.

The rest of the paper is organized as follows. Section 2 focuses the discussion and establishes the notation by describing the traditional and backward bootstrap procedures used for constructing forecast densities. Both procedures are illustrated in the context of a non-Gaussian bivariate VAR(2) model. In Section 3, the asymptotic validity of the forward bootstrap procedure is established. Section 4 reports Monte Carlo results on several bivariate VAR models with different parameter configurations, including stationary, persistent and near-cointegrated models. The finite sample performances of the forward bootstrap forecast regions are compared with those of

¹ [Pascual et al. \(2011\)](#) is a previous version of the present paper by the same authors.

Gaussian and backward bootstrap procedures. Section 5 implements the forward bootstrap procedure in data simulated via a VAR-DCC model, while Section 6 illustrates the results through an empirical application. Finally, Section 7 concludes the paper with suggestions for further research.

2. Multi-step forecast densities and regions for VAR models

This section establishes our notation and briefly describes the traditional and backward bootstrap procedures for constructing forecast densities in stationary VAR models.

Consider the following multivariate VAR(p) model of finite lag order p :

$$Y_t = \mu + \Phi_1 Y_{t-1} + \dots + \Phi_p Y_{t-p} + \varepsilon_t, \quad t = -p + 1, \dots, T, \quad (1)$$

where Y_t is the $N \times 1$ vector of observations at time t , μ is a $N \times 1$ vector of constants, and Φ_i , $i = 1, \dots, p$, are $N \times N$ parameter matrices that satisfy the stationarity restriction. Finally, ε_t is a sequence of $N \times 1$ independent white noise vectors with distribution function F_{ε} , positive definite contemporaneous covariance matrix Σ_{ε} , and finite fourth order moments.

The point predictor of Y_{T+h} that minimizes the Mean Squared Forecast Error (MSFE) is its conditional mean, which, in practice, is obtained by substituting the unknown parameters with consistent estimates, as follows:

$$\widehat{Y}_{T+h|T} = \widehat{\mu} + \widehat{\Phi}_1 \widehat{Y}_{T+h-1|T} + \dots + \widehat{\Phi}_p \widehat{Y}_{T+h-p|T}, \quad (2)$$

where $\widehat{Y}_{T+j|T} = Y_{T+j}$, $j \leq 0$, and, in this paper, $\widehat{\theta} = (\widehat{\mu}, \widehat{\Phi}_1, \dots, \widehat{\Phi}_p)$ denotes the LS estimator of the parameters. The MSFE of $\widehat{Y}_{T+h|T}$, denoted by $\widehat{\Sigma}_{\widehat{Y}}^A(h)$, can be obtained using the asymptotic distribution to approximate the parameter uncertainty with all unknown parameters substituted by their sample estimates; see Lütkepohl (1991) for a detailed description. Obviously, the contribution of the parameter uncertainty to the MSFE of $\widehat{Y}_{T+h|T}$ depends on the dimension of the system, N , the VAR order, p , and the sample size, T ; see for instance Baillie (1979) and Reinsel (1980). As long as N and/or p are big enough relative to T , the effect of the parameter uncertainty can be substantial. However, granted that a consistent estimator is used, the importance of this uncertainty could be small in systems consisting of only a few variables if T is relatively large; see Riise and Tjostheim (1984).

If ε_t is further assumed to be Gaussian for the model in Eq. (1), then the h -step-ahead forecast density can be estimated by

$$Y_{T+h} \sim N(\widehat{Y}_{T+h|T}, \widehat{\Sigma}_{\widehat{Y}}^A(h)). \quad (3)$$

From Eq. (3), it is possible to obtain h -step-ahead joint ellipsoids for the variables within the system. Constructing these ellipsoids can be quite demanding when N is larger than two or three. As a consequence, Lütkepohl (1991) proposes the construction of forecast regions using Bonferroni cubes based on marginal forecast intervals for each of the variables in the system.

The forecast density in Eq. (3) incorporates the uncertainty due to parameter estimation through the asymptotic distribution, but still relies on Gaussian forecast errors. Therefore, the corresponding intervals and regions could be inadequate when this assumption is not satisfied. In addition, in this case, the shape of the densities for $h \geq 2$ is generally unknown. Finally, note that the Gaussian forecast densities in Eq. (3) can be misleading in cases where the asymptotic approximation is unreliable; see Dufour and Jouini (2006).

As an illustration, we consider the following non-Gaussian stationary bivariate VAR(2) model previously considered by Kim (2001):

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} 0.9 & 0 \\ -0.5 & -0.7 \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} -0.2 & 0 \\ 0.8 & -0.1 \end{bmatrix} \begin{bmatrix} y_{1,t-2} \\ y_{2,t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}, \quad (4)$$

where $\varepsilon_t = (\varepsilon_{1,t}, \varepsilon_{2,t})'$ has a χ_4^2 distribution, standardized so that $\text{vech}(\Sigma_{\varepsilon}) = (1, 0.5, 1)'$, with vech denoting the lower-diagonal column stacking operator of a symmetric matrix; see Kilian (1998a) for the adequacy of this distribution for representing some macroeconomic time series. The dominant root of the VAR(2) model in Eq. (4) is 0.5, so the model is far from the non-stationary boundary. Panel (a) of Fig. 1 displays the true joint one-step-ahead density of Y_{T+1} . After generating a time series of size $T = 100$, the VAR(2) parameters are estimated by LS, assuming that the lag order is known. Panel (b) of Fig. 1 plots the corresponding bivariate density, obtained under the assumption that the forecast errors are jointly Gaussian, as given in Eq. (3) with $\widehat{Y}_{T+1|T} = (-2.21, -0.27)'$ and $\text{vech}(\widehat{\Sigma}_{\widehat{Y}}^A(1)) = (1.13, 0.59, 1.02)'$. Comparing panels (a) and (b), it is obvious that the Gaussian density fails to capture the asymmetry of the error distribution.

Alternatively, bootstrap procedures can be implemented so as to obtain forecast densities that incorporate the parameter uncertainty without relying on Gaussian forecast errors. In order to take into account the conditionality of VAR forecasts on past observations, Kim (1999) proposes to obtain bootstrap replicates of the series based on the backward recursion as follows

$$Y_t^* = \widehat{\omega} + \widehat{\Lambda}_1 Y_{t+1}^* + \dots + \widehat{\Lambda}_p Y_{t+p}^* + \widehat{v}_t^*, \quad t = T - p, \dots, 1, \quad (5)$$

where $Y_t^* = Y_t$ for $t = T - p + 1, \dots, T$, $(\widehat{\omega}, \widehat{\Lambda}_1, \dots, \widehat{\Lambda}_p)$ are LS estimates of the backward parameters and \widehat{v}_t^* are random draws with replacement from the empirical distribution function of the backward residuals, re-scaled by the factor $[(T - p) / (T - (N + 1)p - 1)]^{0.5}$. The bootstrap replicates in Eq. (5) are then used to obtain bootstrap replicates of the parameters in Eq. (1), and finally, these are implemented to obtain bootstrap replicates of Y_{T+h} .

Kim (1999) justifies the use of the backward bootstrap procedure in finite samples by suggesting that the asymptotic results of Thombs and Schucany (1990) can be extended to a multivariate framework. When using the backward representation, one can bootstrap from the forward residuals and use the relationship between the

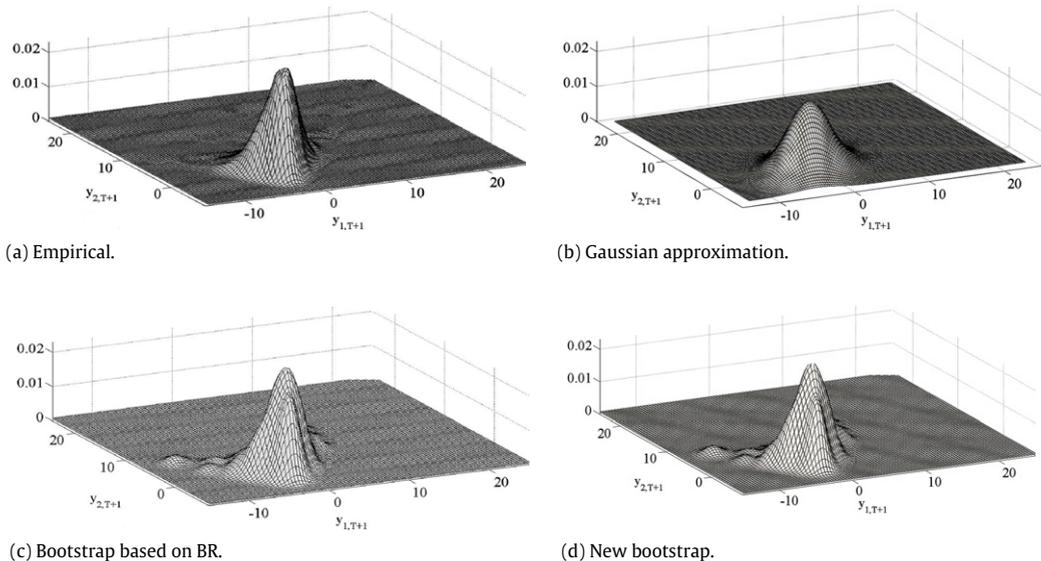


Fig. 1. Kernel estimates of one-step-ahead forecast densities of a simulated bivariate series with $T = 100$ observations, generated by a stationary VAR(2) model with χ_4^2 errors.

backward and forward residuals to obtain the bootstrap replicates of the former; see Kim (1997, 1998) for the expression of the backward representation.² In this case, the forward residuals need to be serially independent and identically distributed, but not necessarily Gaussian, for the bootstrap procedure to be asymptotically valid. However, the relationship between forward and backward residuals can be rather complicated in relatively simple VAR models, and many authors resample directly from the backward residuals. The backward errors are only serially independent if the forward errors are Gaussian. Consequently, the asymptotic validity of the bootstrap procedures based on the backward representation relies on the assumption of Gaussian innovations; see Kim (2001).

Another important disadvantage of the backward bootstrap is that it cannot be implemented in models without this representation, which compromises the flexibility and applicability of the procedure.

The backward bootstrap procedure is illustrated by again considering the time series of size $T = 100$ simulated by the non-Gaussian bivariate VAR(2) model in Eq. (4). Panel (c) of Fig. 1 plots a kernel estimate of the joint backward bootstrap density of Y_{T+1} based on $B = 4999$ bootstrap replicates. When comparing this density with its Gaussian counterpart in panel (b), it is clear that the bootstrap can reproduce the asymmetry and is closer to the true density plotted in panel (a) of the same figure.

