Universidad
Carlos III de Madrid
www.uc3m.es

TESIS DOCTORAL

Bootstrap forecasts of multivariate time series

Autor:

DIEGO E FRESOLI

Director/es:

ESTHER RUIZ

DEPARTAMENTO DE ESTADISTICA

Getafe, Junio 2014

University Carlos III of Madrid

PH.D. THESIS

Bootstrap forecasts of multivariate time series

Author:

Diego E Fresoli

Advisor:

Esther Ruiz

DEPARTMENT OF STATISTICS

Madrid, Spain, June 2014

A mi papá, Iván y Huguito.

Agradecimientos

No puedo sino empezar diciendo algo a Esther. Sin vos esta tesis y yo estarímos todavía andando sin rumbo. Estoy enormemente agradecido por serme de guía, por el crecimiento intelectual que pudieras haber propiciado en mí y tu apoyo incondicional durante estos años. Pero sobre todo estoy muy feliz de haberte conocido. Y sí, capaz esté perdiendo una directora de tesis pero ojalá esté ganando una persona para toda la vida.

Mi vieja, Norma, y mis hermanas, Pao y Luli, empujan mucho, todo el tiempo. Tienen una fuerza que se siente a un oceano de distancia. Sin ese empuje, yo no me muevo nada. Gracias por estar siempre y soportarme, los años de tesis y todos los otros que llevamos juntos. Y no puedo pasar por alto a Santi y Bianca, porque ellos sin saberlo, me vienen acompañdo mucho en este final de proceso.

Hay un grupo de personas sin quienes me es imposible imaginar mi vida pasada y actual en Madrid. A Pau, Peter, Jona, Guadi, Juancho fueron mi familia Madrileña por unos buenos y lindos años. A Raqui, Carlo, Berto y Maria, por muchas tardes de cañas y paseos para reponer energías. A las familias Sebastián Lago y Fernández Folgueiras que siempre me han hecho sentir en casa. A Sabi, Demian, Gabi, Nacho, Madda y Marianna, ustedes siempre están ahí para pasarla bien y también cuando lo necesito. Y a Uxía, por acompañarme días enteros este último tiempo y tirar mucho de mí.

Y tanta otra gente con la que he compartido las gracias y desgracias de la tesis: abuelos, tios, primos y amigos. A todos ustedes, que nos lo menciono por miedo a olvidar alguno, y mi adversión al riesgo me inmoviliza las manos, les debo mucho de todo esto.

"In god we trust, all others bring data."

Anonymous

Resumen

En esta tesis se estudia el desempeño de procedimientos que tienen por objetivo la aproximación de densidades de predicción y sus respectivos intervalos y regiones de confianza en series de tiempos multivariantes. En concreto, desarrollamos procedimientos bootstrap para predecir los modelos VAR y DCC, utilizados a menudo en la modelización y predicción de series temporales macroeconómicas y financieras. La metodología bootstrap analizada en esta tesis es atractiva debido a que no necesita supuestos distribucionales y es apropiada para incorporar la incertidumbre de los parámetros y del modelo.

En el Capítulo 1 se describen los modelos VAR y DCC y el enfoque tradicional para construir densidades de predicción con los mismos. Los problemas que surgen con este enfoque nos motiva a considerar alternativas, algunas de ellas basadas en bootstrapping. Para entonces será un momento propicio para presentar brevemente la metodología bootstrap en el marco de series de tiempo así como su aplicación en problemas de predicción.

En el Capítulo 2 se establece la validez asintótica y se analiza el desempeño en muestras pequeñas de un procedimiento bootstrap propuesto para construir densidades de predicción multivariante en el contexto de modelos VAR no Gausianos. Este procedimiento bootstrap no necesita de la representación backward usada por las alternativas existentes en la literatura y, por lo tanto, se puede utilizar para obtener densidades de predicción multivariante en, por ejemplo, modelos VARMA o VAR-GARCH. En el contexto de un VAR bivariado y estacionario, desarrollamos varios experimentos de Monte Carlo con el objetivo de estudiar sus propiedades en muestras pequeñas, obteniendo que las mismas son comparables a las de las alternativas basadas en la representación

backward. Por tanto, nada se pierde cuando se abandonan los procedimientos bootstrap que hacen uso the esta última. Este resultado también lo sugiere una de nuestras aplicaciones empíricas en la que construímos densidades de predicción conjunta para la inflación, el desempleo y el crecimiento del producto trimestrales de EEUU y sus correspondientes regiones de predicción con sus coberturas empíricas, éstas últimas obtanidas en base a un esquema rolling window. Por último, reproducimos un ejemplo de libro que utiliza la metodología de Gaussiana para predecir la inversión, el consumo y los ingresos en Alemania Occidental para después añadir los intervalos de predicción bootstrap con fines de comparación.

El modelo, los parámetros y la distribución del error rara vez son conocidos con exactitud por el usuario y, por tanto, la incertidumbre causada por la implementación del modelo estimado debe tenerse en cuenta la hora de hacer predicciones. Distintos métodos bootstrap se han desarrollado con éxito para hacer frente a las diferentes fuentes de incertidumbre en el contexto de los modelos VAR. Por esta razón, en el Capítulo 3 se compara el desempeño de las regiones de predicción construídas en base a la metodología tradicional Gaussiana y diversas variantes del procedimiento bootstrap que incorporan sucesivamente la incertidumbre en la distribución del error, del parámetro, la corrección por sesgo y la incertidumbre del modelo. Nuestro experimento de Monte Carlo sugiere que la incertidumbre de los parámetros juega un papel prepondetante cuando se predicen modelos VAR altamente persistentes.

En cuanto a las predicciones de los modelos DCC, hay dos problemas que merecen atención. En primer lugar, la no Gaussianidad de los rendimientos requiere de formas alternativas para aproximar sus densidades de predicción. En segundo lugar, sólo se pueden obtener predicciones puntuales de volatilidades, covarianzas y correlaciones en cada momento del tiempo. Por tanto, en el Capítulo 4 se presenta un procedimiento bootstrap para predecir los rendimientos, volatilidades, covarianzas y correlaciones en los modelos DCC. También llevamos a cabo simulaciones de Monte Carlo para evaluar en muestras pequeñas las propiedades del procedimiento propuesto. Los resultados muestran un buen desempeño con diferentes tamaños de muestras y distribuciones del error. Finalmente, empleamos el algoritmo bootstrap en dos sistemas de rendimien-

tos financieros. En primer lugar, construímos predicciones fuera de la muestra de rendimientos, volatilidades, covarianzas y correlaciones para un sistema de rendimientos de tipos de cambio diarios, Euro, Yen japonés y Dólar australiano respecto el Dólar de EE.UU. Asimismo, obtuvimos intervalos de predicción bootstrap, dentro de la muestra, de la correlación condicional de los rendimientos obtenidos con los índices de mercados S&P500 y NASDAQ. Ambas aplicaciones a datos reales sugieren que nuestro procedimiento brinda información adicional que enriquece la predicción de modelos DCC.

Por último, en el Capítulo 5 se concluye y presentan las líneas de investigación que quedan abiertas.

Abstract

In this thesis we study the performance of bootstrap procedures to approximate forecast densities and their intervals and regions for multivariate time series data. In particular, we develop bootstrap procedures for VAR and DCC models which are often implemented when modeling and forecasting macroeconomic and financial time series. The bootstrap methodology considered in this thesis is attractive since it is free of distributional assumptions and well suited to incorporate the parameter uncertainty and can be even designed to deal with the model uncertainty.

Chapter 1 introduces the VAR and DCC models and the standard approach to construct forecast densities within them. The problems that arise when forecasting these models push us to consider alternatives, some of them based on bootstrapping. By then, it will be a propitious time to briefly introduce the bootstrap methodology for time dependent data and review its implementation to forecast problems.

Chapter 2 establishes the asymptotic validity and analyses the finite sample performance of a simple bootstrap procedure to construct multi-step multivariate forecast densities in the context of non-Gaussian unrestricted VAR models. This bootstrap procedure avoids the backward representation used by existing alternatives and, consequently, can be implemented to obtain multivariate forecast densities in, for example, VARMA or VAR-GARCH models. In the context of bivariate stationary VAR models, we carry out several Monte Carlo experiments to study its finite sample properties, finding that these are comparable to those of alternatives based on the backward representation. Hence we remark that nothing is lost when we abandon the more complicated backward representation. This result is also suggested by one of the empirical appli-

cations, in which we construct joint forecast densities of US quarterly inflation, unemployment and GDP growth and the corresponding forecast regions with their empirical coverages obtained using a rolling window scheme. Finally, we reproduce a textbook example that applies the standard Gaussian methodology to forecast West German investment, consumption and income and then, for comparison purposes, we include the bootstrap forecast intervals.

The model, the parameters and the error distribution are rarely known without uncertainty by the forecaster and, thus, the sampling variability caused by the use of the estimated model needs to be taken into account. Bootstrap methods are successfully designed to deal with different sources of uncertainties in the context of forecasting VAR models. For this reason, Chapter 3 compares the forecast performance of the regions constructed using the traditional Gaussian methodology and several variants of the bootstrap procedure that successively incorporate error distribution, parameter uncertainty, bias correction and lag order uncertainty. Our Monte Carlo study suggests that the parameter uncertainty plays a prominent role when forecasting highly persistent VAR models.

Regarding DCC forecast, there are two problem that deserve attention. First, the non-Gaussianity of returns demands alternative ways of approximating its forecast density. Second, only point forecasts of volatilities, covariances and correlations can be obtained at each moment of time. These issues encouraged us to developed bootstrap procedure to forecast returns, volatilities, covariances and correlations in corrected DCC models, which is described in Chapter 4. We conduct Monte Carlo simulations in order to evaluate its finite sample properties, which show a rather good performance of the bootstrap procedure under different sample sizes and error distributions. We apply the proposed bootstrap algorithm to two systems of returns. First, we obtain out-of-sample forecast of returns, volatilities, covariances and correlations in the context of a system of daily exchange rates returns of the Euro, Japanese Yen and Australian Dollar against the US Dollar. Second, we construct within sample forecast intervals of the conditional correlation of S&P500 and NASDAQ returns. Both empirical applications point out that our bootstrap algorithm can provide additional information that enriches the DCC forecast approach.

Finally, Chapter 5 concludes and presents the research lines that are still open.

Contents

List of Figures

List of Tables

Chapter 1

Forecasting multivariate time series

1.1 Introduction

The main topic of this thesis is forecasting multivariate densities in systems of economics and financial variables. Increasingly, analysts are interested in assessing the uncertainty around a point forecasts by obtaining the full probability distribution. Furthermore, multivariate forecast densities capture the dynamic interrelationships among variables in the system. For example, nowadays it is common practice by Central Banks to construct forecast densities of systems containing both policy goal and instruments variables; see, for instance, the Bank of England inflation reports which focus on forecast densities of inflation, gross domestic product and interest rates. Likewise, risk managers rely on forecast densities of returns portfolios and since the simultaneous risk is tied to certain features of the joint distribution; see, for instance, Prékopa (2012) who highlights the relevance of obtaining multivariate forecast densities of portfolios for the implementation of multivariate Value-at-Risk. These examples, although simple, are eloquent of the increasing importance of multivariate forecast densities for many applications in economics and finance; see Tay and Wallis (2000) and Corradi and Swanson (2006) for complete surveys of forecast densities with their applications.

For multivariate time series, one essential tool to forecast macroeconomic and financial series

are Vector Autoregressive (VAR) models; see Stock and Watson (2001) who highlight the ability of VAR models specially in data description and forecasting tasks. Traditionally, the forecast density of future values of the variables in the system for a given horizon is taken to be Gaussian with mean and variance equal to the point forecast and Mean Square Forecast Error, respectively. Of course, these two moments depend on unknown parameters which are substituted, in practice, by their estimates. For forecast purposes, this is particularly relevant since the sampling variability can be understated if the parameter uncertainty is ignored, especially when the system is large relative to the sample size. Furthermore, the appropriateness of the Gaussianity assumption is often questioned when forecasting macroeconomic and financial series; see, for instance, Lam and Veall (2002) who illustrate how the standard formula behaves when the errors depart from Gaussianity, noticing a bad performance that does not improve even in large samples. Consequently, the Gaussian forecast density and its corresponding forecast regions could be misleading to describe what analysts could expect in the future. Furthermore, the true model is usually unknown by analysts who rely on data to approximate it. The question then arises as to whether consider the approximated model as it were the true one and proceed to forecast or, alternatively, to incorporate the variability due to model uncertainty into the forecast method. Moreover, the parameters estimates may be subject to small sample biases, especially in the case of highly persistent autoregressive models, an issue that should not be ignored since it may lead to a very poor forecast performance of large macroeconomic systems; see Tjøstheim and Paulsen (1983).

In the context of financial time series, it is of interest to forecast conditional means but also conditional variances and covariances. One of the most popular models implemented with this purpose are Multivariate Generalized Autoregressive Conditional Heteroscedasticity (MGARCH) models, which are crucial for financial market participants when dealing with, for example, risk management, derivative pricing models, hedging strategies or portfolio allocation models. For many MGARCH models, if the errors distribution is known, then the distribution of returns for the next time period is fully characterized. However, due to the non-linearity of returns, it is impossible to compute analytically the density forecast when we project more further into the future.

On the other hand, by using the standard approach, only point forecasts of conditional volatilities, covariances and correlations can be provided, so that efforts need to be directed to alternative methods of approximating their uncertainties. This is particularly important if users demand conditional correlations because, unlike the case with the conditional variances and covariances, their point forecasts are biased; see Engle (2009). Despite these issues regarding forecasts of MGARCH models, it is striking that there have been no attempts in literature to implement alternative methods to approximate the density forecasts of non-Gaussian returns and its second moments.

Several authors propose obtaining forecast densities using Bayesian procedures which are attractive because they not only generate directly forecast densities that incorporate the parameter uncertainty, but can also be designed to reduce this uncertainty; see, for instance, Koop (2013) and Wright (2013) just to mention a few recent references. However, Bayesian forecast densities also rely on the assumption of Gaussian errors which, as mentioned above, can be violated when dealing with real macroeconomic and financial systems. Alternatively, in the last decades there has been an increasing interest in procedures that allow approximating the conditional probability distribution of a multivariate system at a given future period without imposing any assumption on the distribution of the errors while simultaneously incorporating the parameter uncertainty. Some of them are based on bootstrap techniques.

The bootstrap methodology was introduced by Efron (1979) who proposed it as an alternative to assess the statistical accuracy or to approximate the sampling distribution of an estimator in the case of independent and identically distributed (i.i.d.) data. Thenceforth, taking advantage of the the fast computers era, the bootstrap has found scope in many different statistical problems; see Efron and Tibshirani (1993) for a broad review about the bootstrap and its scope. In particular, Freedman (1981) developed the first application of bootstrap to regression analysis, which facilitated its posterior use for dependent data, either for estimation or forecast problems. Bootstrap procedures have several advantages in the context of forecasting. First, they account for possible non-Normality of the errors, providing yet asymptotically valid results for the sampling distribution of future values of the series. Second, bootstrap forecast intervals are well suited

to incorporate the sampling variability due to parameters uncertainty. In addition, they can be designed to tackle small-sample biases and the sampling variability due to model uncertainty. Finally, bootstrap confidence intervals have shown to give valid results in those cases in which Normal forecast intervals only serve as rough approximations, due to the lack of justification, such as GARCH models.

In multivariate VAR setting, bootstrap procedures have been implemented to obtain forecast intervals for a single variable or even forecast path or regions for a set of variables; see Kim (1999, 2001, 2004) and Grigoletto (2005), Staszewska-Bystrova (2011). These proposals are based on the backward representation. Pascual et al. (2004) show, for univariate ARIMA models, that this backward representation is not needed and only adds computational complexity. Furthermore, when using the backward representation, the asymptotic validity of the bootstrap can only be proved assuming Gaussian errors. In this thesis we extend the bootstrap procure of Pascual et al. (2004) to multivariate VAR(p) models to obtain the forecast densities and their corresponding regions. The main advantages of this extension is that it does not rely on the backward representation, thus avoiding its high computational costs, and that its asymptotic validity is not grounding on the Gaussianity assumption. Moreover, it can deal with models that do not have a backward representation such as VAR models with MGARCH errors. In particular, the bootstrap procedure is adapted to forecast returns, volatilities, covariances and correlations in the context of Dynamic Conditional Correlation models.

The rest of this chapter is organized as follows. In Section 2, we describe the VAR(p) model and the issue of forecasting within this models. In Section 3, we introduce the Dynamic Conditional Correlation (DCC) model and how it can be implemented to forecast returns and correlations. Section 4 discusses the idea of bootstrapping time dependent data and its implementation to forecast problems. Finally, Section 5 summaries the objectives of this thesis.

1.2 Forecasting with VAR models

In this section we briefly describe the forecast problem in a VAR framework.

1.2.1 Description of VAR models

Consider a multivariate discrete stochastic process y_t of dimension N and assume it has the following VAR(p) representation

$$y_t = \mu + \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + ... + \Phi_p y_{t-p} + \varepsilon_t, \quad t = ..., -1, 0, 1, 2... , \tag{1.1}$$

where Φ_j are $N \times N$ matrices of coefficients, μ is a $N \times 1$ vector containing the intercept terms, ε_t is a vector of i.i.d innovations such that $E(\varepsilon_t) = 0$, $E(\varepsilon_t \varepsilon_t') = \Sigma_\varepsilon$ is a $N \times N$ symmetric positive definite matrix with finite elements, and $E(\varepsilon_t \varepsilon_s') = 0$ for all $t \neq s$. The model is stationary so that all the roots of the characteristic equation $det(I_N - \Phi_1 L - \Phi_2 L^2 - ... - \Phi_p L^p) = 0$ lie outside the unit circle, where I_N is a $N \times N$ identity matrix and $det(\cdot)$ denotes the determinant of a matrix. Stationary assures that the VAR process has the following moving average representation (MA)

$$y_t = \sum_{j=0}^{\infty} \Pi_j \varepsilon_{t-j} \tag{1.2}$$

where $\Pi_0 = I_N$ and $\Pi_j = \sum_{i=1}^{j} \Pi_{j-i} \Phi_i$ for $i = 1, 2, ...$

Let y_{T+h} for $h > 0$ be a future value of the process and $\{y_1, ..., y_T\}$ the information set available at time T. Given $\{y_1, ..., y_T\}$, the future of the process, y_{T+h}, is fully characterized by its conditional multivariate distribution function $F_T(\cdot)$. When focusing on a point forecast, if the criteria is to minimize the Mean Square Forecast Error (MSFE) then the optimal forecast turns to be the conditional expectation at T, which is given by

$$y_{T+h|T} = E_T\left(y_{T+h}\right) = \mu + \Phi_1 y_{T+h-1|T} + \Phi_2 y_{T+h-2|T} + ... + \Phi_p y_{t+h-p|T} \tag{1.3}$$

where $y_{T+j|T} = y_{T+j}$ for $j \leq 0$.[1] Note that, if the errors ε_t are not independent but just uncorrelated, then in general $E_T(\varepsilon_{T+h}) \neq 0$ and thus the expression in (1.3) is not the conditional mean, even though it is still the best linear forecast. Henceforth, expression (1.3) will be referred as the optimal predictor irrespective of the properties of the white noise ε_t; see Lütkepohl (1991). The corresponding MSFE is given by

$$\Sigma_y(h) = E_T[(y_{T+h} - y_{T+h|T})(y_{T+h} - y_{T+h|T})'] = \sum_{j=0}^{h-1} \Pi_j \Sigma_\varepsilon \Pi_j'. \tag{1.4}$$

If ε_t is further assumed to be Gaussian, then the h-steps-ahead forecast error

$$\epsilon_{T+h} = y_{T+h} - y_{T+h|T} = \sum_{j=0}^{h-1} \Pi_j \varepsilon_{T+h-j} \tag{1.5}$$

is also Gaussian with zero mean and variance $\Sigma_y(h)$. Consequently,

$$y_{T+h} \sim N(y_{T+h|T}, \Sigma_y(h)). \tag{1.6}$$

This Gaussian distribution can be used to construct h-steps- ahead forecast regions with a predetermined probability content. However, in practice, it is not possible to obtain the forecast densities in (1.6) given that the autoregressive order p and the model parameters are unknown. Therefore, when dealing with real data, one substitutes p by the selected order, $\widehat{p}$, and the parameters by estimates. In this thesis, we consider the AIC to choose p and the LS estimator of the VAR parameters. As a result, the estimated h-steps-ahead forecast of y_{T+h} is given by

$$\widehat{y}_{T+h|T} = \widehat{\mu} + \sum_{j=1}^{\widehat{p}} \widehat{\Phi}_j y_{T+h-j|T}, \tag{1.7}$$

where $\widehat{y}_{T+j|T} = y_{T+j}$ for $j \leq 0$. The standard textbook approach is to construct the forecast

[1] The subscript in the expectation operator denotes condition on the data set available, $\{y_1, ..., y_T\}$.

density assuming Gaussian errors as follows

$$y_{T+h} \sim N(\widehat{y}_{T+h|T}, \widehat{\Sigma}_y(h)). \tag{1.8}$$

where $\widehat{\Sigma}_y(h)$ is obtained as (2.3) substituting the matrices Π_j obtained from the estimated model. The Gaussian forecast densities can be inadequate to construct forecast regions for several reasons. First, it is well known that the LS estimator is biased; see Tjøstheim and Paulsen (1983). Therefore, this bias affects both $\widehat{y}_{T+h|T}$ and $\widehat{\Sigma}_y(h)$. The biased parameters can be corrected using, for example, the formula proposed by Pope (1990). Furthermore, $\widehat{\Sigma}_y(h)$ is not the true MSFE of $\widehat{y}_{T+h|T}$ because it does not incorporate the uncertainty associated with the selection of $\widehat{p}$ and with the estimation of the parameters. Finally, the Gaussianity assumption is often violated in macroeconomic and financial systems.

The sampling variability due to the LS estimator can be incorporated by following the arguments by Lütkepohl (1991). First, assume that the estimation is done using a realization of the multivariate process independent of that used to forecast and with the same stochastic structure. Then it is possible to incorporate the parameter uncertainty by approximating the sample distribution of the LS estimator by its asymptotic distribution. In this case, the MSFE matrix with asymptotic parameter uncertainty is given by

$$\Sigma_y^A(h) = \Sigma_y(h) + \frac{1}{T}\Omega(h) \tag{1.9}$$

where $\Omega(h) = E\left[\frac{\partial y_{T+h}}{\partial \beta'}\Sigma_\beta \frac{\partial y'_{T+h}}{\partial \beta}\right]$ with β being the vector of unknown parameters in (1.1); see Lütkepohl (1991). Note that, as this correction implies just an additional term in the MSFE matrix, the Gaussianity of y_{T+h} still holds if the errors are assumed to be Gaussian.

For the time being, it is worth noting that the traditional approach to forecast depends crucially on the Gaussianity of the errors. However, in fields such as macroeconomics and finance, several authors have found evidence of departure from Gaussianity; see, for instance, Kilian (1998b) and Harvey and Newbold (2003). Therefore, the Gaussianity assumption of y_{T+h} may

be misleading and give a poor depiction of the future density. Moreover, Lütkepohl (1991)'s formula seems to understate the sampling variability due to parameter uncertainty in finite sample, especially in highly persistent models; see Kim (1999). For this reason, in this thesis we consider the implementation of resampling methods as the bootstrap since it is completely free of distributional assumptions and well suited to deal with the parameter uncertainty in small samples.

Next, we discuss the bias-correction formula and the way of selecting the lag order implemented in this thesis.

1.2.2 Bias-correction

The LS estimator is biased in small sample, specially when the VAR system is highly persistent; see, for instance, Tjøstheim and Paulsen (1983), Kilian (1998c), Kim (2004) and Engsted and Pedersen (2011), among others. Of course, forecast densities constructed using parameters estimates can be inadequate to properly represent the future projection of the system if they are distorted due to bias. Furthermore, parameter bias can induce a double bias to bootstrap algorithms since the the bootstrap estimates of the parameters may also be subject to bias; see Clements and Kim (2007). For these reasons, it might be desirable to introduce a method for small sample bias correction. There are many alternative to incorporate bias-correction in a VAR setting. For example, one may used bootstrap methods as that proposed by Kilian (1998c) or an analytic formula as those proposed by Yamamoto and Kunitomo (1984) or Pope (1990). Here we describe the analytic formula proposed by Pope (1990) because there is evidence that bootstrap based bias-correction do not provide any advantage over the analytic formulas and, that the expression due to Yamamoto and Kunitomo (1984) is equivalent to Pope's one; see for details Engsted and Pedersen (2011).

To describe Pope's analytic formula, consider the VAR(p) model in (1.1) which can alternatively be expressed as a VAR(1) without intercept, $(Y_t - \mu_Y) = \Phi (Y_{t-1} - \mu_Y) + \epsilon_t$, where

$Y_t = [y_t', ..., y_{t-p+1}']'$, $\mu_Y = \left[\mu_y', ..., \mu_y'\right]'$, $\epsilon_t = [\varepsilon_t, 0, ..., 0]'$ are $Np \times 1$ vectors and

$$\Phi = \begin{pmatrix} \Phi_* & \Phi_p \\ I_{N(p-1)} & \mathbf{0} \end{pmatrix}$$

with $\mu_y = (I_N - \Phi_1 - \cdots - \Phi_p)^{-1}\mu$ being a $N \times 1$ vector of unconditional means, $\mathbf{0}$ being a $N(p-1) \times p$ matrix of zeros and $\Phi_* = [\Phi_1, ..., \Phi_{p-1}]$ being a $N \times N(p-1)$ matrix. Pope (1990) shows that the small-sample bias of the LS estimator of the autoregressive parameters has a closed form expression given by

$$\Omega_\Phi = E(\Phi) - \Phi = \frac{-b_\Phi}{T} + O(T^{-3/2}) \tag{1.10}$$

where

$$b_\Phi = J[(I - \Phi')^{-1} + \Phi\prime(I - \Phi'^2)^{-1} + \sum_{\lambda \in Spec(\Phi)} \lambda(I - \lambda\Phi')^{-1}]\Gamma(0)^{-1},$$

$\Gamma(0) = E\left[(Y_t - \mu_y)(Y_t - \mu_y)'\right]$, $J = E(\epsilon_t\epsilon_t')$ and $Spec(\Phi)$ is the set of eigenvalues of the matrix Φ. This formula for the bias can be implemented to obtain bias-corrected parameters, as follows

$$\Phi^c = \Phi - \Omega_\Phi. \tag{1.11}$$

The analytic formula in (1.10) has some features that are worth mentioning. First, note that it only provides a close expression for the bias in the autoregressive matrices, notwithstanding bias-correction of the intercept can be obtained by $(I_N - \Phi_1^c\mu_y - ... - \Phi_p^c)\mu_y$. Second, the error in (1.10) goes to zero at a rate $T^{-3/2}$, so it is negligible even for moderate sample sizes. Third, since the bias formula depends on unknown parameters, in practice they are replaced by their LS estimates and higher order terms are ignored. The last issue to bring out is that the bias correction can push the corrected parameters into the nonstationary region and, thus, an additional adjustment is needed in order to prevent this from happening. One method widely implemented for this purpose is that proposed by Kilian (1998a,c).

1.2.3 Lag order estimation

Up to now, we have assumed that the lag order p is known. However, in practice when dealing with real data, the analysts is unlikely to know the true lag order p and will estimate it by using an Information Criteria (IC). These IC select the lag order that minimizes the following expression

$$IC(\widehat{p}) = ln\left(|\Sigma_\varepsilon(\widehat{p})|\right) + P_T\widehat{p} \quad 0 \leq \widehat{p} \leq p_u\,, \tag{1.12}$$

where $\Sigma_\varepsilon(\widehat{p}) = \sum_{t=\widehat{p}+1}^{T} \varepsilon_t(\widehat{p})\varepsilon_t(\widehat{p})'/(T)$ with

$$\varepsilon_t(\widehat{p}) = y_t - \mu - \Phi_1 y_{t-1} - ... - \Phi_{\widehat{p}} y_{t-\widehat{p}}.$$

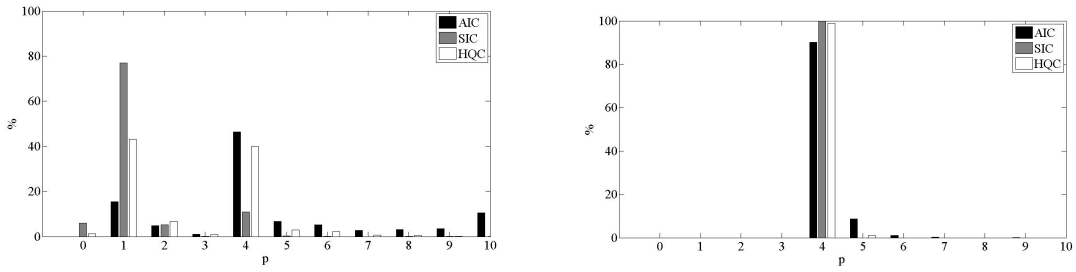
P_T is a function that depends on N and T, and playing the role of penalizing large $\widehat{p}$, and p_u is the maximum lag order. The Aikaike (AIC), Hannan-Quinn (HQC) and Schwarz (SIC) differ in the expression given to P_T. The AIC considers it equals to $2N^2/T$, HQC to $2N^2 ln[ln(T)]/T$ and SIC to $N^2 ln(T)/T$; see Lütkepohl (1991) for details.

HQC and SIC are strongly consistent while AIC tends to overestimate the true lag order asymptotically; see Lütkepohl (1991). However, Kilian (1998a) finds that, when the sample size is small or moderate, HQC and SIC tend to underestimate the true lag order more frequently than the AIC. With the purpose of checking this fact, we simulate 1000 series for a stationary VAR(4) with Gaussian errors and $T = 50$ and $T = 300$. [2] Then, for each of the simulated series, we select the lag order using the HQC, SIC and AIC with $0 \leq \widehat{p} \leq 10$. Figure 1.1, which plots the relative frequency of estimated lag orders, shows that, when the sample size is small, HQC and SIC select more frequently lag orders below 4 than AIC. On the other hand, when the sample size is large, all the criteria perform reasonable well, though AIC is more likely to choose a lag order greater than 4, with a relative frequency of about 0.10. As regards forecasting, the lag order underestimation may be more harmful than its overestimation and, consequently, AIC shows up to be preferable;

[2] Details about this VAR(4) model can be found in the Appendix A, model (a).

see, for example, Müller and Stock (2011). Furthermore, Burnham and Anderson (2004) point out that AIC works better in those cases in which there is not a "true" model.

Beyond the procedure used to estimate the lag order, for our purpose is crucial to recognize that most of the approaches to forecast VAR models are conditioned on the estimated lag order which is treated as it were the true unknown lag order. As a result, these forecasting approaches do not tackle the variability that can be attributed to the model uncertainty. In contrast, in this thesis we apply bootstrap algorithm to incorporate the lag order variability.



(a) $T = 50$ (b) $T = 300$

Figure 1.1: Relative frequency (in percent) of estimated lag order using AIC, HQC and SIC and $0 \leq \widehat{p} \leq 10$, based on 1000 replicates of a stationary VAR(4) model with Gaussian errors and samples sizes $T = 50$ and 300.

1.3 Forecasting with DCC models

MGARCH models are designed to model the second order moments of multivariate time series and, consequently, they play an important role in, for example, asset pricing, risk management and portfolio allocation. One of the most popular MGARCH is the DCC model. In this section, we briefly describe the DCC and how it can be implemented to forecast returns and correlations.

1.3.1 Description of DCC models

We consider a vector of returns y_t which is given by

$$y_t = H_t^{1/2} a_t \tag{1.13}$$

where y_t is a $N \times 1$ vector , a_t is a $N \times 1$ independent white noise vector with identity covariance matrix, H_t is a $N \times N$ positive definite conditional covariance matrix.

Many MGARCH models have been proposed in the literature with different specifications of the conditional covariance matrix in (1.13). Some of them focus directly on the covariance matrix as, for example, the Diagonal VEC of Bollerslev et al. (1988) or the BEKK of Engle and Kroner (1995), while others are based on the decomposition of the conditional covariance matrix into conditional standard deviations and dynamic conditional correlations as, for example, the Constant Conditional Correlation (CCC) model of Bollerslev (1990) and DCC model of Engle (2002); see Bauwens et al. (2006) and Silvennoinen and Teräsvirta (2009) for comprehensive surveys on these parametric MGARCH models and their extensions. Among these alternative parametric specification of the covariance matrix, the most-widely implemented in applied research is the DCC model.

The DCC model makes use of the following decomposition of the covariance matrix

$$H_t = D_t R_t D_t \tag{1.14}$$

where $D_t = [\sigma_{i,t}]$, $i = 1, ...N$, is a $N \times N$ diagonal matrix containing the univariate GARCH-type conditional standard deviations of each of the variables in y_t and $R_t = [\rho_{ij,t}]$, $i, j = 1, ...N$, is the $N \times N$ matrix of conditional correlations. On the other hand, in its simplest scalar version, the DCC model assumes that the correlation matrix R_t in the system evolve over time in a fashion similar to that of univariate GARCH conditional variances, according to the following dynamic

process

$$R_t = dg(Q_t)^{-\frac{1}{2}} Q_t dg(Q_t)^{-\frac{1}{2}} \tag{1.15}$$

where $dg(Q_t)$ gives a diagonal matrix with the diagonal elements of Q_t. In the scalar DCC model, Q_t is given by

$$Q_t = (1 - \alpha - \beta)S + \alpha\varepsilon_{t-1}\varepsilon'_{t-1} + \beta Q_{t-1} \tag{1.16}$$

where S is a $N \times N$ matrix of constants, $\varepsilon_t = D_t^{-1} y_t$ are the standardized errors and α and β are scalars. Positiveness of Q_t and R_t is guaranteed if S is positive definite together with α, $\beta > 0$ and $\alpha + \beta < 1$; see Ding and Engle (2001). Note that the scalar DCC model in (1.16) restricts the dynamics of all the correlations to be governed by the same parameters. Furthermore, as it happens with the conditional variances, the conditional correlations are observable one-step-ahead.

DCC models own their popularity to the fact that they can be easily estimated by a two-steps target correlation procedure proposed by Engle and Sheppard (2001) and Engle (2002), allowing their implementation to estimate the volatilities and correlations of very large systems of financial returns as those often encountered in portfolio management problems. However, as pointed out by, for example, Caporin and McAleer (2012), the basic scalar DCC model has some pitfalls. One has to do with the estimator proposed by Engle and Sheppard (2001) and Engle (2002) which is not consistent; see Caporin and McAleer (2012) and Aielli (2013). Nevertheless, the correlation driving process in the DCC model can be reformulated in order to assure a consistent estimator, giving rise to the corrected DCC (cDCC) model; see Aielli (2013). Moreover, the basic DCC model has been criticized because it does not deal with many empirical features usually observed in financial returns. For this reason, it has been extended to consider asset-specific features in the correlation dynamics; see, for instance, Cappiello et al. (2006), Billio et al. (2006) and Hafner and Franses (2009). Also, Cappiello et al. (2006) and Audrino and Trojani (2011) propose extensions to account for asymmetry in correlations which often react more abruptly to joint bad news than to good ones and Nakatani and Teräsvirta (2009) and Kasch and Caporin (2012) incorporate feed-

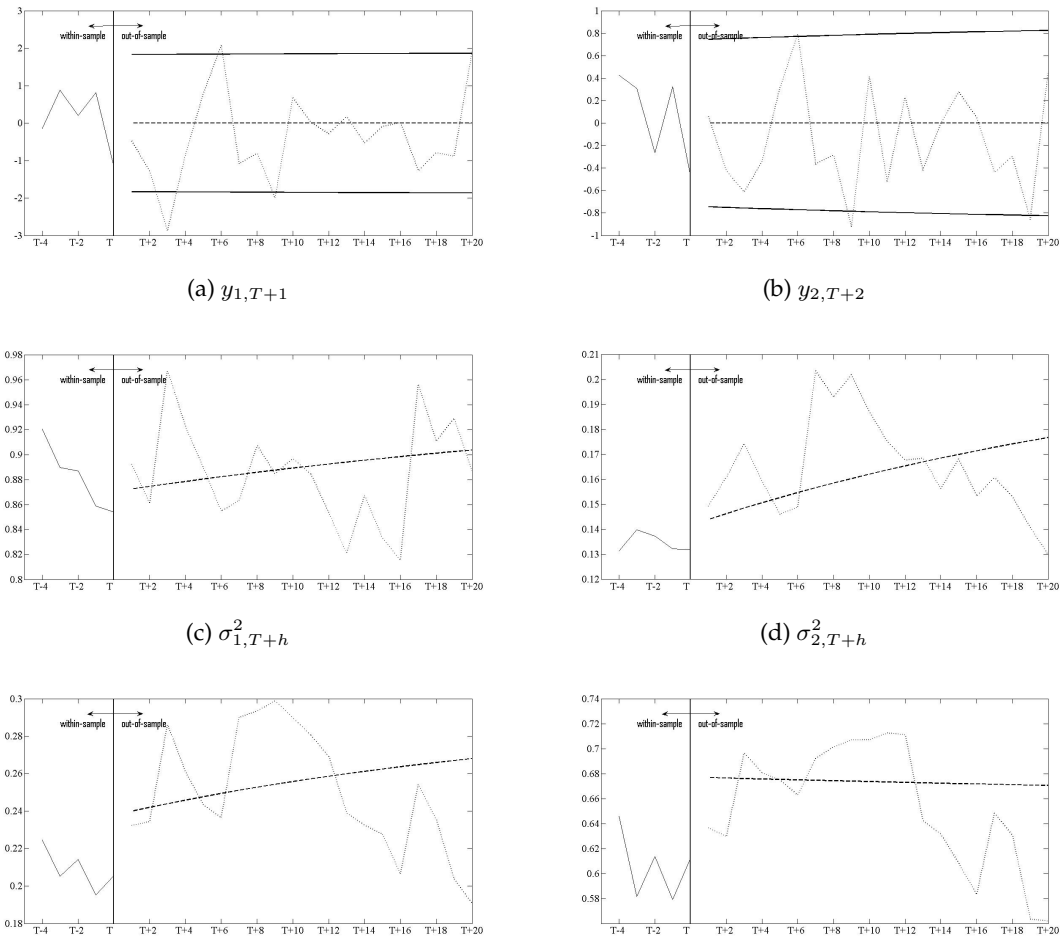
back among volatilities and between volatilities and correlations. Finally, focusing on the variance side, added flexibility has recently been proposed by Aielli and Caporin (in press).

Given the parameters, the DCC model could be used to forecast correlations. As the correlation driving process is non-linear, (1.16) cannot be solved forward to construct exact forecasts. For this reason, Engle and Sheppard (2001) assume that that $E_T[\varepsilon_{t+h}\varepsilon'_{t+h}] \approx E_T[Q_{t+h}]$, in which case the following forecast of R_{T+h} is obtained

$$R_{T+h|T} = (1-\alpha-\beta)S\sum_{j=0}^{h-2}(\alpha+\beta)^j + (\alpha+\beta)^{h-1}R_{T+1|T}, \tag{1.17}$$

where $R_{T+1|T} = dg(Q_{T+1|T})^{\frac{1}{2}}Q_{T+1|T}dg(Q_{T+1|T})^{\frac{1}{2}}$ with $Q_{T+1|T} = (1-\alpha-\beta)S + \alpha\varepsilon_T\varepsilon'_T + \beta Q_T$. Alternatively, one can implement the cDCC for which the h-steps-ahead correlation process can be solved forward.