² For the simpler expression of the backward representation in which the lag values of the variables in Eq. (1) are substituted by forward values, Chan, Ho, and Tong (2006) and Tong and Zhang (2005) show that a necessary condition for the VAR(p) model to have this backward representation is that the covariance matrices $\Upsilon(h) = E[(Y_t - E(Y_t))(Y_{t-h} - E(Y_{t-h}))']$ are symmetric for all h . This is a very strong restriction which is not likely to be satisfied in real data systems.

3. Forward bootstrap procedure

In this section, we describe the forward bootstrap procedure and establish its asymptotic validity. Its performance is illustrated for the non-Gaussian stationary VAR(2) model considered above.

3.1. Description of the algorithm

The forward bootstrap forecast density of Y_{T+h} is based on the same assumption that is used to construct forecast densities when incorporating the parameter uncertainty using the asymptotic distribution, namely, that the sample used to estimate the parameters is independent of the sample used to forecast. In this way, we follow Runkle (1987) when generating the bootstrap replicates used to estimate the parameters and use the forward instead of the BR. The forward algorithm for obtaining bootstrap replicates of Y_{T+h} is a direct generalization of the algorithm proposed by Pascual et al. (2004a); see Lütkepohl et al. (2015), Staszewska-Bystrova and Winker (2013) and Wolf and Wunderli (2012) for implementations of the forward bootstrap for obtaining path forecasts in VAR models.

For clarity, we next describe the algorithm.

Step 1. After selecting the order p , obtain LS estimates of the parameters of model (1) and the corresponding vector of residuals. Denote by \hat{F}_ε the empirical distribution function of the re-scaled residuals.

Step 2. Construct a bootstrap series $\{Y_1^*, \dots, Y_T^*\}$ as follows:

$$Y_t^* = \hat{\mu} + \hat{\Phi}_1 Y_{t-1}^* + \dots + \hat{\Phi}_p Y_{t-p}^* + \hat{\varepsilon}_t^*, \quad t = 1, \dots, T, \quad (6)$$

where $\hat{\varepsilon}_t^*$ are random draws with replacement from \hat{F}_ε and $Y_t^* = Y_t$, for $t = -p + 1, \dots, 0$. Obtain $\hat{\theta}^* = (\hat{\mu}^*, \hat{\Phi}_1^*, \dots, \hat{\Phi}_p^*)$, a bootstrap replicate of the LS estimates,

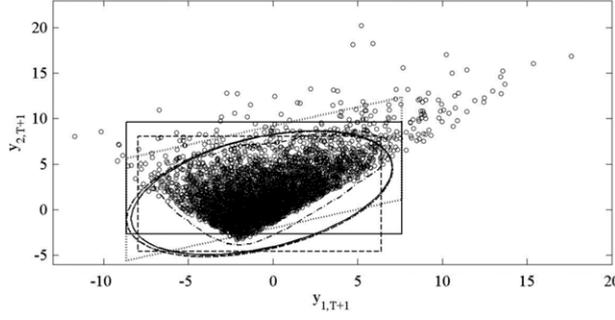


Fig. 2. A realization of size $R = 5000$ of Y_{T+1} (\circ), generated by a stationary bivariate VAR(2) model with χ_4^2 errors, together with 95% one-step-ahead forecast regions based on a sample size $T = 100$: Gaussian ellipsoid and cube (discontinuous lines), bootstrap ellipsoid and cube (continuous line), corrected bootstrap cube (dotted line) and High Density Region (dotted-discontinuous line).

by fitting a VAR(p) model to the bootstrap replicate $\{Y_1^*, \dots, Y_T^*\}$.

Step 3. Using the model in Eq. (1), with the parameters substituted by their bootstrap estimates, and fixing the last p observations of the original series, obtain recursively a bootstrap replicate of Y_{T+h} as follows:

$$\begin{aligned} \widehat{Y}_{T+h|T}^* &= \widehat{\mu}^* + \widehat{\Phi}_1^* \widehat{Y}_{T+h-1|T}^* \\ &+ \dots + \widehat{\Phi}_p^* \widehat{Y}_{T+h-p|T}^* + \widehat{\varepsilon}_{T+h}^*, \end{aligned} \quad (7)$$

with $\widehat{\varepsilon}_{T+h}^*$ being a random draw with replacement from \widehat{F}_ε , and $\widehat{Y}_{T+h|T}^* = Y_{T+h}$, $h \leq 0$.

Step 4. Repeat steps 2 and 3 B times.

It is important to point out that the bootstrap replicates used in step 2 to obtain bootstrap estimates of the parameters are obtained as proposed by Runkle (1987) using the forward expression in Eq. (6) instead of the backward expression in Eq. (5). However, the last p observations in the series are still fixed when forecasting the future values in Eq. (7).

To illustrate the implementation of the forward bootstrap procedure and the differences between it and the backward procedure, we again consider the bivariate time series generated by the non-Gaussian stationary VAR(2) model in Eq. (4). Panel (d) of Fig. 1, showing the results for the forward procedure, displays a kernel estimate of the bootstrap joint density of Y_{T+1} , which is very similar to the density obtained by implementing the backward bootstrap procedure.

Using the algorithm described above, B bootstrap replicates of Y_{T+h} , denoted by $\{\widehat{Y}_{T+h|T}^{*(1)}, \dots, \widehat{Y}_{T+h|T}^{*(B)}\}$, are obtained. Their empirical bootstrap distribution can be used to obtain the corresponding bootstrap prediction ellipsoid with probability content $(1 - \delta)$ 100% as follows:

$$\begin{aligned} E_{T+h} &= \{Y_{T+h} | [Y_{T+h} - \bar{Y}_{T+h|T}^*] S_{\widehat{Y}^*}(h)^{-1} \\ &\times [Y_{T+h} - \bar{Y}_{T+h|T}^*] < Q_K^*\}, \end{aligned} \quad (8)$$

where $\bar{Y}_{T+h|T}^*$ is the sample mean of the B bootstrap replicates, $S_{\widehat{Y}^*}(h)$ is the corresponding sample covariance, and Q_K^* is the $(1 - \delta)$ 100% percentile of the empirical bootstrap distribution of the quadratic form $[\widehat{Y}_{T+h|T}^* -$

$\bar{Y}_{T+h|T}^*] S_{\widehat{Y}^*}(h)^{-1} [\widehat{Y}_{T+h|T}^* - \bar{Y}_{T+h|T}^*]$. Furthermore, the Bonferroni cube with at least $(1 - \delta)$ 100% nominal coverage is given by

$$C_{T+h} = \left\{ Y_{T+h} | Y_{T+h} \in \bigcup_{i=1}^N [q_i^*(\tau), q_i^*(1 - \tau)] \right\}, \quad (9)$$

where $q_i^*(\tau)$ is the τ th quantile of the empirical bootstrap distribution of the i th element of $\widehat{Y}_{T+h|T}^*$. The Bonferroni cubes are defined as in Eq. (9) because they are better suited for dealing with potential asymmetries of the error distribution than the percentile- t intervals; see Hall (1992). In addition, Kilian (1999) shows that, in the absence of pivotal statistics, as is the case when the VAR process is close to the nonstationary region, bootstrap percentile methods that do not rely on studentized statistics have better coverage accuracies than those based on the percentile- t .

To illustrate the construction of ellipsoids and cubes, Fig. 2 plots the forward bootstrap 95% ellipsoid and the Bonferroni cube in Eqs. (8) and (9) respectively, together with the corresponding regions constructed from the Gaussian density in Eq. (3), both obtained for a particular one-step-ahead realization of the VAR(2) model in Eq. (4) that is also displayed. We observe that, although the forward bootstrap density is very different from the Gaussian density, there are no big differences among the corresponding ellipsoids, due to the fact that the first two moments involved in their definition do not differ significantly among the procedures, which estimate similar centers and dispersions of the future values. Note that $\bar{Y}_{T+1|T}^* = (-2.19, -0.26)'$ and $\text{vech}(S_{\widehat{Y}^*}(1)) = (1.11, 0.58, 1)'$ are very similar to the corresponding quantities used to compute the Gaussian ellipsoid. Also, when looking at the bootstrap Bonferroni cube, we can observe that it is located above and to the right of the corresponding Gaussian cube, representing the asymmetries of the joint distribution of Y_{T+1} . Note that, even though the bootstrap Bonferroni cube is somehow more adequate to represent the asymmetry in the forecast error distribution, it is not satisfactory when constructing forecast regions for systems of correlated non-Gaussian variables, such as those considered in the illustration. Consequently, we explore two further alternatives for the construction of these regions. First, we consider the High Density Regions (HDR)

proposed by Hyndman (1996) based on kernel estimates of the joint bootstrap density. Fig. 2 also plots the 95% HDR computed from the bootstrap replicates, $\hat{Y}_{T+1|T}^*$. We can see that the shape of the HDR seems to be a more adequate representation of the realization of Y_{T+1} than either the ellipsoid or the Bonferroni cube. However, HDR are unfeasible when the dimension of the system is large, as, in this case, there are no satisfactory kernel estimators of the bootstrap densities. Consequently, we also explore a simple modification of the Bonferroni cube that takes into account the correlation between the variables in the system. The modified Bonferroni cube is defined by the following four points: $[q_1^*(\tau), q_2^*(\tau) + p_{21,h}q_1^*(\tau)]$, $[q_1^*(1-\tau), q_2^*(\tau) + p_{21,h}q_1^*(1-\tau)]$, $[q_1^*(\tau), q_2^*(1-\tau) + p_{21,h}q_1^*(\tau)]$, and $[q_1^*(1-\tau), q_2^*(1-\tau) + p_{21,h}q_1^*(\tau)]$, where $p_{21,h} = \hat{\sigma}_{21,T+h}^*/\hat{\sigma}_{1,T+h}^{2*}$, with $\hat{\sigma}_{21,T+h}^*$ and $\hat{\sigma}_{1,T+h}^{2*}$ being elements of $S_{\hat{Y}^*}(h)$. Note that the proposed transformation re-expresses the original cube in a direction defined by the association between $y_{1,T+h}$ and $y_{2,T+h}$, as measured by $p_{21,h}$. Furthermore, the volume of the modified cube remains unchanged, since the coordinates for the first variable stay the same, while those of the second variable are transformed by the same amount, either $p_{21,h}q_1^*(\tau)$ or $p_{21,h}q_1^*(1-\tau)$. Fig. 2 plots the modified Bonferroni cube, which is rotated in the direction of the correlation observed between $\hat{y}_{1,T+1|T}^*$ and $\hat{y}_{2,T+1|T}^*$, and, as a consequence, gives a more appropriate picture of the values of $y_{1,T+1}$ and $y_{2,T+1}$ that can be expected one step ahead.

Before establishing the asymptotic validity of the forward bootstrap procedure, we should mention that it can be modified easily by introducing asymptotic stationarity bias corrections of the bootstrap parameters and the endogenous lag order bootstrap algorithm, as was proposed by Kilian (1998c); see Staszewska-Bystrova and Winker (2013) for an implementation of the forward bootstrap algorithm using both modifications.

3.2. Asymptotic validity

Consider the stationary VAR(p) model in Eq. (1), where the errors are given by

$$\begin{aligned} \varepsilon_t(\theta) &= Y_t - \mu - \Phi_1 Y_{t-1} - \dots - \Phi_p Y_{t-p}, \\ t &= 1, \dots, T. \end{aligned} \quad (10)$$

In Eq. (10), the errors depend explicitly on the unknown parameters contained in θ . If θ is estimated by $\hat{\theta}$, the corresponding estimated residuals are given by

$$\begin{aligned} \hat{\varepsilon}_t(\hat{\theta}) &= Y_t - \hat{\mu} - \hat{\Phi}_1 Y_{t-1} - \dots - \hat{\Phi}_p Y_{t-p}, \\ t &= 1, \dots, T, \end{aligned} \quad (11)$$

which have $\hat{F}_{\hat{\varepsilon}}(\hat{\theta})$ as the empirical distribution function.

The following theorem establishes the asymptotic validity of the empirical bootstrap distribution of $\hat{Y}_{T+h|T}^*$, as given in Eq. (7), to approximate the distribution of a future value Y_{T+h} .

Theorem. Let $\{Y_t, t = -p+1, \dots, 1, 2, \dots, T\}$ be a realization of a stationary VAR(p) process, defined as in Eq. (1), $\hat{\theta}$ be the LS estimator of θ , and $\hat{Y}_{T+h|T}^*$ be obtained by following steps 1–4 in the previous subsection. Then, $\hat{Y}_{T+h|T}^*$, conditioned on $\{Y_t, t = -p+1, \dots, 1, 2, \dots, T\}$, converges weakly in probability to Y_{T+h} as $T \rightarrow \infty$.

Proof. Following the arguments of Pascual et al. (2004a), consider first the one-step-ahead bootstrap future value given by

$$\hat{Y}_{T+1|T}^* = \hat{\mu}^* + \hat{\Phi}_1^* Y_T + \dots + \hat{\Phi}_p^* Y_{T-p+1} + \hat{\varepsilon}_{T+1}^*. \quad (12)$$

For $h = 2$, we have

$$\hat{Y}_{T+2|T}^* = \hat{\mu}^* + \hat{\Phi}_1^* \hat{Y}_{T+1|T}^* + \dots + \hat{\Phi}_p^* Y_{T-p+2} + \hat{\varepsilon}_{T+2}^*. \quad (13)$$

Replacing $\hat{Y}_{T+1|T}^*$ in Eq. (13) with its expression in Eq. (12), it follows that

$$\begin{aligned} \hat{Y}_{T+2|T}^* &= N_0(\hat{\theta}^*) + N_1(\hat{\theta}^*) Y_T + \dots + N_p(\hat{\theta}^*) Y_{T-p+1} \\ &\quad + M_1(\hat{\theta}^*) \hat{\varepsilon}_{T+1}^* + \hat{\varepsilon}_{T+2}^*, \end{aligned} \quad (14)$$

where $N_i(\hat{\theta}^*)$ and $M_i(\hat{\theta}^*)$ are appropriately defined continuous functions of the estimated parameters.

Proceeding in this way, the following expression is obtained for the h -step-ahead bootstrap forecast:

$$\begin{aligned} \hat{Y}_{T+h|T}^* &= N_0(\hat{\theta}^*) + N_1(\hat{\theta}^*) Y_T + \dots + N_p(\hat{\theta}^*) Y_{T-p+1} \\ &\quad + M_1(\hat{\theta}^*) \hat{\varepsilon}_{T+1}^* + M_2(\hat{\theta}^*) \hat{\varepsilon}_{T+2}^* + \dots + \hat{\varepsilon}_{T+h}^*, \end{aligned} \quad (15)$$

where the functions $N_i(\hat{\theta}^*)$ and $M_i(\hat{\theta}^*)$ are different for different horizons. Eq. (15) defines the bootstrap future values as a function of the observed realization $\{Y_{-p+1}, \dots, Y_T\}$, the independent random draws $\hat{\varepsilon}_{T+h}^*$, and continuous functions of the bootstrap parameter estimates $\hat{\theta}^*$.

In order to establish the asymptotic convergence of $\hat{Y}_{T+h|T}^*$, we start by considering the terms involving $N_i(\hat{\theta}^*)$ for which the asymptotic validity of $\hat{\theta}^*$ is needed. The asymptotic validity of the bootstrap LS estimator is established by Bose (1988), who proves the almost sure convergence in probability of $\hat{\theta}^*$ to θ . Therefore, given that $N_i(\hat{\theta}^*)$ are continuous functions of the parameters, it follows that $N_i(\hat{\theta}^*) \xrightarrow{p} N_i(\theta)$ almost surely. Moreover, note that Y_{T-i+1} are fixed values, and consequently, the terms involving $N_i(\hat{\theta}^*) Y_{T-i+1} \xrightarrow{d} N_i(\theta) Y_{T-i+1}$ in probability. Second, using the same arguments as before, we can see that $M_i(\hat{\theta}^*) \xrightarrow{p} M_i(\theta)$ almost surely. Finally, consider the terms $\hat{\varepsilon}_{T+i}^*$, which are random draws with replacement from $\hat{F}_{\hat{\varepsilon}}(\hat{\theta})$. Using the results of Bickel and Freedman (1981) and Freedman (1984), it is straightforward to prove that $d_2(\hat{F}_{\hat{\varepsilon}}(\hat{\theta}), F_{\varepsilon}(\theta)) \rightarrow 0$ in probability as $T \rightarrow \infty$, where d_2 is a Mallow's metric. Given that convergence in d_2 implies weak convergence of the corresponding random variables, it follows that $\hat{\varepsilon}_{T+i}^* \xrightarrow{d} \varepsilon_{T+i}$ in probability. On the other hand, $\hat{\varepsilon}_{T+i}^*$ are independent of $M_i(\hat{\theta}^*)$. Consequently, $M_i(\hat{\theta}^*) \hat{\varepsilon}_{T+i}^* \xrightarrow{d} M_i(\theta) \varepsilon_{T+i}$ by the independence of $\hat{\varepsilon}_{T+i}^*$ and the bootstrap version of Slutsky's Theorem. Consequently, all terms in Eq. (15) converge weakly in probability, and, as a result, $Y_{T+h|T}^* \xrightarrow{d} Y_{T+h}$ in probability as $T \rightarrow \infty$. ■

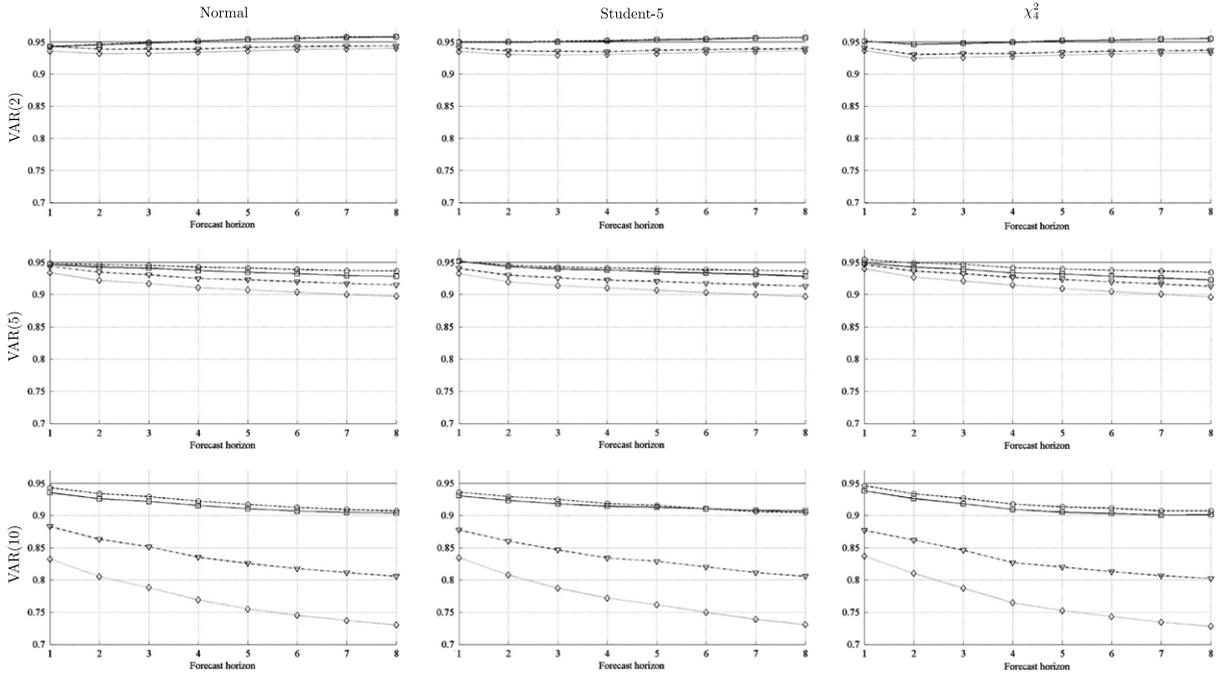


Fig. 3. Monte Carlo averages of the empirical coverages of Bonferroni forecast cubes based on the: (i) Gaussian (\diamond), (ii) Gaussian with asymptotic MSFE (∇), (iii) bootstrap with BR (\circ), and (iv) new bootstrap (\square) densities, for a stationary VAR(2) model (first row), a persistent VAR(5) model (second row) and a near-cointegrated VAR(10) model (third row) with $T = 100$ and Gaussian (first column), Student-5 (second column) and χ_4^2 (third column) errors. Nominal coverage: 95%.

Before concluding this section, it is important to remark that, even though asymptotic validity is established for centered residuals, [Stine \(1987\)](#) shows that it is still valid if they are also re-scaled. Furthermore, the asymptotic bias correction of the parameters proposed by [Pope \(1990\)](#) does not alter the asymptotic validity of our procedure, since the bias and its bootstrap version are $O_p(T^{-1})$; see [Kilian \(1998c\)](#) for further details. Finally, [Kilian \(1998a\)](#) also shows that the endogenous lag order bootstrap algorithm is still asymptotically valid for standard lag order selection criteria, such as the AIC considered in this paper.