For the purposes of this thesis several aspects of the DCC forecasts deserve consideration. First, it is widely recognized that returns distribution show fat-tails. For this reason, using a Gaussian distribution to approximate the conditional future distribution of returns might be a poor depiction of their uncertainty, even if a_t were truly Gaussian. To illustrate this issue, we consider a bivariate DCC with stationary GARCH(1,1) conditional variances, correlation parameters $(\alpha = 0.10, \beta = 0.88)$, an unconditional correlation of 0.5 and Gaussian errors distribution. We simulate a series of size 1020, use $T = 1000$ and keep 20 as out-of-sample realizations. In panel (a) and (b) of Figure 1.2 we plot the h-steps-ahead point forecast of ith return together their corresponding 95% Gaussian forecast intervals. It seems that the 95% Gaussian forecast intervals fail to capture the out-of-sample realization more than we might expect. Second, note that, given the model parameters, equation (1.17) provides only point forecasts of correlations but it does not provide any hint about the uncertainty around them. As we mentioned before, the correlation matrix is observable one-step-ahead and thus it contains only parameter uncertainty which is usually omitted. Third, in addition to the parameter uncertainty, when the forecast horizon exceeds one, forecasts of correlations omit the the error uncertainty. The same arguments apply to the condi-



(a) $y_{1,T+1}$ (b) $y_{2,T+2}$

(c) $\sigma^2_{1,T+h}$ (d) $\sigma^2_{2,T+h}$

(e) $\gamma_{12,T+h}$ (f) $\rho_{12,T+h}$

Figure 1.2: h-steps-ahead forecasts (discontinuous balck lines) of (a)-(b) returns, $y_{i,T+h}$, (c)-(d) conditional variances, $\sigma^2_{i,T+h}$, (e) conditional covariance, $\gamma_{12,T+h}$, and (f) conditional correlation, $\rho_{12,T+h}$, for a simulated series from a bivariate DCC model with $T = 1000$, Gaussian errors and unconditional correlation of 0.5, for $h = 1, ..., 20$, together with out-of-sample true realization (discontinuous grey lines). Panels (a)-(b) show the 95% Gaussian forecast intervals (continuous balck lines).

tional variances and covariances. Middle of Figure 1.2 plot h-steps- ahead point forecasts of $\sigma^2_{i,T+h}$ for $i, j = 1, 2$, $i \neq j$ and $h = 1, 2, ...20$, while lower panels plot $\gamma_{12,T+h} = \sigma_{1,T+h}\sigma_{2,T+h}\rho_{12,T+h}$ and $\rho_{12,T+h}$, for $h = 1, 2, ...20$. We observe only point forecasts of conditional variances, covariances and correlations and no measure of uncertainty around them.

In any case, the parameter variability needs to be incorporated into DCC forecasts as it might be large; see, for instance, Aielli (2013). Moreover, practitioners are not interested only in point forecasts but also in assessing the uncertainty around them, a problem that is unsolved when forecasting conditional variances, covariances and correlations with DCC models, in particular, or any other MGARCH model, in general.

1.4 A short review of bootstrap methods for time series

The original idea of bootstrap was introduced by Efron (1979) who was concerned about obtaining the standard error of an estimator or even its sampling distribution in the context of i.i.d. data. Generally speaking, the bootstrap can be considered resampling method in which no parametric assumptions are made about the underlying population from where the random sample is drawn. Instead, the random sample is used as an estimator of the population; see Efron and Tibshirani (1993). Of course, in the case of time series data, resampling directly from the time series $Y_T = (y_1, y_2, ..., y_T)$ is not a good choice because the random sampling scheme destroys the underlying dependence in the data. Consequently, the statistics obtained from samples with replacement from Y_T turn out to be inconsistent; see Efron and Tibshirani (1993) and Cao (1999), among others. However, two bootstrap approaches are proposed to preserve the time dependence between observations in the bootstrap samples; see Berkowitz and Kilian (2000) for a review of bootstrap methodology in time series.

If no parametric model is available or it is inadvisable to apply it, then we can implement bootstrap methods thought to generate replicates that mimic the time dependence structure. The most widely used are block and stationary bootstrap. Broadly speaking, the bock bootstrap starts

with l non-overlapping blocks of length s, $y_i^* = y_{i,s}^* = (y_i, y_{i+1}, ..., y_{i+s-1})$, for $i = 1, ..., n+s-1$. It is assumed that $T \approx ls$. Then, l bootstrap samples are drawn with replacement from y_i until a bootstrap series of length T is obtained, $Y_T^* = \{y_1^*, ..., y_l^*\}$. This bootstrap series can be used to compute the statistic of interest. The problem with the block bootstrap is that the generated data is non-stationary even if the originally data is. The easiest way to see this point is to note that consecutive observations in different blocks are independent while consecutive observations within a block are dependent. Furthermore, the block size s affects the performance of the block bootstrap method significantly and there is not guidance about how to determine it optimally. On the other hand, the stationary bootstrap starts with a random observation from Y_T, say y_t^*, and the next observation is given by y_{t+1}^* with probability $1-p$, where p is small, or by a new randomly selected observation with probability p. Note that in this way the length of the block is randomly determined. The main advantage of the stationary bootstrap is that its performance does not depends strongly on p as the block bootstrap does on s; see Bühlmann (2002) and Politis (2003) for reviews of non-parametric bootstrap procedures for time series.

Alternatively, if a given parametric model is assumed for the dynamic dependence, then the problem of the dependent structure of the data is reduced to an i.i.d. structure by bootstrapping the corresponding residuals. Next, we briefly describe the steps to follow in the case bootstrapping a VAR(p) model as proposed by Runkle (1987).

Step 1. Estimate the parameters of the model using a consistent estimator, as for example, the LS estimator. Denote it by $\{\widehat{\mu}, \widehat{\Phi}_1, ..., \widehat{\Phi}_p\}$.

Step 2. Obtain the residuals

$$\widehat{\varepsilon}_t = y_t - \widehat{\mu} - \sum_{j=1}^{p} \widehat{\Phi}_j y_{t-j} \quad t = p+1, p+2, ..., T.$$

Denote the empirical distribution function of the residuals by $F_{\widehat{\varepsilon}}$

Step 3. Choose an initial condition $Y_p^* = (y_1^*, ..., y_p^*)$ (for instance, by either drawing a random

block of length p or fixing the first observations to y_t for $t = -p+1, ..., 0$) and generate bootstrap replicates of the series by

$$y_t^* = \widehat{\mu} + \sum_{j=1}^{p} \widehat{\Phi}_j y_{t-j}^* + \varepsilon_t^*$$

where ε_t^* are drawn with replacement from the empirical distribution function of the residuals, $F_{\widehat{\varepsilon}}$.

Step 4. Calculate the bootstrap parameter estimate, $\{\widehat{\mu}^*, \widehat{\Phi}_1^*, ..., \widehat{\Phi}_p^*\}$.

Step 5. Repeat steps 1-4 B times.

The B bootstrap replicates $\widehat{\mu}$ and $\widehat{\Phi}_j^*$ can be used to approximate the distribution of $\widehat{\Phi}_j$, for $j = 1, ..., p$. Bose (1988) establishes, under the general assumptions of section 1.2.1, that for almost every y_t,

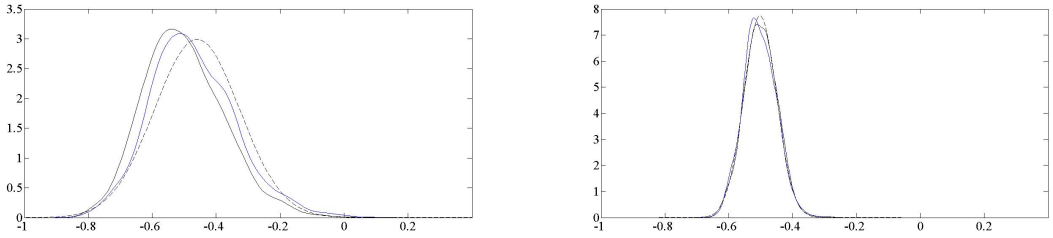
$$\sup_x |P^* \left(T^{1/2} \widehat{\Sigma}_\phi^{1/2} (\widehat{\phi}^* - \widehat{\phi}) \leq x \right) - P \left(T^{1/2} \Sigma_\phi^{1/2} (\widehat{\phi} - \phi) \leq x \right)| = o(T^{-1/2})$$

and

$$\sup_x |P^* \left(T^{1/2} \widehat{\Sigma}_\sigma^{1/2} (\widehat{\sigma}^* - \widehat{\sigma}) \leq x \right) - P \left(T^{1/2} \Sigma_\sigma^{1/2} (\widehat{\sigma} - \sigma) \leq x \right)| = o(T^{-1/2}),$$

where $\Sigma_\phi^{1/2}$ and $\Sigma_\sigma^{1/2}$ are Cholesky decompositions of their corresponding covariance matrices. Therefore, the bootstrap distribution of the parameters of a VAR(p) model is $o(T^{1/2})$ and, consequently, it improves the accuracy of the Gaussian approximation obtained by the standard asymptotic theory which is $O(T^{1/2})$. If the true model is not of finite order, the bootstrap can still be successfully implemented if the bootstrap replicates are generates according to a finite p chosen as a function of T, such that $p^{3.5}/T^{0.5}$ goes to zero. Paparoditis (1996) provides the asymptotic validity of the bootstrap joint distribution of the autoregressive coefficients of an infinite order VAR model.

To illustrate Runkle (1987) procedure we simulate 2000 series of a stationary VAR(1) model with autoregressive matrix $vec\,(\Phi_1)$ equal to $(\phi_{11} = -0.5, 0, 0.5, 0.5)$, χ_4^2 errors distribution and



(a) $T = 50$ (b) $T = 300$

Figure 1.3: Empirical (continuous black line), asymptotic (discontinuous black line) and bootstrap (continous blue line) densities for ϕ_{11} of a simulated VAR(1) with $T = 50$ and 300 and χ^2_4 errors. The number Monte Carlo replications for the empirical density raises to 2000 while for the bootstrap counterparts 1000.

sample sizes $T = 50$ and 300. For each of the series, the parameters are estimated using LS. Panels (a) and (b) of Figure 1.3 plot the empirical density of $\widehat{\phi}_{11}$. Then we select one of the series and approximate the distribution of $\widehat{\Phi}_1$ by using its asymptotic Gaussian distribution and by implementing the bootstrap procedure; see Lütkepohl (1991) for details about the LS estimator. A quick view of panel (a) of Figure 1.3 shows that, when the sample size is $T = 50$, the bootstrap density of $\widehat{\phi}_{11}$ is closer to the empirical than the asymptotic Normal approximation. Of course, when the sample size increases to 300, as it is plotted in panel (b) Figure 1.3, both the asymptotic and the bootstrap densities are rather similar and close to the empirical. This result serves to highlight the value of bootstrap methods since, even when there exist a well-established asymptotic theory, it can be provide desirable approximations in finite sample.

The performance of the basic bootstrap procedure for a VAR(p) model proposed by Runkle (1987) does not recognize that the LS estimates of the autoregressive matrices can be biased in small samples. However, this basic procedure can be modified to tackle this issue by using the formula of Pope (1990) in steps 1 and 2 of the algorithm to correct parameters estimates and the bootstrap parameters. This modification has not effect on the asymptotic validity of the bootstrap since the estimated bias is of order $O(T^{-1})$ while the LS estimates converge at a rate $T^{-1/2}$; see Kilian (1998a,b).

Furthermore, the performance of the basic bootstrap method can also be undermined if the

estimated lag order offers a poor representation of the true model. Thus, Kilian (1998a) suggests incorporating the lag order uncertainty by re-estimating the lag order in step 4. This gives rise to what is called the endogenous bootstrap algorithm. To estimate the lag order, it may be advisable to implement AIC since, as we have previously mentioned, it shows less tendency to choose a lag order below the true one than HQC and SIC.

Next, we briefly review the implementation of bootstrap methods to density forecasting in univariate and multivariate settings.

1.4.1 Bootstrap forecast of time series

Forecast users are interested in the future distribution of the process conditional to the information up to the forecast origin. This distribution can be obtained using the bootstrap procedure proposed by Thombs and Schucany (1990) for univariate AR(p) models. This procedure is grounded on the backward representation in order to obtain bootstrap replicates of the parameters conditional on the last p observations of the series; see Masarotto (1990), McCullough (1994), Grigoletto (1998), Lam and Veall (2002) and Liu et al. (2014) for univariate applications of bootstrap forecast based on the backward representation. However, using the backward representation has two main drawbacks. First, the backward residuals are i.i.d. only if the errors are Normal, but just uncorrelated if the errors depart from Normality; see Breidt and Davis (1992). In the latter situation, it turns out that resampling with replacement from the backward residuals might not be appropriate and, therefore, it would be necessary to use the relationship between the backward and forward representation of the model in order to generate bootstrap replicates; see Breidt and Davis (1992) who show that the relationship between the backward and forward errors can be rather complicated. Consequently, it is common practice to resample directly from the dependent backward residuals. Note also that the asymptotic validity of the backward bootstrap procedures can only be proved by imposing i.i.d. and, as a result, they require Normal errors. Second, these bootstrap alternatives can only be applied to models with a backward representation which excludes their implementation in, for example, multivariate models with Moving Average (MA)

components or with GARCH disturbances.

To overcome these drawbacks, Pascual et al. (2004) proposes a boootstrap procedure that avoids the backward representation; see, for example, Clements and Kim (2007) and Kim et al. (2010) for implementations of this proposal to forecast univariate autoregressive processes. The main advantage of this bootstrap procedure is that its asymptotic validity can be established irrespective of the error distribution. Furthermore, this bootstrap procedure is attractive because it can be adapted to deal with non-linear time series; see for instance, Reeves (2005) and Pascual et al. (2006) who use the procedure to forecast returns and conditional variances in GARCH models and Li (2011) who implements a slight modification of the procedure to forecast SETAR models.

In the context of multivariate time series, there are not many papers implementing bootstrap procedures to construct future conditional distributions and their corresponding forecast regions. The first attempt to bootstrapping a VAR(p) model to approximate the future conditional distribution of the series is proposed by Kim (1999), who extends the univariate algorithm based on the backward representation to multivariate setting; see Kim (2001, 2004) Grigoletto (2005), Staszewska-Bystrova (2011) for implementations to forecast VAR(p). Also, Lütkepohl (2006) cites Kim (1999) as a bootstrap alternative to forecast VAR(p) models when errors are non-Gaussian. More recently, Guerbyenne and Hamdi (in press) implement the multivariate bootstrap procedure of Kim (1999) to obtain forecast regions in periodic state-space models. Nonetheless, all these bootstrap procedures previously proposed in the literature to forecast VAR models depend on the backward representation of the process.

1.5 Objectives of the Thesis

In this thesis we describe bootstrap methods to construct forecast density for multivariate time series. Since forecast densities in multivariate VAR models are obtained using the backward representation, we propose an alternative solution which generalizes the procedure due to Pascual

et al. (2004) to a multivariate setting. Its main advantage is that it does not use the backward representation and, consequently, its asymptotic validity does not depend on the Gaussianity assumption of the error. Moreover, we gather enough evidence through Monte Carlo experiments showing that the bootstrap method based on the backward and that propose in this thesis behave rather similar in finite sample. This is also the main result when we compare the forecast performance of both alternatives in a system containing US quarterly series inflation, unemployment rate and growth. Hence, nothing is gained when we use the more complicated bootstrap algorithm based on the backward representation. This is the objective of Chapter 2.

Several variants of the basic bootstrap procedure are intended to evaluate the quality of the density forecasts. As we have mentioned, it is possible to correct the potential bias often found in LS estimates of the autoregressive matrices and/or to incorporate the lag order uncertainty into forecasts. Yet a direct comparison of these variants is missing in the literature. In this thesis we fill this gap by looking simultaneously the standard Gaussian methodology and several bootstrap alternatives which are the basic bootstrap with and without parameter uncertainty, that incorporating model uncertainty as well as adjusting the parameter estimates for their bias. Such a comparison is useful to asses the extent to which biased parameter and error distribution, model and parameters uncertainty might affect the forecast performance. Our simulation results enhance the relative importance of parameter uncertainty when forecasting highly persistent VAR models. This comparison is carried out in Chapter 3.

MGARCH models contain undoubtedly many challenges for applied researcher, some of them dealing directly with forecasting. First, one has to deal with the non-Gaussianity of returns, which requires alternative ways of approximating its forecast density. Second, there are no methods in the literature to assess the uncertainty of volatility, covariances and correlation forecasts obtained with these models. For this reason, in this thesis, we describe a bootstrap procedure that allows approximating the forecast density of returns, volatilities and covariances in DCC models, though it can be readily adapted to others MGARCH models. Our simulation study and empirical application suggest that bootstrapping ideas can be successfully implemented to approximate density

forecasts in DCC models. This is the topic considered Chapter 4.

Finally, Chapter 5 concludes the thesis and proposes possible lines of future research.

Chapter 2

Bootstrap forecast of non-Gaussian VAR models[1]

2.1 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 chapter, we focus on the forecasting ability of VAR models. It is well known that, in practice, VAR forecasts of large macroeconomic systems could 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 acceptable precision. Some selected examples of useful VAR forecasts are Marcellino et al. (2003) and D'Agostino et al. (2011) for unemployment, Batchelor et al. (2007) for international freight prices, Gupta et al. (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.

[1]This chapter, with some slight variations, first came out as an article with the same title that has been recently accepted for publication in the *International Journal of Forecasting*. In the subsequent chapters, we refer directly to the article version of this chapter.

Most of the literature dealing with VAR forecasts focus 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 Doan et al. (1984), Waggoner and Zha (1999) and Baumeister and Kilian (2012). In order to define spillover measures Diebold and Yilmaz (2009) also consider multivariate multiperiod-ahead forecasts; see also Klößner 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 Diebold et al. (1999), Jore et al. (2010), Wolden Bache et al. (2011) and Clark (2011) among others. Traditionally, a multivariate forecast density for a given horizon can be obtained assuming Gaussian forecast errors and known lag order and model parameters. However, since long, it has been recognized that the parameter uncertainty could be an important issue when dealing with VAR forecasts in practice; see Lewis and Reinsel (1985) and Fair and Shiller (1990) 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, Kilian (1998b) and Harvey and Newbold (2003). 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 et al. (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 on 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 to deal with different issues in the context of forecasting using multivariate VAR and periodic state-space models; see, for instance, Grigoletto (2005, 2012), Staszewska-Bystrova (2011) and Guerbyenne and Hamdi (in press). 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 attractiveness of bootstrap procedures is their ability to predict 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 representation which excludes, for example, multivariate models with Moving Average (MA) components or with 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 to obtain 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 et al. (2004) in the context of univariate ARIMA models, one can implement simple bootstrap procedures that incorporate the parameter uncertainty without requiring the BR. For example, Wolf and Wunderli (2012), Staszewska-Bystrova and Winker (2013) implement the bootstrap procedure originally described by Pascual et al. (2011)[2] to construct 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) to construct confidence bands in the context of impulse response functions.

[2] Pascual et al. (2011) is a previous version of the article Fresoli et al. (in press) by the same authors.

In this chapter, 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 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 if the errors are non-Gaussian the improvement in coverage is moderate. In any case, the bootstrap procedures provide similar coverage accuracy regardless of whether they are based on the backward representation or not. The importance of constructing forecast regions taking into account the non-Gaussianity of the variables is illustrated by implementing the forward bootstrap to obtain joint forecast densities of US quarterly inflation, unemployment and growth rates.

The rest of the chapter is organized as follows. Section 2 focuses the discussion and establishes the notation by describing the traditional and backward bootstrap procedures to construct 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 performance of the forward bootstrap forecast regions are compared with those of Gaussian and backward bootstrap procedures. Section 5 illustrates the results with two empirical applications. First, forecast densities of US quarterly future inflation, unemployment and GDP growth are obtained. Second, we implement the proposed bootstrap procedure to a textbook example in Lütkepohl (1991) to forecast West German investment, income and consumption. Finally, Section 6 concludes the chapter with suggestions for further research.

2.2 Multi-step forecast densities and regions for VAR models

This section establishes notation and briefly describes the traditional method for the construction of Gaussian forecast densities in stationary VAR models using asymptotic approximations of the finite sample distribution of the parameter estimator. The bootstrap procedure based on the BR is also described. Both procedures are illustrated by implementing them to obtain forecast regions of a bivariate series generated by a stationary non-Gaussian VAR(2) model.

2.2.1 Gaussian forecast densities

Consider the VAR(p) model of finite lag order p

$$y_t = \mu + \Phi_1 y_{t-1} + ... + \Phi_p y_{t-p} + \varepsilon_t, \;\; t = -p+1, ..., T \tag{2.1}$$

described in the previous chapter. 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 by consistent estimates as follows

$$\widehat{y}_{T+h|T} = \widehat{\mu} + \widehat{\Phi}_1 \widehat{y}_{T+h-1|T} + ... + \widehat{\Phi}_p \widehat{y}_{T+h-p|T} \tag{2.2}$$

where $\widehat{y}_{T+j|T} = y_{T+j}$, $j \leq 0$ and, in this paper, $\widehat{\theta} = (\widehat{\mu}, \widehat{\Phi}_1, ..., \widehat{\Phi}_p)$ denotes the LS estimator of the parameters. The MSFE of $\widehat{y}_{T+h|T}$, calculated by substituting the estimated parameters in the expression of the MSFE of $y_{T+h|T}$, does not incorporate the parameter uncertainty and, consequently, underestimates the true MSFE of $\widehat{y}_{T+h|T}$; see Baillie (1979) and Reinsel (1980). Consider, for simplicity, $h = 1$ and a VAR(1) model without constant. In this case, given the information available at time T, the MSFE of $\widehat{y}_{T+1|T}$ can be decomposed as follows

$$MSFE(\widehat{y}_{T+1|T}) = \Sigma_y(1) + E_T[(\Phi_1 - \widehat{\Phi}_1) y_T y_T' (\Phi_1 - \widehat{\Phi}_1)'] \tag{2.3}$$

where $\Sigma_y(1) = E_T\left[(y_{T+1} - y_{T+1|T})(y_{T+1} - y_{T+1|T})'\right]$ and the T under the expectation denotes that it is conditional on the information available at time T. The estimator $\widehat{\Phi}_1$ is a function of the available sample $\{y_1, ..., y_T\}$ and, consequently, the expectation of the second term of expression (2.3) can be dropped. Conditioning on the whole past of the process implies that we fix the values of the parameters estimates; see Dufour (1985). As a result, in order to incorporate the parameter uncertainty in the MSFE of $\widehat{y}_{T+1|T}$, many authors assume that the sample used to estimate the parameters is independent from the sample used to forecast; see, for example, Reinsel (1980), Bhansali (1981) and Yamamoto (1981). In this case, the second term of expression (2.3) can be computed using the asymptotic distribution of $\widehat{\Phi}_1$; see Reinsel (1980) for the corresponding expression. The MSFE of $\widehat{y}_{T+h|T}$ obtained by using the asymptotic distribution to approximate the parameter uncertainty and with all unknown parameters substituted by their sample estimates, will be denoted by $\widehat{\Sigma}_y^A(h)$.

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, or both, are big enough relative to T, the effect of the parameter uncertainty can be substantial. However, granted that a good estimator is used, the importance of this uncertainty could be small in systems consisting in few variables; see Riise and Tjøstheim (1984). Obviously, as the sample size gets larger, the parameter uncertainty contribution to the MSFE vanishes. If, in model (2.1), ε_t is further assumed to be Gaussian, then the h-steps-ahead forecast density is estimated by

$$y_{T+h} \sim N(\widehat{y}_{T+h|T}, \widehat{\Sigma}_y^A(h)). \tag{2.4}$$

From (2.4) it is possible to obtain h-steps-ahead joint ellipsoids for the variables within the system. Nonetheless, constructing these ellipsoids can be quite demanding when N is larger than two or three. Consequently, Lütkepohl (1991) proposes using the following Bonferroni cubes with

coverage at least (1-δ)100%

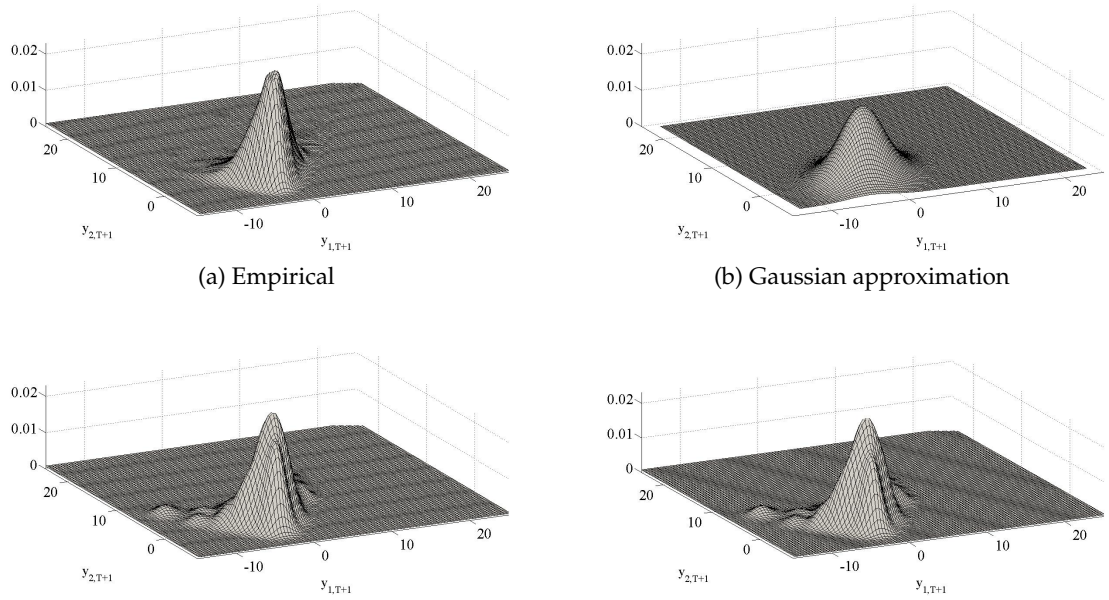
$$GC_{T+h} = \left\{ y_{T+h} | y_{T+h} \in \cup_{i=1}^{N} \left[\widehat{y}_{i,T+h|T} \pm z_\tau \widehat{\sigma}_{i,h} \right] \right\}, \tag{2.5}$$

where $\widehat{y}_{i,T+h|T}$ is the ith element of $\widehat{y}_{T+h|T}$, z_τ is the τth quantile of the standard Normal distribution with $\tau = 0.5(\delta/N)$ and $\widehat{\sigma}^2_{i,h}$ is the ith element of the main diagonal of $\widehat{\Sigma}^A_y(h)$. The Bonferroni cubes in (2.5) depend on the marginal distributions and, consequently, the information on the dependence among the variables in the system is lost.

The forecast densities in (2.4) incorporate the uncertainty due to parameter estimation throughout the asymptotic distribution but still rely on Gaussian forecast errors. Therefore, the corresponding intervals and regions could be inadequate when this assumption is not satisfied. Even more, in this latter case, the shape of the densities for h >2 is unknown in general. Finally, the forecast densities in (2.4) can be missleading in cases in which the asymptotic approximation is unreliable; see Dufour and Jouini (2006). As an illustration, consider the following non-Gaussian stationary bivariate VAR(2) model previously considered by Kim (2001),

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} 0.9 & 0 \\ -0.5 & -0.7 \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} -0.2 & 0 \\ 0.8 & -0.1 \end{bmatrix} \begin{bmatrix} y_{1,t-2} \\ y_{2,t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix} \tag{2.6}$$

where $\varepsilon_t = (\varepsilon_{1,t}, \varepsilon_{2,t})'$ is a serially independent white noise vector with contemporaneous covariance matrix given by $vech(\Sigma_\varepsilon) = (1, 0.5, 1)'$ where $vech$ denotes the lower-diagonal column stacking operator of a symmetric matrix. The distribution of ε_t is a χ^2_4 which has been adequately standardized to have variance one; see Kilian (1998b) for the adequacy of this distribution to represent some macroeconomic time series. The dominant root of $|I_2 - \Phi_1(z^{-1}) - \Phi_2(z^{-2})| = 0$ in the VAR(2) model in (2.6) is 0.5 so the model is far from the non-stationary boundary. Panel (a) of Figure 2.1 displays the true joint one-step-ahead density of $y_{1,T+1}$ and $y_{2,T+1}$, which shows clear asymmetries. 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 Figure 2.1 plots the corresponding joint



(a) Empirical (b) Gaussian approximation

(c) Bootstrap based on BR (d) New bootstrap

Figure 2.1: Kernel estimates of one-step-ahead forecast densities of a simulated bivariate series generated by a stationary VAR(2) model with χ^2_4 errors based on $T = 100$ observations.

density obtained assuming that the forecast errors are jointly Gaussian as given in expression (2.4) with the following estimated MSFE that incorporates the asymptotic parameter uncertainty, $vech\left(\widehat{\Sigma}^A_y(1)\right) = (1.13, 0.59, 1.02)'$. Comparing panels (a) and (b), it is obvious that the Gaussian density fails to capture the asymmetry of the error distribution. This inadequacy is reflected in Figure 2.2 which plots a realization of y_{T+1} of size $R = 5000$ together with the 95% ellipsoid and the Bonferroni region of y_{T+1} obtained from the density in (2.4). We can observe that none of both regions are appropriate to construct a satisfactory forecast region for y_{T+1}. Figure 2.2, also plots the ellipsoid and the Bonferroni cube obtained when the MSFE of $\widehat{y}_{T+h|T}$ is computed without incorporating the parameter uncertainty which are nearly identical to those obtained when the parameter uncertainty is incorporated. Therefore, in this particular example, it seems that the parameter uncertainty is not a big issue while the non-Gaussianity of the forecast errors is a serious concern.

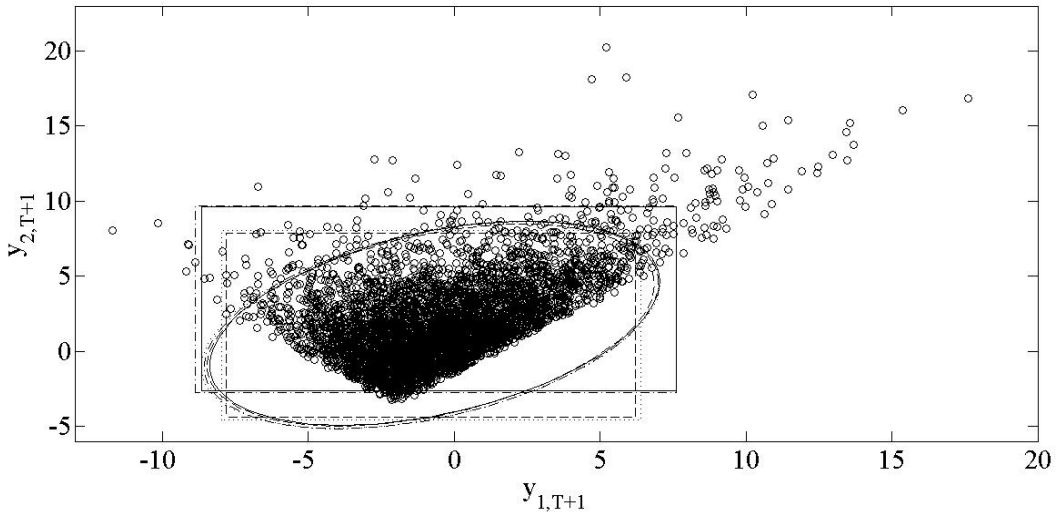


Figure 2.2: A realization of size $R = 5000$ of y_{T+1} ($\circ$) generated by a stationary bivariate VAR(2) model with χ^2_4 errors together with 95% one-step-ahead forecast Bonferroni cubes and elliptical regions constructed using $T = 100$ observations and based on the following densities (a) Gaussian without parameter uncertainty (- -), (b) Gaussian with asymptotic parameter uncertainty ($\cdot\cdot$), (c) Bootstrap with BR (- $\cdot$) and (d) New Bootstrap (-).

2.2.2 Bootstrap forecast densities based on BR

As mentioned in the introduction, bootstrap procedures can be implemented to obtain forecast densities that incorporate the parameter uncertainty without relying on Gaussian forecast errors. The first bootstrap procedure proposed in the context of multivariate VAR models is due to Kim (1999) who, in order to take into account the conditionality of VAR forecasts on past observations, follows Thombs and Schucany (1990), and proposes to obtain bootstrap replicates of the series based on the following backward recursion

$$y_t^* = \widehat{\omega} + \widehat{\Lambda}_1 y_{t+1}^* + ... + \widehat{\Lambda}_p y_{t+p}^* + \widehat{v}_t^*, \; t = T-p, ..., 1, \tag{2.7}$$

where $y_t^* = y_t$, for $t = T-p+1, ..., T$, are p starting values which coincide with the last values of the original series, $\widehat{\omega}, \widehat{\Lambda}_1, ..., \widehat{\Lambda}_p$, 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 $[(T-p)/(T-2p)]^{0.5}$.[3] Then, bootstrap LS estimates of the parameters of the forward representation are obtained by estimating the forward VAR(p) model in (2.1) using $\{y_1^*, ..., y_T^*\}$. Denote these estimates by $\widehat{\theta}^* = (\widehat{\mu}^*, \widehat{\Phi}_1^*, ..., \widehat{\Phi}_p^*)$. The h-steps-ahead bootstrap forecast is given by

$$\widehat{y}_{T+h|T}^* = \widehat{\mu}^* + \widehat{\Phi}_1^* \widehat{y}_{T+h-1|T}^* + ... + \widehat{\Phi}_p^* \widehat{y}_{T+h-p|T}^* + \widehat{\varepsilon}_{T+h}^* \tag{2.8}$$

where $\widehat{y}_{T-i|T}^* = y_{T-i}$, for $i > 0$, and $\widehat{\varepsilon}_{T+h}^*$ are random draws with replacement from the empirical distribution function of the rescaled forward residuals. Having obtained B bootstrap replicates of $\widehat{y}_{T+h|T}^*$, Kim (1999) proposes to construct the following bootstrap forecast ellipsoid with probability content $(1-\delta)100\%$ as follows

$$KE_{T+h} = \left\{ y_{T+h} | \left[y_{T+h} - \bar{y}_{T+h|T}^* \right]' S_{y^*}^K(h)^{-1} \left[y_{T+h} - \bar{y}_{T+h|T}^* \right] < Q_K^* \right\} \tag{2.9}$$

where $\bar{y}_{T+h|T}^*$ is the sample mean of the B bootstrap replicates, $\widehat{y}_{T+h|T}^{*(b)}$, $S_{y^*}^K(h)$ is the corresponding sample covariance [4] and Q_K^* is the $(1-\delta)100\%$ percentile of the empirical bootstrap distribution of the quadratic form $[\widehat{y}_{T+h|T}^* - \bar{y}_{T+h|T}^*]' S_{y^*}^K(h)^{-1} [\widehat{y}_{T+h|T}^* - \bar{y}_{T+h|T}^*]$. Furthermore, the Bonferroni cube with at least (1-δ)100% nominal coverage is given by

$$KC_{T+h} = \left\{ y_{i,T+h} | y_{i,T+h} \in \cup_{i=1}^N \left[q_i^*(\tau), q_i^*(1-\tau) \right] \right\} \tag{2.10}$$

where $q_i^*(\tau)$ is the τth quantile of the empirical bootstrap distribution of $\widehat{y}_{i,T+h}^*$ which is the ith element of the vector $\widehat{y}_{T+h|T}^*$. Kim (1999) uses the percentile and percentile-t methods of Hall (1992) to define the Bonferroni cubes. However, Bonferroni cubes are defined as in (2.10) because they are better suited to deal with potential asymmetries of the error distribution than the percentile-t intervals; see Hall (1992). In addition, Kilian (1999) shows that, in absence of pivotal statistics, as it

[3] The residuals should be centered if no intercept is included in the model.

[4] Kim (1999) does not explicitly show how $S_{y^*}^K(h)$ should be defined. One can also obtain $S_{y^*}^K(h)$ by substituting the expression of the MSFE matrix of $y_{T+h|T}$ by their corresponding bootstrap estimates and computing the average through all bootstrap replicates. By calculating it with the sample covariance or by substituting the bootstrap parameters in the corresponding expression we get similar results.

is the case of VAR process close to the nonstationary region, bootstrap percentile methods that do not rely on studentized statistics have better coverage accuracy than those based on percentile-t.

The bootstrap procedure proposed by Kim (1999) is illustrated by considering again the same time series of size $T = 100$ simulated by the non-Gaussian bivariate VAR(2) model in (2.6) considered above. Panel (c) of Figure 2.1 plots a kernel estimate of the joint bootstrap density of $y_{1,T+1}$ and $y_{2,T+1}$ 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. Figure 2.2 plots the corresponding 95% ellipsoid and Bonferroni cube defined in (2.9) and (2.10), respectively. First of all, observe that the bootstrap and Gaussian ellipsoids are very similar and not adequate to represent the shape of the realization of y_{T+1} plotted in Figure 2.2. This similarity could be expected given that both ellipsoids only differ in the way the MSFE is computed and the bootstrap MSFE is given by $vech(S^K_{y^*}(1)) = (1.11, 0.58, 1)'$ which is very close to the MSFE computed using the asymptotic approximation. Figure 2.2 clearly illustrates that, when dealing with non-Gaussian forecast errors, the forecast regions constructed from the bootstrap joint densities cannot be based on ellipsoids as they assume a symmetric distribution; see Wolf and Wunderli (2012). On the other hand, when the forecast region is constructed using the Bonferroni cube, Figure 2.2 shows that the bootstrap cube is located towards the northeast so it is more adequate than the regions based on the Gaussian forecast density. This is in fact reflecting that the quantiles of the marginal densities used to construct the Bonferroni cube can cope with the asymmetry while the ellipsoids use a quadratic form based on the wrong Gaussianity assumption. The bootstrap Bonferroni cube is better suited to deal with asymmetries but, as mentioned above, cannot cope with the dependence between the variables.

Kim (1999) justifies the use of his bootstrap procedure in finite samples by suggesting that the asymptotic results of Thombs and Schucany (1990) can be extended to a multivariate framework. However, as mentioned in the introduction, the asymptotic validity of the bootstrap procedures based on the backward representation relies on the assumption of Gaussian innovations; see Kim

(2001). Note that alternatively one could use the relationship between the forward and backward residuals and resampling from the former to obtain the latter. However, obtaining the backward representation can be very complicated in VAR(p) models with large lag order; see Kim (1997, 1998) for the expression of the backward representation.[5] Furthermore, it is obvious that any bootstrap procedure based on the BR cannot be implemented in models without this representation, compromising the flexibility and applicability of the procedure.

2.3 A new bootstrap procedure

In this section, we propose obtaining the joint forecast density of y_{T+h} in VAR(p) models using an extension of the bootstrap procedure proposed by Pascual et al. (2004) for univariate ARIMA models, which avoids using the BR when incorporating the parameter uncertainty. We describe the proposed procedure and prove its asymptotic validity. Its performance is illustrated for the same non-Gaussian stationary VAR(2) model considered above.

2.3.1 Description of the algorithm

The forward procedure proposed in this chapter to obtain the bootstrap forecast density of y_{T+h} is based on the same assumption 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 from the sample used to forecast. In this way, we avoid the BR when generating the bootstrap replicates used to estimate the parameters. However, the last p observations in the series are still fixed when forecasting the future values. The proposed algorithm to obtain bootstrap replicates of y_{T+h} is as follows.

Step 1. After selecting the order p, estimate by LS the parameters of model (2.1) and obtain the corresponding vector of residuals. Denote by $\widehat{F}_{\widehat{\varepsilon}}$ the empirical distribution function of the

[5] Tong and Zhang (2005) and Chan et al. (2006) show that a necessary condition for the VAR(p) model to have this backward representation is that the covariance matrices $\Upsilon(h) = E\left[(y_t - E(y_t))(y_{t-h} - E(y_t))'\right]$ are symmetric for all h. This is a very strong restriction not likely to be satisfied in real data systems.

re-scaled residuals.

Step 2. Construct a bootstrap series $\{y_1^*, \ldots, y_T^*\}$ as follows

$$y_t^* = \widehat{\mu} + \widehat{\Phi}_1 y_{t-1}^* + \cdots + \widehat{\Phi}_p y_{t-p}^* + \widehat{\varepsilon}_t^*, \; t = 1, \ldots, T, \tag{2.11}$$

where $\widehat{\varepsilon}_t^*$ are random draws with replacement from $\widehat{F}_{\widehat{\varepsilon}}$ and $y_t^* = y_t$, for $t = -p+1, ..., 0$. Obtain $\widehat{\theta}^* = (\widehat{\mu}^*, \widehat{\Phi}_1^*, ..., \widehat{\Phi}_p^*)$, a bootstrap replicate of the LS estimates by fitting a VAR(p) model to the bootstrap replicate $\{y_1^*, ..., y_T^*\}$.

Step 3. Using (2.1) 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

$$\widehat{y}_{T+h|T}^* = \widehat{\mu}^* + \widehat{\Phi}_1^* \widehat{y}_{T+h-1|T}^* + \cdots + \widehat{\Phi}_p^* \widehat{y}_{T+h-p|T}^* + \widehat{\varepsilon}_{T+h}^*, \tag{2.12}$$

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

Step 4. Repeat steps 2 and 3 B times.

We obtain B bootstrap replicates of y_{T+h}, denoted by $\{\widehat{y}_{T+h|T}^{*(1)}, ..., \widehat{y}_{T+h|T}^{*(B)}\}$, by using this procedure. Their empirical bootstrap distribution can be used to obtain the corresponding ellipsoids and Bonferroni cubes as described before for the procedure based on the BR. It is important to point out that in step 2, the bootstrap replicates used to obtain bootstrap estimates of the parameters are obtained as proposed by Runkle (1987). This is the main difference with the procedure based on the BR and described in the previous section, which obtains replicates of y_t, $t = 1, ..., T$, using expression (2.7) while we propose using the forward representation in expression (2.11).

To illustrate the implementation of the new bootstrap procedure proposed in this chapter, we consider again the bivariate time series generated by the non-Gaussian stationary VAR(2) model in (2.6). Panel (d) of Figure 2.1 displays a kernel estimate of the bootstrap joint density of y_{T+1} which is very similar to the density obtained by implementing the bootstrap procedure based on the BR. Figure 2.3 plots the bootstrap ellipsoid and Bonferroni cube obtained from the new bootstrap density. We observe that, although the forward bootstrap density is

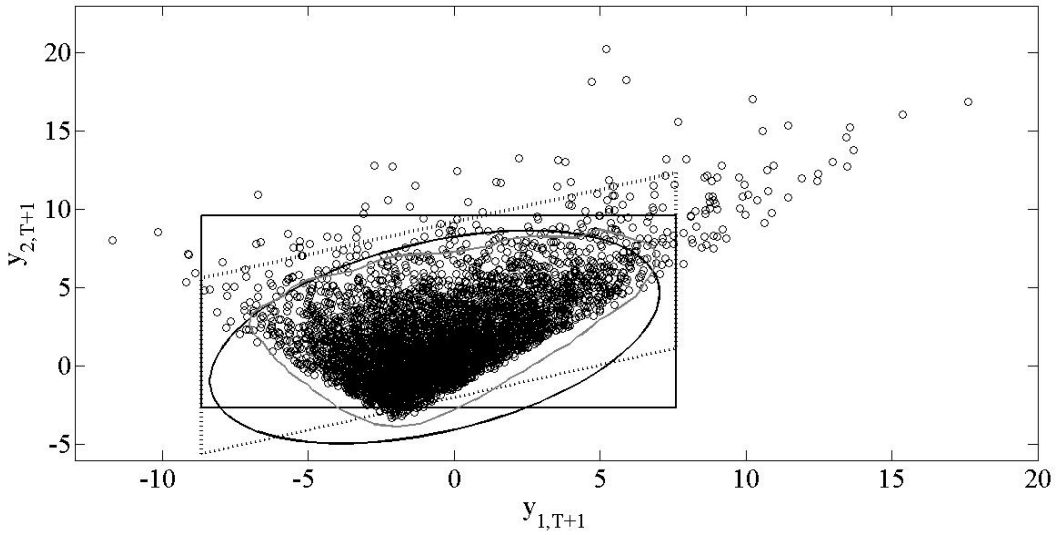


Figure 2.3: A realization of size $R = 5000$ of y_{T+1} ($\circ$) generated by a stationary bivariate VAR(2) model with χ^2_4 errors together with 95% one-step-ahead forecast bootstrap regions based on a sample size $T = 100$ (a) ellipsoid (-) (b) Bonferroni cube (-), (c) corrected Bonferroni cube ($\cdot\cdot$) and (d) High Density Regions (grey curve).

very different from the Gaussian density, there are not 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 dispersion of the future values. Note that $\bar{y}^*_{T+1|T} = (-2.19, -0.26)'$ and $vech\,(S_{y^*}(1)) = (1.11, 0.58, 1)'$ 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 towards the right and up with respect to the corresponding Gaussian cube representing the assymetries 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 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. Figure 2.1 also plots

the 95% HDR computed from the bootstrap replicates, $\widehat{y}^*_{T+1|T}$. We can observe that the shape of the HDR seems to be more adequate to represent 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 not 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 $p_{21,h} = \widehat{\sigma}^*_{21,T+h}/\widehat{\sigma}^{2*}_{1,T+h}$ with $\widehat{\sigma}^*_{21,T+h}$ and $\widehat{\sigma}^{2*}_{1,T+h}$ being elements of $S_{y^*}(h)$. 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 are not changed while those of the second variable are transformed by the same amount, either $p_{21,h}q^*_1(\tau)$ or $p_{21,h}q^*_1(1-\tau)$. Figure 2.1 plots the modified Bonferroni cube which is rotated in the direction of the correlation observed between $\widehat{y}^*_{1,T+1|T}$ and $\widehat{y}^*_{2,T+1|T}$ and, consequently, it 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 easily modified by introducing asymptotic stationarity bias corrections of the bootstrap parameters and the endogenous lag order bootstrap algorithm as proposed by Kilian (1998a,c); see Staszewska-Bystrova and Winker (2013) for an implementation of the forward bootstrap algorithm using both modifications.

2.3.2 Asymptotic validity

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

$$\varepsilon_t(\theta) = y_t - \mu - \Phi_1 y_{t-1} - ... - \Phi_p y_{t-p},\ t = 1, ..., T. \tag{2.13}$$

In expression (2.13), it is explicit that the errors depend on the unknown parameters contained in θ. If θ is estimated by $\widehat{\theta}$, the corresponding estimated residuals are given by

$$\widehat{\varepsilon}_t(\widehat{\theta}) = y_t - \widehat{\mu} - \widehat{\Phi}_1 y_{t-1} - ... - \widehat{\Phi}_p y_{t-p},\ t = 1, ..., T, \tag{2.14}$$

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

The following theorem establishes the asymptotic validity of the empirical bootstrap distribution of $\widehat{y}^*_{T+h|T}$ as given in (2.12) to approximate the distribution of a future value y_{T+h}.

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

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

$$\widehat{y}^*_{T+1|T} = \widehat{\mu}^* + \widehat{\Phi}^*_1 y_T + ... + \widehat{\Phi}^*_p y_{T-p+1} + \widehat{\varepsilon}^*_{T+1}. \tag{2.15}$$

For $h = 2$ we have

$$\widehat{y}^*_{T+2|T} = \widehat{\mu}^* + \widehat{\Phi}^*_1 \widehat{y}^*_{T+1|T} + ... + \widehat{\Phi}^*_p y_{T-p+2} + \widehat{\varepsilon}^*_{T+2}. \tag{2.16}$$

Replacing in (2.16) $\widehat{y}^*_{T+1|T}$ by its expression in (2.15), it follows that

$$\widehat{y}^*_{T+2|T} = N_0(\widehat{\theta}^*) + N_1(\widehat{\theta}^*) y_T + ... + N_p(\widehat{\theta}^*) y_{T-p+1} + M_1(\widehat{\theta}^*)\widehat{\varepsilon}^*_{T+1} + \widehat{\varepsilon}^*_{T+2}, \tag{2.17}$$

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

Proceeding in this way the following expression is obtained for the h-steps-ahead bootstrap

forecast

$$\begin{aligned} \widehat{y}^*_{T+h|T} = N_0(\widehat{\theta}^*) + N_1(\widehat{\theta}^*)y_T + ... + N_p(\widehat{\theta}^*)y_{T-p+1} \\ + M_1(\widehat{\theta}^*)\widehat{\varepsilon}^*_{T+1} + M_2(\widehat{\theta}^*)\widehat{\varepsilon}^*_{T+2} + ... + \widehat{\varepsilon}^*_{T+h}, \end{aligned} \tag{2.18}$$

where the functions $N_i(\widehat{\theta}^*)$ and $M_i(\widehat{\theta}^*)$ are different for different horizons. Expression (2.18) defines the bootstrap future values as a function of the observed realization $\{y_{-p+1}, ..., y_T\}$, the independent random draws $\widehat{\varepsilon}^*_{T+h}$ and continuous functions of the bootstrap parameter estimates $\widehat{\theta}^*$.

In order to establish the asymptotic convergence of $\widehat{y}^*_{T+h|T}$, we start by considering the terms involving $N_i(\widehat{\theta}^*)$ for which the asymptotic validity of $\widehat{\theta}^*$ is needed. The asymptotic validity of the bootstrap LS estimator is established by Bose (1988) who proves the convergence in probability almost surely of $\widehat{\theta}^*$ to θ. Therefore, given that $N_i(\widehat{\theta}^*)$ are continuous functions of the parameters, it follows that $N_i(\widehat{\theta}^*) \xrightarrow{p} N_i(\theta)$ almost surely. Moreover, note that y_{T-i+1} are fixed values and, consequently, the terms involving $N_i(\widehat{\theta}^*)y_{T-i+1} \xrightarrow{d} N_i(\theta)y_{T-i+1}$ in probability. Second, using the same arguments as before, we can see that $M_i(\widehat{\theta}^*) \xrightarrow{p} M_i(\theta)$ almost surely. Finally, consider the terms $\widehat{\varepsilon}^*_{T+i}$ which are random draws with replacement from $\widehat{F}_{\widehat{\varepsilon}}(\widehat{\theta})$. Using the results in Bickel and Freedman (1981), it is straightforward to prove that $d_2(\widehat{F}_{\widehat{\varepsilon}}(\widehat{\theta}), F_{\varepsilon}(\theta)) \to 0$ in probability as $T \to \infty$, 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 $\widehat{\varepsilon}^*_{T+i} \xrightarrow{d} \varepsilon_{T+i}$ in probability.[6] On the other hand, $\widehat{\varepsilon}^*_{T+i}$ are independent from $M_i(\widehat{\theta}^*)$. Consequently, $M_i(\widehat{\theta}^*)\widehat{\varepsilon}^*_{T+i} \xrightarrow{d} M_i(\theta)\varepsilon_{T+i}$ by the independence of $\widehat{\varepsilon}^*_{T+i}$ and the bootstrap version of Slutsky's Theorem. Consequently, all terms in expression (2.18) converge weakly in probability and, as a result, $y^*_{T+h|T} \xrightarrow{d} y_{T+h}$ as $T \to \infty$, in probability.

■

Before concluding this section it is important to remark that, even though the asymptotic

[6]The d_2 metric is given by $d_2(G_Z, G_Y) = \inf_C \left[E_G(|Z-Y|)^2\right]^{\frac{1}{2}}$, where C is the set of pairs of random variables (Z, Y) with cumulative distribution function G such that the marginal cumulative distribution functions are G_Z and G_Y, respectively.

validity is established for centered residuals, Stine (1987) shows that it is still valid if they also are 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 as the AIC considered in this chapter.

2.4 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 bivariate data generating processes (DGP) with different configuration of parameters and lag orders which reproduce either stationary, persistent or near-cointegrated processes. DGP1 is the stationary VAR(2) model defined in (2.6). 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 A. In each DGP, we consider three distributions of the errors, namely Gaussian, Student-5 and χ^2_4 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 order is equal to 12 and 16 for $T = 100$ and 300, respectively. After estimating the lag order, the VAR parameters are estimated by LS and are bias-corrected using the asymptotic correction of Pope (1990) and taking into account the stationarity restriction as proposed by Kilian (1998c). Forecast densities for horizons $h = 1, ..., 8$ steps-ahead are constructed assuming Gaussian errors without parameter uncertainty and computing the MSE as in (2.4) using the asymptotic approximation. Forecast densities are also constructed using $B = 1999$ bootstrap replicates obtained by the backward bootstrap and by the

forward bootstrap procedure. The bootstrap procedures are implemented using the asymptotic bias and endogenous lag order corrections as proposed by Kilian (1998a,c). In each of case, we construct the corresponding 95% ellipsoids and Bonferroni cubes.

Based on $R = 5000$ future values of the process y_{T+h}, we calculate the empirical coverage of each forecast region. Figure 2.4 plots the average through the Monte Carlo replicates of the empirical coverages of the Bonferroni regions constructed for the stationary VAR(2) model. We can observe that, regardless of the distribution of the error, when $T = 100$, the average coverages of the cubes constructed using the Gaussian approximation without incorporating the parameter uncertainty, are clearly smaller than the nominal 95%. When incorporating the parameter uncertainty through the asymptotic approximation, the coverages get slightly closer to nominal. Using the bootstrap Bonferroni cubes, we get very similar coverages regardless of whether the BR is used. The bootstrap coverages are clearly closer to the nominal, irrespective of the errors distribution. When the errors are non-Gaussian, we can observe that the Gaussian coverages are under the nominal while the bootstrap cubes maintain their accuracy. On the other hand, in large sample sizes, when $T = 300$, we observe that all procedures have similar coverages which are very close to the nominal if the errors are Gaussian. Yet the Gaussian cubes are slightly below the nominal coverages when the errors distribution departs from Gaussianity. When the sample size is small, one would expect that the Gaussian approach to forecast is more likely to understate the true sampling uncertainty as one portion of it, that due the parameter uncertainty, is not taken into account. Upper panels of Figure 2.7 display the average volumes of each approach to forecasting when $T = 100$. We observe that the Gaussian standard cubes are below the empirical true volumes, and this happen especially when the error distribution is non-Gaussian. Of course, when the MSFE forecast is augmented to consider the sampling variability of the parameters, Gaussian cubes get larger and, as a result, the corresponding empirical coverages improve. But, when the sample size increases to $T = 300$, the differences in average volumes between Gaussian bootstrap cubes are comparatively minor, which in part reflect the quite similar performance in empirical coverages. The average volume differences between bootstrap procedures are negligible.

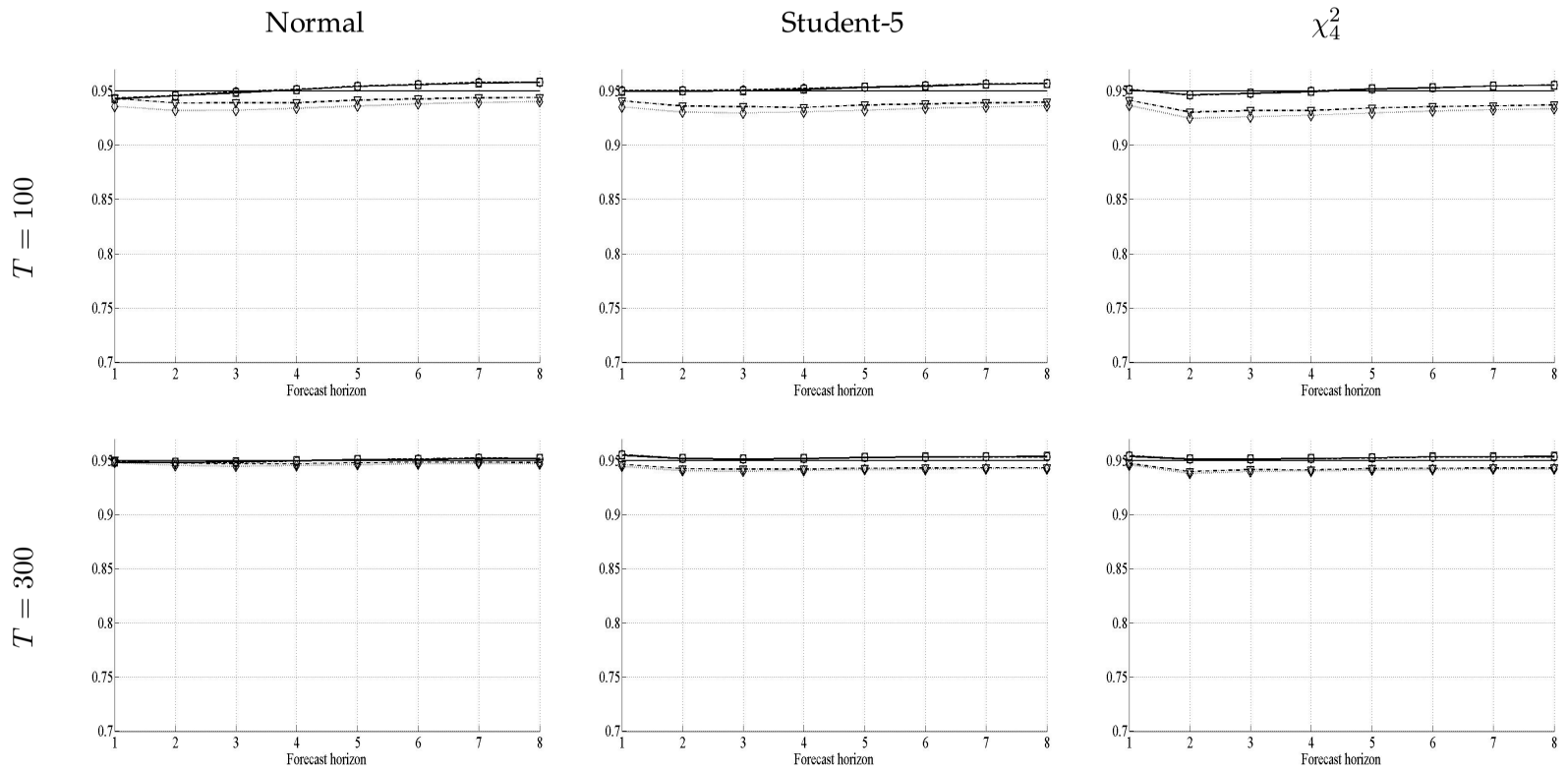


Figure 2.4: Monte Carlo averages of empirical coverages of Bonferroni forecast cubes based on: i) Gaussian (- - $\diamondsuit$), ii) Gaussian with asymptotic MSFE ($\cdot\cdot$ $\triangledown$), iii) bootstrap with BR (- $\circ$) and iv) new bootstrap (- $\cdot$ $\square$) densities, for a stationary VAR(2) model with sample sizes $T = 100$ (first row) and 300 (second row) and Gaussian (first column), Student-5 (second column) and χ^2_4 (third column) errors. Nominal coverage 95%.

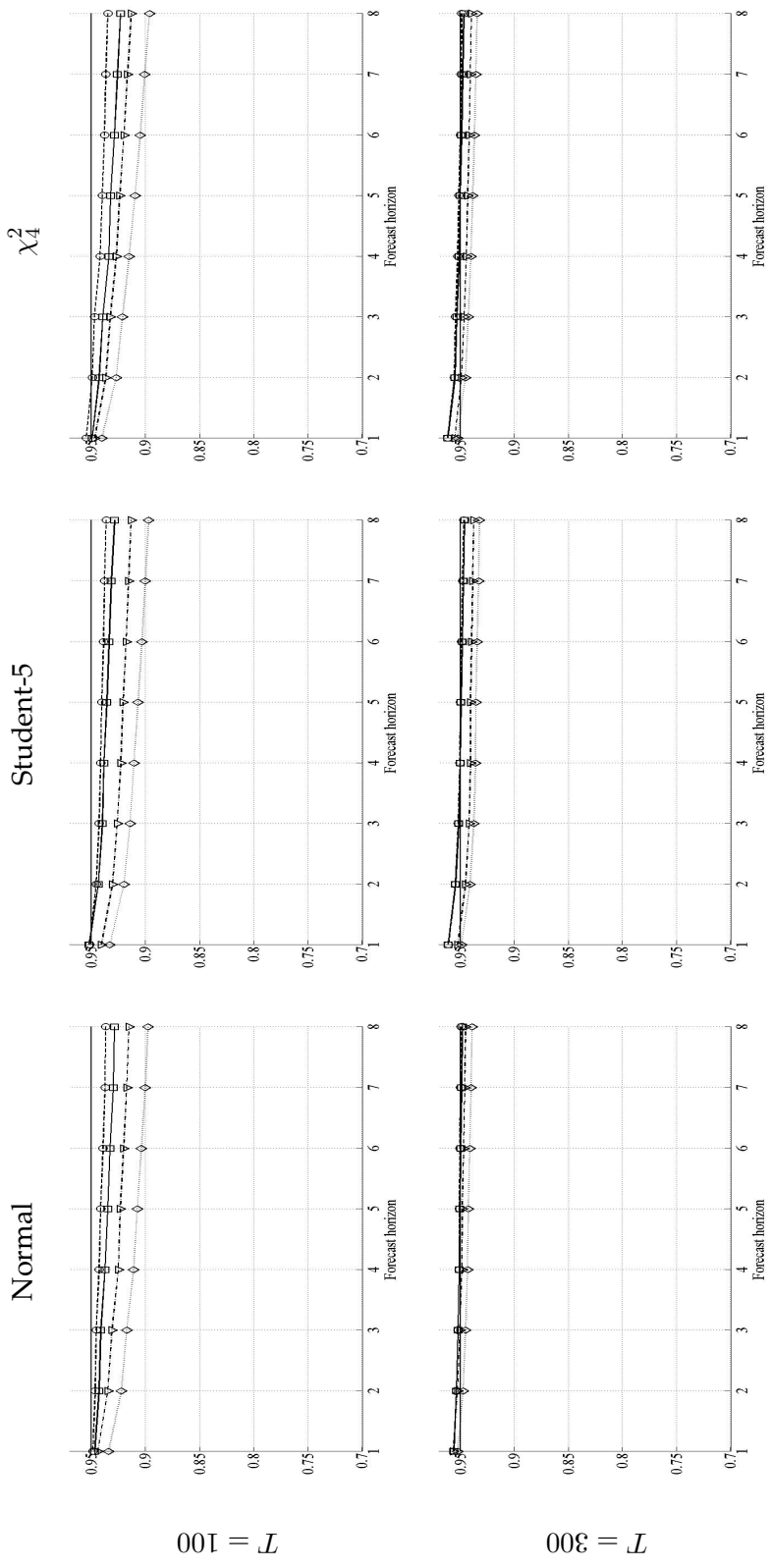


Figure 2.5: Monte Carlo averages of empirical coverages of Bonferroni forecast cubes based on: i) Gaussian (- - $\diamond$), ii) Gaussian with asymptotic MSFE ($\cdot\cdot\ \triangledown$), iii) bootstrap with BR (- $\circ$) and iv) new bootstrap (- $\cdot$ $\square$) densities, for a persistent VAR(5) model with sample sizes $T = 100$ (first row) and 300 (second row) and Gaussian (first column), Student-5 (second column) and χ^2_4 (third column) errors. Nominal coverage 95%.

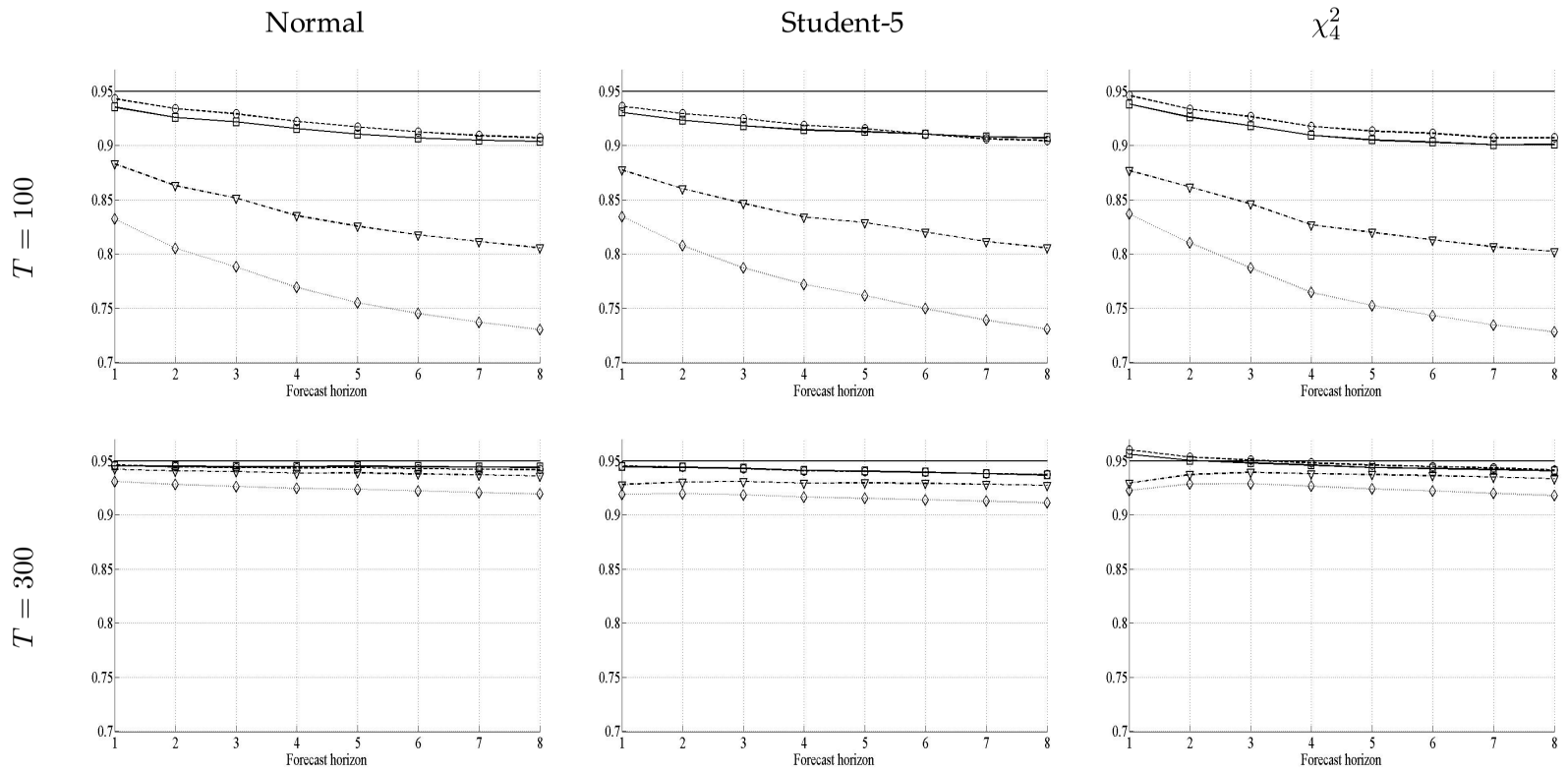


Figure 2.6: Monte Carlo averages of empirical coverages of Bonferroni forecast cubes based on: i) Gaussian (- - ♢), ii) Gaussian with asymptotic MSFE (·· ▽), iii) bootstrap with BR (- ○) and iv) new bootstrap (- · □) densities, for a near-cointegrated VAR(10) model with sample sizes $T = 100$ (first row) and 300 (second row) and Gaussian (first column), Student-5 (second column) and χ^2_4 (third column) errors. Nominal coverage 95%.

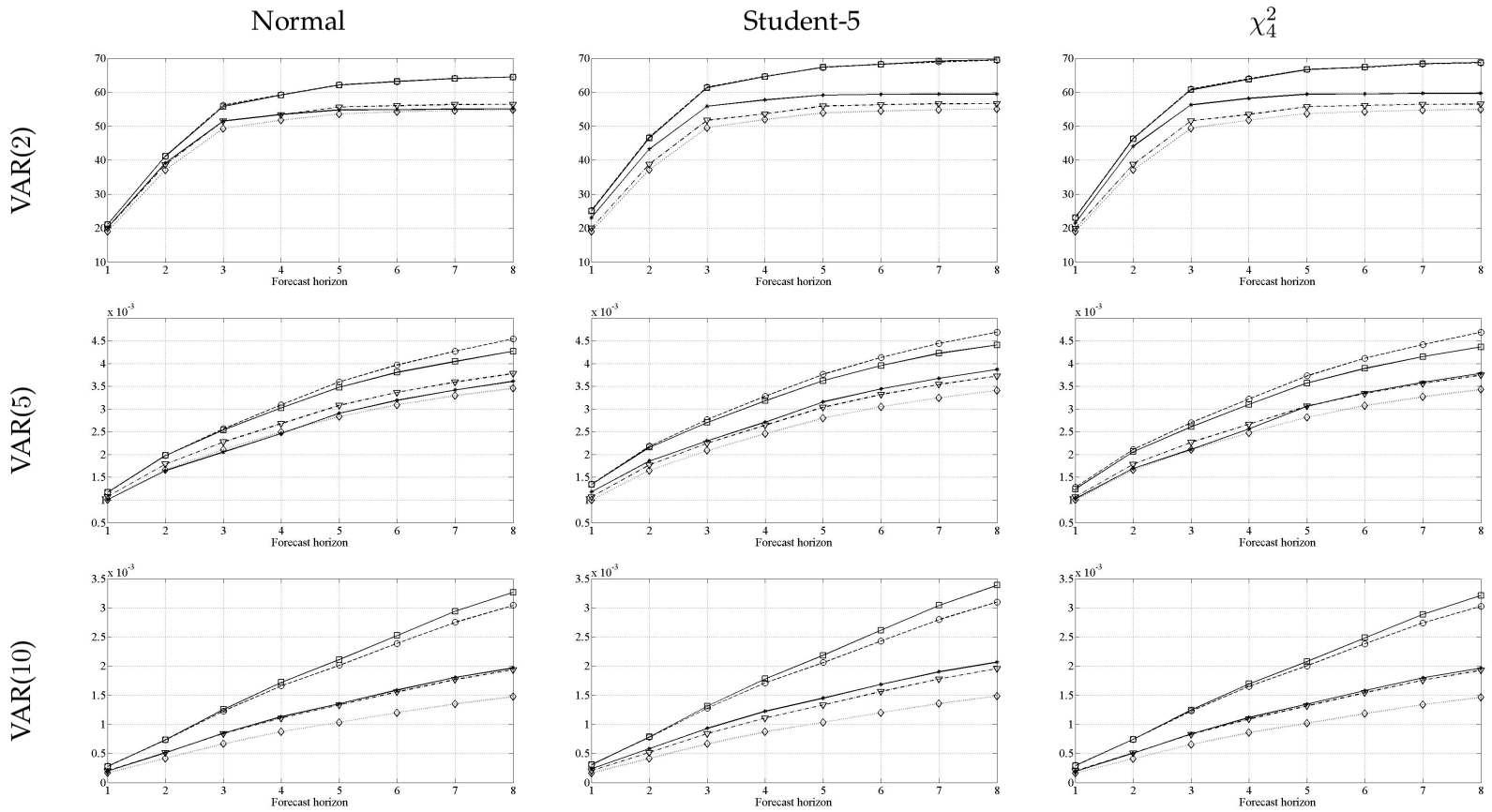


Figure 2.7: Monte Carlo averages of empirical volumes of Bonferroni forecast cubes based on: i) Gaussian ($\diamondsuit$), ii) Gaussian with asymptotic MSFE ($\triangledown$), iii) bootstrap with BR ($\circ$) and iv) new bootstrap ($\square$) densities, for a stationary VAR(2) (first row), persistent 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^2_4 (third column) errors. Nominal coverage 95%.

Note that, for this model, 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 of the LS estimator is not important when constructing the forecast regions.

Figures 2.5 and 2.6 plot the coverages corresponding to the persistent VAR(5) and the 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. Consider first $T = 100$. We observe a clear difference the coverage performance between standard and bootstrap procedures. For instance, for the VAR(10), is striking the bad performance of the Gaussian cube even in the presence of Gaussian errors, with a coverages going from 0.85, when $h = 1$, to 0.77, when $h = 8$. Incorporating the parameter uncertainty into the sampling variability improves the coverage performances of the Gaussian cubes, though they still drop away from the 0.95 line. The bootstrap cubes are considerably better than the Gaussian, for both the VAR(5) and VAR(10), though they deteriorate with the forecast horizon in the both models, a feature that is more pronounced in the VAR(10) model. For instance, for the VAR(10) and Student-5 error distribution, bootstrap coverages goes from 0.93 to 0.91, when $h = 1$ and 8, respectively. This deterioration may be due to the fact that this model is close to the non-stationary regions and, consequently, it may have more difficulties when forecasting in the long run. Furthermore, it is worth noting that the coverages of both bootstrap cubes are rather similar regardless of whether they use the BR or the forward representation. Also, there seem not to be differences between coverage accuracy of the Gaussian cubes across the error distribution. Under a large sample situation, when $T = 300$, empirical coverages for all the procedures recover to a level close the 0.95 line and the differences between the Gaussian and the bootstrap procedures reduce, around 0.025 in average, when $h = 1$.

Finally, with respect to average volumes, in middle and lower panels Figure 2.7 we observe the same qualitative patterns highlighted before. The only difference is that, for the VAR(5), the bootstrap cubes based on the BR are larger than that of the forward bootstrap procedure, a feature that is reflected in the slightly better performance of the former in therm of coverage accuracy.

Nonetheless, the opposite occurs for the VAR(10). To sum up, both bootstrap procedures provide cubes that are generally larger than those obtained by the standard methodology, even when the latter incorporates the sample variability due the parameter estimation. Lager bootstrap volumes that reflect the true extent of sample variability are more likely to capture, in average, more future realizations of the process.

After all, the simulations conducted in this section show that the forward bootstrap procedure performs better than traditional methods based on Gaussian densities and does not worse than the bootstrap procedure based on the BR. In addition, they suggest that, when the sample size is small, as the persistence of the system increases, it is important to take into account the parameter uncertainty to obtain coverages close to the nominal ones.[7]

2.5 Empirical application

2.5.1 Forecasting quarterly US inflation, unemployment and GDP growth

In this section, we implement the forward bootstrap procedure to construct forecast densities of quarterly US inflation (π_t), unemployment rate (u_t) and GDP growth (g_t) observed from 1948Q1 to 2011Q3.[8] 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 has been transformed into quarterly 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 2.1 reports the sample mean, standard deviation (sd), skewness and kurtosis

[7]For completeness, Appendix B we report the empirical average of Monte Carlo volumes when $T = 300$ and Appendix C includes results for forecast intervals of the first variable of the systems. Results for ellipsoids are not reported to save space. For the same reason, we do not report the results with bias-corrected parameters and without endogenous lag-order corrections and without both, bias and endogenous lag order. All these results are available upon request.

[8]The data was obtained from the Federal Reserve Bank of St. Louis, webpage: www.stlouisfed.org.

Table 2.1: Descriptive statistics of quarterly US inflation (π), unemployment (u) and GDP growth (g) observed from 1948Q1 to 2009Q3 with p-values in parenthesis.

Series	Mean	Sd	Skewness	Kurtosis	KD	ADF[1]	$Q(8)$	$Q_2(8)$
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)			

[1] MacKinnon's p-values approximation.

of each of the series during the estimation period together with the joint measures of skewness and kurtosis proposed by Kilian and Demiroglu (2000). Table 2.1 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 both individually or jointly at 10%. Table 2.1 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 2.1. 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. Following Kilian (1998a, 2001) and Marcellino et al. (2006), the lag order of the VAR is selected by the AIC with maximum lag order being equal to 14, which chooses $\widehat{p} = 4$. Given that the 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 procedure proposed in this chapter is implemented to construct out-sample forecast densities for $h = 1, ..., 8$.

Figure 2.8 plots bivariate one-step-ahead Gaussian and kernel estimates of the bootstrap densities for the three variables in the system taken two by two. The bootstrap densities clearly reflect

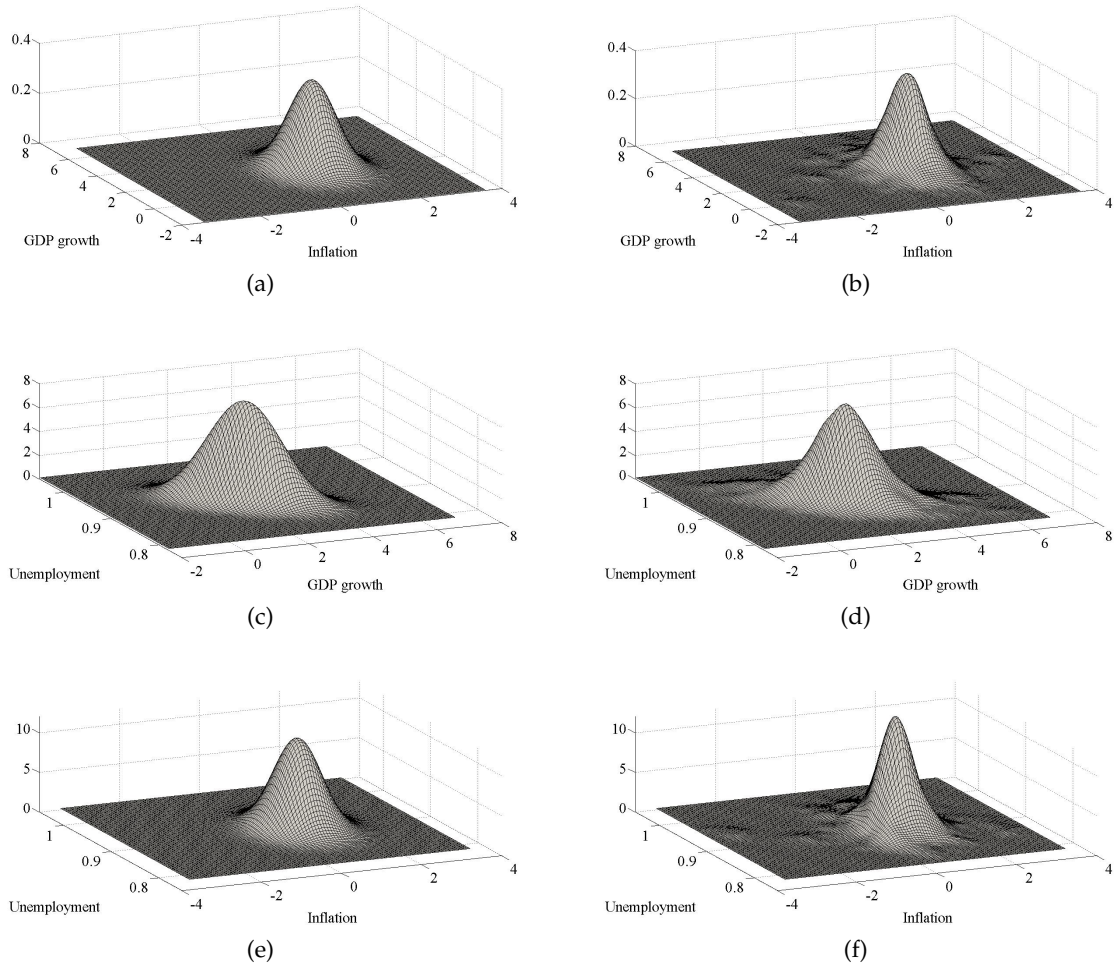
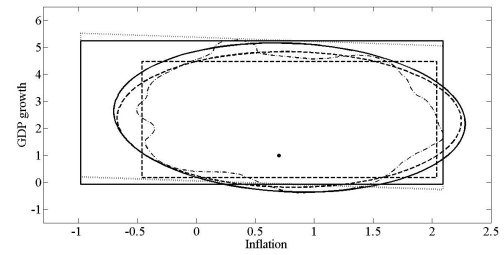
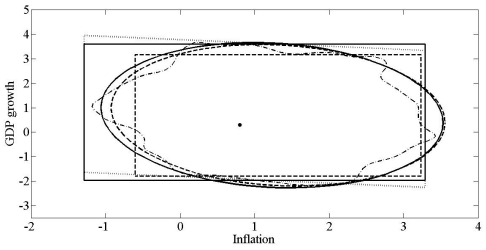


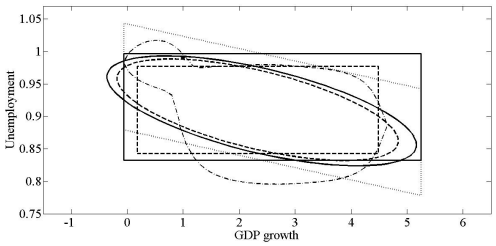
Figure 2.8: Kernel estimates of the bivariate densities of one-step-ahead forecast of inflation, unemployment and GDP growth and computed using the Gaussian approximation with parameter uncertainty (first column) and the bootstrap procedure proposed in this chapter (second column).