4. Small sample properties

In this section, we carry out Monte Carlo experiments to analyse the finite sample properties of the forward bootstrap procedure and compare them with those of the alternatives. We consider three different bivariate data generating processes (DGP), with different configurations of parameters and lag orders which reproduce stationary, persistent or near-cointegrated processes. DGP1 is the stationary VAR(2) model defined in Eq. (4). DGP2 and DGP3 are a persistent VAR(5) model and a near-cointegrated VAR(10) model, respectively; see [Kilian \(1998a\)](#) for similar specifications which are described in detail in the [Appendix](#). In each DGP, we consider three distributions of the errors, namely Gaussian, Student-5 and χ_4^2 , which are adequately centered and re-scaled. For each of the resulting nine specifications, we generate $M = 2000$ replicates of sizes $T = 100$ and 300. The sample sizes have been chosen to be in concordance with those usually encountered

in practice when forecasting with real macroeconomic series. For each generated series, the lag order is estimated according to the AIC. The maximum lag orders are equal to 12 and 16 for $T = 100$ and 300, respectively. After estimating the lag order, the VAR parameters are estimated by LS and bias-corrected using the asymptotic correction of [Pope \(1990\)](#), taking into account the stationarity restriction proposed by [Kilian \(1998a\)](#). Forecast densities for horizons of $h = 1, \dots, 8$ steps ahead are constructed under the assumption of Gaussian errors without parameter uncertainty, and computing the MSFE as in Eq. (3) using the asymptotic approximation. Forecast densities are also constructed using $B = 1999$ bootstrap replicates obtained by the backward and forward bootstrap procedures. The bootstrap procedures are implemented using the asymptotic bias and endogenous lag order corrections proposed by [Kilian \(1998a,c\)](#). In each case, we construct the corresponding 95% ellipsoids and Bonferroni cubes.

We calculate the empirical coverage of each forecast region based on $R = 5000$ future values of the process Y_{T+h} . [Fig. 3](#) plots the averages through the Monte Carlo replicates of the empirical coverages of the Bonferroni regions constructed for all models considered when $T = 100$.³ We can observe that, regardless of the distribution of the error and the procedure used to construct the cubes, the average coverages are rather close to the nominal when

³ Results for $T = 300$ and the ellipsoids are not reported, to save space. For the same reason, we do not report the results without bias and endogenous lag-order corrections. They are qualitatively similar to those reported in this paper. All of these results are available upon request.

dealing with the VAR(2) models considered in this paper. Note that, in this case, the number of parameters is rather small and the roots of the model are far from the non-stationary region. As a consequence, the uncertainty in the LS estimator is not important when constructing the forecast regions. When the errors are non-Gaussian, we can see that the Gaussian coverages are slightly under the nominal, while the bootstrap cubes maintain their accuracy.

The second and third rows of Fig. 3 plot the coverages corresponding to the persistent VAR(5) and near-cointegrated VAR(10) models, respectively. These models are interesting because the number of parameters is rather large and they are close to the non-stationarity bounds. We can observe that the coverages of the two bootstrap cubes are rather similar, regardless of whether they use the BR or the forward representation. The bootstrap coverages are very close to the nominal, although they deteriorate with the forecast horizon in both models, a feature that is more pronounced in the VAR(10) model. This deterioration may be due to the fact that this model is close to the non-stationary bounds, and consequently, it may have more difficulties when forecasting in the long run. Fig. 3 also shows that, in the VAR(10) model, incorporating the parameter uncertainty may be important even when the errors are Gaussian. The coverages of the Gaussian cubes in which the MSFEs do not incorporate the parameter uncertainty are smaller than those of cubes constructed by any of the procedures that take into account this uncertainty. Also, when the errors are non-Gaussian, the coverages are under-estimated by both cubes based on Gaussian densities.

To summarize, the simulations carried out in this section show that the forward bootstrap procedure performs better than traditional methods based on Gaussian densities, and performs no worse than the bootstrap procedure based on the BR. In addition, they suggest that, as the persistence of the system increases, it becomes important to take the parameter uncertainty into account in order to obtain coverages which are close to the nominal ones.

5. Bootstrap forecasts of returns, volatilities and correlations in the DCC model

Multivariate GARCH models (MV-GARCH) are useful for representing the dynamic dependence in the second order moments of multivariate time series. Forecasting correlations is a key issue in financial management, derivative pricing models or hedging strategies; see for example Engle (2009). The asymptotic validity of the forward bootstrap procedure has been established above for linear VAR models. However, given that the bootstrap procedure considered in this paper does not rely on the BR, it can be extended to deal with conditionally heteroscedastic models; see Pascual, Romo, and Ruiz (2006) and Reeves (2005) for bootstrap procedures in the context of forecasting univariate GARCH models. In this section, we show how to use the forward bootstrap procedure to obtain forecast intervals and regions for returns, volatilities and correlations in the context of the Dynamic Conditional Correlation

(DCC) model proposed by Engle (2002). This implementation can also be seen as a multivariate extension of the bootstrap procedure proposed by Pascual et al. (2006) for univariate GARCH models. The DCC model is a popular MV-GARCH model that simplifies the estimation of the conditional covariance by, first, estimating univariate GARCH models for each variable in the system, and, second, estimating the conditional correlation matrix using the resulting standardized residuals.

In order to simplify the exposition, we will present the results for the following VAR(p)-DCC(1, 1) model:

$$Y_t = \mu + \Phi_1 Y_{t-1} + \dots + \Phi_p Y_{t-p} + \varepsilon_t$$

$$\varepsilon_t = H_t^{1/2} a_t, \quad (16)$$

where a_t is a sequence of $N \times 1$ independent white noise vectors with an identity covariance matrix, and H_t is a $N \times N$ positive definite conditional covariance matrix given by

$$H_t = D_t R_t D_t, \quad (17)$$

with D_t being a diagonal matrix containing the univariate GARCH(1, 1) conditional standard deviations of each variable in the system, given by

$$\sigma_{i,t} = \sqrt{\omega_i + \alpha_i \varepsilon_{i,t-1}^2 + \beta_i \sigma_{i,t-1}^2}, \quad i = 1, \dots, N. \quad (18)$$

The matrix R_t is the conditional correlation matrix of the standardized errors $\varepsilon_t^s = D_t^{-1} \varepsilon_t$. In order to ensure the positiveness of the correlation matrix, R_t is defined as follows:

$$R_t = Q_t^s Q_t Q_t^s, \quad (19)$$

where

$$Q_t = (1 - \alpha - \beta)Q + \alpha \varepsilon_{t-1}^s \varepsilon_{t-1}^{s'} + \beta Q_{t-1}, \quad (20)$$

with Q being the unconditional correlation matrix. Finally, Q_t^s is a diagonal matrix, with its elements being the inverse square root of the elements in the main diagonal of Q_t . All of the parameters are assumed to satisfy the stationarity and positivity conditions.

The evolution of the conditional correlation matrix in Eq. (19) is a nonlinear process where

$$R_{T+h|T} = Q_{T+h|T}^s Q_{T+h|T} Q_{T+h|T}^s \quad (21)$$

and

$$Q_{T+h|T} = (1 - \alpha - \beta)Q + \alpha E_T(\varepsilon_{T+h-1}^s \varepsilon_{T+h-1}^{s'}) + \beta Q_{T+h-1|T}, \quad (22)$$

with $E_T(\varepsilon_{T+h-1}^s \varepsilon_{T+h-1}^{s'}) = E_T(R_{T+h-1})$. Consequently, the h -step-ahead forecast of the correlation matrix cannot be solved forward directly so as to provide a convenient method for forecasting. To overcome this problem, Engle and Sheppard (2001) propose using the approximation $E_T[R_{t+h}] = E_T[Q_{t+h}]$. In this case, R_{T+h} can be forecast directly by

$$R_{T+h|T} = (1 - \alpha - \beta)Q \sum_{j=0}^{h-2} (\alpha + \beta)^j + (\alpha + \beta)^{h-1} R_{T+1|T}. \quad (23)$$

Note that the one-step-ahead forecast of the correlation matrix, $R_{T+1|T}$, can be solved backward, which results in

$$R_{T+1|T} = Q + \alpha \sum_{j=0}^{T-1} \beta^j (\varepsilon_{T-j} D_{T-j}^{-2} \varepsilon'_{T-j} - Q). \quad (24)$$

On the other hand, the h -step-ahead forecast of the volatility of each of the returns is given by

$$\sigma_{i,T+h|T}^2 = \omega_i \sum_{j=1}^{h-2} (\alpha_i + \beta_i)^j + (\alpha_i + \beta_i)^{h-1} \sigma_{i,T+1|T}^2, \quad (25)$$

$$i = 1, \dots, N.$$

An analogous expression is valid for the one-step-ahead forecast of the i th variance, $\sigma_{i,T+1|T}^2$, which is given by

$$\sigma_{i,T+1|T}^2 = \frac{\omega_i}{1 - \alpha_i - \beta_i} + \alpha_i \sum_{j=0}^{T-1} \beta_i^j \left(\varepsilon_{i,T-j}^2 - \frac{\omega_i}{1 - \alpha_i - \beta_i} \right), \quad (26)$$

$$i = 1, \dots, N.$$

Eqs. (24) and (26) show that, given the model parameters, the one-step-ahead forecasts of the conditional variances and correlations depend only on the observed data $\{Y_1, \dots, Y_T\}$.

Using Eqs. (23) and (25), it is possible to construct $H_{T+h|T}$, after which, assuming that a_t is Gaussian, one can obtain the forecast density of Y_{T+h} , which is given by

$$Y_{T+h}|Y_1, \dots, Y_T \sim N(Y_{T+h|T}, H_{T+h|T}). \quad (27)$$

In practice, the parameters in Eqs. (23) and (25) are unknown and must be estimated. The density in Eq. (27) does not incorporate the parameter uncertainty, and, as a consequence, will underestimate the uncertainty associated with the forecast of Y_{T+h} . In order to estimate the parameters, in this paper, we first estimate the VAR(p) parameters using LS, assuming that p is known. Then, as was proposed by Engle (2002), the parameters of the DCC model are estimated in two steps using the VAR residuals as follows: (i) the parameters involved in each conditional variance equation are estimated in a univariate fashion using Quasi Maximum Likelihood (QML) by maximizing the Gaussian log-likelihood, and (ii) the parameters governing the correlation dynamics are also estimated by QML using the standardized residuals.⁴

On the other hand, note that, even if the errors were truly Gaussian, the forecast density in Eq. (27) is valid for $h = 1$ and provides an approximation for $h \geq 2$. Furthermore, when the errors are non-Gaussian, the future densities of returns predicted using Eq. (27) could be inappropriate. Finally, even assuming Gaussian errors, it is not straightforward to obtain forecast intervals and regions for future volatilities and correlations. The forward bootstrap

procedure can be implemented to deal with the parameter uncertainty and non-Gaussianity of the errors.