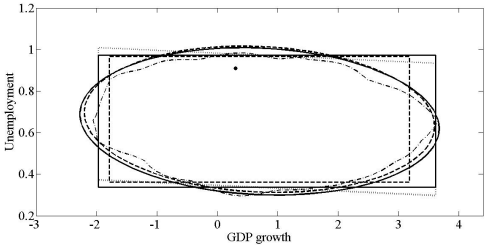
(a) Inflation-GDP growth, $h = 1$



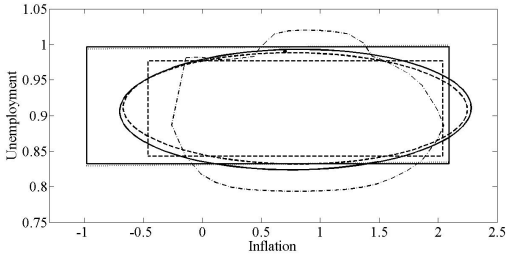
(b) Inflation-GDP growth, $h = 8$



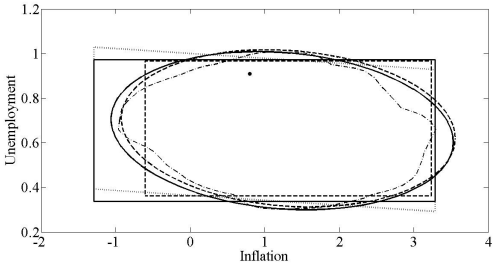
(c) GDP growth-Unemployment, $h = 1$



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



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



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

Figure 2.9: 95% Gaussian ellipsoids and cubes (discontinuous lines), bootstrap ellipsoid and cubes (continuous lines), corrected bootstrap cubes (dotted lines) and HDRs (dotted-discontinuous lines) for one-step-ahead (first column) and eight-steps-ahead (second column) forecast of Inflation-GDP growth (first row), GDP growth-Unemployment (second row) and Inflation-Unemployment (third row).

the non-Gaussianity which is more evident in the inflation-unemployment density. Figure 2.9 plots, for each of the three variables considered two-by-two, 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 assuming Gaussian errors. Figure 2.9 also plots the HDRs and the Bonferroni cubes modified to take into account the correlation between the variables. In each case, the corresponding observed out-of-sample values are displayed by a dot. First, we can observe 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 assuming Gaussianity. Also note that, when $h = 1$, the HDRs suggest non-elliptical densities. But, when $h = 8$, HRDs 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 by using the forward bootstrap procedure. Finally, when looking at the modified Bonferroni cubes, we can observe that they only differ from the corresponding original cube in the case of the GDP growth-unemployment region when $h = 1$. It is worth mentioning that, when looking at whether the regions contain the true observations, we observe that, when $h = 8$, all regions contain them. However, when $h = 1$, the GDP growth-unemployment observation falls outside the Gaussian regions and but is it close to the boundary of the bootstrap ellipsoid and cube, but 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 out-of-sample 95% ellipdois and Bonferroni cubes for the three variables in the system for $h = 1, ..., 8$. 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

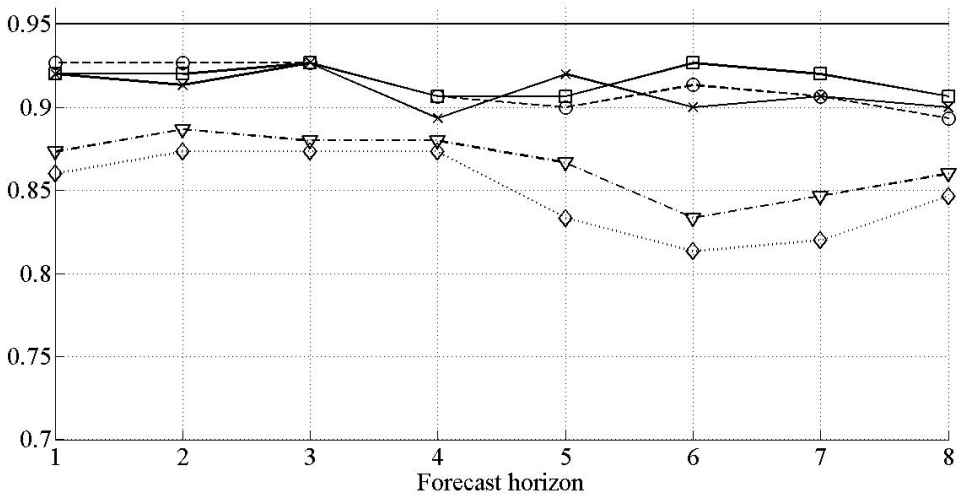


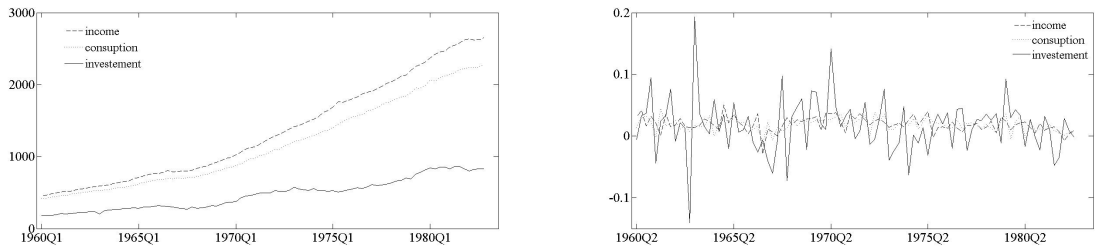
Figure 2.10: Empirical coverages of Bonferroni cubes for US inflation, unemployment and GDP growth based on rolling window estimates with $T = 100$ observations: i) Gaussian without parameter uncertainty ($\diamondsuit$), ii) Gaussian with asymptotic parameter uncertainty ($\triangledown$), iii) backward bootstrap ($\circ$), iv) forward bootstrap ($\square$) and v) corrected cube ($\times$). Nominal coverage 95%.

observation. The results are plotted in Figure 2.10. 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. Also note that the performances of all bootstrap regions are remarkably similar. It is obvious that there is not any gain in 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 Figure 2.10 are in line with those obtained with simulated data in the sense that there are no large differences between using the forward bootstrap procedure or that based on the BR.

2.5.2 Forecasting West German investment, income and consumption

We carry out an empirical implementation using the example published by Lütkepohl (1991) so we are able to show how the forecast intervals are different when using the bootstrap instead

of the asymptotic correction of the MSFE. We take the West German data set used in Lütkepohl (1991) to illustrate the estimation and analysis of VAR models. The data consist of quarterly seasonally adjusted fixed investment, disposable income and consumption expenditure from 1960Q1 to 1982Q4. Panel (a) of Figure 2.11 plots the level series which seem to have trend. Thus, to removed the trend, we take first difference which are plotted in Panel (b) of Figure 2.11. Then, the data set is split into an estimation period from 1960Q1-1978Q4 with $T = 75$ quarters and out-of-sample period from 1979Q1-1982Q4 with $H = 16$ quarters. Finally, we estimate a VAR(2) for the period 1960Q1-1978Q4; see Lütkepohl (1991) for details about the selection of the lag order in this empirical application.

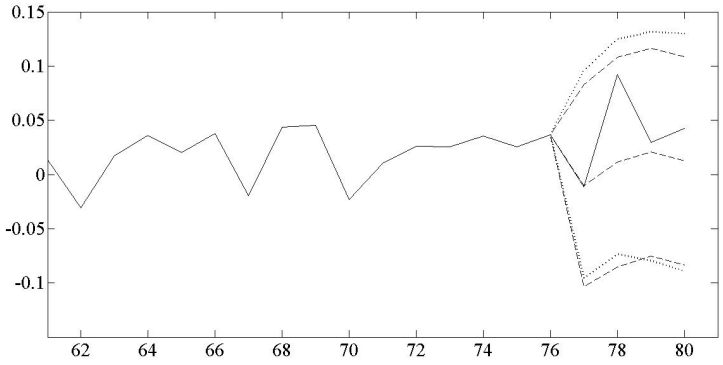


(a) Level Series (b) First difference

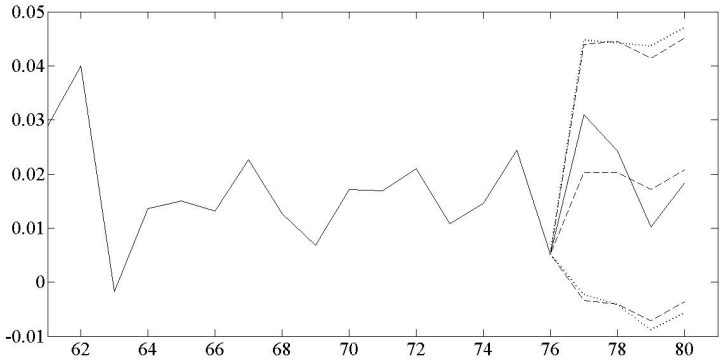
Figure 2.11: (a) level and (b) first difference of West German investment, income, and consumption series for the period 1960Q1-1982Q4.

The proposed bootstrap procedure with bias-corrected parameters and endogenous lag order selection is implemented to forecast during the out-of-sample period based on $B = 1999$ replicates. Figure 2.12 plots the 1979 out-of-sample realization, the point forecast and the 95% intervals obtained by using the standard approach and the proposed bootstrap procedure.[9] Panels (a) and (b) of Figure 2.12 show that the the out-of-sample values for 1979 ($t = 77, ..., 80$) are included within the 95% bounds of the standard and bootstrap forecast intervals for the investment and income. However, the out-of-sample realization of consumption falls outside the 95% standard interval in 1979Q2 ($t = 78$) and 1979Q3 ($t = 79$). On the other hand, we observe that the 95% bootstrap forecast intervals are usually wider, especially for the consumption series for which the

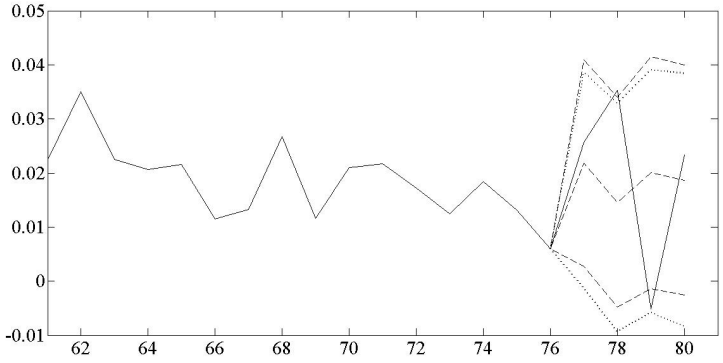
[9]The axis are chosen to reproduce the Figure 3.3 of Lütkepohl (1991).



(a) Investement



(b) Income



(c) Consumption

Figure 2.12: 95% standard with AMSE (dashed lines) and bootstrap (dotted lines) forecast intervals together with the out-of-sample realizarion (continuous line) of the first difference of West German (a) investment, (b) income and (c) consumption for the out-of-sample year 1979 ($t = 77, ..., 80$).

1979Q3 ($t = 79$) observation is well-inside its corresponding bootstrap bounds.

2.6 Conclusions

In this chapter, we establish the asymptotic validity and analyze the finite sample performance of the multivariate extension of the bootstrap procedure proposed by Pascual et al. (2004) to construct forecast densities in multivariate VAR(p) models. The main attractiveness 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. Finally, our Monte Carlo experiments show that the procedure works properly in 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 off constructing 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 between the Gaussian and bootstrap coverages. Furthermore, the forward procedure is computationally simple and asymptotically valid and, consequently, an attractive alternative when dealing with density forecast in multivariate models.

The flexibility of the procedure proposed in this chapter suggests its implementation in other multivariate models. For example, as proposed by Pascual et al. (2005) in 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 Bårdsen and Lütkepohl (2011).

Finally, further effort should be directed to the construction of bootstrap forecast regions. In

this sense, it is worth noting that the forecast 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 approximation deteriorates. The Bonferroni cubes do provide a better solution 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 either from the elliptical or rectangular shapes.

Chapter 3

Bootstrap forecasts in VAR models: the effect of parameters bias, parameter and model uncertainties

3.1 Introduction

When analyzing a multivariate systems of time series it is often of interest to forecast their future values. Traditionally, multivariate forecasts were provided as point forecasts. However, recently, the attention of both econometricians and practitioners has moved to obtain probability forecast distributions. A popular model to represent and forecast the dynamic evolution of multivariate time series is the VAR(p) model which is the focus of this paper. The standard procedure for obtaining forecast densities in VAR models is based on a given lag order with known parameters and Gaussian errors; see, for instance, Lütkepohl (1991). However, in practice, the lag order needs to be selected, the parameters should be estimated and the error distribution is rarely known. Consequently, the model implemented to forecast is an estimated VAR model which is treated as if it were the true data generating process (DGP) and the forecast densities are conditional on the selected lag order, the estimated parameters and the chosen error distribution with the sampling

variability due to model and parameter uncertainty being ignored. It is important to take into account that forecasts obtained with the estimated model can be inadequate to represent the future projection of the system if the parameter estimates are biased; see, for instance, Tjøstheim and Paulsen (1983) and Clements and Kim (2007). Even worse, Chatfield (1996) suggests that ignoring the model uncertainty may lead to over-optimism regarding the adequacy of the forecasts in the sense that the resulting forecast intervals and regions may overstate the number of future observations lying outside them. Finally, although the bias and the parameter uncertainty are of small order of magnitude, their effect could be substantial in large and highly persistent systems when the sample size is small; see Riise and Tjøstheim (1984). The objective of this chapter is to analyze the contribution of the parameter bias and the model and parameter uncertainties when building forecast regions in the context of multivariate VAR(p) models. For this purpose, we first construct forecast regions using the standard textbook approach based on estimated VAR models with Gaussian errors. Second, we implement the bootstrap procedure proposed by Fresoli et al. (in press) incorporating the uncertainty about the error distribution but with fixed parameters. Third, the bootstrap including the parameter uncertainty. Then, we add to the bootstrap algorithm bias-correction. Finally, we also implement the bootstrap incorporating the lag order uncertainty.

The remainder of this chapter is organized as follows. Section 2 establish notation by briefly describing the VAR(p) model and the construction of forecast densities and regions using the standard Gaussian and bootstrap procedures. In Section 3, we present the Monte Carlo study based on bivariate VAR models with different order and persistence properties, carried out to compare the Gaussian forecast regions with bootstrap forecast regions. Finally, Section 4 concludes.

3.2 Forecasting VAR models

Consider the VAR(p) model introduced in the previous chapters and denote by $\Pi = [\Phi_1, ..., \Phi_p]$ a $N \times Np$ matrix called companion matrix. The standard textbook approach is to construct the

forecast density assuming Gaussian errors as follows

$$y_{T+h} \sim N(\widehat{y}_{T+h|T}, \widehat{\Sigma}_y(h)). \tag{3.1}$$

where $\widehat{y}_{T+h|T}$ and $\widehat{\Sigma}_y(h)$ are obtained by substituting the unknown parameters in the expression of the conditional mean of y_{T+h} given $\{y_1, ..., y_t\}$ and its corresponding MSFE matrix by the LS estimates. The forecast densities in (3.1) are denoted here as standard (STD). The STD forecast densities can be inadequate to construct forecast regions for several reasons. First, they do not tackle the parameter bias which affects both $\widehat{y}_{T+h|T}$ and $\widehat{\Sigma}_y(h)$. Second, $\widehat{\Sigma}_y(h)$ partially reflects the uncertainty around the point forecast since other sources of uncertainty, such as those associated with the selection of $\widehat{p}$ and with the estimation of the parameters, are omitted. Finally, the Gaussianity assumption may be misleading for many macroeconomic and financial systems.

Bias-corrected parameters can be obtained by using the expression proposed by Pope (1990). The parameter uncertainty can be incorporated into the MSFE by using its asymptotic distribution; see Reinsel (1980) and Lütkepohl (1991). Alternatively, Bayesian forecast densities are able to incorporate the parameter and lag order uncertainties; see, for instance, Koop (2013) and Wright (2013) just to mention two recent references. However, Bayesian forecast densities usually rely on assumptions about errors distribution . Finally, bootstrap procedures can approximate the forecast density of a VAR(p) model without relying on any distributional assumptions for the errors and are well suited to tackle the parameter and lag order uncertainty; see, for instance, Kilian (1998a,b). From a practical point of view, it is important to find out the sources that are explaining the bad performance of STD approach in Monte Carlo experiments and empirical applications; see, for example, Kim (1999) and Fresoli et al. (in press). To achieve this goal, we rely on several bootstrap alternatives that successively tackle these issues.

Next we describe the bootstrap procedure implemented to construct h-steps-ahead forecast densities for VAR(p) models; see Fresoli et al. (in press) for its asymptotic validity.

Step 1. Select to lag order, $\widehat{p}$, by using AIC and estimate the parameters of the model by LS.

Step 2. Correct the LS estimate by subtracting its bias, $\widehat{\Pi}^c = \widehat{\Pi} - \widehat{\Omega}$. Then, the largest root of the resulting bias-corrected parameter is computed, $r(\widehat{\Pi}^c)$. If $r(\widehat{\Pi}^c) < 1$, then $\widehat{\Pi}^c$ is considered as the bias-corrected parameter estimate. Otherwise, set $\widehat{\Omega}_1$ and $\delta_1 = 1$ and define $\widehat{\Omega}_{i+1} = \delta_i \widehat{\Omega}_i$, $\delta_{i+1} = \delta_i - 0.01$ and $\widehat{\Pi}^c = \widehat{\Pi} - \widehat{\Omega}_i$, and iterate, for $i = 1, 2, ...$, until the largest root of the bias-corrected parameter falls in the stationary region, $r(\widehat{\Pi}^c) < 1$. Obtain also the bias-corrected version of the intercept, $\widehat{\mu}^{*c} = (I_N - \widehat{\Phi}_1^c - ... - \widehat{\Phi}_p^c)\bar{y}$, where $\bar{y}$ is a $N \times 1$ vector of sample means. Get the residuals , $\widehat{\varepsilon}_t$, for $t = \widehat{p}+1, ...T$, by using the bias-corrected parameters and denote the empirical distribution function from the centered and rescaled residuals as $\widehat{F}_{\hat{\varepsilon}}$.

Step 3. Generate a bootstrap replicate $\{y_1^*, ..., y_T^*\}$ by using the estimated lag order, $\widehat{p}$, and the bias-corrected parameters, as follows

$$y_t^* = \widehat{\mu}^c + \widehat{\Phi}_1^c y_{t-1}^* + \widehat{\Phi}_2^c y_{t-2} + ... + \widehat{\Phi}_p^c y_{t-\widehat{p}}^* + \widehat{\varepsilon}_t^* \tag{3.2}$$

where $\widehat{\varepsilon}_t^*$ are random draws with replacement from $\widehat{F}_{\hat{\varepsilon}}$. The first p observations of $\{y_1, ..., y_p\}$ are set as initial conditions.

Step 4. Use the bootstrap replicate of the process to select the lag order of the model, $\widehat{p}^*$. Obtain $\{\widehat{\mu}^{*c}, \widehat{\Phi}_1^{*c}, ..., \widehat{\Phi}_p^{*c}\}$, after adjusting a bootstrap replicate of the LS estimates by fitting a VAR($\widehat{p}^*$) model to the bootstrap series $\{y_1^*, ..., y_T^*\}$ and correct the bootstrap parameters estimates as in *Step 2*.

Step 5. Generate a bootstrap replicate of the VAR forecast by using the following recursion

$$\widehat{y}_{T+h|h}^* = \widehat{\mu}^{*c} + \widehat{\Phi}_1^{*c}\widehat{y}_{T+h-1|h}^* + \widehat{\Phi}_2^{*c}\widehat{y}_{T+h-2|h}^* + ... + \widehat{\Phi}_p^{*c}\widehat{y}_{T+h-\widehat{p}^*|T}^* + \widehat{\varepsilon}_{T+h}^* \tag{3.3}$$

where $\widehat{y}_{T+j}^* = y_{T+j}$ if $j \leq 0$ and $\widehat{\varepsilon}_{T+h}^*$ is a random draw with replacement from $\widehat{F}_{\hat{\varepsilon}}$.

Step 6. Repeat steps 1-5 B times.

As a result of implementing this procedure, we obtain $\{\widehat{y}_{T+h|T}^{*(1)}, ..., \widehat{y}_{T+h|T}^{*(B)}\}$ which can be used to approximate the h-steps-ahead forecast distribution of the process. More specifically, if $G^*(x) =$

$P(y^*_{T+h} < x)$ is the distribution function of y^*_{T+h}, then its Monte Carlo estimate is given by

$$\widehat{G}^*_h(x) = \sum_{b=1}^{B} \frac{I\left(\widehat{y}^{*(b)}_{T+h|T} < x\right)}{B},$$

where $I(\cdot)$ takes the value 1 when the argument is true and 0 otherwise. The estimated bootstrap distribution is then used to construct forecast intervals and regions with appropriate probability content.

The bootstrap procedure can be run with fixed lag order and parameters by avoiding steps 2, 3 and 4 and using the parameter estimates in step 5. In this case, the forecast densities, called here as distribution bootstrap (Dist-B), incorporate the uncertainty about the error distribution. For example, STATA implements the Dist-B to construct bootstrap forecast intervals for VAR models (StataCorp, 2013). Second, we consider the bootstrap procedure that avoids the step 2 and does not re-estimate the lag order for the bootstrap replicates in step 4. This bootstrap procedure, which is called *basic bootstrap*, tackles, in addition to the error distribution uncertainty, the sample variability due to parameter uncertainty; see, for instance, Wolf and Wunderli (2012) who construct bootstrap forecast regions using this method. Third, if we avoid only re-estimation of the lag order in step 4, we have in hand a bootstrap alternative that incorporates the bias-correction formula proposed by Pope (1990); see Clements and Kim (2007) who highlight that in presence of biased parameters the resulting bootstrap forecast regions may be double biased since the bootstrap estimates of the parameters might also suffer from bias. We called this bootstrap procedure *bias-corrected bootstrap* (BCB). Finally, we consider the *bias-corrected and endogenous bootstrap* (BCEB) in which the uncertainty about the lag order is tackled by re-estimating it for each bootstrap series; see, for example, Staszewska-Bystrova and Winker (2013).

3.3 Monte Carlo experiments

In order to compare the role played by the error distribution, the bias, the parameter uncertainty and the lag order uncertainty on the construction of forecast regions, in this section we carry out the Monte Carlo experiments. We consider the same VAR(5) and VAR(10) models of the previous chapter with three three error distributions, namely Gaussian, Student-5 and χ^2_4, which are adequately centered and re-scaled.[1] $M = 1000$ replicates of sizes $T = 100$ and 300 are generated by each of the 6 model considered. For each replicate, the AIC is implemented to choose the lag order with the maximum lag order being equal to 12 and 16 for $T = 100$ and 300, respectively. After estimating the parameters by LS, h-steps-ahead forecast densities, for $h = 1, ..., 8$, are constructed assuming Gaussian errors (STD) as in (3.1). Furthermore, bootstrap forecast densities are also constructed based on $B = 1999$ by implementing the DIST-B, BB, BCB and BCEB procedures. In each case, we construct 95% Bonferroni cubes and compute their empirical coverages after generating 2000 future true values of y_{T+h} and counting the number of them lying inside the 95% forecast Bonferroni cubes.[2]

Figure 3.1 plots the empirical coverages for the VAR(10) model which are similar regardless of the error distribution. Consider first the case when $T = 100$. The upper panels of Figure 3.1 show that coverages of the Bonferroni cubes are always below the 0.95 nominal and decreases with the forecast horizon, a pattern usually observed in VAR models with roots close to the non-stationary region; see, for instance, Kim (1999, 2004) and Fresoli et al. (in press). Furthermore, the STD regions show the largest undercoverage for all forecast horizon. For example, when the error is Student-5, empirical coverages of the Gaussian cubes are 0.834 and 0.719, when h =1 and 8, respectively. The regions constructed using the Dist-B procedure with fixed lag order and estimated parameter attain a smaller undercoverage. For example, when the errors are Student-5 and $h = 1$, the uncercoverage of the STD region is 12.23% while that of the Dist-B region is

[1] See Appendix A, models (b) and (c), for details on this VAR specifications.

[2] Monte Carlo results for the persistent VAR(5) are qualitatively the same. For his reason we include them in the Appendix D.

8.21%, and these values rise to 24.31% and 19.77% when $h = 8$. However, the largest reduction in undercoverage is obtained when incorporating the parameter uncertainty. For the same case considered above, the regions constructed using the BB procedure understate the nominal by 2.8%, when $h = 1$, and by 6.73%, when $h = 8$. It is worth noting that the difference between Dist-B and BB coverages increases as h increases. This finding suggests that incorporating the parameter uncertainty is important for the accuracy of long-run forecasts. Including further bias-correction only varies marginally the extent to which the coverages are below the nominal. Finally, tackling the lag order uncertainty provide additional gains in terms of empirical coverages, though these are less pronounced that those obtained when incorporating the parameter uncertainty. Continuing with the example above, the BCB procedure is below the nominal by 2.93% while the BCEB methods by 1.99%, when $h = 1$, and the undercoverages rise to 6.11% and 5.00% for BCB and BCEB, respectively, when $h = 8$. Therefore, in contrast to what is often posited in the literature, this finding points out that incorporating bias-correction and dealing with lag order uncertainty is not as prominent as the parameter uncertainty, at least in term of the accuracy of forecast regions; see, for instance, Chatfield (1996) and Kim (2004).

Consider now the resutls when $T = 300$ which are plotted in lower panels of Figure 3.1. We observe that, regardless of the error distribution and the forecast horizon, the empirical coverages are close to the 0.95 nominal. Also, the differences between coverages attained by the STD and bootstrap cubes are less pronounced. Yet bootstrap cubes attain slightly better empirical coverages than the Gaussian cubes, except when the parameters are fixed. On one hand, we have STD and DIST-B, with coverages that are below the nominal by about 2 or 3%, for all forecast horizon. On the other hand, we have BB, BCB and BCEB which have empirical coverages that are far from the nominal by about 0.5% when $h = 1$, and by 1 or 1.5% when $h = 8$.

In any case, these results suggest that the better empirical coverages of bootstrap cubes than Gaussian ones may be attributable, to large extent, to the sampling variability due to parameter uncertainty, though model uncertainty also plays a moderate role. The reason may be attributed to the fact that the VAR(10) model implemented in this simulation study is highly persistent with

a number of parameter to be estimated rather large in comparison to the sample size and, thus, parameter uncertainty turns out to be large; see Tjøstheim and Paulsen (1983). Notwithstanding, it is important to point out that these features can be less visible in VAR models with fewer parameters relative to sample size and lower persistence in their dynamics.

3.4 Conclusion

In this chapter we asses the effect on coverages of Bonferroni forecast regions constructed for VAR(p) models, of error distribution uncertainty, bias and parameter uncertainty and lag order uncertainty. With this purpose, we carry out Monte Carlo experiments and construct the regions using bootstrap procedures which are able to differentiate each of these sources of uncertainty. We show that the better coverages of bootstrap cubes are obtained when taking into account the parameter estimation. Moreover, our Monte Carlo experiments shows that there are valuable gains, in terms of coverages, of adding the model uncertainty at least for short forecast horizons. The benefits of implementing the bias-correction formula seems to be less visible. Finally, even when the model and parameter uncertainties and biased estimates are not a serious concern due to large sample, bootstrap forecast cubes are preferable.

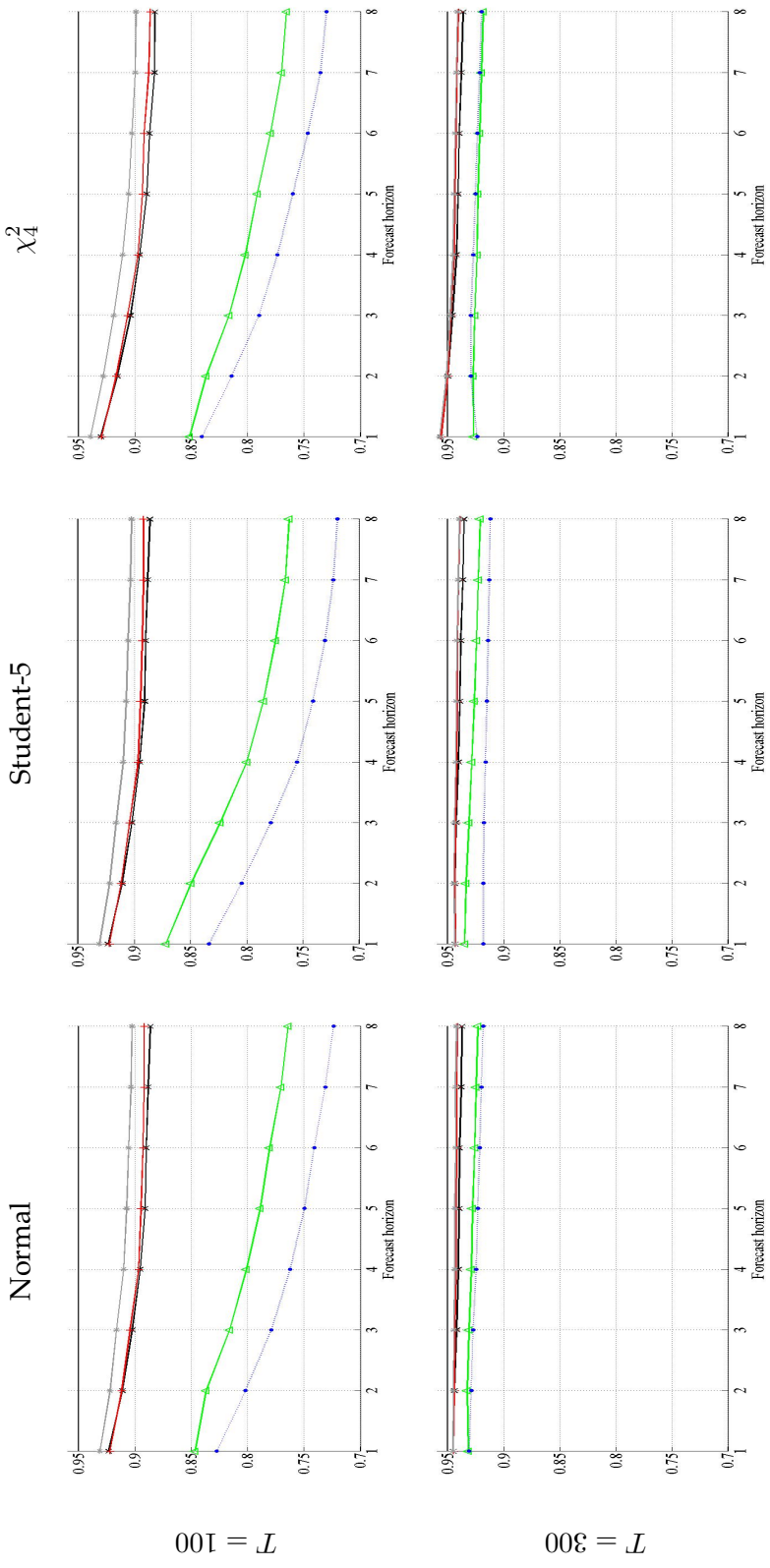


Figure 3.1: Monte Carlo averages of coverages of 95% Bonferroni cubes based on STD (blue $\cdot$), DIST-B (green $\triangle$), BB (black $\times$), BCB (red $+$) and BCEB (grey $\star$), for a near-cointegrated VAR(10) model with sample sizes $T = 100$ (first row) and 300 (second row) and Gaussian (first column), Student-5 (second column) and χ^2_4 (third column) errors.

Chapter 4

The uncertainty in forecasting returns, volatilities and correlations in DCC models

4.1 Introduction

Forecasting conditional correlations is a key issue for financial market participants when dealing, for example, with risk management, derivative pricing models, hedging strategies or portfolio allocation models; see, for example, details in Engle (2009). Multivariate GARCH (MGARCH) models are also widely implemented to macroeconomic time series. Modeling the time variation in macroeconomic volatility is important to the accuracy of inference and to explain, for example, the sources of the Great Moderation; see, for example, Clark and Ravazzolo (in press) for a very recent reference. Moreover, MGARCH are becoming popular when modeling non-financial time series as, for example, wind speed; see Jeon and Taylor (2012) among others. As a result, modeling the second order moments of multivariate time series has become a prevailing field of research and many MGARCH models have been developed with this purpose; see Bauwens et al. (2006) and Silvennoinen and Teräsvirta (2009) for comprehensive surveys. Among the many MGARCH

models available in the literature, the Dynamic Conditional Correlation (DCC) model of Engle (2002) has become one of the most popular models for the estimation of conditional correlations; see Audrino (in press) and Laurent et al. (2012) for the forecasting usefulness of DCC models. The DCC model assumes that the dynamic evolution of the pairwise correlations is similar to that of univariate GARCH conditional variances. Recently, Aielli (2013) proposes a reformulation of the original specification of the correlation driving process so that it is possible to obtain a consistent Quasi-Maximum Likelihood (QML) estimator of the parameters, giving rise to the corrected Dynamic Conditional Correlation (cDCC) model. Note that the consistency of the sample estimator of the unconditional correlations is crucial for the adequate performance of the bootstrap procedure considered in this chapter. The cDCC model is now the benchmark for empirical applications in the context of multivariate conditional heterocedastic times series of financial returns and it will be the focus of this chapter; see, for example, Engle and Kelly (2012), Hafner and Reznikova (2012), Bauwens et al. (2013), Aielli and Caporin (in press) and Audrino (in press) for applications of the cDCC model.

It is important to note that the conditional correlations in cDCC models are observable one-step-ahead; see Caporin and McAleer (2013) who suggest that, in spite of its limitations, the DCC model can be considered as a filter for estimating and forecasting conditional correlations. Consequently, the only uncertainty associated with one-step-ahead correlations is that attributable to the parameter estimation. However, when the correlations are forecast more than one-step-ahead, they also have uncertainty associated with the future forecast errors. As far as we know, there have been no attempts in the literature to measure the uncertainty associated with conditional correlations forecast by DCC models; see, for example, Laurent et al. (2012) and Caporin and McAleer (2010) who compare point forecasts from various MGARCH models without even mentioning the associated uncertainty. The only attempts of measuring the uncertainty of conditional correlations appear in the context of high frequency realized correlations, where Barndorff-Nielsen and Shephard (2004) provide asymptotic intervals and Dovonon et al. (2013) propose bootstrap intervals.

In the context of univariate GARCH models, Pascual et al. (2006) propose using bootstrap procedures to obtain forecast densities of returns and conditional volatilities; see Hartz et al. (2006), Gaglianone et al. (2011), Grigoletto and Lisi (2011), Huang and Wang (2012) and Wang et al. (2012) for empirical implementations. In a multivariate setting, Fresoli et al. (in press) illustrate how to adapt this bootstrap procedure to obtain forecast distributions of the correlations in the context of a VAR-DCC model. The proposed bootstrap procedure is appealing because it allows the construction of multivariate forecast densities for returns, volatilities and correlations that incorporate the parameter uncertainty without relying on any particular assumption about the distribution of standardized returns. This chapter analyzes the finite sample performance of the bootstrap procedure described by Fresoli et al. (in press) when implemented to obtain conditional forecast regions for future returns, volatilities and correlations in the context of the cDCC model. It is important to remark that this bootstrap procedure can be easily adapted to deal with other MGARCH models. We illustrate the procedure by implementing it to construct time-varying forecast Bonferroni regions for returns and forecast intervals for volatilities, covariances and correlations in a three-dimensional system containing daily exchange rate returns of the Euro, Japanese Yen and Australian Dollar against the US Dollar. Moreover, we implement the bootstrap algorithm to forecast the within sample conditional correlation of S&P500 and NASDAQ returns.

The rest of this chapter is structured as follows. Section 2 describes the cDCC model and the bootstrap algorithm proposed to approximate the conditional forecast densities of returns, volatilities and correlations. In Section 3, we carry out Monte Carlo experiments to analyze the finite sample properties of the bootstrap procedure. In Section 4, the bootstrap algorithm is implemented to forecast out-of-sample returns, volatilities, covariances and correlations of a three-dimensional system of aily exchange rates returns and to forecast within sample correlations of daily returns of two US market indexes. Finally, Section 5 concludes the chapter.

4.2 Bootstrap densities for conditional correlations

In this section, we establish briefly notation by describing the scalar cDCC model. The bootstrap procedure proposed by Fresoli et al. (in press) to obtain forecast densities of returns, volatilities and correlations in the context of the cDCC model is also described.

4.2.1 The DCC model

The scalar cDCC model as originally proposed by Engle (2002) and modified by Aielli (2013) is given by

$$y_t = H_t^{1/2} a_t \tag{4.1}$$

$$H_t = D_t R_t D_t \tag{4.2}$$

where y_t is a $N \times 1$ vector of returns observed at time t, a_t is a $N \times 1$ serially independent vector with zero mean and identity covariance matrix, H_t is a $N \times N$ positive definite conditional covariance matrix, D_t is a $N \times N$ diagonal matrix containing the univariate GARCH-type conditional standard deviations of each of the variables in y_t and R_t is the $N \times N$ matrix of conditional correlations. The most popular model for univariate conditional variances is the basic GARCH(1,1) of Engle (1982) and Bollerslev (1986) given by

$$\sigma_{i,t}^2 = \omega_i + \alpha_i y_{i,t-1}^2 + \beta_i \sigma_{i,t-1}^2, \qquad i = 1, ..., N\ , \tag{4.3}$$

where $\sigma_{i,t}$ is the ith element in the main diagonal of D_t. The parameters satisfy the conditions for positivity and weak stationarity of $y_{i,t}$; see Teräsvirta (2009) for a recent survey on univariate GARCH models.[1]

[1] For simplicity, we focus on the simplest cDCC(1,1) specification with GARCH(1,1) conditional variances.

The correlation matrix R_t is defined as follows

$$R_t = dg(Q_t)^{-\frac{1}{2}} Q_t dg(Q_t)^{-\frac{1}{2}} \tag{4.4}$$

where $dg(Q_t)$ is a diagonal matrix with the same diagonal elements as Q_t which determines the dynamics of the conditional correlations as follows

$$Q_t = (1 - \alpha - \beta)S + \alpha \upsilon_{t-1}\upsilon'_{t-1} + \beta Q_{t-1} \tag{4.5}$$

where $\upsilon_t = dg(Q_t)^{\frac{1}{2}}\varepsilon_t$ with $\varepsilon_t = D_t^{-1}y_t$ being the vector of standardized errors, S is the unconditional correlation matrix of υ_t and α and β are scalars. The scalar cDCC model in equations (4.1)-(4.5) restricts the dynamics of all the correlations to be governed by the same parameters, namely α and β. Furthermore, positiveness is guaranteed if S is positive definite and α, $\beta > 0$ and $\alpha + \beta < 1$.

The parameters of the cDCC model can be estimated by maximizing the Gaussian likelihood (QML). First, Denote by $\theta = (\Upsilon_1, \Upsilon_2, s)$ where $\Upsilon_1 = [(\omega_1, \alpha_1, \beta_1), ..., (\omega_N, \alpha_N, \beta_N)]$ is a vector containing the parameters in the conditional variances, $\Upsilon_2 = (\alpha, \beta)$ is a vector whose components are the parameters that govern the dynamics in the conditional correlation, and $s = vech(S)$, is a vector stacking the lower off-diagonal elements of S, with $vech$ being the off-diagonal stacking operator. Without taking into account the constants, the Gaussian log-likelihood is given by

$$\begin{aligned} L(\theta) &= -\tfrac{1}{2}\textstyle\sum_{t=1}^{T}\left[log(|H_t|) + y'_t H_t^{-1} y_t\right] \\ &= -\tfrac{1}{2}\textstyle\sum_{t=1}^{T}\left[2log(|D_t|) + log(|R_t|) + \varepsilon'_t R_t^{-1}\varepsilon_t\right] \end{aligned} \tag{4.6}$$

which is maximized with respect to all the parameters in θ in order to get the QML estimator. Under some regularity conditions, the QML estimator is consistent and asymptotically Gaussian; see McAleer et al. (2008) and Hafner and Preminger (2009). Furthermore, if the error distribution is Gaussian, the QML estimator is also efficient.

However, the maximization of the Gaussian likelihood in (4.6) with respect to θ can be computationally costly for medium to large systems. Alternatively, the estimation of the parameters can be done in three steps. The log-likelihood in (4.6) can be split into two parts and maximized sequentially. In particular, $L(\theta) = L_v(\Upsilon_1) + L_c(\Upsilon_1, \Upsilon_2, s)$, where

$$\begin{aligned} L_v(\Upsilon_1) &= -\tfrac{1}{2}\sum_{t=1}^{T}\left[2log(|D_t|) + y_t' D_t^{-2} y_t\right] \\ &= \sum_{i=1}^{N}\left(-\tfrac{1}{2}\sum_{t=1}^{T}\left[log(\sigma_{i,t}) + \tfrac{y_{i,t}^2}{\sigma_{i,t}^2}\right]\right) \end{aligned} \tag{4.7}$$

is the volatility part of the likelihood, and

$$L_c(\Upsilon_1, \Upsilon_2, s) = -\frac{1}{2}\sum_{t=1}^{T}\left[log(|R_t|) + \varepsilon_t' R_t^{-1}\varepsilon_t - \varepsilon_t'\varepsilon_t\right] \tag{4.8}$$

is the correlation part with $\varepsilon_t = D_t^{-1} y_t$. First, the parameters involved in the conditional variance equations, Υ_1, are estimated by maximizing (4.7), obtaining $\widehat{\Upsilon}_1$ and the corresponding standardized returns, $\widehat{\varepsilon}_t = \widehat{D}_t^{-1} y_t$. Note that the last term in (4.8), $\varepsilon_t'\varepsilon_t$, only depends on the conditional variances parameters and can be neglected when miximizing L_c with respect to Υ_2 and s, giving rise to

$$L_c(\Upsilon_1, \Upsilon_2, s) = -\frac{1}{2}\sum_{t=1}^{T}\left[log(|dg(Q_t)^{-\frac{1}{2}} Q_t dg(Q_t)^{-\frac{1}{2}}|) + \upsilon_t' dg(Q_t)^{-\frac{1}{2}} Q_t^{-1} dg(Q_t)^{-\frac{1}{2}}\upsilon_t\right]. \tag{4.9}$$

Furthermore, υ_t does not depend on S, thus it is possible to substitute S within expression (4.9) by

$$\widetilde{S}_{(\widehat{\Upsilon}_1, \Upsilon_2)} = T^{-1}\sum_{t=1}^{T} dg(Q_t)^{\frac{1}{2}}\widehat{\varepsilon}_t\widehat{\varepsilon}_t' dg(Q_t)^{\frac{1}{2}}, \tag{4.10}$$

obtaining $L_c(\widehat{\Upsilon}_1, \Upsilon_2, \widetilde{s})$ which can be maximized with respect to Υ_2. Once both Υ_1 and Υ_2 are estimated, one can obtain $\widehat{\upsilon}_t$ and estimate S by the corresponding moment estimator. This three-steps estimation procedure is known as correlation targeting. Aielli (2013) derives the asymptotic distribution of the cDCC three-steps correlation targeting estimator which is consistent and

asymptotically Gaussian under some standard assumptions.

The cDCC model delivers point forecasts of correlations and volatilities. The h-steps-ahead forecast of the volatility of each return in y_t is obtained easily by iterating forward as follows

$$\sigma^2_{i,T+h|T} = \omega_i \sum_{j=1}^{h-2}(\alpha_i + \beta_i)^j + (\alpha_i + \beta_i)^{h-1}\sigma^2_{i,T+1|T} \tag{4.11}$$

where the one-step-ahead forecast of the conditional variance is determined by the observed series of returns as follows

$$\sigma^2_{i,T+1|T} = \frac{\omega_i}{1-\alpha_i-\beta_i} + \alpha_i \sum_{j=0}^{T-1} \beta_i^j \left(y^2_{i,T-j} - \frac{\omega_i}{1-\alpha_i-\beta_i} \right). \tag{4.12}$$

Similarly, the point forecast of the conditional correlation matrix R_{T+h} is given by

$$R_{T+h|T} = dg(Q_{T+h|T})^{-\frac{1}{2}} Q_{T+h|T} dg(Q_{T+h|T})^{-\frac{1}{2}}. \tag{4.13}$$

where

$$Q_{T+h|T} = (1-\alpha-\beta)S \sum_{j=0}^{h-2}(\alpha+\beta)^j + (\alpha+\beta)^{h-1} Q_{T+1|T} \tag{4.14}$$

and

$$Q_{T+1|T} = S + \alpha \sum_{j=0}^{T-1} \beta^j \left[(dg(Q_{T-j})^{-\frac{1}{2}} D^{-1}_{T-j} y_{T-j} y'_{T-j} D^{-1}_{T-j} dg(Q_{T-j})^{-\frac{1}{2}} - S \right]. \tag{4.15}$$

Using $D_{T+h|T}$ and $R_{T+h|T}$, it is possible to construct $H_{T+h|T}$ using equation (4.2). Finally, assuming further that ε_t is conditionally Gaussian, one can approximate the joint forecast density of y_{T+h} as follows

$$y_{T+h}|y_1, ..., y_T \sim N(0, H_{T+h|T}). \tag{4.16}$$

Note that, even if the errors were truly Gaussian, the forecast density in (4.16) is only valid for $h = 1$ and it should be considered as an approximation for $h > 1$; see Drost and Nijman (1993) in

the context of univariate GARCH models. In any case, in practice, standardized financial returns often depart from Gaussianity usually due to fat-tails; see, for instance, Bauwens and Laurent (2005), Pesaran and Pesaran (2007) and Rossi and Spazzini (2010). If the errors are non-Gaussian, future densities predicted as in (4.16) might be a poor approximation of the conditional distribution of returns.

Furthermore, the parameters in equations (4.11) and (4.14) are unknown and must be estimated using, for example, the three-steps correlation target estimator described by Engle (2009) and Aielli (2013); see Hafner and Reznikova (2012) for a comparison of alternative estimators of DCC models. Aielli (2013) derives heuristically the asymptotic distribution of the cDCC three-steps correlation target estimator and shows that it is consistent and asymptotically Gaussian under standard assumptions. The density in (4.16) does not incorporate the parameter uncertainty and, consequently, it will underestimate the uncertainty associated with future returns.

Finally, note that equations (4.11) and (4.14) only yield point forecasts of future volatilities and correlations. The bootstrap procedure described next allows constructing forecast densities for future returns, volatilities and correlations that incorporate the parameter uncertainty without relying on specific assumptions on the error distribution.

4.2.2 Bootstrap forecasts in the cDCC model

Fresoli et al. (in press) propose a bootstrap procedure to approximate the forecast density of future returns, volatilities and correlations for the DCC model. In this section, we describe how to adapt this procedure to forecast in the context of the cDCC model.

The algorithm to obtain bootstrap replicates of returns, volatilities and correlations is as follows.

Step 1. Estimate the model parameters θ by the three-steps correlation target procedure as described by Aielli (2013). Denote the estimated parameter by $\widehat{\theta}$. Obtain $\widehat{a}_t = \widehat{H}_t^{-\frac{1}{2}} y_t$ and the corresponding empirical distribution function denoted by $\widehat{F}_{\widehat{a}}$.

Step 2. Compute $Q_1^* = \widehat{S}$ and the correlation matrix $R_1^* = dg(Q_1^*)^{-\frac{1}{2}} Q_1^* dg(Q_1^*)^{-\frac{1}{2}}$. Obtain $\varepsilon_1^* = R_1^{*\frac{1}{2}} a_1^*$ and $v_1^* = dg(Q_1^*)^{\frac{1}{2}} \varepsilon_1^*$, where a_1^* is a random draw with replacement from $\widehat{F}_{\widehat{a}}$. Construct recursively for $t = 2, ..., T$, a bootstrap replicate of ε_t^*, v_t^* and R_t^* as follows

$$Q_t^* = (1 - \widehat{\alpha} - \widehat{\beta})\widehat{S} + \widehat{\alpha} v_{t-1}^* v_{t-1}^{*\prime} + \widehat{\beta} Q_{t-1}^*,$$

$$R_t^* = dg(Q_t^*)^{-\frac{1}{2}} Q_t^* dg(Q_t^*)^{-\frac{1}{2}},$$

$$\varepsilon_t^* = R_t^{*\frac{1}{2}} a_t^*, \tag{4.17}$$

$$v_t^* = dg(Q_t^*)^{\frac{1}{2}} \varepsilon_t^*, \tag{4.18}$$

where a_t^* are random draws with replacement from $\widehat{F}_{\widehat{a}}$. Consider $\sigma_{i,1}^{*2} = \widehat{\omega}_i / (1 - \widehat{\alpha}_i - \widehat{\beta}_i)$ and $y_{i,1}^* = \varepsilon_{i,1}^* \sigma_{i,1}^*$, for $i = 1, ..., N$. Obtain recursively for $t = 2, ..., T$, a bootstrap replicate of y_t and their conditional variances as follows

$$\sigma_{i,t}^{2*} = \widehat{\omega}_i + \widehat{\alpha}_i y_{i,t-1}^{*2} + \widehat{\beta}_i \sigma_{i,t-1}^{*2}, \tag{4.19}$$

$$y_{i,t}^* = \varepsilon_{i,t}^* \sigma_{i,t}^*, \tag{4.20}$$

where $\varepsilon_{i,t}^*$ is the ith element of ε_t^* obtained in (4.17).

Step 3. Obtain a bootstrap estimate of the parameters $\widehat{\theta}^*$ by fitting the cDCC model to the bootstrap replicate $\{y_1^*, ..., y_T^*\}$. Construct $\widehat{D}_t^*$ and $\widehat{Q}_t^*$, for $t = 1, ..., T$, which contain the in-sample estimates of the univariate conditional standard deviations and correlations obtained by using the bootstrap estimates of the parameters and the returns series as follows

$$\widehat{\sigma}_{i,t}^* = \left[\widehat{\omega}_i^* + \widehat{\alpha}_i^* y_{i,t-1}^2 + \widehat{\beta}_i^* \widehat{\sigma}_{i,t-1}^{*2} \right]^{\frac{1}{2}}, \tag{4.21}$$

$$\widehat{Q}_t^* = (1 - \widehat{\alpha}^* - \widehat{\beta}^*)\widehat{S}^* + \widehat{\alpha}^* \left[dg(\widehat{Q}_{t-1}^{*\frac{1}{2}}) \widehat{D}_{t-1}^{*-1} y_{t-1} y_{t-1}' \widehat{D}_{t-1}^{*-1} dg(\widehat{Q}_{t-1}^{*\frac{1}{2}}) \right] + \widehat{\beta}^* \widehat{Q}_{t-1}^*. \tag{4.22}$$

Step 4. Compute one-step-ahead bootstrap forecasts of conditional correlations, variances and returns according to

$$Q^*_{T+1|T} = \widehat{S}^* + \widehat{\alpha}^* \sum_{j=0}^{T-1} \widehat{\beta}^{*j} \left[dg(\widehat{Q}^*_{T-j})^{\frac{1}{2}} \widehat{D}^{*-1}_{T-j} y_{T-j} y'_{T-j} \widehat{D}^{*-1}_{T-j} dg(\widehat{Q}^*_{T-j})^{\frac{1}{2}} - \widehat{S}^* \right] \tag{4.23}$$

$$R^*_{T+1|T} = dg(Q^*_{T+1|T})^{-\frac{1}{2}} Q^*_{T+1|T} dg(Q^*_{T+1|T})^{-\frac{1}{2}}, \tag{4.24}$$

$$\varepsilon^*_{T+1} = R^{*\frac{1}{2}}_{T+1|T} a^*_{T+1},$$

$$\upsilon^*_{T+1} = dg(Q^*_{T+1|T})^{\frac{1}{2}} \varepsilon^*_{T+1}$$

$$\sigma^{2*}_{i,T+1|T} = \frac{\widehat{\omega}^*_i}{1 - \widehat{\alpha}^*_i - \widehat{\beta}^*_i} + \widehat{\alpha}^*_i \sum_{j=0}^{T-1} \widehat{\beta}^{*j}_i \left(y^2_{i,T-j} - \frac{\widehat{\omega}^*_i}{1 - \widehat{\alpha}^*_i - \widehat{\beta}^*_i} \right), \tag{4.25}$$

$$y^*_{i,T+1} = \varepsilon^*_{i,T+1} \sigma^*_{i,T+1|T}. \tag{4.26}$$

where a^*_{T+1} is a random draw with replacement from $\widehat{F}_{\widehat{a}}$. Likewise, obtain future values of the correlations, volatilities and returns through the following recursion for $h = 2, ..., H$

$$Q^*_{T+h|T} = (1 - \widehat{\alpha}^* - \widehat{\beta}^*)\widehat{S}^* + \widehat{\alpha}^* \upsilon^*_{T+h-1} \upsilon^{*\prime}_{T+h-1} + \widehat{\beta}^* Q^*_{T+h-1|T}, \tag{4.27}$$

$$R^*_{T+h|T} = dg(Q^*_{T+h|T})^{-\frac{1}{2}} Q^*_{T+h|T} dg(Q^*_{T+h|T})^{-\frac{1}{2}}, \tag{4.28}$$

$$\varepsilon^*_{T+h} = R^{*\frac{1}{2}}_{T+h|T} a^*_{T+h}$$

$$\upsilon^*_{T+h} = dg(Q^*_{T+h|T})^{\frac{1}{2}} \varepsilon^*_{T+h}$$

$$\sigma^{2*}_{i,T+h|T} = \widehat{\omega}^*_i + \widehat{\alpha}^*_i y^{*2}_{i,T+h-1} + \widehat{\beta}^*_i \sigma^{*2}_{i,T+h-1|T}; \tag{4.29}$$

$$y^*_{i,T+h} = \varepsilon^*_{i,T+h} \sigma^*_{i,T+h|T}. \tag{4.30}$$

where a^*_{T+h} are random draws with replacement from $\widehat{F}_{\widehat{a}}$.

Step 5. Repeat steps 2 to 4, B times, obtaining $y^{*(b)}_{T+h|T}$, $D^{*(b)}_{T+h|T}$ and $R^{*(b)}_{T+h|T}$ for $h = 1, ..., H$ and

$b = 1, ..., B$.

It is worth mentioning that the one-step-ahead forecast of Q_t in (4.23) depends on $\widehat{D}_t^*$ and $\widehat{Q}_t^*$, which are the in-sample estimates of D_t and Q_t evaluated at the bootstrap estimates of the parameters, $\widehat{\theta}^*$, and the original observations. Therefore, although the one-step-ahead bootstrap forecasts of the correlation matrix is conditional on the observed sample, their variability is affected by the parameter uncertainty. The only component which varies from one bootstrap replicate to another is the bootstrap estimates of the parameters, $\widehat{\theta}^*$, while $\{y_1, ..., y_T\}$ are kept fixed throughout all bootstrap replicates. When forecasting further into the future, the bootstrap forecasts of the correlation matrix also includes the error uncertainty by sampling with replacement from $\widehat{F}_{\widehat{a}}$ to obtain $a_{T+h}^{*(b)}$. A similar argument holds for the bootstrap forecasts of the conditional variances; see also the discussion in Pascual et al. (2006) for a similar argument in the univariate context.

Using the bootstrap replicates, one can obtain estimates of the multivariate forecast densities of returns, volatilities and correlations. Also, it is possible to obtain their forecast regions and intervals. Consider first the construction of time-varying forecast regions for returns. First, we can obtain bootstrap forecast ellipsoids with probability content $(1 - \alpha)100\%$ which are given by

$$BE_y(T+h) = \left\{ y | y \in \left[y_{T+h} - \overline{\widehat{y}^*_{T+h|T}} \right] H^{*-1}_{T+h|T} \left[y_{T+h} - \overline{\widehat{y}^*_{T+h|T}} \right] < C^* \right\} \tag{4.31}$$

where $\overline{\widehat{y}^*_{T+h}}$ is the sample mean of the B bootstrap replicates of returns $\widehat{y}^{*(b)}_{T+h}$, $H^*_{T+h|T}$ is the corresponding sample covariance of y^*_{T+h}, and C^* is the $(1 - \alpha)100\%$ percentile of the bootstrap distribution of the following quadratic form

$$\left[\widehat{y}^{*(b)}_{T+h} - \overline{\widehat{y}^{*(b)}_{T+h}} \right] H^{*(b)-1}_{T+h|T} \left[\widehat{y}^{*(b)}_{T+h} - \overline{\widehat{y}^{*(b)}_{T+h}} \right].$$

It is well-known that the ellipsoids are only appropriate when the distribution is Gaussian; see Fresoli et al. (in press) and Wolf and Wunderli (2012). Furthermore, constructing ellipsoids can be computationally complicated when the dimension of the system is very large. Alternatively, one

can construct Bonferroni cubes with probability content at least $(1-\alpha)100\%$ for future returns which are given by

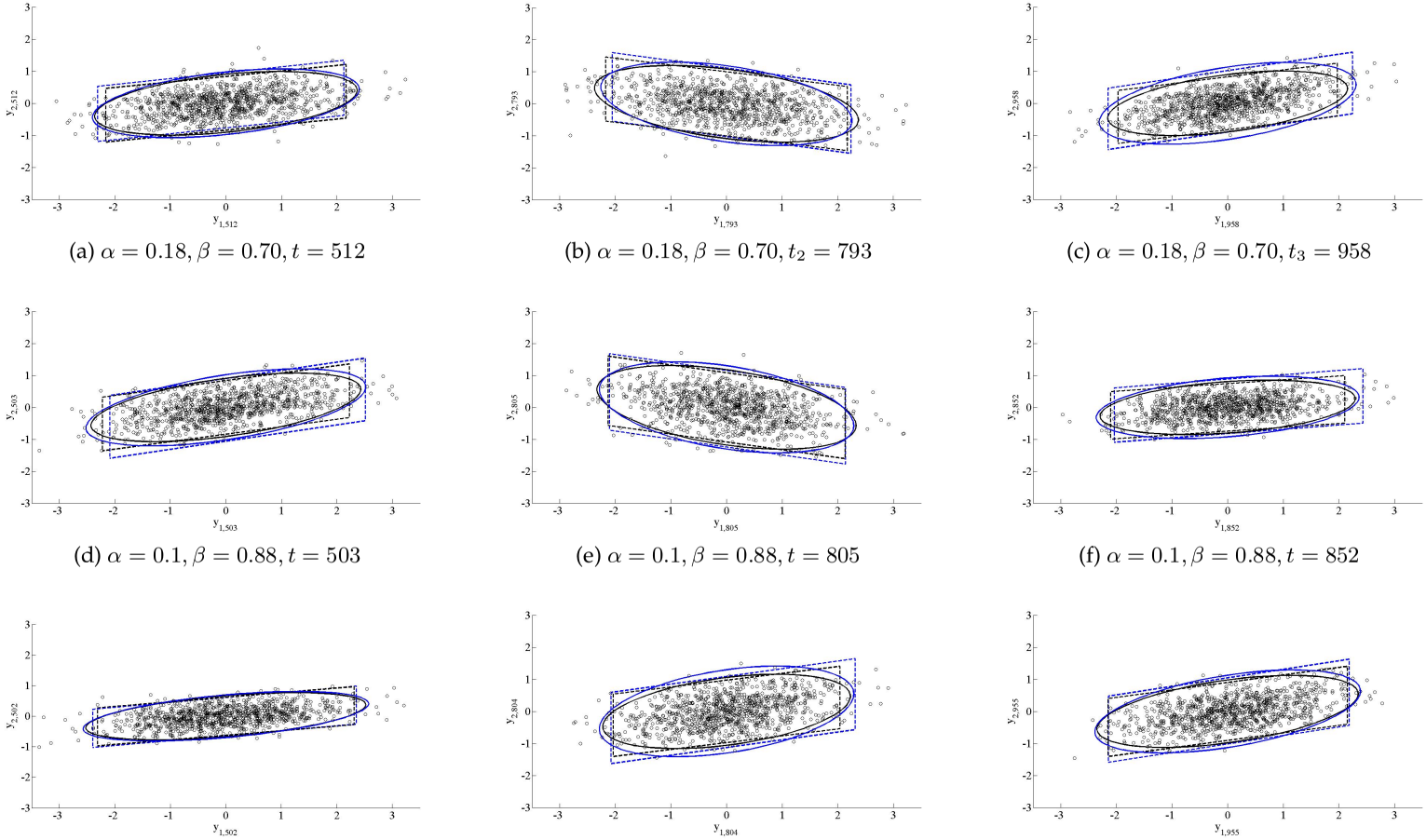
$$BC_y(T+h) = \left\{ y | y \in \cup_{j=1}^{N} \left[q^{y*}_{j,T+h}\left(\frac{\alpha}{2K}\right), q^{y*}_{j,T+h}\left(1-\frac{\alpha}{2K}\right)\right]\right\} \tag{4.32}$$

where $q^{y*}_{j,T+h}(\zeta)$ is the ζth quantile of the empirical bootstrap distribution of the jth return, $G^{*}_{y_{j,T+h}}(x) = \#(y^{*(b)}_{j,T+h|T} \leq x)/B$. As the Bonferroni cubes do not offer information about the association between returns, they can be corrected to be expressed in the direction of the correlations; see Fresoli et al. (in press).

As an illustration, we consider several bivariate cDCC models obtained combining three sets of correlation parameters, $\Upsilon_2 = (0.18, 0.70)$, $(0.1, 0.88)$ and $(0.03, 0.95)$, two error distributions, namely Gaussian and Student-7, two values for the unconditional correlation, 0.5 and 0.9, and univariate GARCH models with parameters $\Upsilon_1 = (0.05, 0.05, 0.90,\ 0.01, 0.10, 0.85)$. So we consider 12 different models. For each model we generate a time series of size $T = 1000$ and estimate the cDCC model parameters by the three-steps correlation target estimator.[2] Then, the bootstrap procedure is implemented with $B = 1000$ bootstrap replicates to obtain one-steps-ahead forecast densities of returns. We also obtain the corresponding Gaussian forecast densities computed as in equation (4.16) with $H_{T+1|T}$ substituted by $\widehat{H}_{T+1|T}$. From each of these conditional densities, we construct the 95% one-step-ahead ellipsoids and corrected Bonferroni cubes. The results are plotted in Figure 4.1 for the three models with Gaussian errors and unconditional correlation equal to 0.5 for three selected within-sample periods. In each, a realization of 1000 returns, y_{T+1}, has been represented by dots. A quick inspection of all panels in Figure 4.1 reveals that, for a particular model, both the shape and the slope of the ellipsoids and Bonferroni cubes change over time. Also, it is noticeable that the bootstrap ellipsoids have larger volumes than the corresponding

[2]The computations have been carried out by using a MATLAB code developed by the first author in a workstation with processor Intel Core i5-2.50GHz and 8GB of RAM. Correlation target estimates are obtained numerically by using *fminunc* with the *interior point* algorithm. For each GARCH(1,1) process, the intercept, ω_i, is restricted to be greater than e^{-6} while the elements in S are obtained as centered correlations. All other parameters, $\alpha_i, \beta_i, \alpha$ and β, are restricted to lie in the interval $(e^{-6}, 1-e^{-6})$ with $\alpha_i + \beta_i < 1$ and $\alpha + \beta < 1$ being less than $1 - e^{-6}$.

Gaussian ellipsoids and, in some cases, the true observation is included in the former but not in the latter. The discrepancy between bootstrap and standard regions is even more pronounced for the corrected Bonferroni cubes. Given that the models considered in Figure 4.1 have Gaussian errors and the forecast horizon is one, the differences between the regions based on the Gaussianity assumption and the bootstrap regions can be attributed to the parameter uncertainty.



(a) $\alpha = 0.18, \beta = 0.70, t = 512$ (b) $\alpha = 0.18, \beta = 0.70, t_2 = 793$ (c) $\alpha = 0.18, \beta = 0.70, t_3 = 958$

(d) $\alpha = 0.1, \beta = 0.88, t = 503$ (e) $\alpha = 0.1, \beta = 0.88, t = 805$ (f) $\alpha = 0.1, \beta = 0.88, t = 852$

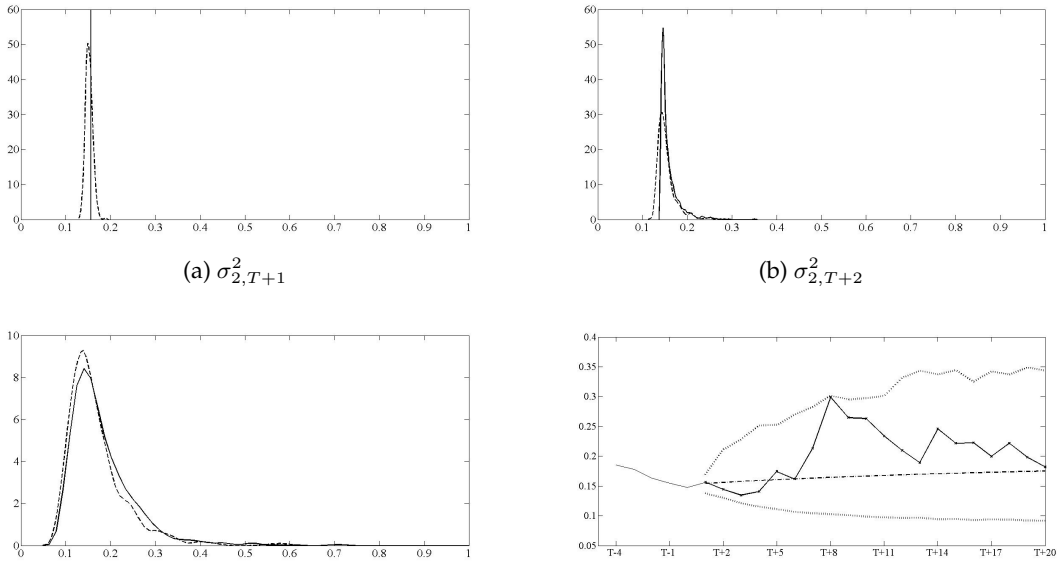
(g) $\alpha = 0.03, \beta = 0.95, t = 502$ (h) $\alpha = 0.03, \beta = 0.95, t = 805$ (i) $\alpha = 0.03, \beta = 0.95, t = 956$

Figure 4.1: 95% one-step ahead Gaussian ellipsoids and corrected cubes (black discontinuous), bootstrap ellipsoids and corrected cubes (blue discontinuous) together with the return realization (dot) for three bivariate cDCC models with $T = 1000$ and Gaussian errors. Rows show different configurations of parameters while columns different periods of times.

Following the same ideas as in Pascual et al. (2006) for univariate conditional variances, we can use the bootstrap replicates of the jth univariate conditional variance to construct its empirical bootstrap distribution, $G^*_{\sigma^2_{j,T+h}}(x) = \#(\sigma^{2*(b)}_{j,T+h|T} \leq x)/B$. Given that, in the cDCC model, the conditional variances are assumed to be contemporaneously independent, here we only consider the construction of the (1-α)100% marginal forecast intervals which are given by

$$BI_{\sigma^2}(T+h) = \left\{\sigma^2 | \sigma^2 \in \left[q^{\sigma^2*}_{j,T+h}\left(\frac{\alpha}{2}\right), q^{\sigma^2*}_{j,T+h}\left(1-\frac{\alpha}{2}\right)\right]\right\} \tag{4.33}$$

where $q^{\sigma^2*}_{j,T+h}(\zeta) = G^{*-1}_{\sigma^2_{j,T+h}}$ is the ζth percentile of the empirical bootstrap distribution of the jth conditional variance. To illustrate the construction of bootstrap forecast densities and their corresponding forecast intervals for conditional variances, we consider a bivariate time series generated by the cDCC with correlation parameters $\Upsilon_2 = (0.03, 0.95)$, the unconditional correlation being 0.9 and the error distribution being Student-7. For the sake of brevity, we just focus on the conditional variance for the second variable in the system. Figure 4.2 displays kernel estimates of the h-steps-ahead bootstrap conditional densities together with the empirical conditional densities of $\sigma^2_{2,T+h}$, for $h = 1, 2$ and 20. The empirical densities have been obtained using a kernel estimate implemented to 2000 simulated future values of the series, conditional on $\{y_1, ..., y_T\}$. First, observe that the one-step-ahead forecast of $\sigma^2_{2,T+1}$ plotted in panel (a) of Figure 4.2 is concentrated on a fixed value. As we stated before, no uncertainty is associated with forecasting one-step-ahead conditional variances. Yet the conditional bootstrap distribution of $\sigma^2_{2,T+1}$ has some degree of variability reflecting the parameter uncertainty. Second, note that the conditional 2-steps-ahead forecast of $\sigma^2_{2,T+2}$ is bounded by the value of the intercept, ω_2 and, consequently, the empirical density is asymmetric to the right. However, the bootstrap replicates, $\sigma^{2*}_{2,T+2}$, might be smaller than this value depending on the bootstrap estimates of the parameters, $\left(\widehat{\omega}^*_2, \widehat{\alpha}^*_2, \widehat{\beta}^*_2\right)$. Third, the uncertainty associated with $\sigma^2_{2,T+h}$ clearly increases as we forecast further into the future due to the addition of error uncertainty, a fact that is reflected by both, the empirical and the bootstrap conditional distributions of $\sigma^2_{2,T+h}$. After all, the first three panels of Figure 4.2 display



(a) $\sigma^2_{2,T+1}$ (b) $\sigma^2_{2,T+2}$

(c) $\sigma^2_{2,T+20}$ (d) $\sigma^2_{2,T+h}$

Figure 4.2: Kernel estimates of empirical (continuous line) and bootstrap densities (discontinuous line) of $\sigma^2_{2,T+h}$ within the bivariate cDCC model with sample size $T = 1000$ and Student-7 errors for (a) $h = 1$, (b) $h = 2$ (c) $h = 20$ (d) 95% h-steps-ahead forecast interval, for $h = 1, ..., 20$, (discontinous lines) together with point forecasts (dash-dot line) and out-of-sample true realization (continuous line).

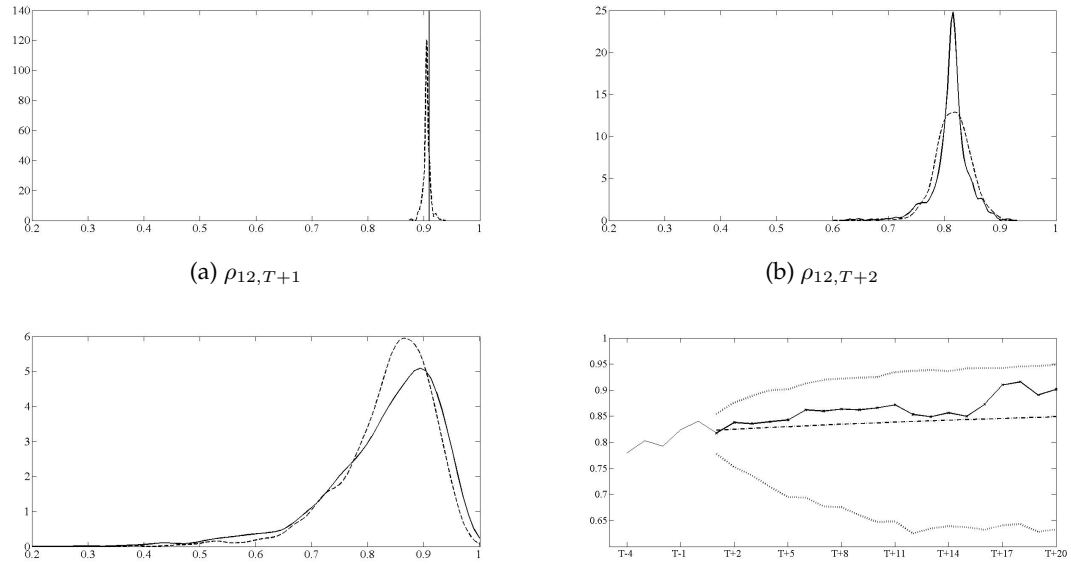
bootstrap densities that approximate quite well the true empirical ones with the approximation being better for longer forecast horizon than short ones. This is due to the fact that the model implies known one-step-ahead forecast.

Panel (d) of Figure 4.2 plots one realization of $\sigma^2_{2,T+h}$, for $h = 1, ..., 20$, together with the corresponding point forecasts and the 95% bootstrap intervals. It can be observed that the true out-of-sample realizations of the volatilities are included within the bootstrap forecast intervals.

Finally, we implement the bootstrap procedure to construct h-steps-ahead bootstrap forecast densities for conditional correlations. Panels (a), (b) and (c) of Figure 4.3 plot kernel estimates of the conditional empirical h-steps-ahead forecast densities of $\rho_{12,T+h}$, for $h = 1, 2$ and 20 together with kernel estimates of the corresponding bootstrap densities for the bivariate cDCC model, with correlation parameters $\Upsilon_2 = (0.10, 0.88)$, unconditional correlation equal to 0.9 and Student-

7 errors. The conclusions are similar to those obtained when forecasting volatilities. However, the densities are asymmetric to the left because in this particular illustration the marginal correlation is close to one. After all, it is clear from the first three panels of Figure 4.3 that the bootstrap procedure it approximates adequately the shape of the conditional empirical densities of $\rho_{12,T+h}$.

Once more, the h-steps-ahead bootstrap densities can be used to construct the corresponding intervals for the forecasts of the correlations. Panel (d) of Figure 4.3 plots an out-of-sample realization of the correlations, $\rho_{12,T+h}$, for $h = 1, 2..., 20$, together with the point forecasts obtained after estimating the model and the 95% bootstrap forecast intervals. We observe that the true correlations fall within the bounds of the forecast intervals. Further, we observe that the point forecasts underestimate the corresponding the out-of-sample correlations; see Engle and Sheppard (2001) and Engle (2009).



(a) $\rho_{12,T+1}$ (b) $\rho_{12,T+2}$

(c) $\rho_{12,T+20}$ (d) $\rho_{12,T+h}$

Figure 4.3: Kernel estimates of empirical densities (continuous line) and bootstrap densities (discontinuous line) of $\rho_{12,T+h}$ within of a bivariate cDCC with $T = 1000$ and Student-7 error for (a) $h = 1$, (b) $h = 2$ (c) $h = 20$ (d) 95% h-steps-ahead forecast interval, for $h = 1, ..., 20$, (discontinuous lines) together with point forecasts (dash dot line) and out-of-sample true realization (continuous line).

4.3 Monte Carlo evidence

In this section, we carry out Monte Carlo experiments to analyse the finite sample properties of the bootstrap procedure described in the previous section. We focus on two of the bivariate cDCC models, used for illustration in the previous section. The first model (model 1) has $\Upsilon_2 = (0.10, 0.88)$ and an unconditional correlation of 0.5 while the second model (model 2) has $\Upsilon_2 = (0.03, 0.95)$ with an unconditional correlation of 0.9. Note that these two models bring up two different situations. In both models the persistence is 0.98. However, the first model is characterized by a large response-to-shock parameter, α while the second model resembles closely the parameters found when fitting cDCC models to financial series, with α being relatively small; see, for example, the empirical application in next section. Three alternative distributions for a_t are assumed, namely, Gaussian, Student-7 and χ^2_5. The second distribution is chosen to reproduce the heavy tails often observed in the distribution of standardized financial returns while the third is chosen to represent potential asymmetries. The sample sizes considered are $T = 500$, 1000 and 2000. The number of Monte Carlo replicates is 500; see Hafner and Franses (2009) and Hafner and Reznikova (2012) for similar number of Monte Carlo replications. For each replicate, we generate 2000 conditional future values of y_{T+h}, σ^2_{T+h} and $\rho_{12,T+h}$, for h =1,...,20, to approximate the empirical distribution of returns, conditional variances and correlations, respectively. The parameters are estimated by the three-steps correlation target estimator of Aielli (2013). The number of bootstrap replicates is $B = 1000$, and these are used to construct 95% h-steps-ahead bootstrap forecast intervals for each return, conditional variance and correlation in the system. In order to assess the small sample properties of the bootstrap procedure, the empirical coverage is computed by counting the number of future values inside the corresponding intervals. In addition, to measure the adequacy of the bootstrap distributions to approximate the out-of-sample distributions of returns, we compute the Earth Mover's Distance (EMD) between the h-step-ahead bootstrap distribution of the ith return and the corresponding empirical distribution. Also, we compute the EMD when the h-steps-ahead distribution of the ith return is approximated by assuming Nor-

mality; see Rubner et al. (1998) for the definition of the EMD and Arroyo and Maté (2009) for an application of the EDM to forecast histograms.[3]

Figure 4.4 displays the empirical coverages of the 95% bootstrap forecast intervals for $y_{2,T+h}$, $\sigma^2_{2,T+h}$ and $\rho_{12,T+h}$ when considering model 1.

First, we consider the empirical coverages of $y_{2,T+h}$ which are plotted in the first row of Figure 4.4. Although the empirical coverages of the bootstrap forecast intervals are below the nominal of 95%, they are always above 94.5%, and this happen irrespective of the forecast horizon. We also observe that, as the sample size increases, the coverage rates of the bootstrap intervals are closer to the nominal, a feature that is more evident when the error distribution departs from Gaussianity. For the purpose of comparing alternative methods, the first row of Figure 4.4 also includes the empirical coverage of the Gaussian forecast intervals. As expected, the difference between the Gaussian and bootstrap intervals are less pronounced when the error distribution is Gaussian. In contrast, in the case of the Student-7 error distribution, the coverage rates of the Gaussian intervals are well below the nominal as well as below the bootstrap coverages for all sample sizes. Surprisingly, when the error distribution is χ^2_5, the Gaussian intervals provide empirical coverages above the nominal, though this distortion seems to dampen with the forecast horizon.

[3]The EMD can be approximated as follows. Let $x_{(1)} \leq ... \leq x_{(B)}$ and $y_{(1)} \leq ... \leq y_{(B)}$ be ordered realizations of X and Y. Then the EMD is given by $EMD_{xy} = \left[\frac{1}{N}\sum_{i=1}^{R} \mid x_{(i)} - y_{(i)} \mid^2\right]^{\frac{1}{2}}$, which is just the L^2 distance between the ordered vectors. We are thankful to Ruben Zamar for suggesting using this measure.

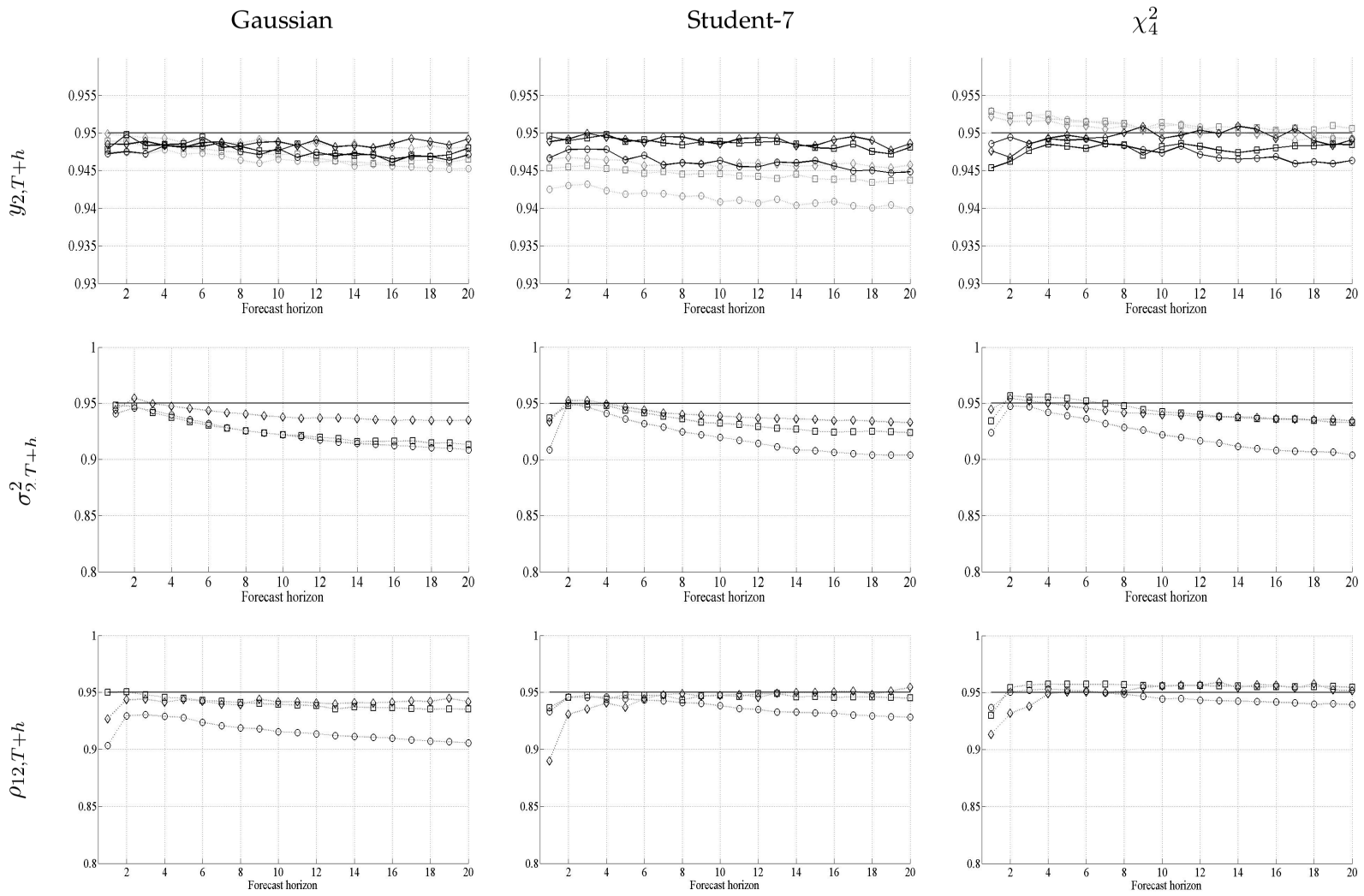


Figure 4.4: Monte Carlo average of coverages of h-steps-ahead bootstrap forecast intervals of $y_{2,T+h}$ (first row), $\sigma^2_{2,T+h}$ (second row) and $\rho_{12,T+h}$ model 1 with sample size $T = 500$ ($\circ$), 1000 ($\square$) and 2000 ($\diamond$) with Gaussian (first column), Student-7 (second column) and χ^2_5 (third column) errors. Nominal coverage 95%.