Next, we describe the bootstrap procedure for obtaining forecast densities of the returns, volatilities and correlations associated with Y_{T+h} .

Step 1. Select the orders of the VAR and GARCH components and estimate the model parameters as described above. Obtain $\widehat{a}_t, \widehat{\sigma}_t^2$, which has an empirical distribution function given by $\widehat{F}_{\widehat{a}_t}$.

Step 2. Recursively obtain bootstrap replicates of the standardized correlated residuals, $\widehat{\varepsilon}_t^{s*}$, and the correlation matrix as follows:

$$\widehat{\varepsilon}_t^{s*} = R_t^{s*} \widehat{a}_t^*, \quad t = 1, \dots, T, \quad (28)$$

where \widehat{a}_t^* are random draws with replacement from $\widehat{F}_{\widehat{a}_t}$ and $R_t^* = Q_t^{s*} Q_t^* Q_t^{s*}$ with

$$Q_t^* = (1 - \widehat{\alpha} - \widehat{\beta}) \widehat{Q} + \widehat{\alpha} \widehat{\varepsilon}_{t-1}^{s*} \widehat{\varepsilon}_{t-1}^{s*'} + \widehat{\beta} Q_{t-1}^*, \quad (29)$$

$$t = 2, \dots, T.$$

The recursion starts with $R_1^* = Q_1^* = \widehat{Q}$, where \widehat{Q} contains the sample correlations of $\{Y_1, \dots, Y_T\}$.

Step 3. Obtain bootstrap replicates of ε_t and their variances as follows:

$$\sigma_{i,t}^{2*} = \widehat{\omega}_i + \widehat{\alpha}_i \varepsilon_{i,t-1}^{2*} + \widehat{\beta}_i \sigma_{i,t-1}^{2*}, \quad (29)$$

$$i = 1, \dots, N, t = 2, \dots, T,$$

$$\widehat{\varepsilon}_{i,t}^* = \widehat{\varepsilon}_{i,t}^{s*} \sigma_{i,t}^{2*}, \quad i = 1, \dots, N, t = 1, \dots, T, \quad (30)$$

where $\sigma_{i,1}^{2*} = \widehat{\omega}_i / (1 - \widehat{\alpha}_i - \widehat{\beta}_i)$.

Step 4. Construct a bootstrap replicate of Y_t as follows:

$$Y_t^* = \widehat{\mu} + \widehat{\Phi}_1 Y_{t-1}^* + \dots + \widehat{\Phi}_p Y_{t-p}^* + \varepsilon_t^*, \quad t = 2, \dots, T,$$

where $\widehat{Y}_t^* = Y_t$ for $t = -p + 1, \dots, 0$. Obtain bootstrap estimates of the parameters $\widehat{\theta}^* = [\text{vec}(\widehat{\mu}^*), \text{vec}(\widehat{\Phi}_1^*), \dots, \text{vec}(\widehat{\Phi}_p^*), \widehat{\alpha}^*, \widehat{\beta}^*, \dots, \widehat{\omega}_1^*, \widehat{\alpha}_1^*, \widehat{\beta}_1^*, \dots, \widehat{\omega}_N^*, \widehat{\alpha}_N^*, \widehat{\beta}_N^*]$ and the sample correlation matrix, \widehat{Q}^* , by fitting Eq. (20) to the bootstrap replicate $\{Y_1^*, \dots, Y_T^*\}$. Given the bootstrap estimates of the parameters, $\widehat{\theta}^*$, and the residuals of the fitted VAR(p) model, $\{\widehat{\varepsilon}_1, \dots, \widehat{\varepsilon}_T\}$, compute the correlations and variances at the forecast origin T as follows:

$$\widehat{Q}_T^* = \widehat{Q}^* + \widehat{\alpha}^* \sum_{j=0}^{T-1} \widehat{\beta}^{*j} (\widehat{\varepsilon}_{T-j-1} D_{T-j-1}^{*2} \widehat{\varepsilon}_{T-j-1} - \widehat{Q}^*)$$

$$\widehat{\sigma}_{i,T}^{2*} = \frac{\widehat{\omega}_i^*}{1 - \widehat{\alpha}_i^* - \widehat{\beta}_i^*} + \widehat{\alpha}_i^* \sum_{j=0}^{T-1} \widehat{\beta}_i^{*j} + \widehat{\beta}_i^{*j}$$

$$\times \left(\widehat{\varepsilon}_{i,T-j-1}^2 - \frac{\widehat{\omega}_i^*}{1 - \widehat{\alpha}_i^* - \widehat{\beta}_i^*} \right), \quad i = 1, \dots, N. \quad (31)$$

Obtain also the value of the standardized error at T , $\widehat{\varepsilon}_T^{s*} = D_T^{*-1} \widehat{\varepsilon}_T$. Keep the values of $\widehat{Q}_T^*, \widehat{\sigma}_{i,T}^{2*}, \widehat{\varepsilon}_T^{s*}$ and $\widehat{\varepsilon}_T$.

Step 5. Obtain future values of the correlated standardized errors and the correlations through the following recursions

$$\widehat{\varepsilon}_{T+h|T}^{s*} = \widehat{R}_{T+h|T}^{s*} \widehat{a}_{T+h}^* \quad (32)$$

$$\widehat{Q}_{T+h|T}^* = (1 - \widehat{\alpha}^* - \widehat{\beta}^*) \widehat{Q}^* + \widehat{\alpha}^* \widehat{\varepsilon}_{T+h-1|T}^{s*} \widehat{\varepsilon}_{T+h-1|T}^{s*'} + \widehat{\beta}^* \widehat{Q}_{T+h-1|T}^*$$

⁴ The procedure can be modified adequately to deal with the cDCC model of Aielli (2013); see Fresoli and Ruiz (2014) for a detailed analysis of the finite sample properties.

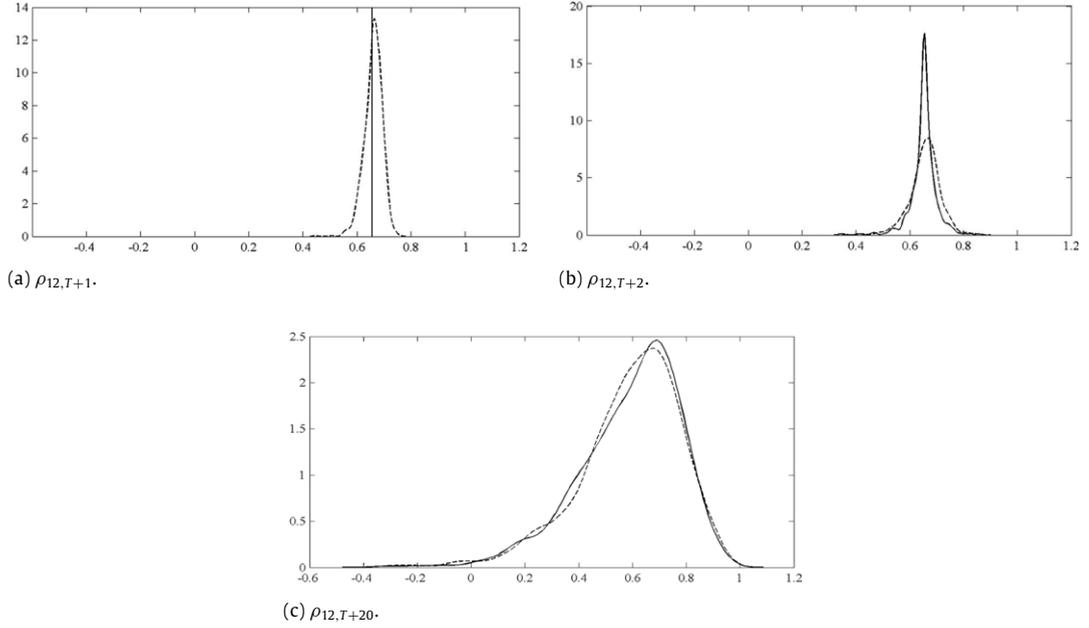


Fig. 4. Kernel estimates of the empirical (continuous line) and bootstrap (discontinuous line) densities of (a) one-step-ahead, (b) two-step-ahead and (c) twenty-step-ahead conditional forecasts of the correlations of a bivariate VAR(1)-DCC(1, 1) model with $T = 1000$ and Student-7 errors.

where \widehat{a}_{T+h}^* are random draws with replacement from \widehat{F}_a and $\widehat{R}_{T+h|T}^* = \widehat{Q}_{T+h|T}^* \widehat{Q}_{T+h|T}^{*2} \widehat{Q}_{T+h|T}^{*3}$ with $\widehat{\varepsilon}_{T|T}^{s*} = \widehat{\varepsilon}_T^{s*}$ and $\widehat{Q}_{T|T}^* = \widehat{Q}_T^*$.

Step 6. Obtain bootstrap replicates of future errors ε_{T+h} and their conditional variances as follows:

$$\widehat{\sigma}_{i,T+h|T}^{2*} = \widehat{\omega}_i^* + \widehat{\alpha}_i^* \widehat{\varepsilon}_{i,T+h-1|T}^{*2} + \widehat{\beta}_i^* \widehat{\sigma}_{i,T+h-1|T}^{2*},$$

$$i = 1, \dots, N, t = 2, \dots, T, \quad (33)$$

$$\widehat{\varepsilon}_{i,T+h|T}^* = \widehat{\varepsilon}_{i,T+h|T}^{s*} \widehat{\sigma}_{i,T+h|T}^{*},$$

$$i = 1, \dots, N, t = 2, \dots, T, \quad (34)$$

where $\widehat{\sigma}_{i,T|T}^{2*} = \widehat{\sigma}_{i,T}^{2*}$ and $\widehat{\varepsilon}_{i,T|T}^{*2} = \widehat{\varepsilon}_{i,T}^{*2}$.

Step 7. The bootstrap replicate of the future value Y_{T+h} is generated by

$$\widehat{Y}_{T+h|T}^* = \widehat{\mu}^* + \widehat{\Phi}_1^* \widehat{Y}_{T+h-1|T}^* + \dots + \widehat{\Phi}_p^* \widehat{Y}_{T+h-p|T}^* + \widehat{\varepsilon}_{T+h|T}^*, \quad (35)$$

where $\widehat{Y}_{T-j|T}^* = Y_{T-j}$ for $j > 0$.