The coverages plotted in the first row of Figure 4.4 are not informative about whether the shape of the density of returns is well approximated by the bootstrap density. This is why, we also compute the EMD. Figure 4.5 plots the average EMD between the Gaussian and the bootstrap conditional densities of $y_{2,t}$ with respect to the corresponding empirical density, when the sample size is $T = 500$. Consider first the results when the errors are Gaussian. In this case, we can observe that the performance of the Gaussian and the bootstrap densities is similar for $h = 1$. Note that, in this case, the true forecast density is Normal. However, if $h > 1$, the forecast densities are not Normal and the EMDs of the Gaussian densities increase with h and are clearly larger than the EMDs of the bootstrap densities. Therefore, even if the errors are truly Gaussian, it is worth to implement the bootstrap procedure to forecast more than one-step-ahead. On the other hand, when the errors are either Student-7 or χ^2_5, the bootstrap EMDs are clearly smaller than the EMDs of the Gaussian densities.

The small sample properties of bootstrap forecast intervals for univariate GARCH(1,1) processes have studied in detail by Pascual et al. (2006). In this chapter, however, we also display some results regarding average coverages of the conditional variance of $y_{2,t}$. The second row of Figure 4.4 plots the empirical coverages of the bootstrap forecast intervals of $\sigma^2_{2,T+h}$ for Gaussian, Student-7 and χ^2_5 errors in model 1. We can observe that the coverages of the bootstrap forecast intervals depend on the forecast horizon. Regardless of the error distribution, the coverages are below the nominal of 95% when $h = 1$. This undercoverage is more pronounced when the errors are non-Gaussian and the sample size is small. In any case, it is important to note that the coverages are always over 0.9. When $h > 2$, the coverages tend to decrease as the forecast horizon increases. We can observe that, as expected, the coverages are closer to the nominal as T increases suggesting that the procedure is consistent. For instance, when $T = 2000$, the coverages are below the nominal by less than 2.5%, and this happen for all the forecast horizons and error distributions; these results are qualitatively similar to those reported by Pascual et al. (2006), though there is a slight difference in the decline of the coverage curve as the forecast horizon increases.

The third row of Figure 4.4 plots the Monte Carlo coverages of the bootstrap forecast intervals

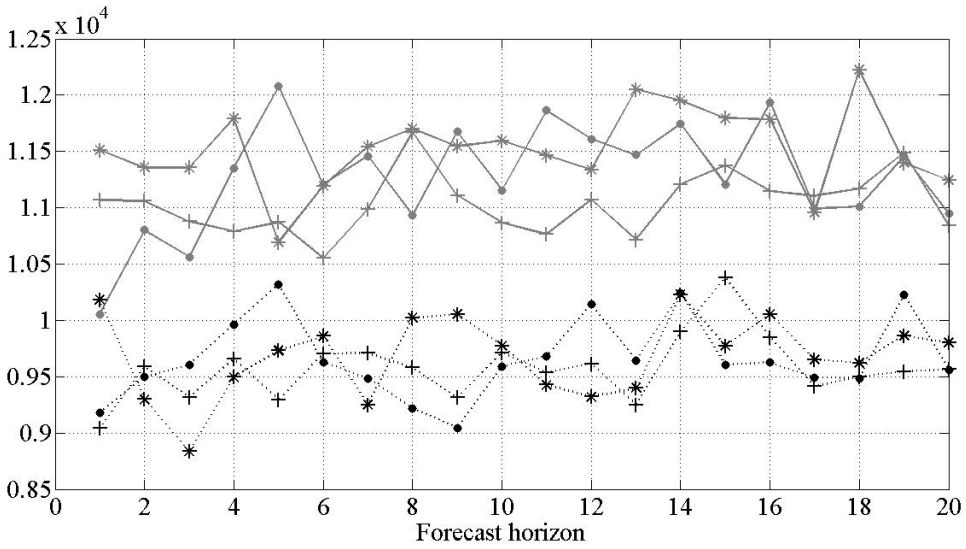
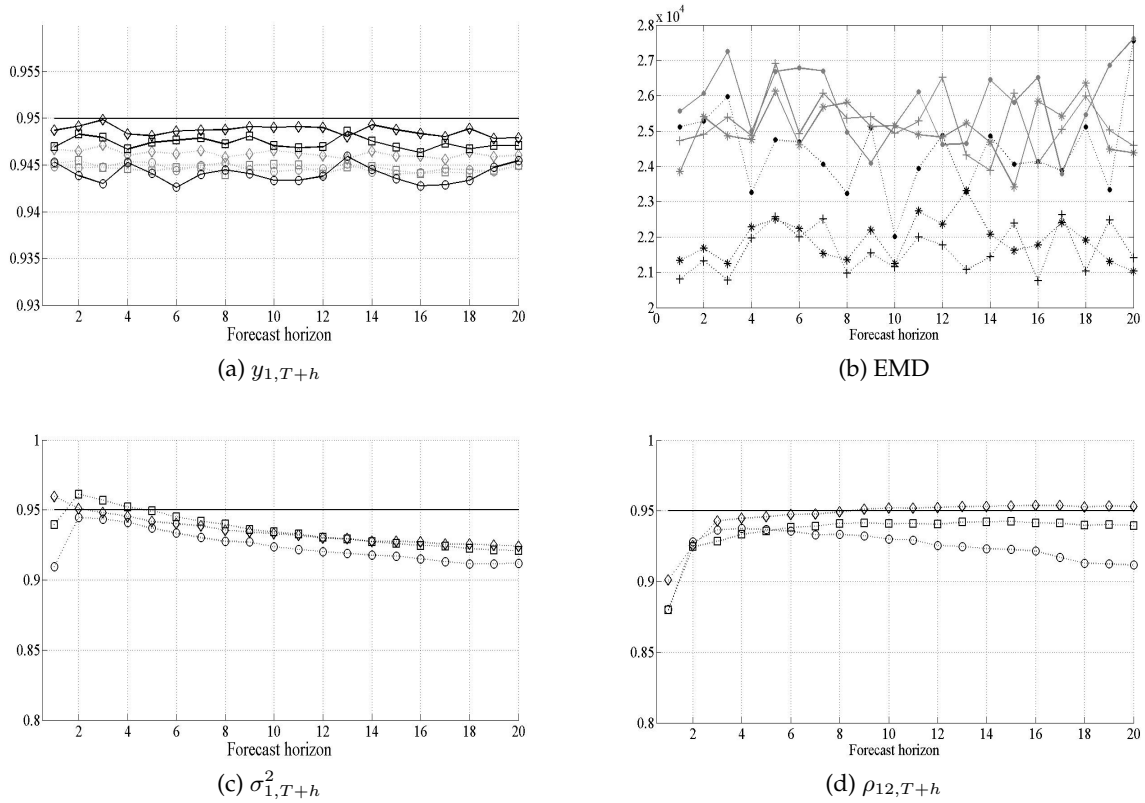


Figure 4.5: Monte Carlo averages of EMD distances between the h-steps-ahead Gaussian forecast densities (grey lines) and bootstrap densities (black lines) of $y_{2,T+h}$ and the corresponding empirical conditional density for $y_{2,t}$ in model 1 with $T = 500$ and Gaussian ($\bullet$), Student-7 ($+$) and χ^2_5 ($*$) errors.

for the conditional correlation, $\rho_{12,t}$ in model 1 when the nominal coverage is 95%. The patterns observed are roughly similar to those observed for the conditional variance of $y_{2,t}$. Once more, the coverages of the one-step-ahead bootstrap forecast are under the nominal for all error distribution and sample sizes. This undercoverage is slightly larger than that observed for conditional variances. When $h \geq 2$, we also observe that the coverages decrease with the forecast horizon. However, this undercoverage is only observed for the smallest sample size when $T = 500$. When $T = 1000$ and 2000, the coverages of the bootstrap forecast intervals of $\rho_{12,t}$ are approximately equal to the nominal.

Finally, Figure 4.6 plots the empirical coverages of the bootstrap forecast intervals for $y_{1,T+h}$, $\sigma_{1,T+h}$, $\rho_{12,T+h}$ and average EMD between the Gaussian forecast densities and bootstrap densities of $y_{1,T+h}$ and the corresponding empirical density, when the error distribution is Student-7 and $T = 500$, in model 2. Much of what we commented about Figures 4.4 and 4.5 is still valid for Figure 4.6. In panel (a) of Figure 4.6, we observe that the empirical coverages of the bootstrap forecast intervals of $y_{1,T+h}$ are below the nominal. Also, bootstrap intervals of $y_{1,T+h}$ improve



(a) $y_{1,T+h}$ (b) EMD

(c) $\sigma^2_{1,T+h}$ (d) $\rho_{12,T+h}$

Figure 4.6: Monte Carlo average of coverages of h-steps-ahead bootstrap forecast intervals of (a) $y_{1,T+h}$, (b) $\sigma^2_{1,T+h}$ and (c) $\rho_{12,T+h}$ model 2 with sample size $T = 500$ ($\circ$), 1000 ($\square$) and 2000 ($\diamond$) with Student-7 errors and nominal coverage 95%, and (d) averages of EMD distances between the h-steps-ahead Gaussian forecast densities (grey lines) and bootstrap densities (black lines) of $y_{1,T+h}$ and the corresponding empirical conditional density for $y_{1,t}$ in model 1 with $T = 500$ and Gaussian ($\bullet$), Student-7 ($+$) and χ^2_5 ($*$) errors.

their coverage with the sample size and they provide better accuracy than the corresponding Gaussian intervals for moderate and large sample sizes. Moreover, panel (b) of Figure 4.6 shows that the bootstrap densities are clearly closer to the corresponding true density than the Gaussian densities when the error distributions depart from Gaussianity. Still, when the error distribution is Gaussian, the bootstrap seems to be as good as the Gaussian density when $h = 1$. Panels (c) and (d) of Figure 4.6, that plot the empirical coverages $\sigma_{1,T+h}$ and $\rho_{12,T+h}$, respectively, look qualitatively similar to those plotted in the second and third rows of Figures 4.4. Quantitatively, comparing the empirical coverages of Figures 4.4 and Figure 4.6 for $\rho_{12,T+h}$ we observe that are slightly closer to the nominal in model 1 than in model 2. This seems to be caused by the greater uncertainty encountered when α is larger relative to the β. In such a case, the the bootstrap forecast intervals become considerably wider and, thus, they are more likely to capture future realization of the conditional correlation.

After all, the Monte Carlo simulation results show that the bootstrap algorithm can reasonably deal with forecasting returns, volatilities and correlations cDCC models.

4.4 Empirical application

In this section we implement the bootstrap procedure to forecast out-of-sample returns, volatilities, covariances and correlations of a three dimensional system of daily exchange rates returns, Euro , Japanese Yen and Australian Dollar against the US Dollar. Moreover, we illustrate the bootstrap algorithm within sample by forecasting the conditional correlation between S&P500 and NASDAQ returns.

4.4.1 Bootstrap out-of-sample forecast of exchange rates returns

In this section, the bootstrap algorithm is implemented to forecast future returns, volatilities, covariances and correlations of a system of three currencies, Euro (EUR), Japanese Yen (JPY) and Australian Dollar (AUD) against the US Dollar (USD). Note that this section is purely illustrative

of the application of the bootstrap procedure and limited attention is given to the economic interpretation of the results. Daily exchange rates of the three currencies, $P_{i,t}$ $i = 1, 2$ and 3, have been observed from 3/1/2003 to 13/9/2013 with 2693 observations. The full sample period have been split in a estimation period spanning from 3/1/2003 to 1/8/2013 with $T = 2663$ and a forecast out-of-sample period from 2/8/2013 to 13/9/2013 with $H = 30$. The currencies are transformed to log-returns as usual by $y_{i,t} = 100 \times \triangle log(P_{i,t})$. Several descriptive statistics are reported in Table 4.2 for the overall period as well as for the estimation and out-of-sample periods. Note that, for all three currencies, the univariate skewness and excess kurtosis are significantly different from zero in the within-sample and overall periods, though they do not differ significantly from zero in the out-of-sample period. Table 4.2 also reports results for the multivariate extensions of skewness and excess kurtosis as proposed by Kilian and Demiroglu (2000). It worth noting that the multivariate distribution of returns is positive skewed in the within-sample and overall periods, but seems to be symmetric around zero in the out-of-sample period. Similarly, heavy tails are considerable in the within-sample and overall periods, but not in the out-of-sample period. These features of the return distribution across time are corroborated by the univariate Normality and multivariate Jarque-Bera tests which reject the null of individual Normality of each of the series and joint Gaussianity, respectively, in the within-sample and overall periods, though there is not strong evidence against marginal and joint Gaussianity in the out-of-sample period; see Fiorentini et al. (2004) for the univariate Normality test and Kilian and Demiroglu (2000) for the multivariate Jarque-Bera test and the bootstrap approximation of their distributions. With respect to the dependence of returns, Table 4.2 reports the Ljung-Box statistics adjusted to account for conditional heteroscedasticity as proposed by Diebold (1988) and denoted by $Q(l)$, for $l = 10$. It may be observed that the returns show no linear dependence in any of the sub-periods considered. We also report the multivariate portmanteau statistic of order 10 proposed by Patilea and Raïssi (2013) to test for multivariate linear dependence in the presence of conditional heterocedasticity, which suggests an uncorrelated vector of returns, irrespective of the sub-period considered. However, marginal squared returns are characterized by significant correlations according to the Ljung-Box

statistics of order 10 for the individual squared returns in the within-sample and overall periods. Likewise, the multivariate portmanteau statistic of order 10 implemented to $vech(y_t y_t')$ shows strong dependence in these periods. When we focus on the out-of-sample period, we note that, even though $Q_2(10)$ is not significant, it is noticeably larger than $Q(10)$ for each of the series, both individually and jointly, thus supporting the existence of a weak dependence in the square returns in this period. Finally, the last column of Table 4.2 reports the Ljung-Box statistics for the first ten cross-correlations between returns and future squared returns, $Q_{12}(10)$, suggesting the presence of leverage effect only in the case of AUD. In the within-sample and overall periods, the descriptive results are in line with the stylized facts of returns, which are less visible in the out-of-sample period, a fact that might be attributed to the calm in financial markets during 2013.

With respect to the correlation structure, the sample pairwise unconditional correlations for the EUR-AUD and EUR-JPY are 0.63 and 0.25, respectively, and both are significant according to the corrected Breusch and Pagan (1980) test to account for heterocedasticity; see Halunga et al. (2011). On the other hand, the unconditional correlation between AUD and JPY is not significantly different from zero. We also test the constancy of the conditional correlations, implementing the test suggested by Tse (2000). First, we implement the test to the pairwise correlations, rejecting the null of constant conditional correlations in all cases. Note that although the unconditional JPY-AUD correlation is not different from zero, the conditional correlation seems to evolve over time. Then, we test for the joint constancy of the three correlations in the system. The LM statistic is 360.52 with a p-value equal to 0.00. Therefore, the constant conditional correlation hypothesis of the whole system is also rejected. Finally, we test for asymmetries in conditional correlations, implementing the exceedance test of Hong et al. (2007) which is based on the exceedance correlations above or below a threshold. For instance, when the threshold is the median of each exchange rate return, the exceedance test statistics are 0.30, 0.53 and 0.17 for EUR-JPY, EUR-AUD and JPY-AUD, respectively, with the corresponding p-values being all larger than 0.40. Considering different thresholds does not change this result. Therefore, the presence of asymmetries in

Table 4.1: Summary statistics of daily returns for Euro (EUR), Japanese Yen (JPY) and Australian Dollar (AUD) against the US Dollar observed for the within-sample period from 3/1/2003 to 1/8/2013, the out-of-sample period 2/8/2013 to 13/9/2013 and the overall period from 3/1/2003 to 1/8/2013. Mean, standard deviation (Std), skewnees (Skew) and excess kurtosis (Ek), Normality test, Ljung-Box for returns ($Q(10)$), squared returns ($Q_2(10)$) and cross-correlations between returns and future squared returns ($Q_{12}(10)$). Asymptotic p-values in parenthesis.

	Mean	Std	Skew	Ek	NT[1]	$Q(10)$	$Q_2(10)$	$Q_{12}(10)$
Within-sample								
EUR	0.00	0.28	-0.10 (0.02)	2.45 (0.00)	669.67 (0.00)	0.41 (0.99)	356.96 (0.00)	0.50 (0.99)
JPY	0.00	0.29	-0.29 (0.00)	4.55 (0.00)	2324.29 (0.00)	0.53 (0.99)	164.39 (0.00)	3.25 (0.97)
AUD	-0.01	0.40	0.74 (0.00)	12.22 (0.00)	16761.15 (0.00)	2.88 (0.98)	1665.31 (0.00)	80.95 (0.00)
Multivariate			339.02 (0.00)	18999.12 (0.00)	19338.13 (0.00)	188.55 (0.46)	4438.06 (0.00)	
Out-of-sample								
EUR	-0.01	0.16	0.24 (0.30)	-0.78 (0.19)	0.85 (0.65)	0.04 (0.99)	8.29 (0.60)	0.00 (0.99)
JPY	0.00	0.31	0.10 (0.41)	-0.28 (0.38)	0.12 (0.94)	0.32 (0.99)	14.66 (0.15)	0.99 (0.02)
AUD	-0.05	0.30	-0.48 (0.14)	-0.42 (0.32)	1.14 (0.56)	0.55 (0.99)	7.85 (0.64)	0.02 (0.99)
Multivariate			4.19 (0.13)	1.31 (0.66)	5.50 (0.18)	98.37 (0.11)	350.76 (0.62)	
Overall period								
EUR	0.00	0.28	-0.10 (0.02)	2.47 (0.00)	689.65 (0.00)	0.40 (0.99)	366.79 (0.00)	0.50 (0.99)
JPY	0.00	0.29	-0.28 (0.00)	4.48 (0.00)	2285.50 (0.00)	0.54 (0.99)	164.05 (0.00)	3.17 (0.98)
AUD	-0.01	0.40	0.74 (0.00)	12.21 (0.00)	16945.48 (0.00)	2.82 (0.99)	1683.26 (0.00)	80.21 (0.00)
Multivariate			332.72 (0.10)	18973.98 (0.00)	19306.71 (0.00)	184.03 (0.49)	4486.00 (0.00)	

[1] Bootstrap p-values as proposed by Kilian and Demiroglu (2000).

Table 4.2: Estimates of the parameters of the cDCC(1,1) model and diagnosis for daily returns of Euro (EUR), Japanese Yen (JPY) and Australian Dollar (AUD) exchange rates observed from 3/1/2003 to 1/8/2013. Lagrange Multiplier test statistics for the univariate and multivariate standardized residuals, ($LM_2^*(10)$). Asymptotic p-values in parenthesis.

Conditional variances	ω_i	α_i	β_i	$LM_2^*(10)$
EUR	$2.29e^{-4}$ (0.00)	0.03 (0.00)	0.97 (0.00)	3.50 (0.97)
JPY	$8.37e^{-4}$ (0.00)	0.03 (0.00)	0.96 (0.00)	10.23 (0.42)
AUD	$1.06e^{-3}$ (0.00)	0.06 (0.00)	0.93 (0.00)	10.13 (0.43)
Conditional correlations	$s_{i,j}$	α	β	
EUR-JPY	0.26 (0.00)	—	—	—
EUR-AUD	0.63 (0.00)	0.04 (0.00)	0.95 (0.00)	278.35 (0.99)
JPY-AUD	0.11 (0.02)	—	—	—

correlations is rejected.

The results above suggest fitting the symmetric cDCC model.[4] Table 4.2 reports the three-steps target correlation estimates of the parameters. Observe that all parameters are significant and close to the values usually estimated when the cDCC model is fitted to systems of financial returns. Also, observe that the constant of the equation of the correlation between the Japanese Yen and the Australian Dollar exchange rate returns is not significantly different from zero when α =1%. The last column of Table 4.2 reports the LM test statistics for the univariate squared standardized residuals, $\widehat{a}_{i,t}$, and the multivariate standardized residuals, $vech(\widehat{a}_t)$; see Engle (1982) and Lütkepohl (2004).[5] The null hypothesis of no further dynamics in standardized returns is not rejected. Therefore, the cDCC model seems to be adequate to represent the dynamic dependence in the system of exchange rate returns considered.

Using the bootstrap replicates of the parameters and the original observations, we obtain one-

[4]Further analysis shows that the presence of leverage effect in AUD is rather weak.

[5]When dealing with the standardized residuals, we report results for the Lagrange Multiplier test as proposed by Engle (1982) and Lütkepohl (2004) as its validity is guaranteed under general conditions. However, asymptotic validity of the portmanteau test is still unsolved in the case the test is implemented to residuals. Yet, results for the multivariate portmanteau tests of Hosking (1980) and Duchesne and Lalancette (2003) lead to the same conclusions.

step-ahead densities of returns, volatilities and correlations during the within-sample period. Figure 4.7 plots the one-step-ahead forecasts and the 95% bootstrap intervals for correlations for the period between 2/7/2006 and 1/8/2013 together with the sample estimate of the unconditional correlations. We can observe that the intervals are rather tiny so that the forecasts of correlations are very reliable. The second column of Figure 4.7 highlights this conclusion by plotting the one-step-ahead forecasts of the correlations for the last nineteen days, covering the period from 8/7/2013 to 1/8/2013. We can observe large and frequent changes in the EUR-JPY and JPY-AUD conditional correlations while the correlations between EUR-AUD are more stable. Note that, although the unconditional correlation between AUD and JPY is not significantly different from zero, Figure 4.7 shows that the conditional correlations fluctuate around zero, with only 1.04% of the bootstrap intervals containing the zero within their corresponding bounds.

The proposed bootstrap procedure is also implemented to forecast returns, volatilities and correlations out-of-sample. Figure 4.8 plots h-steps-ahead Gaussian and bootstrap pairwise forecast ellipsoids and cubes for $h = 1, 20$ and 30. In each figure, the true observed return is represented by a dot. First, we observe that the Gaussian and bootstrap ellipsoids are very similar. This result is not surprising as the means and conditional covariances used to construct them do not differ significantly. Consider, for example, the EUR and AUD returns. In this case, the Gaussian ellipsoid is centered on $(0.00, 0.00)$ and the one-step-ahead covariance matrix is given by $vech(\widehat{H}_{T+1|T}) = (0.05, 0.04, 0.13)$, while the bootstrap ellipsoid is centered at $(0.00, 0.01)$ and the bootstrap conditional covariance matrix is given by $vech(\widehat{H}^*_{T+1|T}) = (0.06, 0.05, 0.13)$. Clearly, these two moments are rather similar in both cases and the corresponding ellipsoids are also similar. Second, as expected, we note that the bootstrap corrected cubes are larger than those based on Gaussianity, a fact that might be attributed to the parameter uncertainty and/or the non-Gaussianity of the errors. In order to differentiate the effect of both features, we implement a slight modification of the proposed bootstrap procedure, which consists in omitting the parameter uncertainty by fixing the parameters at their three-steps correlation target estimates throughout all bootstrap forecast replicates. As a result, we avoid the sampling variability due to the pa-

rameter uncertainty while we are able to capture the asymmetries in the empirical distribution of the error. In this case, the difference between both bootstrap corrected cubes can be attributed to the parameter uncertainty. Figure 4.8 shows that both bootstrap cubes are rather similar. Consequently, the difference between the Gaussian and bootstrap cubes is, to a large extent, due to the non-Gaussianity of the errors with the parameter uncertainty playing a minor role. Once more, this is not surprising as the sample size in this empirical application is considerably large. Finally, observe that the financial market have reached a period of calm, as highlighted also in the descriptive analysis, extreme returns are less frequent during the out-of-sample period and, consequently, they tend to fall within the boundaries of all the regions plotted in Figure 4.8.

Finally, in order assess the adequacy of the out-of-sample h-steps-ahead bootstrap forecast intervals for correlations, we compute realized correlations; see Andersen et al. (2001), Andersen et al. (2003) and Barndorff-Nielsen and Shephard (2004). It is widely recognized that the estimation of realized covariances and correlations suffers from asynchronous trading and market microstructure noise, causing the covariance estimator to be biased and inconsistent; see, for instance, McAleer and Medeiros (2008) and Corsi and Audrino (2012). In our empirical example, the intra-day data is sampled using 1-minute intervals, avoiding the effect of non-synchronous data. The sample data spans from 2/8/2013 to 13/9/2013, which corresponds to 30 weekdays, and consists of transaction prices for the three exchange rates considered. In order to avoid the effect of non official trading we only consider prices between 9:00am and 4:00pm. After computing the intra-daily returns, the realized covariance matrix is obtained as follows

$$H_t^r = \sum_{i=j}^{m} y_{(m)t+j/m} y'_{(m)t+j/m} \tag{4.34}$$

for $j = 1, ..., m$ where $y_{(m)t+j/m}$ is the returns vector during the interval $\frac{1}{m}$. Barndorff-Nielsen and Shephard (2004) establish the asymptotic distribution of the realized covariance and correlation matrices defined in (4.34).[6]

[6] Dovonon et al. (2013) propose an i.i.d. bootstrap for realized correlations which is shown to outperform the asymptotic theory of Barndorff-Nielsen and Shephard (2004). Comparing the bootstrap intervals for daily correlations based

Figure 4.9 plots out-of-sample forecasts for the volatilities, covariances and correlations of the cDCC model. The diagonal corresponds to volatilities, the lower-off-diagonal to covariances and the upper-off-diagonal to correlations. All figures include the 95% and 99% h-steps-ahead bootstrap forecast intervals of volatilities, covariances and correlations together with the point forecast of the cDCC model and their corresponding realized measures. The latter have been plotted together with the 95% asymptotic intervals obtained as suggested by Barndorff-Nielsen and Shephard (2004). We can observe that the path depicted by the realized measures and their 95% asymptotic intervals is, in general, well-captured by their corresponding 95% and 99% bootstrap forecast intervals during the out-of-sample period. Note also that, when $h = 1$, the realized volatilities and covarianes show a large shifts and, as a result, they fall far outside their corresponding bootstrap intervals bounds. However, the corresponding realized correlations correlation show up to be more stable and only two of them are not included in the corresponding 99% one-step-ahead bootstrap interval. After all, the bootstrap intervals seem to depict reasonably well the uncertainty surrounding the conditional estimates obtained with the cDCC model.

on daily or high-frequency data is left for further research.

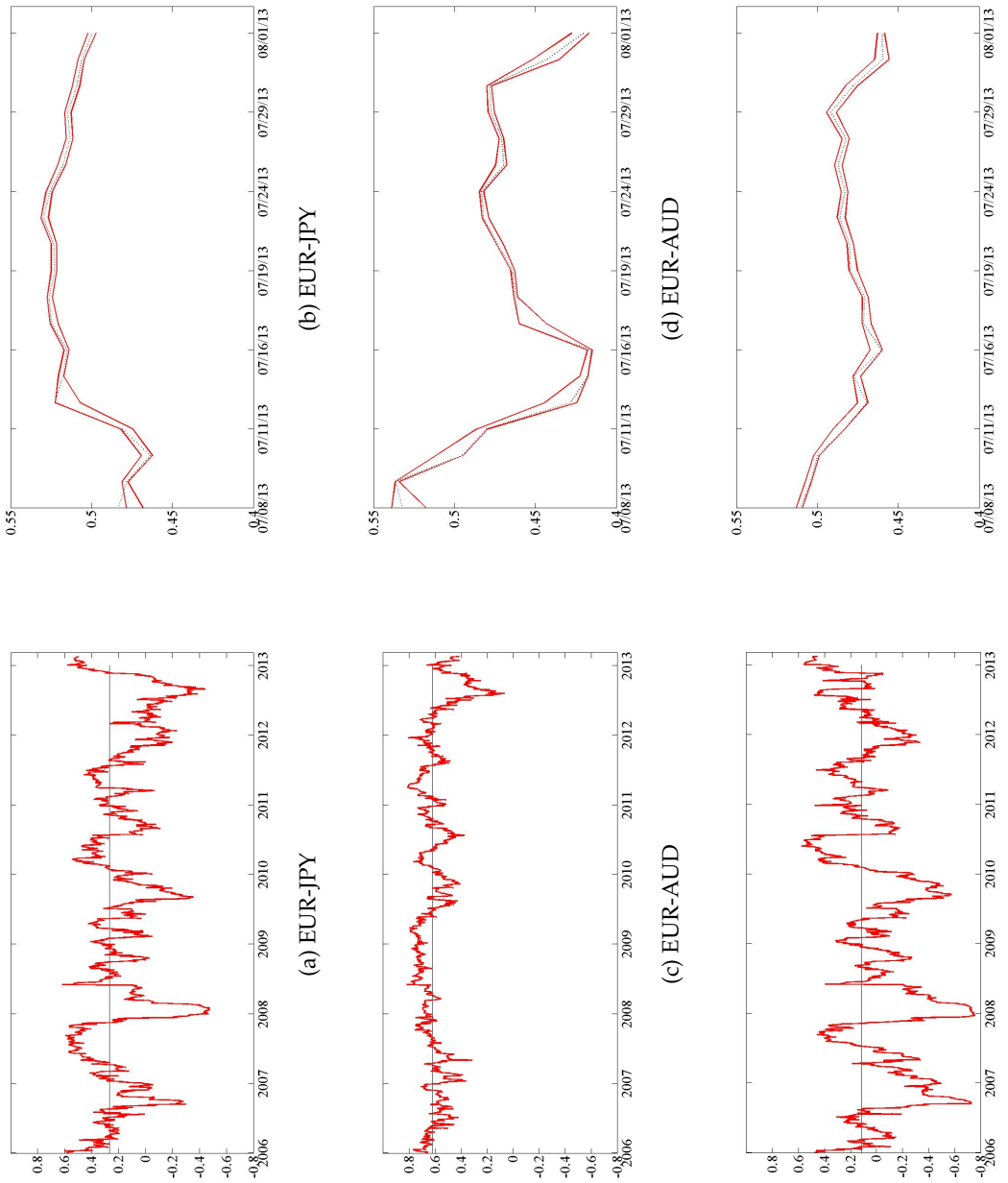


Figure 4.7: Conditional correlation estimates for EUR-JPY, EUR-AUD and JPY-AUD together with the 95% one-step-ahead bootstrap forecast intervals (red lines) for the periods from 2/1/2006 to 1/8/2013 (first column) and from 8/7/2013 to 8/1/2013 (second column).

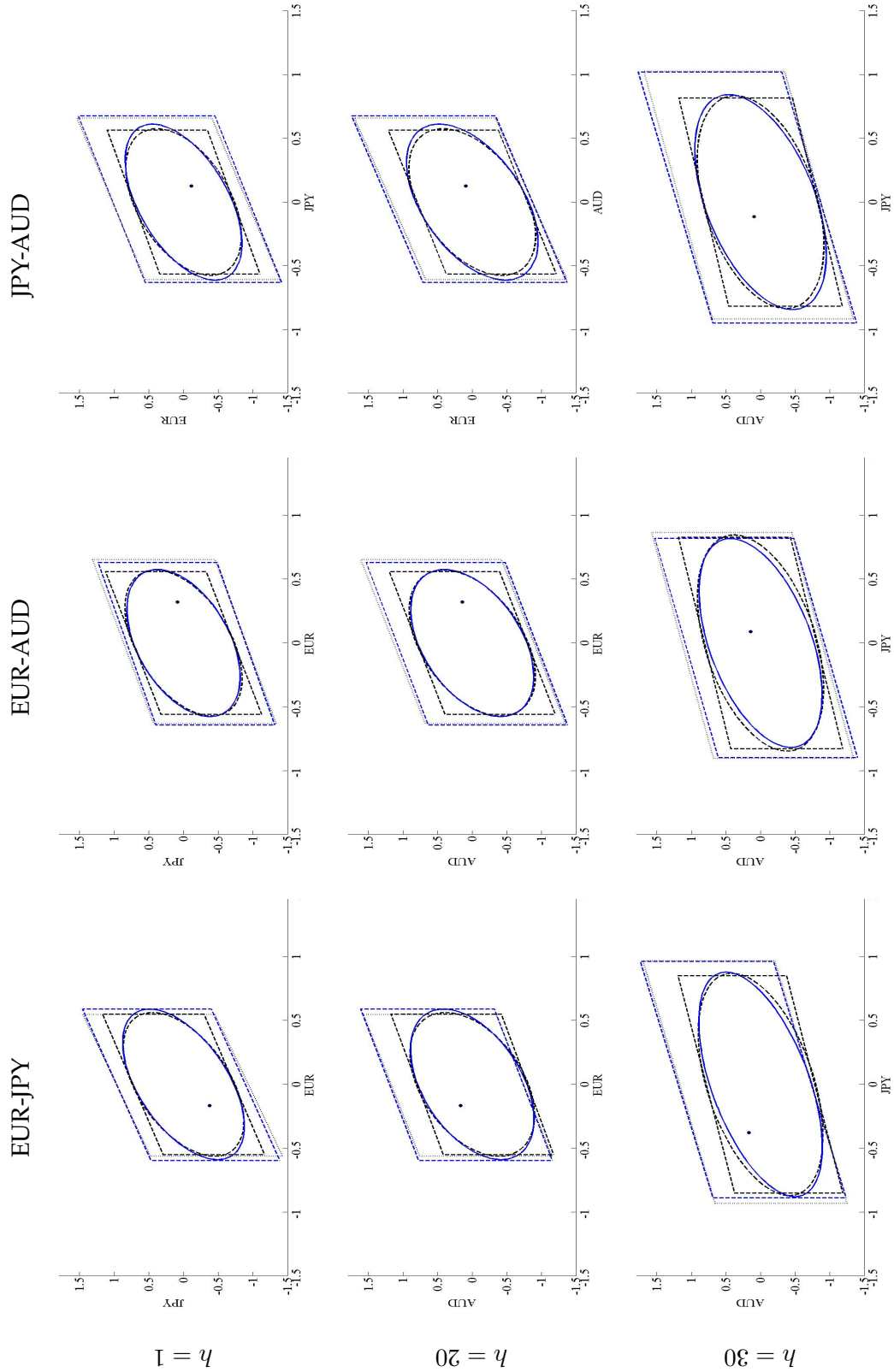


Figure 4.8: Pairwise 95% h-steps-ahead Gaussian ellipsoids and corrected cubes (black discontinuous line), bootstrap ellipsoids and corrected cubes (blue discontinuous line), and bootstrap corrected cube without parameter uncertainty (dotted line), for $h = 1$ (first row), $h = 20$ (second row) and $h = 30$ (third row), of EUR-JPY (first column), EUR-AUD (second column) and JPY-AUD (third column) returns together with the corresponding out-of-sample realizations.

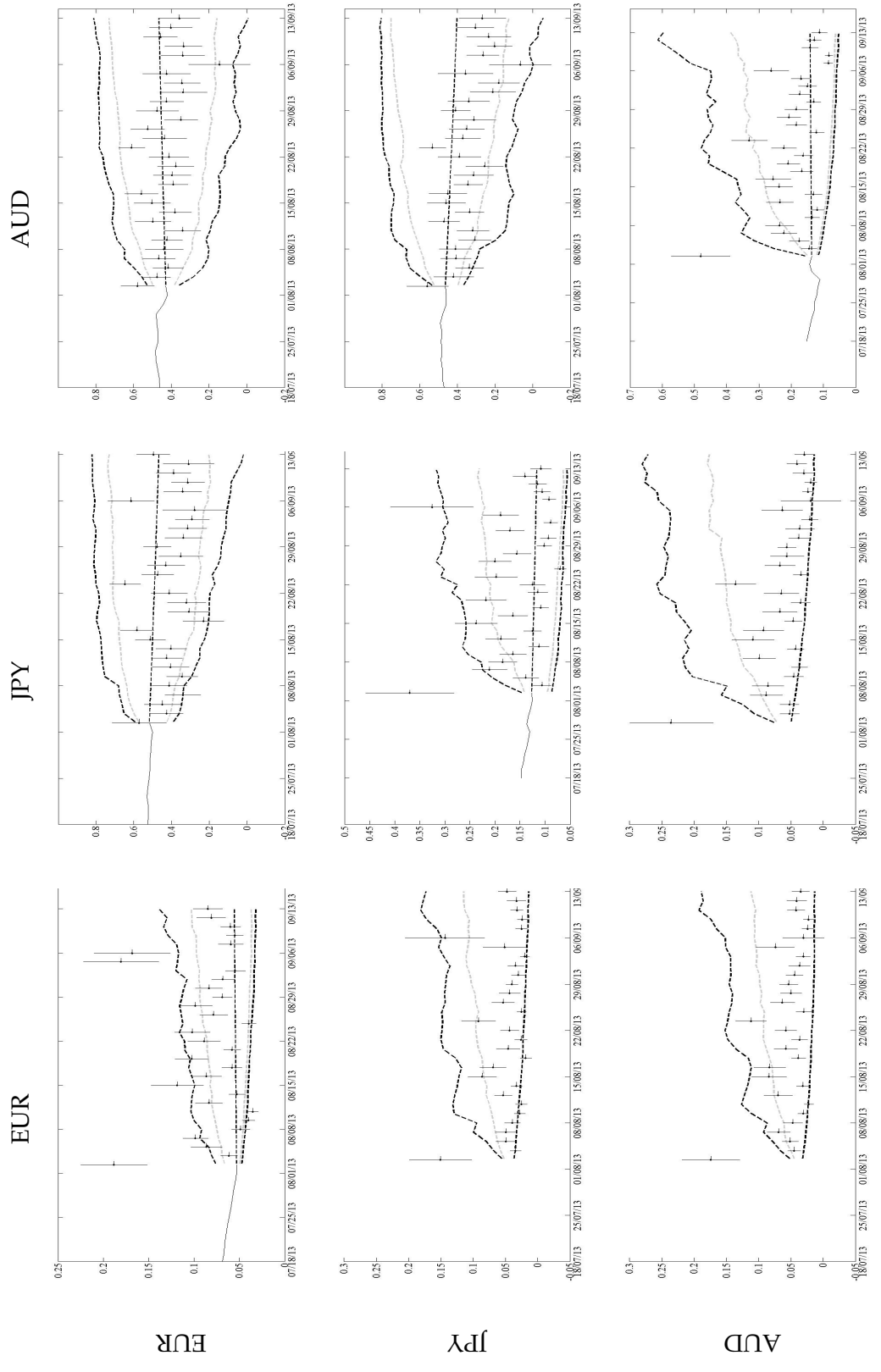


Figure 4.9: Intra-daily realized variances (diagonal), covariances (lower off-diagonal) correlations (upper off diagonal) for EUR, JPY and AUD exchange rates returns together with 95% (grey discontinuos line) and 99% bootstrap h-steps-ahead (black discontinuos line) forecast intervals for the out-of-sample period from 2/8/2013 to 13/9/2013. The vertical lines represent 95% asymptotic intervals for realized correlations.

4.4.2 Bootstrap within sample forecasts of conditional correlations of US market indexes returns

In this application we implement the bootstrap algorithm to construct within 95% bootstrap intervals for conditional correlations of daily log returns of S&P500 (SP_t) and NASDAQ (NQ_t) observed from 30/1/2004 to 1/8/2013. Table 4.3 provides summary statistics for SP_t and NQ_t. They have roughly the same qualitative characteristics as those commented for exchange rate returns system, with the exception that now the cross-correlation between returns and squared returns which are significant. This suggests that a different model for the conditional variances may be needed to deal with asymmetries. Also, the exceedence test of Hong et. al (2007) is not so conclusive about the nonexistence of asymmetries in conditional correlations. Finally, Tse (2001) test strongly rejects the null of constant conditional correlation. For simplicity, we consider a simple cDCC(1,1) with GARCH(1,1) volatilities and recognize that there could be scope to improve it due to asymmetries in volatilities and correlations. The estimated cDCC model with the asymptotic p-values in parentheses is as follows

$$H_{SP,t} = \underset{(0.00)}{3.35e^{-3}} + \underset{(0.00)}{0.09}\, SP_{t-1}^2 + \underset{(0.00)}{0.89}\, H_{SP,t-1},$$

$$H_{NQ,t} = \underset{(0.00)}{4.41e^{-3}} + \underset{(0.00)}{0.07}\, NQ_{t-1}^2 + \underset{(0.00)}{0.91}\, H_{NQ,t-1},$$

$$Q_t = (1 - 0.03 - 0.96)\widehat{S} + \underset{(0.00)}{0.03}\, \upsilon_{t-1}\upsilon_{t-1}' + \underset{(0.00)}{0.96}\, Q_{t-1},$$

$$\widehat{S} = \begin{bmatrix} 1 & . \\ \underset{(0.00)}{0.93} & 1 \end{bmatrix}. \tag{4.35}$$

The LM test statistics implemented to $\widehat{a}_{i,t}^2$, for $i = SP$ and NQ, and to $vech(\widehat{a}_t\widehat{a}_t')$, put forward that the model behaves reasonable well in modeling the dynamics. [7]

[7] The LM statistic of order 20, with the p-values in parentheses, are 6.65 (0.75) and 11.63 (0.75) for the $\widehat{a}_{SP,t}^2$ and $\widehat{a}_{NP,t}^2$, respectively. The multivariate LM of order 20 implemented to $vech(\widehat{a}_t\widehat{a}_t')$ gives a statistic equal to 217.96 (0.03).

Table 4.3: Summary statistics of daily returns for S&P500 and NASDAQ observed from 30/1/2004 to 1/8/2013. Mean, standard deviation (Std), skewnees (Skew) and excess kurtosis (Ek), test of Fiorentini et al. (2004) (Normality), Ljung-Box for returns ($Q(20)$), squared returns ($Q_2(20)$) and cross-correlations between returns and future squared returns ($Q_{12}(20)$). p-values in parenthesis.

	Mean	Std	Skew	Ek	Normality	$Q(20)$	$Q_2(20)$	$Q_{12}(20)$
S&P500	0.01	0.57	-0.08 (0.06)	13.86 (0.00)	19117.93 (0.00)	8.92 (0.53)	3085.93 (0.00)	66.41 (0.00)
NASDAQ	0.01	0.61	-0.21 (0.00)	6.95 (0.00)	4824.50 (0.00)	6.98 (0.72)	3324.01 (0.00)	36.98 (0.00)
Multivariate			7.14 (0.03)	20021.84 (0.00)	20028.40 (0.00)	295.15 (0.45)	7979.00 (0.00)	

The bootstrap procedure is implemented to approximate the within sample conditional correlations using (22) based on 1000 bootstrap replicates. Once more, we gather high frequency mid-quotes prices for the US indexes sampled using a calendar time scheme with 1-minute intervals and compute the intra-day returns and their corresponding 95% asymptotic confidence intervals just as before. For this application, intra-daily data spans from 12/10/2012 to 1/8/2013, which means 389 days. Figure 4.10 plots the one-step-ahead 95% and 99% bootstrap intervals for the conditional correlations together with the intra-day point measures and their 95% confidence intervals for the period from 12/10/2012 to 1/8/2013. First, we observe that the conditional correlations fluctuate in a narrow interval. Of course, this dynamic is in line with those estimated parameters reported above. Second, we note that the bootstrap forecast densities of the conditional correlations are negative skew. The reason is that, positive related shocks increase correlation but only gradually since the unconditional conditional correlation is close to upper bound of 1. On the other hand, negative related shocks have a substantial effect in reducing the conditional correlation. Finally, we can observe that the path depicted by the realized correlations and their 95% asymptotic intervals is roughly capture by the corresponding bootstrap intervals.[8]

[8]Note that the realized correlation is substantially low and its precision large in the day 12/3/2013. The reason is that intra-day returns during this day contains a lot of non-overlapping zeros, which probably is causing the downward in the realized covariance estimates.

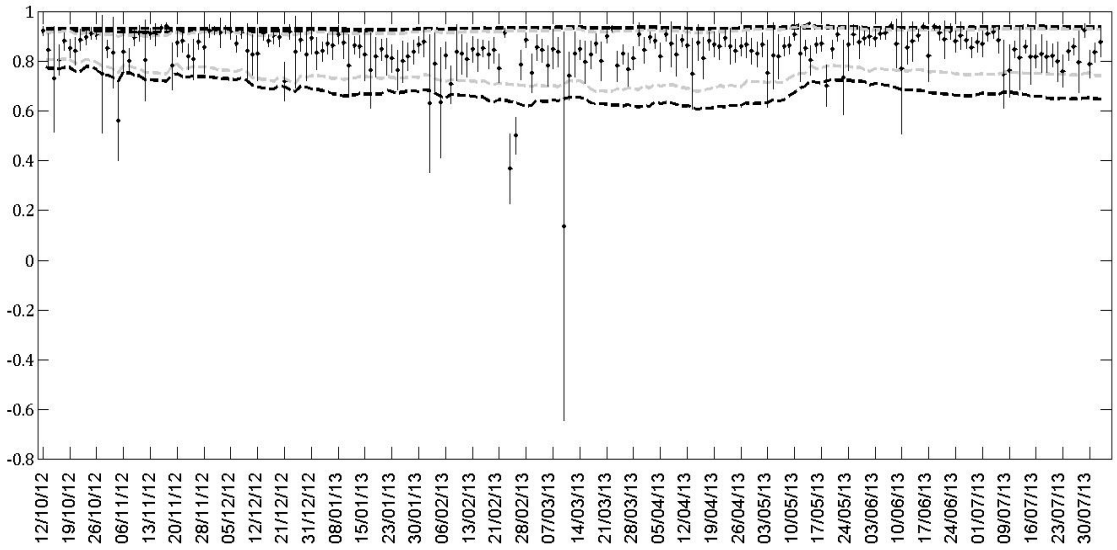


Figure 4.10: One-step-ahead bootstrap intervals and realized measures together with its confidence interval for S&P500-NASDAQ conditional correlation from 12/10/2012 to 1/8/2013. Nominal coverages are 95% (dashed grey) and 99% (dashed black) for the bootstrap intervals and 95% for the realized correlations.

4.5 Conclusion

DCC models deliver within-sample and out-of-sample point forecasts of conditional correlations. However, uncertainty measures related to these forecasts are not available. In this chapter, we propose to approximate the uncertainty of conditional correlations in the cDCC model by implementing the bootstrap procedure proposed by Fresoli et al. (in press) for multivariate models. We analyze its finite sample properties and show that it is adequate in situations similar to those encountered when fitting cDCC models to real systems of financial returns.

The bootstrap procedure is implemented to obtain predictive densities of returns, volatilities and correlations of a three-dimensional system of daily exchange rates returns. The bootstrap procedure seems to work adequately when forecasting out-of-sample in the sense that the bootstrap forecast intervals capture most of the realized correlations during the out-of-sample period.

Moreover, the bootstrap procedure considered in this paper can be easily adapted to deal with alternative MGARCH models; see Bauwens et al. (2006) and Silvennoinen and Teräsvirta (2009)

for alternatives. In the context of VAR models, Fresoli et al. (in press) show that the advantages of bootstrapping are larger when the roots are close to the non-stationary regions. Therefore it is of interest to analize the uncertainty associated with forecasting correlations in the context of non-stationary models as those suggested by Amado and Teräsvirta (in press).

Another important issue left for further research is the feasibility of the considered bootstrap procedure in large systems of returns. In this case, Hafner and Reznikova (2012) show that, in high dimensions, there is a problem of biased parameter estimates when implementing the three-steps target estimator considered in this paper due to the ill-conditioning of the sample correlation matrix used for correlation targeting. They suggest using a shrinkage technique to solve this problem and compare it with alternative estimators. Alternatively, one can also use a bootstrap-after-bootstrap approach, though the computational burden is largely deepened. Whether this computational burden can be reduced using the proposals of Giacomini et al. (2013) is also in our future research agenda.

Chapter 5

Conclusions and extensions

5.1 Conclusions

In this thesis we investigated the performance of bootstrap procedures to construct forecast densities, intervals and regions for multivariate time series data. In particular, we focused on bootstrap procedures developed to forecast VAR and DCC models which are useful in modeling and forecasting economic and financial data. The bootstrap methodology is attractive since it is free of distributional assumptions about the errors. Moreover, bootstrap procedures are well suited to incorporate the sampling variability due to parameter uncertainty and can be designed to tackle also the model uncertainty.

In a VAR setting, we established the asymptotic validity of a bootstrap procedure that it is not based on the popular backward representation. We conducted several Monte Carlo experiments to assess its finite sample properties, and observed that the bootstrap procedure considered exhibits as good properties as bootstrap alternatives based on the backward representation. Therefore, consistent with Occam's razor, it is worth to insist in the good properties of the bootstrap forecasts obtained without using the backward representation.

Bootstrap methods are successfully designed to deal with different sources of uncertainty in the context of forecasting multivariate VAR models. In this thesis, we also carried out a direct

comparison of the traditional Gaussian approach and several variants of the bootstrap procedure that successively incorporates error distribution, parameter uncertainty, bias correction and lag order uncertainty. One may expect the role of parameter uncertainty to be of small order of magnitude. Notwithstanding our Monte Carlo study posited that it plays a prominent role, at least in highly persistent VAR models. Although there are additional gains of adding lag order uncertainty and correcting biased parameters, contrary to what is pointed out in the literature, these seem to be less pronounced.

Regarding DCC forecasts, the standard approach only provides point forecasts of conditional variances, covariances and correlations. In this thesis, we described the implementation of a bootstrap procedure to approximate the conditional distribution of returns for any forecast horizon without assuming any particular error distribution. Also, the bootstrap procedure can be implemented to construct forecast densities of volatilities, covariances and correlations and their corresponding forecast intervals. Whilst the asymptotic validity of our proposal for the DCC model is not formally established, we conducted Monte Carlo simulations that showed a rather good performance of our bootstrap procedure under different sample sizes and error distributions. Similarly, two empirical applications, which interposes realized measures with bootstrap forecast intervals of conditional second moments, suggested that our procedure can approximate rather well their uncertainties and, thus, it comes up as an appealing alternative that enriches the DCC forecast methodology.

5.2 Further research

Now we turn to discuss various extensions of the ideas proposed in this thesis.

5.2.1 Conditional forecast

Analysts may be interested in forecasting a subset of variables in the system given the values of the rest of endogenous variables over the forecast period. This situation is referred as conditional

forecasting to distinguish it from the forecast problem treated in this thesis.[1] Conditional forecast can be useful in the context of policy analysis, when a policy maker has in hand a system containing some variables that may be considered as instruments or policy goals; see Waggoner and Zha (1999), Jarociński (2010) and Luciani (in press). For instance, for a Central Banker it may be crucial to forecast next year GDP growth, unemployment and inflation given a path for the interest rate over the next year. There are basically two types of constraints on the future values of the endogenous variables. Hard conditions refers to future values of the endogenous variables that are fixed amounts while soft conditions set the future values in the form of an interval. For instance, the Central Banker may announce a hard condition of 5% next year interest rate or a soft range of (2.5%, 5%) interest rate.

Bootstrap forecasts offer a promising framework to construct such conditional densities. For example, Figure 5.1 plots the h-steps-ahead bootstrap replicates obtained for a simulated series from a bivariate VAR(1) model with autoregressive matrix $vec\,(\Phi_1) = (-0.5, 0, 0.5, 0.5)$, χ^2_4 errors with zero mean and covariance matrix given by $vech\,(\Sigma_\varepsilon) = (1, 0.8, 1)$ and sample size $T = 75$. It is clear that a soft condition on $y_{1,T+h}$ can be represented as a region for $[y_{1,T+h}, y_{2,T+h}]$ such that $y_{1,T+h}$ falls between the bound given by (a, b). On the other hand, hard conditioning corresponds either to $y_{1,T+h} = a$ or $y_{1,T+h} = b$.

Figure 5.1 suggests that the case of soft conditional forecasts can be straightforwardly handle in the context of bootstrap procedure proposed in this thesis. After applying it, B bootstrap realizations of the h-steps-ahead forecast are generated which approximate the forecast distribution of system. Then, in order to impose soft conditions what can be done is to consider those that meet the constraint and get the conditional replicate set which is defined as follow

$$CR = \{\widehat{y}^*_{T+h|T} | \widehat{y}^*_{1,T+h|T} \in [a, b]\}$$

where a and b are the lower and upper bound for $y_{1,T+h}$. The main drawback of this approach

[1] Note that is just terminology since we recognize that both forecasts are conditional and what is actually changing is the information set.

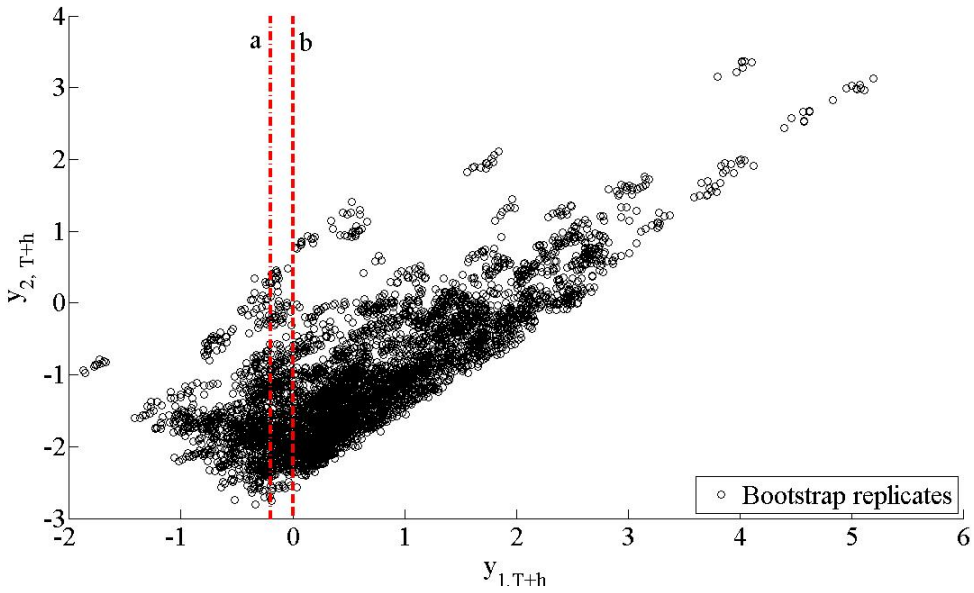


Figure 5.1: Bootstrap realizations of h-steps-ahead forecasts for a bivariate series generated by a stationary VAR(1) process with $T = 75$ and $\chi^2(4)$ errors.

is that the number of elements in CR is not controlled by the analyst. It might be the case that only few unconditional bootstrap replicates fulfill the future restriction, then only few bootstrap realizations make up the CR, and thus the ability of the procedure to approximate the conditional distribution might be drastically worsen. Therefore, to assure that a CR contains as many replication as desired, it would be necessary to repeat steps 1 to 3 in section 2.3.1 until we reach a that number, expressed typically in the form of a lower bound $\bar{B}$. As a result, the bootstrap procedure to soft conditioning forecasts can be describe as a sequence of 5 steps of which the first three are exactly the same as those of section 2.3.1.

Step 1. Repeat steps 1 to 3 in section 2.3.1.

Step 2. After observing the bootstrap replicate of h-steps-ahead density, keep it if it satisfies the restriction $a < \widehat{y}^*_{1,T+h|T} < b$.

Step 3. Repeat 1 and 2 until the number of realization in CR reaches $\bar{B}$.

To illustrate this algorithm we simulate a bivariate series obtained from following VAR(1) model

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} 0.9 & 0 \\ 0.2 & 0.5 \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix} \tag{5.1}$$

where $\varepsilon_t = (\varepsilon_{1,t}, \varepsilon_{2,t})'$ is an independent χ^2_4 white noise vector with covariance matrix given by

$$\Sigma_\varepsilon = \begin{bmatrix} 1 & \rho_{12} \\ \rho_{12} & 1 \end{bmatrix}. \tag{5.2}$$

The contemporaneous correlation can take two different values, $\rho_{12} = 0.1$ or $\rho_{12} = 0.8$.[2] We simulate two series of sample size 75 with χ^2_4, one for each value of ρ_{12}, and then used for estimation $T = 70$ while keeping the last four observations as out-of-sample realizations. Afterwards, this basic bootstrap procedure is implemented to obtain $\bar{B} = 1000$ h-steps-ahead conditional forecasts, for $h = 1, ..., 4$. In this particular example, the first variable is constrained to fall between the $\pm 0.10 std(y_{1,t})$ around its actual out-of-sample value.[3]

Figure 5.2 plots the unconditional and conditional fan charts for $y_{2,t}$ over the forecast period. Upper panels of Figure 5.2 plot the results when ρ_{12} is high. We observe that incorporating the information about $y_{1,t}$ tightens the band for $y_{2,t}$. However, the extent to which tightness becomes apparent depends on the contemporaneous correlation. This is a result of an intuitive idea which states that the larger the contemporaneous correlation between the free and fixed variables, the more advantage any forecast approach should obtain from future information. Lower panels of Figure 5.2 confirms this intuition, showing that the knowledge of the future path of $y_{1,t}$ may not improve our forecast if the correlation between $y_{1,t}$ and $y_{2,t}$ is low.

The performance of bootstrapping to obtain conditional forecasts needs to be studied. Of course, one limitation of this approximation is that computational time might rise abruptly if the soft condition are located in regions where the system is unlikely to evolve. This issue demands

[2] This VAR is stationary, though persistent. Its roots are 1.11 and 2.

[3] Given that the error is ε_t is multivariate χ^2_4 distributed, the conditional distribution of $\varepsilon_{1,t}$ given $\varepsilon_{2,t}$ is unknown.

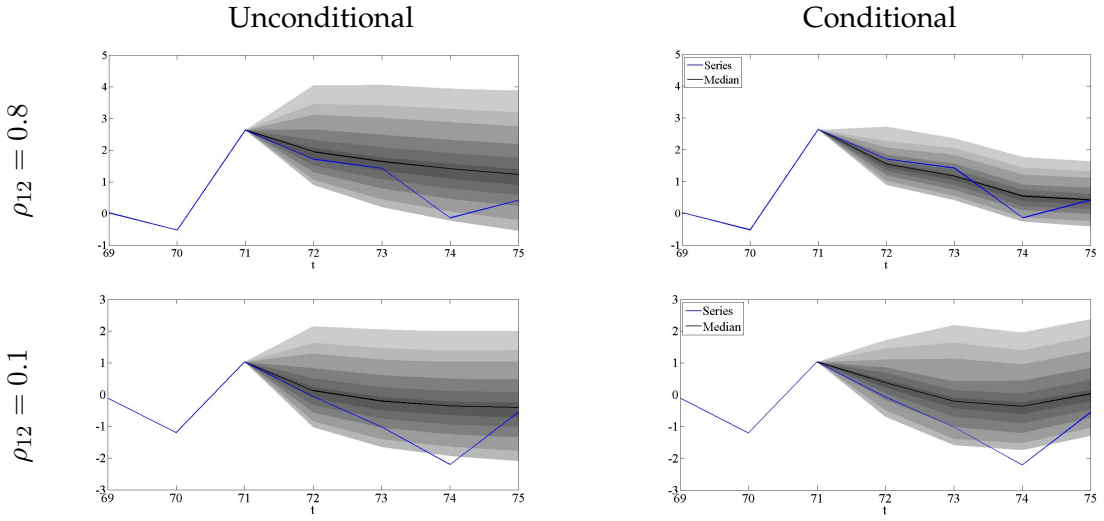


Figure 5.2: Fan chart for unconditional (first column) and conditional forecasts (second column) of the second variable for bivariate series generated by (5.1) with $T = 75$, $\chi^2(4)$ errors and unconditional correlation of $\rho_{12} = 0.8$ (first row) or 0.1 (second row).

further research effort. Furthermore, the problem of hard conditions is trickier and no simple solution can be envisioned straightforwardly. The reason is that to tighten an interval does nothing but deepen the problem associated with soft conditions. Finding a proper way of defining hard conditions in our bootstrap setting represents an interesting challenge that is part of our ongoing research.

5.2.2 Forecasting cointegrated system

Economic theory imposes often long-run equilibrium relationships between variables in levels that are characterized for being nonstationary. For example, the Purchasing Power Parity states that the exchange rate between two counties' currencies is in equilibrium when their domestic price level of a fixed basket of goods and services are the same. Also, financial arbitrage theory tells us that the prices of certain financial assets, which are integrated of order one, must be linked in the long-run and, consequently, any deviation from their equilibrium must vanish through short-run adjustments. Integrated series linked in the long-run through an equilibrium relationship are called cointegrated; see, for instance, Juselius (2006) and Carlucci and Montaruli

(2014) for surveys on cointegration methodology and economic analysis. The connection between long-run equilibrium and the short-run dynamics is usually described under the form of Vector Error Correction (VECM) models, which is given by

$$\Delta y_t = \alpha\beta^{o\prime} y^o_{t-1} + \Gamma_1 \Delta y_{t-1} + ... + \Gamma_{p-1}\Delta y_{t-p+1} + \delta d_t + \varepsilon_t \tag{5.3}$$

where y_t is a $N \times 1$ vector of endogenous variables, $\beta^{o\prime} = [\beta', \delta']$ is $N^* \times r$ with β being a $N \times r$ matrix containing the cointegration relationships, ${y^o_{t-1}}' = [{y_{t-1}}', {d^{co}_{t-1}}']$ is $N^o \times 1$ with N $K^o = N + dim(d^{co}_t)$, d^{co}_t contains all the deterministic components included in the cointegration relationship, d_t contains the remaining deterministic terms, Γ_j and δ are matrices of suitable dimension describing the short-run dynamics of the system and, finally, ε_t is a $N \times 1$ zero mean white noise process with covariance matrix Σ_ε. It is assumed that $0 < rank(\beta) = r < N$; α is a $N \times r$ matrix of loading coefficients with $rank(\alpha) = r$. Also, the column dimension of δ is r. It is worth mentioning that the key assumption regarding the deterministic variable is that they appear just once, either in d^{co}_t or d_t but not in both.

When forecasting cointegrated series, there is an old debate about the effect of imposing the long-run equilibrium on point forecast accuracy over the long-run; see Clements and Hendry (1993) and Christoffersen and Diebold (1998) for different points of view. This debate exemplifies the prominent role of point forecasts when dealing with cointegrated systems; see, for instance, Dreger and Wolters (2014) for money demand, Apergis (2014) for Australian Dollar and Kouwenberg and Zwinkels (2014) for US housing market for recent efforts to obtain point forecast of cointegrated variables. However, little attention has received the construction of future conditional densities in this framework.

An alternative that we would like to explore is the implementation of the bootstrap procedure proposed in this thesis to forecast VECM models. Next we briefly examine this possibility by considering a simple example based on similated data.

We implement the basic bootstrap procedure to forecast cointegrated variables in VECM. For

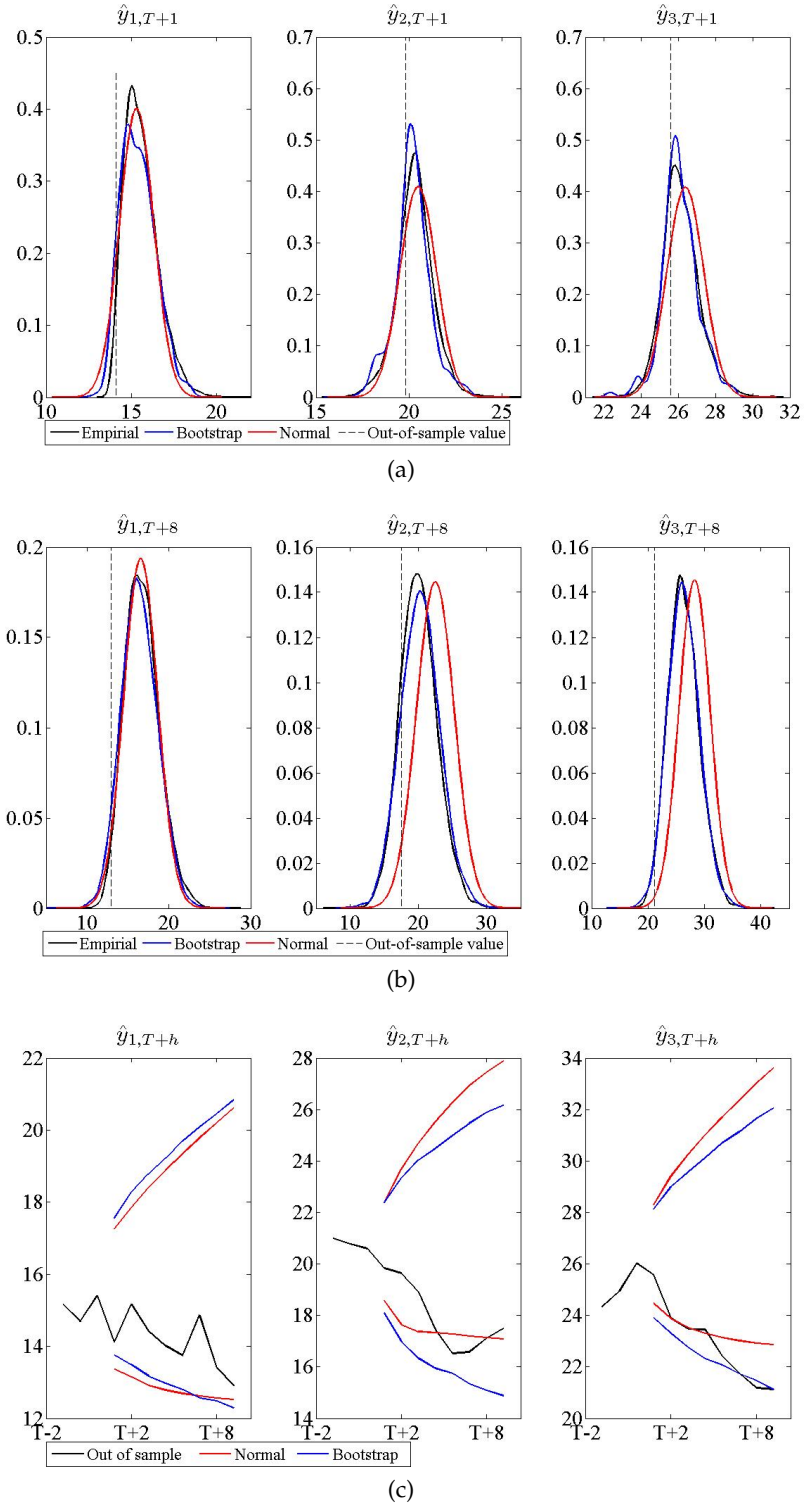


Figure 5.3: Kernel estimates of empirical (black), Gaussian (red) and bootstrap (blue) forecast densities for a simulated trivariate system of a VECM(1) model with $T = 200$ and χ^2_4 errors: (a) $h = 1$, (b) $h = 8$ (c) 95% h-steps-ahead bootstrap (blue) and Gaussian (red) forecast intervals, for $h = 1, ..., 8$, together with the out-of-sample realization (black).

doing this, we simulate data from a trivariate VECM(1) with 208 observations, χ^2_4 errors, $r = 2$, a constant term in (5.3) that becomes a linear trend in the variables in levels and a constant in the cointegration relationship that shifts $y_{2,t}$ and $y_{3,t}$ with respect to $y_{1,t}$.[4] Throughout this example the number of cointegrated relationships is assumed to be known. We keep $T = 200$ observations for estimation and $H = 8$ as out-of-sample realizations. Upper and middle panels of Figure 5.3 plot the empirical one-step-ahead and eight-steps-ahead forecast densities, respectively, which are obtained by replicating 5000 future values given the realization of the process. For the simulated series, we then estimate the parameters by using Feasible Generalized Least Squares; see Lütkepohl (2004). Afterwards, we proceed to forecast the system variables. The standard methodology does not differ from that previously explained in this thesis, it replaces the unknown parameters by their estimates and assumes Gaussian error. The resulting Gaussian forecast densities of $y_{i,T+h}$, for i =1,2,3 and $h = 1$ and 8, are plotted in the upper and middle panels of Figure 5.3. Note that they seem to give a poor description of the corresponding empirical forecast densities. This is evident for $y_{2,T+8}$ and $y_{3,T+8}$, where it can be observed that the Gaussian conditional densities are located towards the right of their corresponding empirical densities; see, for example, Clements and Hendry (2006) who point out that distorted parameters can severely affect the quality of forecast in cointegrated system. The 95% individual forecast intervals are shown in the lower panels of Figure 5.3. Interestingly, the ability of the Gaussian bands to capture the future realization of the process seems to be limited, especially for $y_{2,t}$ and $y_{3,t}$. Finally, we implement a bootstrap algorithm in the spirit of that considered in this thesis to obtain the forecast density of y_{T+h}. The number of bootstrap replicates is set to $B = 1000$. The upper and middle panels of Figure 5.3 add the corresponding one-step-ahead and eight-steps-ahead bootstrap densities. In general, we observe that these densities are able to capture the skewness of the empirical forecast densities and thus they are closer to the latter than the corresponding Gaussian forecast densities. Furthermore, bootstrap densities seem the be better located than the Gaussian densities. A reason for this to happen is that the proposed bootstrap procedure do incorporate parameter uncertainty

[4]See Appendix A, model (d), for details on this specification.

and, thus, do not depend exclusively on one estimate of the parameters that may be distorted, especially when considering the trend parameters of cointegrated systems. Finally, lower panels of Figure 5.3 also plots the 95% bootstrap forecast intervals that do a better job than the Gaussian ones, once more, especially for $y_{2,T+h}$ and $y_{3,T+h}$.

Future research can be directed at establishing the asymptotic validity of the bootstrap procedure to forecast VECM and investigating its finite sample properties. Of course, we also note that the number of cointegrated relations is another source of sampling variability in many applications, and thus we need to incorporate it into the algorithm. Finally, we also recognize that it may be of interest not only to forecast the variables in level but also the long-run equilibrium. These topics are part of the future research agenda.

5.2.3 Forecasting Multivariate GARCH

When dealing with conditionally heteroscedastic systems, the bootstrap procedure considered in this thesis offers several topics that can be investigated. First, although the finite sample properties show that the proposed procedure is rather adequate in terms of coverages of the bootstrap forecast intervals for returns, volatilities and conditional correlations, its asymptotic validity has not been explicitly derived yet. Second, it may be interesting to adapt the algorithm to deal with alternative MGARCH models. For example, to enrich the empirical application, the fitted model might permit asymmetries in conditional variances and correlations and feedback between the former and the latter. Finally, another important issue left for further research is the feasibility of the considered bootstrap procedure in large systems of returns. Of course, as the system becomes larger, the bootstrap is less appealing due to the cost associated with the estimation of the model parameters. Whether the computational burden can be reduced using the proposals of Giacomini et al. (2013) is also in our future research agenda.

5.2.4 Forecast regions and paths

To complete the implementation of bootstrap procedure to multivariate forecast we need to develop statistical methods to delimit forecast regions. For univariate series, this is not a challenge because the construction of intervals is rather straightforward task. However, constructing forecast regions with accurate coverage is not that easy; see, for instance, Yang and Kolassa (2004) and Wolf and Wunderli (2012). One simple alternative is to use Bonferroni regions which give cubical approximation to the ellipsoids, providing a lower bound probability content. The main drawback of Bonferroni cubes is that, as long as they are constructed by using marginal intervals for the variables in the system, the correlations among them are ignored. For this reason, in this thesis we have explored how to correct cubes that use the information provided by correlations. Yet the theoretical justification as well as the small sample properties of the corrected cubes are missing.

References

Aielli, G. P. (2013), "Dynamic conditional correlation: on properties and estimation," *Journal of Business & Economic Statistics*, 31, 282–299.

Aielli, G. P. and Caporin, M. (in press), "Variance clustering improved dynamic conditional correlation MGARCH estimators," *Computational Statistics & Data Analysis*.

Amado, C. and Teräsvirta, T. (in press), "Conditional correlation models of autoregressive conditional heteroskedasticity with nonstationary GARCH equations," *Journal of Business & Economic Statistics*.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Ebens, H. (2001), "The distribution of realized stock return volatility," *Journal of Financial Economics*, 61, 43–76.

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

Apergis, N. (2014), "Can gold prices forecast the Australian dollar movements?" *International Review of Economics & Finance*, 29, 75–82.

Ariño, M. A. and Franses, P. H. (2000), "Forecasting the levels of vector autoregressive log-transformed time series," *International Journal of Forecasting*, 16, 111–116.

Arroyo, J. and Maté, C. (2009), "Forecasting histogram time series with k-nearest neighbours methods," *International Journal of Forecasting*, 25, 192–207.

Athanasopoulos, G. and Vahid, F. (2008), "VARMA versus VAR for macroeconomic forecasting," *Journal of Business & Economic Statistics*, 26, 237–252.

Audrino, F. (in press), "Forecasting correlations during the late-2000s financial crisis: short-run component, long-run component, and structural breaks," *Computational Statistics & Data Analysis*.

Audrino, F. and Trojani, F. (2011), "A general multivariate threshold GARCH model with dynamic conditional correlations," *Journal of Business & Economic Statistics*, 29, 138–149.

Baillie, R. T. (1979), "Asymptotic prediction mean squared error for vector autoregressive models," *Biometrika*, 66, 675–678.

Bårdsen, G. and Lütkepohl, H. (2011), "Forecasting levels of log variables in vector autoregressions," *International Journal of Forecasting*, 27, 1108–1115.

Barndorff-Nielsen, O. E. and Shephard, N. (2004), "Econometric analysis of realized covariation: high frequency based covariance, regression, and correlation in financial economics," *Econometrica*, 72, 885–925.

Batchelor, R., Alizadeh, A., and Visvikis, I. (2007), "Forecasting spot and forward prices in the international freight market," *International Journal of Forecasting*, 23, 101–114.

Baumeister, C. and Kilian, L. (2012), "Real-time forecasts of the real price of oil," *Journal of Business & Economic Statistics*, 30, 326–336.

Bauwens, L., Hafner, C. M., and Pierret, D. (2013), "Multivariate volatility modeling of electricity futures," *Journal of Applied Econometrics*, 28, 743–761.

Bauwens, L. and Laurent, S. (2005), "A new class of multivariate skew densities, with application to generalized autoregressive conditional heteroscedasticity models," *Journal of Business & Economic Statistics*, 23, 346–354.

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

Berkowitz, J. and Kilian, L. (2000), "Recent developments in bootstrapping time series," *Econometric Reviews*, 19, 1–48.

Bhansali, R. (1981), "Effects of not knowing the order of an autoregressive process on the mean squared error of prediction," *Journal of the American Statistical Association*, 76, 588–597.

Bickel, P. J. and Freedman, D. A. (1981), "Some asymptotic theory for the bootstrap," *Annals of Statistics*, 9, 1196–1217.

Billio, M., Caporin, M., and Gobbo, M. (2006), "Flexible Dynamic Conditional Correlation Multivariate GARCH models for asset allocation," *Applied Financial Economics Letters*, 2, 123–130.

Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 307–327.

— (1990), "Modelling the coherence in short-run nominal exchange rates: A Multivariate Generalized ARCH model," *The Review of Economics and Statistics*, 498–505.

Bollerslev, T., Engle, R. F., and Wooldridge, J. M. (1988), "A capital asset pricing model with time-varying covariances," *The Journal of Political Economy*, 116–131.

Bose, A. (1988), "Edgeworth correction by bootstrap in autoregressions," *Annals of statistics*, 16, 1709–1722.

Breidt, F. and Davis, R. (1992), "Time-reversibility, identifiability and independence of innovations for stationary time series," *Journal of Time Series Analysis*, 13, 377–390.

Breusch, T. S. and Pagan, A. R. (1980), "The Lagrange multiplier test and its applications to model specification in econometrics," *The Review of Economic Studies*, 47, 239–253.

Bühlmann, P. (2002), "Bootstraps for time series," *Statistical Science*, 17, 52–72.

Burnham, K. P. and Anderson, D. R. (2004), "Multimodel inference understanding AIC and BIC in model selection," *Sociological Methods & Research*, 33, 261–304.

Cao, R. (1999), "An overview of bootstrap methods for estimating and predicting in time series," *Test*, 8, 95–116.

Caporin, M. and McAleer, M. (2010), "Ranking multivariate GARCH models by problem dimension," *Econometric Institute Research working paper series EI 2010-34*.

— (2012), "Do we really need both BEKK and DCC? A tale of two multivariate GARCH models," *Journal of Economic Surveys*, 26, 736–751.

— (2013), "Ten Things You Should Know About the Dynamic Conditional Correlation Representation," *Econometrics*, 1, 115–126.

Cappiello, L., Engle, R. F., and Sheppard, K. (2006), "Asymmetric dynamics in the correlations of global equity and bond returns," *Journal of Financial Econometrics*, 4, 537–572.

Carlucci, F. and Montaruli, F. (2014), "Co-Integrating VAR Models and Economic Policy," *Journal of Economic Surveys*, 28, 68–81.

Chan, K.-S., Ho, L.-H., and Tong, H. (2006), "A note on time-reversibility of multivariate linear processes," *Biometrika*, 93, 221–227.

Chatfield, C. (1996), "Model uncertainty and forecast accuracy," *Journal of Forecasting*, 15, 495–508.

Christoffersen, P. F. and Diebold, F. X. (1998), "Cointegration and long-horizon forecasting," *Journal of Business & Economic Statistics*, 16, 450–456.

Clark, T. E. (2011), "Real-time density forecasts from Bayesian vector autoregressions with stochastic volatility," *Journal of Business & Economic Statistics*, 29.

Clark, T. E. and Ravazzolo, F. (in press), "Macroeconomic forecasting performance under alternative specifications of time-varying volatility," *Journal of Applied Econometrics*.

Clements, M. and Hendry, D. F. (2006), "Forecasting with structural breaks," Elsevier, vol. 1.

Clements, M. P. and Hendry, D. F. (1993), "On the limitations of comparing mean square forecast errors," *Journal of Forecasting*, 12, 617–637.

Clements, M. P. and Kim, J. H. (2007), "Bootstrap prediction intervals for autoregressive time series," *Computational statistics & data analysis*, 51, 3580–3594.

Corradi, V. and Swanson, N. R. (2006), "Predictive density evaluation," Elsevier, vol. 1, pp. 197–284.

Corsi, F. and Audrino, F. (2012), "Realized covariance tick-by-tick in presence of rounded time stamps and general microstructure effects," *Journal of Financial Econometrics*, 10, 591–616.

D'Agostino, A., Gambetti, L., and Giannone, D. (2011), "Macroeconomic forecasting and structural change," *Journal of Applied Econometrics*, 28, 82–101.

Diebold, F. X. (1988), *Empirical Modeling of Exchange Rate Dynamics*, Springer New York.

Diebold, F. X., Hahn, J., and Tay, A. S. (1999), "Multivariate density forecast evaluation and calibration in financial risk management: high-frequency returns on foreign exchange," *Review of Economics and Statistics*, 81, 661–673.

Diebold, F. X. and Yilmaz, K. (2009), "Measuring Financial Asset Return and Volatility Spillovers, with Application to Global Equity Markets*," *The Economic Journal*, 119, 158–171.

Ding, Z. and Engle, R. (2001), "Large scale conditional covariance matrix modeling, estimation and testing," *NYU working paper No. Fin-01-029*.

Doan, T., Litterman, R., and Sims, C. (1984), "Forecasting and conditional projection using realistic prior distributions," *Econometric reviews*, 3, 1–100.

Dovonon, P., Goncalves, S., and Meddahi, N. (2013), "Bootstrapping realized multivariate volatility measures," *Journal of Econometrics*, 172, 49–65.