Step 8. Repeat steps 2–7 B times.

It is worth noting that the one-step-ahead bootstrap forecasts, $\widehat{Q}_{T+1|T}^*$ and $\widehat{\sigma}_{i,T+1|T}^{2*}$, incorporate only the parameter uncertainty, since the only components which vary from one bootstrap replicate to another are the bootstrap estimates of the parameters $[\widehat{\alpha}^*, \widehat{\beta}^*]$ and $[\widehat{\omega}_i^*, \widehat{\alpha}_i^*, \widehat{\beta}_i^*]$, while $\{\widehat{\varepsilon}_1, \dots, \widehat{\varepsilon}_T\}$ is kept fixed among all bootstrap replicates. As a consequence, all forecasts of returns, volatilities and correlations are conditional on the observations of the system at time T .

The bootstrap procedure for the DCC model is illustrated for a bivariate VAR(1)-DCC(1, 1) model with Student-7 errors, an intercept $\mu = (0, 0)'$, and autoregressive parameters given by $\text{vec}(\Phi_1) = (-0.5, 0, 0.5, 0.5)$, univariate GARCH parameters given by $(\omega_1, \alpha_1, \beta_1) =$

$(0.05, 0.05, 0.90)$ and $(\omega_2, \alpha_2, \beta_2) = (0.01, 0.10, 0.85)$, conditional correlation parameters $(\alpha, \beta) = (0.1, 0.88)$, and an unconditional correlation matrix $\text{vech}(Q) = (1, 0.5, 1)$. After generating a time series of size $T = 1000$ via this model, the bootstrap procedure is used to obtain forecast densities of future returns, volatilities and correlations. Here, we focus on forecast densities of the correlations, because the forecast densities of the levels and their conditional variances have been considered by Pascual et al. (2006) in a univariate context already. Fig. 4 plots h -step-ahead bootstrap densities for $h = 1, 2$ and 20, obtained for the conditional correlation, $\rho_{12,t}$, together with the corresponding empirical densities, where the latter have been obtained by simulating 2000 future values of the process. We can see that the one-step-ahead empirical density has all of its mass concentrated at a fixed point. The reason for this is that, in a DCC model, $\rho_{12,T+1}$ is observable with the information available at time T . However, by incorporating the parameter uncertainty involved in $\widehat{\rho}_{12,T+1}$, one can obtain an estimate of the density of $\rho_{12,t}$ which suggests that the one-step-ahead uncertainty can be rather large. Also note that the observed value of $\rho_{12,T+1}$ is very likely according to this bootstrap density. Panel (b) of Fig. 4 plots the two-step-ahead bootstrap density, together with the empirical density of the realizations of $\rho_{12,T+2}$. We can observe that, as expected, the dispersion of the bootstrap density is larger than that for $T + 1$. However, the empirical density of the two-step-ahead conditional correlation is still more concentrated in the center of the distribution than the bootstrap density. Finally, panel (c) of Fig. 4 plots the bootstrap and empirical densities for twenty-step-ahead conditional correlations. In this case, we observe that the bootstrap density is a good approximation of the empirical density. It is also important to point out that the uncertainty around $\rho_{12,T+20}$ is so large that

Table 1
Descriptive statistics of quarterly US inflation (π), unemployment (u) and GDP growth (g), observed from 1948Q1 to 2009Q3, with p -values in parentheses.

Series	Mean	Sd	Skewness	Kurtosis	KD	ADF ^a	Q(8)	Q ₂ (8)
π	0.91	0.83	0.79	5.43	1478.7 (0.00)	-3.66 (0.00)	130.09 (0.00)	129.45 (0.00)
u	0.56	0.15	0.15	3.45	33.26 (0.08)	-2.68 (0.07)	890.00 (0.00)	839.74 (0.00)
g	0.80	1.02	-0.10	4.20	19.85 (0.00)	-7.25 (0.00)	55.49 (0.00)	29.29 (0.00)
(π, u, g)			1.46	21.19	1669.4 (0.00)			

^a MacKinnon's p -value approximation.

the correlation could be zero or even negative. Note that the quality of the approximation of the bootstrap density to the empirical density improves as the forecast horizon increases. This is due to the fact that, as we forecast further into the future, the role played by the error uncertainty becomes more important, to the detriment of the parameter uncertainty. After all, this simulated example underscores the flexibility of the forward bootstrap procedure for dealing with more complicated models.

6. Empirical application

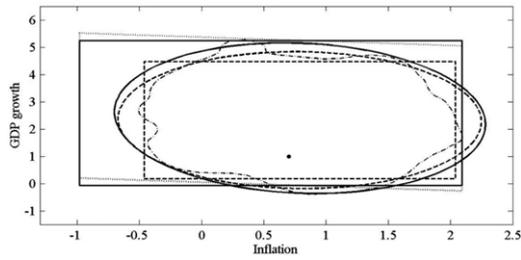
In this section, we use the forward bootstrap procedure to construct forecast densities of quarterly US inflation (π_t), the unemployment rate (u_t) and GDP growth (g_t), observed from 1948Q1 to 2011Q3.⁵ Inflation rates are computed as usual by $\pi_t = \log(IPI_t/IPI_{t-1}) \times 100$, where IPI is the Implicit Price Deflator. Unemployment is measured by the civilian unemployment rate. Finally, the GDP growth is given by $g_t = \log(GDP_t/GDP_{t-1}) \times 100$, where GDP is the Real Gross Domestic Product. Where relevant, monthly data have been transformed into quarterly data by taking the observations of the last month of the quarter. The whole sample period has been split into an estimation period from 1948Q1 to 2009Q3 ($T = 247$) and an out-of-sample period from 2009Q4 to 2011Q3. Table 1 reports the sample mean, standard deviation (sd), skewness and kurtosis of each of the series during the estimation period, together with the joint measures of skewness and kurtosis proposed by Mardia (1970). Table 1 also displays the normality test statistics and corresponding p -values based on the bootstrap procedure proposed by Kilian and Demiroglu (2000) and denoted by KD. The normality is always rejected either individually or jointly at the 10% level. Table 1 also displays the Augmented Dickey-Fuller (ADF) statistics, which reject the non-stationarity hypothesis for all series. Finally, the Box-Ljung statistics of order 8 for the original series and their squares, denoted by $Q(8)$ and $Q_2(8)$ respectively, are displayed in the last two columns of Table 1. We can observe that there is a dynamic dependence in the conditional mean. However, the Box-Ljung statistics of the squared observations are smaller than those of the levels, suggesting that the second order moments do not have significant dependence further to those generated by the conditional mean dependence. Hence, we fit a VAR model.

⁵ The data were obtained from the Federal Reserve Bank of St. Louis webpage: www.stlouisfed.org.

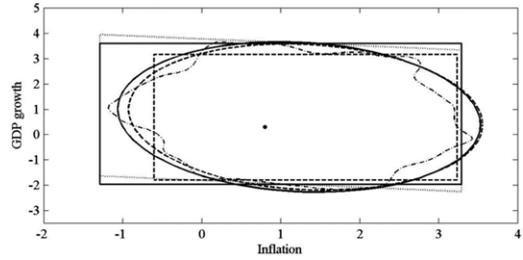
Following Kilian (1998a, 2001) and Marcellino, Stock, and Watson (2006), the lag order of the VAR is selected by the AIC, with the maximum lag order being equal to 14, which chooses $\hat{p} = 4$. Given that normality has been rejected, the traditional approach to forecasting using Gaussian densities may be misleading, and it is advisable to obtain bootstrap forecast densities. Consequently, the forward procedure with the bias and lag-order corrections is implemented for constructing out-of-sample bootstrap forecast densities for $h = 1, \dots, 8$.

For each of the three variables considered two-by-two, Fig. 5 plots 95% one-step-ahead bootstrap forecast ellipsoids and cubes for 2009Q4 (one step ahead) and 2011Q3 (eight steps ahead), together with the corresponding regions obtained under the assumption of Gaussian errors. Fig. 5 also plots the HDRs and the Bonferroni cubes which have been modified to take into account the correlations between the variables. In each case, the corresponding observed out-of-sample values are displayed by a dot. First, we can see that the Gaussian and bootstrap ellipsoids are rather similar, with the exception of the one-step-ahead GDP growth-unemployment ellipsoids, in which case the bootstrap is larger than that obtained under the assumption of Gaussianity. Also note that, when $h = 1$, the HDRs suggest non-elliptical densities, but when $h = 8$, HDRs are closer to the corresponding ellipsoids, suggesting that the Gaussianity may be plausible as the forecast horizon increases. Second, with respect to the Bonferroni cubes, note that the Gaussian cubes are much smaller than those obtained using the forward bootstrap procedure. Finally, when looking at the modified Bonferroni cubes, we can observe that, when $h = 1$, they only differ from the corresponding original cube in the case of the GDP growth-unemployment region. It is worth mentioning that, when looking at whether or not the regions contain the true observations, we observe that all regions contain them when $h = 8$. However, when $h = 1$, the GDP growth-unemployment observation falls outside the Gaussian regions, but is close to the boundary of the bootstrap ellipsoid and cube, though clearly within the bounds of modified Bonferroni cube.

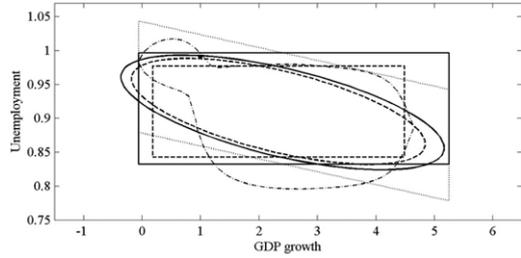
Finally, we compare the empirical coverages obtained when constructing the forecast densities assuming Gaussian errors with and without parameter uncertainty and when implementing the backward and forward bootstrap procedures with the bias and lag-order corrections. For this purpose, we carry out a rolling window estimation with $T = 100$ observations, starting with data from 1948Q1 to 1972Q3. For each estimation period, we construct joint



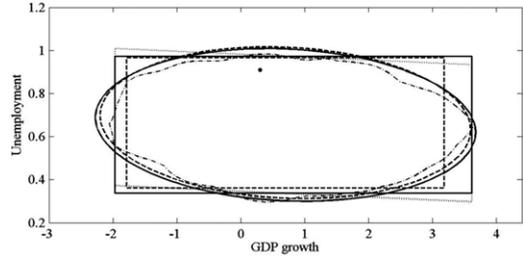
(a) Inflation-GDP growth, $h = 1$.



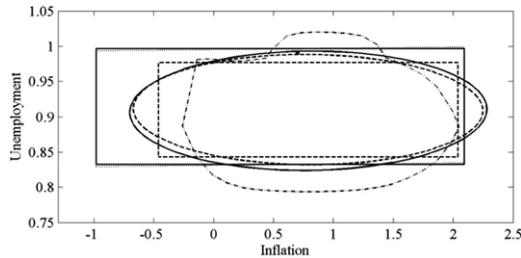
(b) Inflation-GDP growth, $h = 8$.



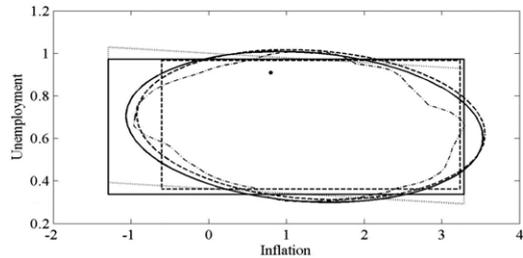
(c) GDP growth-Unemployment, $h = 1$.



(d) GDP growth-Unemployment, $h = 8$.



(e) Inflation-Unemployment, $h = 1$.



(f) Inflation-Unemployment, $h = 8$.

Fig. 5. 95% Gaussian ellipsoids and cubes (discontinuous lines), bootstrap ellipsoids and cubes (continuous lines), corrected bootstrap cubes (dotted lines) and HDRs (dotted-discontinuous lines) for one-step-ahead (first column) and eight-step-ahead (second column) forecasts of Inflation-GDP growth (first row), GDP growth-Unemployment (second row) and Inflation-Unemployment (third row).

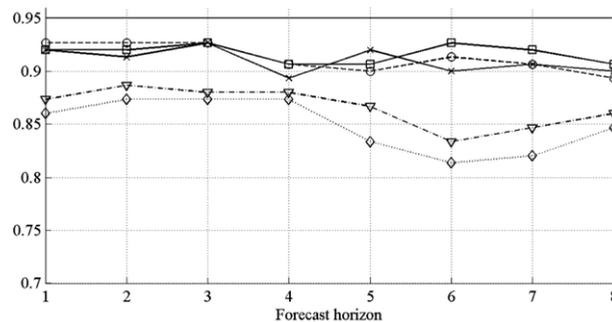


Fig. 6. Empirical coverages of Bonferroni cubes for US inflation, unemployment and GDP growth based on rolling window estimates with $T = 100$ observations: (i) Gaussian without parameter uncertainty (\diamond), (ii) Gaussian with asymptotic parameter uncertainty (∇), (iii) backward bootstrap (\circ), (iv) forward bootstrap (\square) and (v) corrected cube (\times). Nominal coverage: 95%.

out-of-sample 95% ellipsoids and Bonferroni cubes for the three variables in the system for $h = 1, \dots, 8$. Therefore, for each h and procedure, we obtain 148 out-of-sample regions. The empirical coverages for each h and procedure are calculated by counting how many regions contain the true out-of-sample observation. The results are plotted in

Fig. 6. First of all, we observe that the coverages of the Gaussian cubes are clearly smaller than those of the bootstrap ones. In concordance with the Monte Carlo results reported above, this behavior of the Gaussian regions could be expected, due to the non-Gaussianity of the forecast errors and the persistence of the estimated model. Note also

that the performances of all bootstrap regions are remarkably similar. It is obvious that there is no gain from using the BR. Finally, the coverages obtained when using the modified cube are similar to the original ones, though, as we have seen before, their shapes seem to be more informative. After all, the results reported in Fig. 6 are in line with those obtained using simulated data, in the sense that there are no large differences between using the forward bootstrap procedure and that based on the BR.

7. Conclusions

In this paper, we establish the asymptotic validity and analyze the finite sample performance of the multivariate extension of the bootstrap procedure proposed by Pascual et al. (2004a) for constructing forecast densities in multivariate VAR(p) models. The main attraction of the new bootstrap procedure is that it does not require the backward representation. As a result, we establish its asymptotic validity in non-Gaussian models. Furthermore, the forward procedure can be implemented in multivariate models without a BR, as we illustrate by computing forecast intervals for the conditional correlations in a DCC model. Finally, our Monte Carlo experiments show that the procedure works properly for incorporating the parameter uncertainty and is robust in the presence of non-Gaussian errors. When compared with alternatives, the forward bootstrap procedure is comparable with other bootstrap procedures based on the BR, and clearly better than procedures based on Gaussian forecast errors. On the other hand, the performance of the forward bootstrap procedure is comparable with the latter when the errors are Gaussian and the VAR model has roots close to the unit circle, and clearly better in the presence of highly persistent VAR models. Consequently, applied researchers are best to construct forecast densities using the forward bootstrap procedure, regardless of whether there is evidence of fat tails or skewness in the forecast error distribution. Our empirical example shows that there might be important differences between the Gaussian and bootstrap coverages. Furthermore, the forward procedure is computationally simple and asymptotically valid, and is consequently an attractive alternative when dealing with density forecasts in multivariate models.

The flexibility of the procedure proposed in this paper suggests its implementation in other multivariate models. For example, as Pascual, Romo, and Ruiz (2004b) proposed for univariate models, it can also be implemented to obtain forecast regions for the original observations when a VAR model is fitted to log-transformed observations; see Ariño and Franses (2000) and Bardsen and Lütkepohl (2011).

Finally, further effort should be directed into the construction of bootstrap forecast regions. In this sense, it is worth noting that the prediction ellipsoids are only appropriate when the distribution of the future values of the variables in the system is approximately multivariate Gaussian. When the distribution of Y_{T+h} departs from Gaussianity, the quality of such approximations deteriorates. The Bonferroni cubes do provide a better solution by capturing the asymmetry of the distribution. However, the shape of these cubes could not be appropriate when the

variables are correlated, and usually their volumes are too high. Consequently, it would be interesting to obtain regions that depart from either the elliptical or rectangular shapes.

Acknowledgments

The first two authors are grateful for financial support from projects ECO2009-08100 and ECO2012-32401 from the Spanish Government. We are also grateful to two anonymous referees and the editor George Athanasopoulos for very useful comments that led to a substantial improvement in the focus and presentation of this paper. Lutz Kilian helped with the DGP considered in the simulations and Jae Kim made some helpful comments on the construction of the ellipsoids and cubes. Finally, Gloria González-Rivera should also be acknowledged for her useful comments on an earlier version of this paper. The usual disclaimers apply.

Appendix

The models used in the Monte Carlo simulations are the following.

(a) Stationary VAR(5) model

The autoregressive matrices are given by $vec(\Phi_1) = (0.6337, -0.3424, 0.4688, 0.6755)$, $vec(\Phi_2) = (-0.0401, 0.3005, 0.0100, 0.1967)$, $vec(\Phi_3) = (0.4806, -0.5201, 0.2582, -0.2529)$, $vec(\Phi_4) = (0.2952, 0.2041, 0.2734, 0.0241)$ and $vec(\Phi_5) = (-0.2299, 0.1266, -0.1449, 0.3240)$, where vec is the column stacking operator. The dominant root of $|I_2 - \Phi_1(z^{-1}) - \dots - \Phi_5(z^{-5})| = 0$ is 0.89. The intercept is given by $\mu = (0.0053, 0.0018)'$. Finally, the contemporaneous covariance matrix of the error is given by $vech(\Sigma_\varepsilon) = 10^{-3} \times (0.5412, 0.4045, 0.4649)'$.

(b) Near-cointegrated VAR(10)

The autoregressive matrices are given by $vec(\Phi_1) = (1.1763, 0.0542, 0.7126, 1.2874)$, $vec(\Phi_2) = (-0.2624, -0.0668, -0.7618, -0.3596)$, $vec(\Phi_3) = (0.0250, -0.0160, 0.0305, -0.0168)$, $vec(\Phi_4) = (0.0000, 0.0939, 0.0830, -0.0510)$, $vec(\Phi_5) = (0.1485, -0.0512, 0.3678, 0.0656)$, $vec(\Phi_6) = (-0.2795, -0.0278, -0.7140, 0.1410)$, $vec(\Phi_7) = (0.2671, -0.1424, 0.3072, -0.3053)$, $vec(\Phi_8) = (-0.0786, 0.4398, 0.2189, 0.4045)$, $vec(\Phi_9) = (-0.1181, -0.3490, -0.1909, -0.2209)$, $vec(\Phi_{10}) = (0.0915, 0.0828, -0.0229, 0.0071)$. The dominant root of $|I_2 - \Phi_1(z^{-1}) - \dots - \Phi_{10}(z^{-10})| = 0$ is 0.97 and the intercept is given by $\mu = (0.1121, 0.1116)$. Finally, the $vech(\Sigma_\varepsilon) = 10^{-4} \times (0.025, 0.009, 0.387)$.

References

- Aielli, G. P. (2013). Dynamic conditional correlation: on properties and estimation. *Journal of Business and Economic Statistics*, 31(3), 282–289.
- Ariño, M. A., & Franses, P. H. (2000). Forecasting the levels of vector autoregressions log-transformed time series. *International Journal of Forecasting*, 16, 111–116.
- Athanasopoulos, G., & Vahid, F. (2008). VARMA versus VAR for macroeconomic forecasting. *International Journal of Forecasting*, 26(2), 237–252.
- Bache, I. W., Jore, S. A., Mitchell, J., & Vahey, S. P. (2011). Combining VAR and DSGE forecast densities. *Journal of Economic Dynamics and Control*, 35, 1659–1670.

- Baillie, R. T. (1979). Asymptotic prediction mean squared error for vector autoregressive models. *Biometrika*, 66(3), 675–678.
- Bardsen, G., & Lütkepohl, H. (2011). Forecasting levels of log variables in vector autoregressions. *International Journal of Forecasting*, 27(4), 1108–1115.
- Batchelor, R., Alizadeh, A., & Visvikis, I. (2007). Forecasting spot and forward prices in the international freight market. *International Journal of Forecasting*, 23(1), 101–104.
- Baumeister, C., & Kilian, L. (2012). Real-time forecasts of the real price of oil. *Journal of Business and Economic Statistics*, 30(2), 326–336.
- Berkowitz, J., & Kilian, L. (2000). Recent developments in bootstrapping time series. *Econometric Reviews*, 19(1), 1–48.
- Bickel, P., & Freedman, D. (1981). Some asymptotic theory for the bootstrap. *Annals of Statistics*, 9(6), 1196–1217.
- Bose, A. (1988). Edgeworth correction by bootstrap autoregression. *Annals of Statistics*, 16, 1709–1722.
- Chan, K.-S., Ho, L.-H., & Tong, H. (2006). A note on time-reversibility of multivariate linear processes. *Biometrika*, 93(1), 221–227.
- Clark, T. E. (2011). Real-time density forecasts from Bayesian vector autoregression with stochastic volatility. *Journal of Business and Economic Statistics*, 29(3), 327–341.
- D'Agostino, A., Gambetti, L., & Giannone, D. (2013). Macroeconomic forecasting and structural change. *Journal of Applied Econometrics*, 28(1), 82–101.
- Diebold, F. X., Hanh, J., & Tay, A. S. (1999). Multivariate density forecast evaluation and calibration in financial risk management high frequency returns on foreign exchange. *The Review of Economics and Statistics*, 81(4), 661–673.
- Diebold, F. X., & Yilmaz, K. (2009). Measuring financial asset return and volatility spillovers, with application to global equity markets. *Economic Journal*, 119(534), 158–171.
- Doan, T., Litterman, R., & Sims, C. (1984). Forecasting and conditional projection using realistic prior distributions. *Econometric Reviews*, 3(1), 1–100.
- Dufour, J. M., & Jouini, T. (2006). Finite-sample simulation-based tests in VAR models with application to Granger causality testing. *Journal of Econometrics*, 135(1–2), 229–254.
- Eklund, B. (2007). *Forecasting the Icelandic business cycle using vector autoregressive models*. Central Bank of Iceland, WP no. 36.
- Engle, R. (2002). Dynamic conditional correlation: a simple class of multivariate GARCH. *Journal of Business and Economic Statistics*, 20(3), 339–350.
- Engle, R. (2009). *Anticipating correlations: a new paradigm for risk management*. Princeton: Princeton University Press.
- Engle, R., & Sheppard, K. (2001). *Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH*. NBER Working Paper No. 8554.
- Fair, R. C., & Shiller, R. (1990). Comparing information in forecasts from econometric models. *The American Economic Review*, 80(3), 375–389.
- Freedman, D. (1984). On bootstrapping two-stage least-squares estimates in stationary linear models. *Annals of Statistics*, 12(3), 827–842.
- Fresoli, D. E., & Ruiz, E. (2014). *The uncertainty of conditional returns, volatilities and correlations in DCC models*. Working Paper 14-02 (02), Universidad Carlos III de Madrid.
- Grigoletto, M. (2005). Bootstrap prediction regions for multivariate autoregressive processes. *Statistical Methods and Applications*, 14, 179–207.
- Grigoletto, M. (2012). *Forecasting the distribution of aggregated time series: a bootstrap approach*. WP 2012/10, Department of Statistical Science, University of Padua.
- Guerbyenne, H., & Hamdi, F. (2014). Bootstrapping periodic state-space models. *Communications in Statistics—Simulation and Computation*, in press.
- Gupta, R., Kabundi, A., & Miller, S. M. (2011). Using large data sets to forecast house prices: a case study of twenty U. S. states. *Journal of Housing Research*, 20(2), 161–190.
- Hall, P. (1992). *The bootstrap and Edgeworth expansions*. New York: Springer-Verlag.
- Harvey, D. I., & Newbold, P. (2003). The non-normality of some macroeconomic forecast errors. *International Journal of Forecasting*, 19, 635–653.
- Holmes, S., Morris, C., Tibshirani, R., & Efron, B. (2003). Bradley Efron: a conversation with good friends. *Statistical Science*, 18(2), 268–281.
- Hyndman, R. J. (1996). Computing and graphing highest density regions. *The American Statistician*, 50(2), 120–126.
- Jore, A. S., Mitchell, J., & Vahey, S. P. (2010). Combining forecast densities from VARs with uncertain instabilities. *Journal of Applied Econometrics*, 25(4), 621–634.
- Kavussanos, M. G., & Visvikis, I. (2004). Market interactions in returns and volatilities between spot and forward shipping freight markets. *Journal of Banking and Finance*, 28(8), 2015–2047.
- Kilian, L. (1998a). Accounting for lag order uncertainty in autoregressions: the endogenous lag order bootstrap algorithm. *Journal of Time Series Analysis*, 19(5), 531–548.
- Kilian, L. (1998b). Confidence intervals for impulse responses under departures from normality. *Econometric Reviews*, 17, 1–29.
- Kilian, L. (1998c). Small-sample confidence intervals for impulse response functions. *The Review of Economics and Statistics*, 80(2), 218–230.
- Kilian, L. (1999). Finite-sample properties of percentile and percentile-t bootstrap confidence intervals for impulse responses. *Review of Economics and Statistics*, 81(4), 652–660.
- Kilian, L. (2001). Impulse response analysis in vector autoregressions with unknown lag order. *Journal of Forecasting*, 20, 161–179.
- Kilian, L., & Demiroglu, U. (2000). Residual-based tests for normality in autoregressions: asymptotic theory and simulation evidence. *Journal of Business and Economic Statistics*, 18(1), 40–50.
- Kilian, L., & Vigfusson, R. (2013). Do oil prices help forecast U. S. real GDP? The role of nonlinearities and asymmetries. *Journal of Business and Economic Statistics*, 31(1), 78–93.
- Kim, J. H. (1997). Relationship between the forward and backward representation of the stationary VAR model, Problem 97.5.2. *Econometric Theory*, 13, 899–990.
- Kim, J. H. (1998). Relationship between the forward and backward representation of the stationary VAR model, Solution 97.5.2. *Econometric Theory*, 14, 691–693.
- Kim, J. H. (1999). Asymptotic and bootstrap prediction regions for vector autoregression. *International Journal of Forecasting*, 15, 393–403.
- Kim, J. H. (2001). Bootstrap after bootstrap prediction intervals for autoregressive models. *Journal of Business and Economic Statistics*, 19(1), 117–128.
- Kim, J. H. (2004). Bias-corrected bootstrap prediction regions for vector autoregression. *Journal of Forecasting*, 23, 141–154.
- Klobner, S., & Wagner, S. (2014). Exploring all VAR orderings for calculating spillovers? Yes, we can! A note on Diebold and Yilmaz (2009). *Journal of Applied Econometrics*, 29(1), 172–179.
- Komunjer, I., & Owyang, M. T. (2012). Multivariate forecast evaluation and rationality testing. *Review of Economics and Statistics*, 94(4), 1066–1080.
- Lewis, R., & Reinsel, G. C. (1985). Prediction of multivariate time series by autoregressive model fitting. *Journal of Multivariate Analysis*, 16, 393–411.
- Lütkepohl, H. (1991). *Introduction to multiple time series analysis* (2nd ed.). Berlin: Springer-Verlag.
- Lütkepohl, H. (2006). Forecasting with VARMA models. In G. Elliot, C. W. J. Granger, & A. Timmerman (Eds.), *Handbook of economic forecasting*, Vol. 1 (pp. 287–325). Amsterdam: Elsevier, North-Holland.
- Lütkepohl, H., Staszewska-Bystrova, A., & Winker, P. (2015). Comparison of methods for constructing joint confidence bands for impulse response functions. *International Journal of Forecasting*, 31(3), 782–798.
- Marcellino, M., Stock, J. H., & Watson, M. W. (2003). Macroeconomic forecasting in the Euro area: country specific versus Euro wide information. *European Economic Review*, 47(1), 1–18.
- Marcellino, M., Stock, J. H., & Watson, M. W. (2006). A comparison of direct and iterated multistep AR methods for forecasting macroeconomics time series. *Journal of Econometrics*, 135(1), 499–526.
- Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, 57(3), 519–530.
- Pascual, L., Romo, J., & Ruiz, E. (2004a). Bootstrap predictive inference for ARIMA processes. *Journal of Time Series Analysis*, 25, 449–465.
- Pascual, L., Romo, J., & Ruiz, E. (2004b). Bootstrap prediction intervals for power-transformed time series. *International Journal of Forecasting*, 21(2), 219–235.
- Pascual, L., Romo, J., & Ruiz, E. (2006). Bootstrap prediction for returns and volatilities in GARCH models. *Computational Statistics and Data Analysis*, 50(9), 2293–2312.
- Pascual, L., Ruiz, E., & Fresoli, D. (2011). *Bootstrap forecast of multivariate VAR models without using the backward representation*. Working Paper 11-34, Universidad Carlos III de Madrid.
- Polito, V., & Wickens, M. (2012). Optimal monetary policy using unrestricted VAR. *Journal of Applied Econometrics*, 27(4), 525–553.
- Pope, A. L. (1990). Biases of estimators in multivariate non-Gaussian autoregressions. *Journal of Time Series Analysis*, 11, 249–258.
- Reeves, J. J. (2005). Bootstrap prediction intervals for ARCH models. *International Journal of Forecasting*, 21(2), 237–248.
- Reinsel, G. C. (1980). Asymptotic properties of prediction errors for the multivariate autoregressive model using estimated parameters. *Journal of the Royal Statistical Society*, 42(3), 328–333.
- Riise, T., & Tjøstheim, D. (1984). Theory and practice of multivariate ARMA forecasting. *Journal of Forecasting*, 3, 309–317.
- Runkle, D. E. (1987). Vector autoregressions and reality. *Journal of Business and Economic Statistics*, 5(4), 437–442.
- Sims, C. A. (1980). Macroeconomics and reality. *Econometrica*, 48(1), 1–48.

- Staszewska-Bystrova, A. (2011). Bootstrap prediction bands for forecast paths from vector autoregression models. *Journal of Forecasting*, 30(8), 721–735.
- Staszewska-Bystrova, A., & Winker, P. (2013). Constructing narrowest pathwise bootstrap prediction bands using threshold accepting. *International Journal of Forecasting*, 29(2), 221–233.
- Stine, J. H. (1987). Estimating properties of autoregressive forecasts. *Journal of the American Statistical Association*, 82(400), 1072–1078.
- Stock, J. H., & Watson, M. W. (2001). Vector autoregressions. *Journal of Economic Perspectives*, 15(4), 101–115.
- Thombs, L. A., & Schucany, W. R. (1990). Bootstrap prediction intervals for autoregression. *Journal of the American Statistical Association*, 85, 486–492.
- Tong, H., & Zhang, Z. (2005). On time-reversibility of multivariate linear processes. *Statistica Sinica*, 15, 495–504.
- Waggoner, D. F., & Zha, T. (1999). Conditional forecasts in dynamic multivariate models. *The Review of Economics and Statistics*, 81(4), 639–651.
- Wolf, M., & Wunderli, D. (2012). *Bootstrap joint prediction regions*. Working Paper No. 64, University of Zurich Department of Economics.