Dreger, C. and Wolters, J. (2014), "Money demand and the role of monetary indicators in forecasting euro area inflation," *International Journal of Forecasting*, 30, 303–312.

Drost, F. C. and Nijman, T. E. (1993), "Temporal aggregation of GARCH processes," *Econometrica*, 61, 909–927.

Duchesne, P. and Lalancette, S. (2003), "On testing for multivariate ARCH effects in vector time series models," *Canadian Journal of Statistics*, 31, 275–292.

Dufour, J.-M. and Jouini, T. (2006), "Finite-sample simulation-based inference in VAR models with application to Granger causality testing," *Journal of Econometrics*, 135, 229–254.

Efron, B. (1979), "Bootstrap methods: another look at the jackknife," *The annals of Statistics*, 7, 1–26.

Efron, B. and Tibshirani, R. J. (1993), *An Introduction to the Bootstrap*, Chapman and Hall.

Eklund, B. (2007), "Forecasting the Icelandic business cycle using vector autoregressive models," *Central Bank of Iceland, Economics Departmen Working Paper No. 36*.

Engle, R. (2002), "Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models," *Journal of Business & Economic Statistics*, 20, 339–350.

Engle, R. F. (1982), "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation," *Econometrica*, 50, 987–1007.

— (2009), *Anticipating Correlations: A New Paradigm for Risk Management*, Princeton University Press.

Engle, R. F. and Kelly, B. (2012), "Dynamic equicorrelation," *Journal of Business & Economic Statistics*, 30, 212–228.

Engle, R. F. and Kroner, K. F. (1995), "Multivariate simultaneous generalized ARCH," *Econometric theory*, 11, 122–150.

Engle, R. F. and Sheppard, K. (2001), "Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH," *NBER Working Paper No. 8554*.

Engsted, T. and Pedersen, T. Q. (2011), "Bias-correction in vector autoregressive models: A simulation study," *CREATES Research Paper No. 18*.

Fair, R. C. and Shiller, R. J. (1990), "Comparing information in forecasts from econometric models," *The American Economic Review*, 80, 375–389.

Fiorentini, G., Sentana, E., and Calzolari, G. (2004), "On the validity of the Jarque-Bera Normality test in conditionally heteroskedastic dynamic regression models," *Economics Letters*, 83, 307–312.

Freedman, D. A. (1981), "Bootstrapping regression models," *The Annals of Statistics*, 9, 1218–1228.

Fresoli, D., Ruiz, E., and Pascual, L. (in press), "Bootstrap multi-step forecast of non-Gaussian VAR models," *International Journal of Forecasting*.

Gaglianone, W. P., Lima, L. R., Linton, O., and Smith, D. R. (2011), "Evaluating value-at-risk models via quantile regression," *Journal of Business & Economic Statistics*, 29, 150–160.

Giacomini, R., Politis, D. N., and White, H. (2013), "A warp-speed method for conducting Monte Carlo experiments involving bootstrap estimators," *Econometric Theory*, 29, 567–589.

Grigoletto, M. (1998), "Bootstrap prediction intervals for autoregressions: some alternatives," *International Journal of Forecasting*, 14, 447–456.

— (2005), "Bootstrap prediction regions for multivariate autoregressive processes," *Statistical Methods and Applications*, 14, 179–207.

— (2012), "Forecasting the Distribution of Aggregated Time Series: A Bootstrap Approach," *University of Padua Working Paper No. 10*.

Grigoletto, M. and Lisi, F. (2011), "Practical implications of higher moments in risk management," *Statistical Methods & Applications*, 20, 487–506.

Guerbyenne, H. and Hamdi, F. (in press), "Bootstrapping Periodic State-Space Models," *Communications in Statistics-Simulation and Computation*.

Gupta, R., Kabundi, A., and Miller, S. M. (2011), "Using large data sets to forecast house prices: a case study of twenty US states," *Journal of housing research*, 20, 161–190.

Hafner, C. M. and Franses, P. H. (2009), "A Generalized Dynamic Conditional Correlation model: simulation and application to many assets," *Econometric Reviews*, 28, 612–631.

Hafner, C. M. and Preminger, A. (2009), "On asymptotic theory for multivariate GARCH models," *Journal of Multivariate Analysis*, 100, 2044–2054.

Hafner, C. M. and Reznikova, O. (2012), "On the estimation of dynamic conditional correlation models," *Computational Statistics & Data Analysis*, 56, 3533–3545.

Hall, P. (1992), *The bootstrap and Edgeworth expansion*, Springer.

Halunga, A., Orme, C. D., and Yamagata, T. (2011), "A heteroskedasticity robust Breusch-Pagan test for contemporaneous correlation in dynamic panel data models," *University of Manchester Economics Discussion Paper Series No 1118*.

Hartz, C., Mittnik, S., and Paolella, M. (2006), "Accurate value-at-risk forecasting based on the normal-GARCH model," *Computational Statistics & Data Analysis*, 51, 2295–2312.

Harvey, D. I. and Newbold, P. (2003), "The non-normality of some macroeconomic forecast errors," *International Journal of Forecasting*, 19, 635–653.

Holmes, S., Morris, C., Tibshirani, R., and Efron, B. (2003), "Bradley Efron: A conversation with good friends," *Statistical Science*, 18, 268–281.

Hong, Y., Tu, J., and Zhou, G. (2007), "Asymmetries in stock returns: statistical tests and economic evaluation," *Review of Financial Studies*, 20, 1547–1581.

Hosking, J. R. (1980), "The multivariate portmanteau statistic," *Journal of the American Statistical Association*, 75, 602–608.

Huang, T. H. and Wang, Y. H. (2012), "The volatility and density prediction performance of alternative GARCH models," *Journal of Forecasting*, 31, 157–171.

Hyndman, R. J. (1996), "Computing and graphing highest density regions," *The American Statistician*, 50, 120–126.

Jarociński, M. (2010), "Conditional forecasts and uncertainty about forecast revisions in vector autoregressions," *Economics Letters*, 108, 257–259.

Jeon, J. and Taylor, J. W. (2012), "Using conditional kernel density estimation for wind power density forecasting," *Journal of the American Statistical Association*, 107, 66–79.

Jore, A. S., Mitchell, J., and Vahey, S. P. (2010), "Combining forecast densities from VARs with uncertain instabilities," *Journal of Applied Econometrics*, 25, 621–634.

Juselius, K. (2006), *The cointegrated VAR model: methodology and applications*, Oxford University Press.

Kasch, M. and Caporin, M. (2012), "Volatility Threshold Dynamic Conditional Correlations: An International Analysis," *Available at SSRN 968233*.

Kavussanos, M. G. and Visvikis, I. D. (2004), "Market interactions in returns and volatilities between spot and forward shipping freight markets," *Journal of Banking & Finance*, 28, 2015–2049.

Kilian, L. (1998a), "Accounting for lag order uncertainty in autoregressions: the endogenous lag order bootstrap algorithm," *Journal of Time Series Analysis*, 19, 531–548.

— (1998b), "Confidence intervals for impulse responses under departures from normality," *Econometric Reviews*, 17, 1–29.

— (1998c), "Small-sample confidence intervals for impulse response functions," *Review of Economics and Statistics*, 80, 218–230.

— (1999), "Finite-sample properties of percentile and percentile-t bootstrap confidence intervals for impulse responses," *Review of Economics and Statistics*, 81, 652–660.

— (2001), "Impulse response analysis in vector autoregressions with unknown lag order," *Journal of Forecasting*, 20, 161–179.

Kilian, L. and Demiroglu, U. (2000), "Residual-based tests for normality in autoregressions: Asymptotic theory and simulation evidence," *Journal of Business & Economic Statistics*, 18, 40–50.

Kilian, L. and Vigfusson, R. J. (2013), "Do oil prices help forecast US real GDP? the role of nonlinearities and asymmetries," *Journal of Business & Economic Statistics*, 31, 78–93.

Kim, J. H. (1997), "Relationship between the forward and backward representations of the stationary VAR model, Problem 97.5.2.l," *Econometric Theory*, 13, 889–889.

— (1998), "The Relationship between Forward and Backward Representations of the Stationary VAR Models, Solution 97.5.2." *Econometric Theory*, 691–693.

— (1999), "Asymptotic and bootstrap prediction regions for vector autoregression," *International Journal of Forecasting*, 15, 393–403.

— (2001), "Bootstrap-after-bootstrap prediction intervals for autoregressive models," *Journal of Business & Economic Statistics*, 19, 117–128.

— (2004), "Bias-corrected bootstrap prediction regions for vector autoregression," *Journal of Forecasting*, 23, 141–154.

Kim, J. H., Song, H., and Wong, K. K. (2010), "Bias-corrected bootstrap prediction intervals for autoregressive model: new alternatives with applications to tourism forecasting," *Journal of Forecasting*, 29, 655–672.

Klößner, S. and 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, 172–179.

Komunjer, I. and Owyang, M. T. (2012), "Multivariate forecast evaluation and rationality testing," *Review of Economics and Statistics*, 94, 1066–1080.

Koop, G. (2013), "Forecasting with medium and large Bayesian VARs," *Journal of Applied Econometrics*, 24, 177–203.

Kouwenberg, R. and Zwinkels, R. (2014), "Forecasting the US housing market," *International Journal of Forecasting*, 30, 415–425.

Lam, J.-P. and Veall, M. R. (2002), "Bootstrap prediction intervals for single period regression forecasts," *International Journal of Forecasting*, 18, 125–130.

Laurent, S., Rombouts, J. V., and Violante, F. (2012), "On the forecasting accuracy of multivariate GARCH models," *Journal of Applied Econometrics*, 27, 934–955.

Lewis, R. and Reinsel, G. C. (1985), "Prediction of multivariate time series by autoregressive model fitting," *Journal of Multivariate Analysis*, 16, 393–411.

Li, J. (2011), "Bootstrap prediction intervals for SETAR models," *International Journal of Forecasting*, 27, 320–332.

Liu, S., Maharaj, E. A., and Inder, B. (2014), "Polarization of forecast densities: A new approach to time series classification," *Computational Statistics & Data Analysis*, 70, 345–361.

Luciani, M. (in press), "Monetary policy and the housing market: A structural factor analysis," *Journal of Applied Econometrics*.

Lütkepohl, H. (1991), *Introduction to Multiple Time Series Analysis*, Springer-Verlag.

— (2004), "Vector autoregressive and vector error correction models," In ed. by H Lütkepohl, Helmut and H. Krazig: Applied Time Series Econometrics.

— (2006), "Forecasting with VARMA models," Elsevier, vol. 1, pp. 287–325.

Marcellino, M., Stock, J. H., and Watson, M. W. (2003), "Macroeconomic forecasting in the Euro area: Country specific versus area-wide information," *European Economic Review*, 47, 1–18.

— (2006), "A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series," *Journal of Econometrics*, 135, 499–526.

Masarotto, G. (1990), "Bootstrap prediction intervals for autoregressions," *International Journal of Forecasting*, 6, 229–239.

McAleer, M., Chan, F., Hoti, S., and Lieberman, O. (2008), "Generalized autoregressive conditional correlation," *Econometric Theory*, 24, 1554.

McAleer, M. and Medeiros, M. C. (2008), "Realized volatility: A review," *Econometric Reviews*, 27, 10–45.

McCullough, B. (1994), "Bootstrapping forecast intervals: an application to AR (p) models," *Journal of Forecasting*, 13, 51–66.

Müller, U. K. and Stock, J. H. (2011), "Forecasts in a Slightly Misspecified Finite Order VAR," Tech. rep., National Bureau of Economic Research.

Nakatani, T. and Teräsvirta, T. (2009), "Testing for volatility interactions in the constant conditional correlation GARCH model," *The Econometrics Journal*, 12, 147–163.

Paparoditis, E. (1996), "Bootstrapping autoregressive and moving average parameter estimates of infinite order vector autoregressive processes," *Journal of Multivariate Analysis*, 57, 277–296.

Pascual, L., Romo, J., and Ruiz, E. (2004), "Bootstrap predictive inference for ARIMA processes," *Journal of Time Series Analysis*, 25, 449–465.

— (2005), "Bootstrap prediction intervals for power-transformed time series," *International Journal of Forecasting*, 21, 219–235.

— (2006), "Bootstrap prediction for returns and volatilities in GARCH models," *Computational Statistics & Data Analysis*, 50, 2293–2312.

Pascual, L., Ruiz, E., and Fresoli, D. (2011), "Bootstrap forecast of multivariate VAR models without using the backward representation," *UC3M Working Paper No. 11-26*.

Patilea, V. and Raïssi, H. (2013), "Corrected portmanteau tests for VAR models with time-varying variance," *Journal of Multivariate Analysis*, 116, 190–207.

Pesaran, B. and Pesaran, M. H. (2007), "Modelling volatilities and conditional correlations in futures markets with a multivariate t distribution," *CESifo Working Paper No. 2056*.

Politis, D. N. (2003), "The impact of bootstrap methods on time series analysis," *Statistical Science*, 18, 219–230.

Polito, V. and Wickens, M. (2012), "Optimal monetary policy using an unrestricted VAR," *Journal of Applied Econometrics*, 27, 525–553.

Pope, A. L. (1990), "Biases of estimators in multivariate non-Gaussian autoregressions," *Journal of Time Series Analysis*, 11, 249–258.

Prékopa, A. (2012), "Multivariate value at risk and related topics," *Annals of Operations Research*, 193, 49–69.

Reeves, J. J. (2005), "Bootstrap prediction intervals for ARCH models," *International Journal of Forecasting*, 21, 237–248.

Reinsel, G. (1980), "Asymptotic properties of prediction errors for the multivariate autoregressive model using estimated parameters," *Journal of the Royal Statistical Society*, 42, 328–333.

Riise, T. and Tjøstheim, D. (1984), "Theory and practice of multivariate ARMA forecasting," *Journal of Forecasting*, 3, 309–317.

Rossi, E. and Spazzini, F. (2010), "Model and distribution uncertainty in multivariate GARCH estimation: a Monte Carlo analysis," *Computational Statistics & Data Analysis*, 54, 2786–2800.

Rubner, Y., Tomasi, C., and Guibas, L. J. (1998), "A metric for distributions with applications to image databases," in *Sixth International Conference on Computer Vision*, IEEE, pp. 59–66.

Runkle, D. E. (1987), "Vector autoregressions and reality," *Journal of Business & Economic Statistics*, 5, 437–442.

Silvennoinen, A. and Teräsvirta, T. (2009), "Multivariate GARCH models," in *Handbook of Financial Time Series*, Springer.

Sims, C. (1980), "Macroeconomics and reality," *Econometrica*, 48, 1–48.

Staszewska-Bystrova, A. (2011), "Bootstrap prediction bands for forecast paths from vector autoregressive models," *Journal of Forecasting*, 30, 721–735.

Staszewska-Bystrova, A. and Winker, P. (2013), "Constructing narrowest pathwise bootstrap prediction bands using threshold accepting," *International Journal of Forecasting*, 29, 221–233.

StataCorp (2013), "Stata 13 Base Reference Manual," *College Station, TX: Stata Press*.

Stine, R. A. (1987), "Estimating properties of autoregressive forecasts," *Journal of the American Statistical Association*, 82, 1072–1078.

Stock, J. H. and Watson, M. W. (2001), "Vector autoregressions," *Journal of Economic Perspectives*, 15, 101–115.

Tay, A. S. and Wallis, K. F. (2000), "Density forecasting: a survey," *Journal of Forecasting*, 19, 45–68.

Teräsvirta, T. (2009), "An introduction to univariate GARCH models," in *Handbook of Financial Time Series*, Springer.

Thombs, L. A. and Schucany, W. R. (1990), "Bootstrap prediction intervals for autoregression," *Journal of the American Statistical Association*, 85, 486–492.

Tjøstheim, D. and Paulsen, J. (1983), "Bias of some commonly-used time series estimates," *Biometrika*, 70, 389–399.

Tong, H. and Zhang, Z. (2005), "On time-reversibility of multivariate linear processes," *Statistica Sinica*, 15, 495–504.

Tse, Y. K. (2000), "A test for constant correlations in a multivariate GARCH model," *Journal of Econometrics*, 98, 107–127.

Waggoner, D. F. and Zha, T. (1999), "Conditional forecasts in dynamic multivariate models," *Review of Economics and Statistics*, 81, 639–651.

Wang, C.-P., Lin, S.-H., Huang, H.-H., and Wu, P.-C. (2012), "Using neural network for forecasting TXO price under different volatility models," *Expert Systems with Applications*, 39, 5025–5032.

Wolden Bache, I., Sofie Jore, A., Mitchell, J., and Vahey, S. P. (2011), "Combining VAR and DSGE forecast densities," *Journal of Economic Dynamics and Control*, 35, 1659–1670.

Wolf, M. and Wunderli, D. (2012), "Bootstrap joint prediction regions," *University of Zurich Working Paper No. 64*.

Wright, J. H. (2013), "Evaluating Real-Time Var Forecasts With An Informative Democratic Prior," *Journal of Applied Econometrics*, 28, 762–776.

Yamamoto, T. (1981), "Predictions of multivariate autoregressive-moving average models," *Biometrika*, 68, 485–492.

Yamamoto, T. and Kunitomo, N. (1984), "Asymptotic bias of the least squares estimator for multivariate autoregressive models," *Annals of the Institute of Statistical Mathematics*, 36, 419–430.

Yang, B. and Kolassa, J. E. (2004), "Smooth and accurate multivariate confidence regions," *Journal of the American Statistical Association*, 99, 1072–1081.

Appendix A

Models

Model used in the thesis.

a)*Persistent VAR(4) model.* The autoregressive matrices are given by

$$\Phi_1 = \begin{bmatrix} 0.6362 & -0.0012 \\ 0.0190 & 0.5782 \end{bmatrix}, \Phi_2 = \begin{bmatrix} -0.0168 & -0.0285 \\ 0.5211 & -0.3041 \end{bmatrix},$$

$$\Phi_3 = \begin{bmatrix} 0.0273 & 0.1568 \\ 0.2229 & -0.2529 \end{bmatrix}, \Phi_4 = \begin{bmatrix} 0.1517 & -0.0198 \\ 0.2734 & 0.0241 \end{bmatrix},$$

The dominant root of $|I_2 - \Phi_1(z^{-1}) - ... - \Phi_5(z^{-5})| = 0$ is 0.89. The intercept is given by

$$\mu = \begin{bmatrix} 0.001 \\ 0.002 \end{bmatrix}.$$

The contemporaneous covariance matrix of the errors is given by

$$\Sigma_\varepsilon = 10^{-3} \times \begin{bmatrix} 0.0792 & 0.0434 \\ 0.0434 & 0.2940 \end{bmatrix}. \tag{A.1}$$

The models used in the Monte Carlo simulations of Chapter 2 and 3 are the following.

b) *Persistent VAR(5) model.* The autoregressive matrices are given by

$$\Phi_1 = \begin{bmatrix} 0.6337 & -0.3424 \\ 0.4688 & 0.6755 \end{bmatrix}, \Phi_2 = \begin{bmatrix} -0.0401 & 0.3005 \\ 0.0100 & 0.1967 \end{bmatrix},$$

$$\Phi_3 = \begin{bmatrix} 0.4806 & -0.5201 \\ 0.2582 & -0.2529 \end{bmatrix}, \Phi_4 = \begin{bmatrix} 0.2952 & 0.2041 \\ 0.2734 & 0.0241 \end{bmatrix},$$

$$\Phi_5 = \begin{bmatrix} -0.2299 & 0.1266 \\ -0.1449 & 0.3240 \end{bmatrix}.$$

The dominant root of $|I_2 - \Phi_1(z^{-1}) - ... - \Phi_5(z^{-5})| = 0$ is 0.96 while the others are below 0.8. The intercept is given by

$$\mu = \begin{bmatrix} 0.0053 \\ 0.0018 \end{bmatrix}.$$

Finally, the contemporaneous covariance matrix of the error is given by

$$\Sigma_\varepsilon = 10^{-3} \times \begin{bmatrix} 0.5412 & 0.4045 \\ 0.4045 & 0.4649 \end{bmatrix}.$$

c) *Near-cointegrated VAR(10).* The autorregresive matrices are given by

$$\Phi_1 = \begin{bmatrix} 1.1763 & 0.0542 \\ 0.7126 & 1.2874 \end{bmatrix}, \Phi_2 = \begin{bmatrix} -0.2624 & -0.0668 \\ -0.7618 & -0.3596 \end{bmatrix},$$

$$\Phi_3 = \begin{bmatrix} 0.0250 & -0.0160 \\ 0.0305 & -0.0168 \end{bmatrix}, \Phi_4 = \begin{bmatrix} 0.0000 & 0.0939 \\ 0.0830 & -0.0510 \end{bmatrix},$$

$$\Phi_5 = \begin{bmatrix} 0.1485 & -0.0512 \\ 0.3678 & 0.0656 \end{bmatrix}, \Phi_6 = \begin{bmatrix} -0.2795 & -0.0278 \\ -0.7140 & 0.1410 \end{bmatrix},$$

$$\Phi_7 = \begin{bmatrix} 0.2671 & -0.1424 \\ 0.3072 & -0.3053 \end{bmatrix}, \Phi_8 = \begin{bmatrix} -0.0786 & 0.4398 \\ 0.2189 & 0.4045 \end{bmatrix},$$

$$\Phi_9 = \begin{bmatrix} -0.1181 & -0.3490 \\ -0.1909 & -0.2209 \end{bmatrix}, \Phi_{10} = \begin{bmatrix} 0.0915 & 0.0828 \\ -0.0229 & 0.0071 \end{bmatrix}.$$

The largest roots of $|I_2 - \Phi_1(z^{-1}) - ... - \Phi_{10}(z^{-10})| = 0$ are 0.97, 0.97, 0.92 and 0.92. Moreover, the number of roots above 0.8 rises to 12. The intercept is given by

$$\mu = \begin{bmatrix} 0.1121 \\ 0.1116 \end{bmatrix}.$$

The covariance matrix of the error is

$$\Sigma_\varepsilon = 10^{-4} \times \begin{bmatrix} 0.025 & 0.009 \\ 0.009 & 0.387 \end{bmatrix}.$$

VECM model used to illustrate the construction of forecast densities in Chapter 5.

d) Trivariate VECM(1) with $r = 2$ given by

$$\begin{bmatrix} \Delta y_{1,t} \\ \Delta y_{2,t} \\ \Delta y_{3,t} \end{bmatrix} = \begin{bmatrix} -0.2 & 0 \\ 0 & -0.3 \\ 0.1 & 0.1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -0.8 & -0.8 \\ 4 & 1 \end{bmatrix}' \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \\ y_{3,t-1} \end{bmatrix} + \begin{bmatrix} -0.2 & 0 & 0 \\ 0 & 0.4 & 0 \\ 0 & 0 & 0.1 \end{bmatrix} \begin{bmatrix} \Delta y_{1,t-1} \\ \Delta y_{2,t-1} \\ \Delta y_{3,t-1} \end{bmatrix} + \begin{bmatrix} 0.04 \\ 0.03 \\ 0.09 \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \varepsilon_{3,t} \end{bmatrix} \tag{A.2}$$

where ε_t is distributed χ^2_4 with zero mean and covariance matrix

$$\Sigma_\varepsilon = \begin{bmatrix} 1 & 0.3 & 0.4 \\ 0.3 & 1 & 0.3 \\ 0.4 & 0.3 & 1 \end{bmatrix}.$$

Appendix B

Monte Carlo volumes of Bonferroni cubes for VAR(2), VAR(5) and VAR(10) when $T = 300$

We plot the empirical volumes of h-steps-ahead Bonferroni cubes, $h = 1, ..., 8$, for the stationary VAR(2), persistent VAR(5) and a near-cointegrated VAR(10) model when the sample size is $T = 300$.

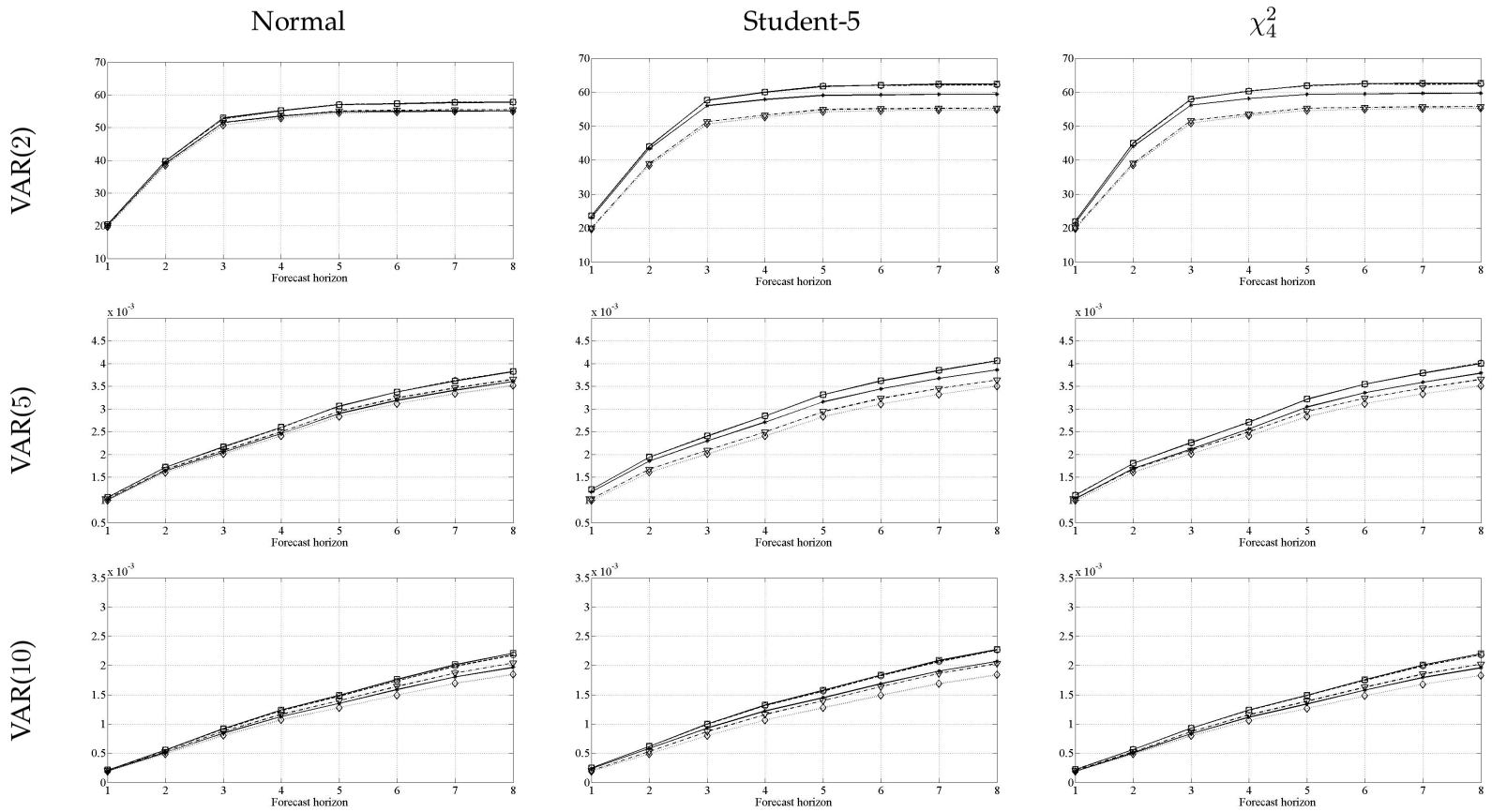


Figure B.1: Monte Carlo averages of empirical volumes of Bonferroni forecast cubes based on: i) Gaussian ($\diamondsuit$), ii) Gaussian with asymptotic MSFE ($\triangledown$), iii) bootstrap with BR ($\circ$) and iv) new bootstrap ($\square$) densities, for a stationary VAR(2) (first row), persistent VAR(5) model (second row) and a near-cointegrated VAR(10) model (third row) with $T = 300$ and Gaussian (first column), Student-5 (second column) and χ^2_4 (third column) errors. Nominal coverage 95%.

Appendix C

Monte Carlo coverages and lengths of marginal forecast intervals for VAR(2), VAR(5) and VAR(10) models

We include the Monte Carlo results for the first variable in the stationary VAR(2), the persistent VAR(5) and the near-cointegrated VAR(10) models since we may also being interested in forecasting only one variable in the system.

Empirical coverages of marginal interval for first component ($y_{1,t}$) of a stationary VAR(2).

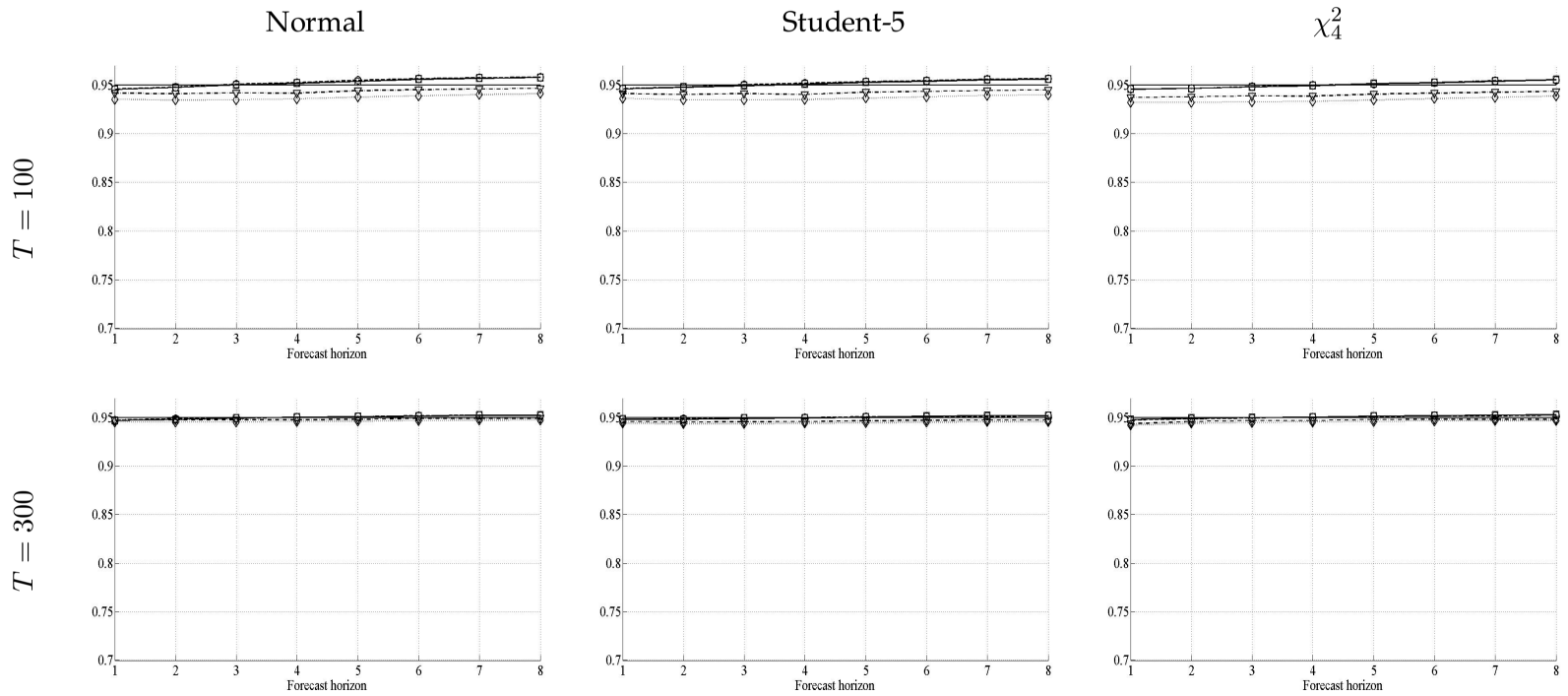


Figure C.1: Monte Carlo averages of empirical coverages of marginal forecast intervals based on : i) Gaussian (- - ♢), ii) Gaussian with asymptotic MSFE (·· ▽), iii) bootstrap with BR (- ○) and iv) new bootstrap (- · □) densities, for a stationary VAR(2) model with samples sizes $T = 100$ (first row) and 300 (second row) and Gaussian (first column), Student-5 (second column) and χ^2_4 (third column) errors. Nominal coverage 95%.

Empirical lengths of marginal interval for first component ($y_{1,t}$) of a stationary VAR(2) model.

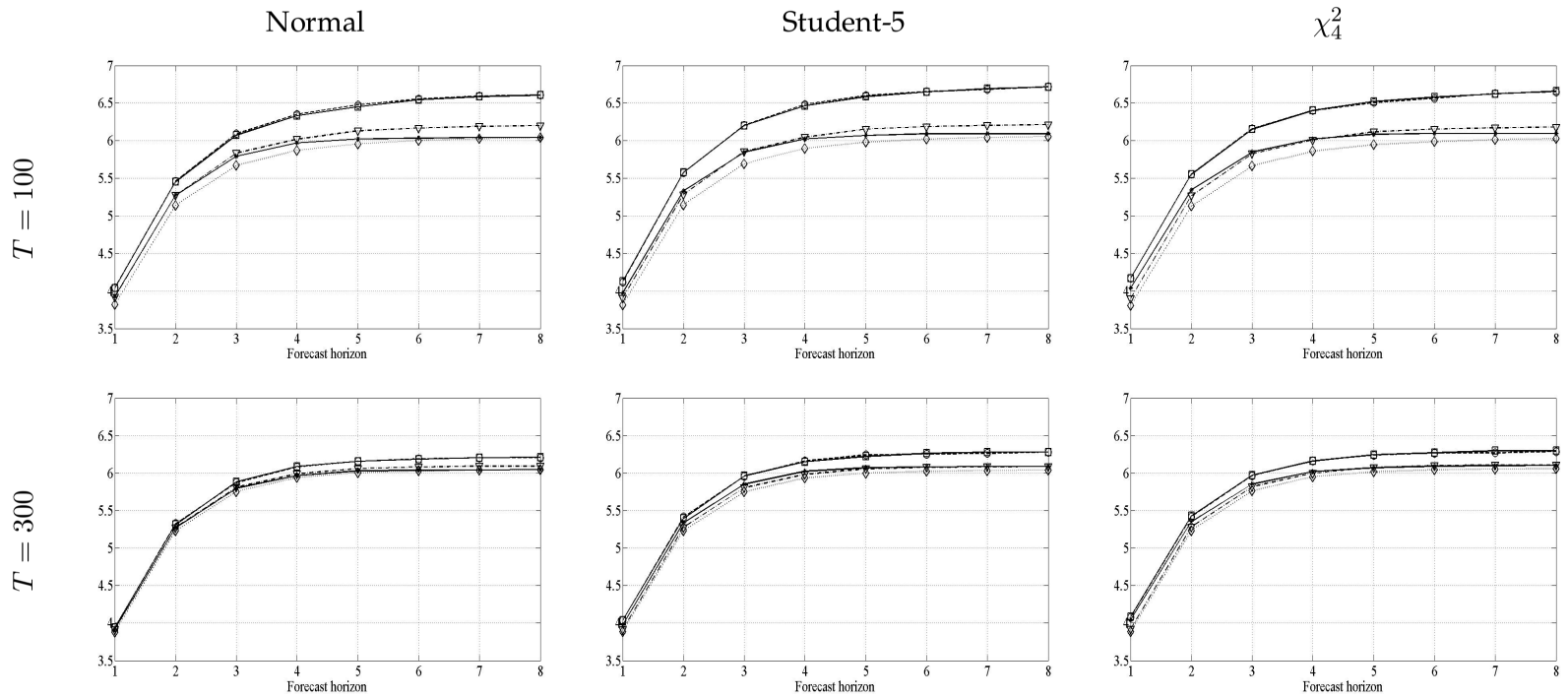


Figure C.2: Monte Carlo averages of empirical lengths of marginal forecast intervals based on : i) Gaussian (- - $\diamondsuit$), ii) Gaussian with asymptotic MSFE ($\cdot\cdot$ $\triangledown$), iii) bootstrap with BR (- $\circ$) and iv) new bootstrap (- $\cdot$ $\square$) densities, for a stationary VAR(2) model with samples sizes $T = 100$ (first row) and 300 (second row) and Gaussian (first column), Student-5 (second column) and χ^2_4 (third column) errors. Nominal coverage 95%.

Empirical coverages of marginal interval for first component ($y_{1,t}$) of a persistent VAR(5) model.

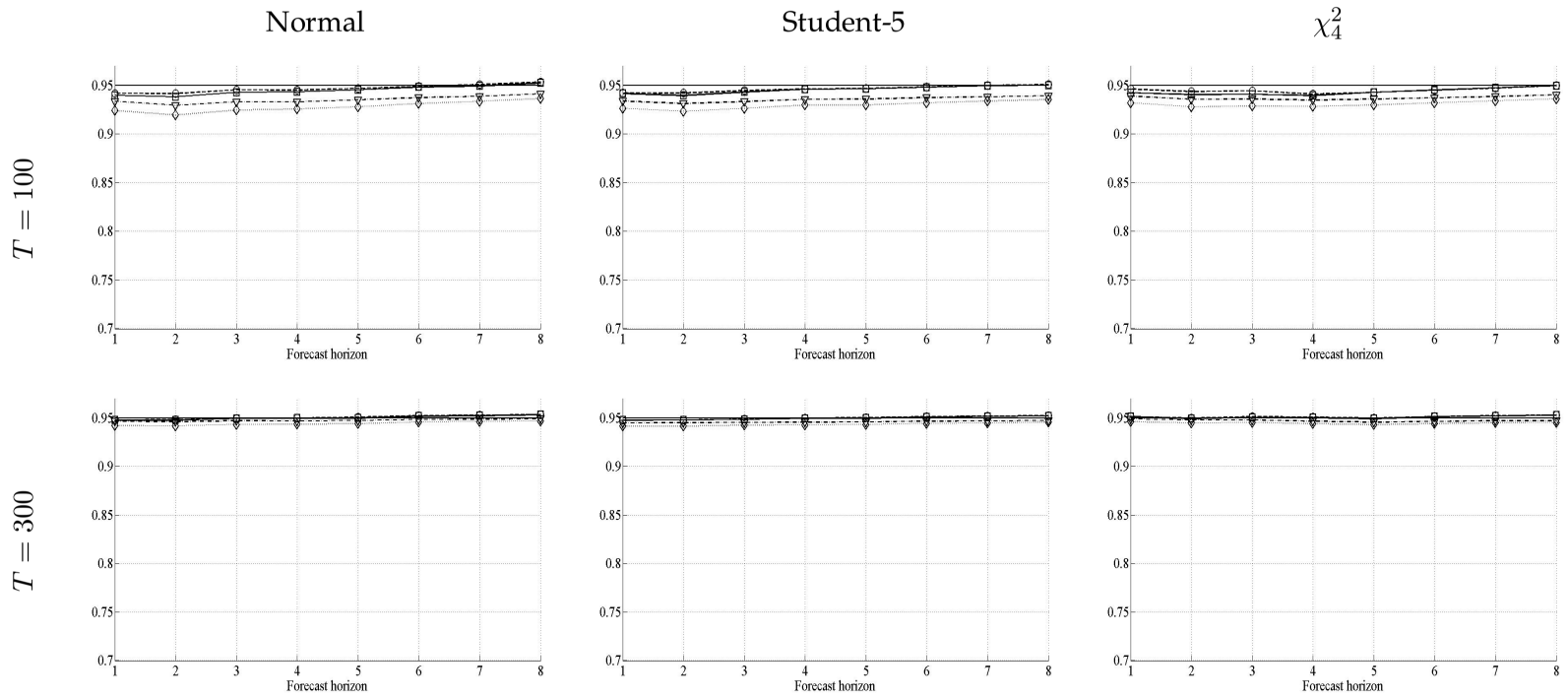


Figure C.3: Monte Carlo averages of empirical coverages of marginal forecast intervals based on : i) Gaussian (- - $\diamondsuit$), ii) Gaussian with asymptotic MSFE ($\cdot\cdot$ $\triangledown$), iii) bootstrap with BR (- $\circ$) and iv) new bootstrap (- $\cdot$ $\square$) densities, for a persistent VAR(5) model with samples sizes $T = 100$ (first row) and 300 (second row) and Gaussian (first column), Student-5 (second column) and χ^2_4 (third column) errors. Nominal coverage 95%.

Empirical lengths of marginal interval for the first component ($y_{1,t}$) of a persistent VAR(5) model.

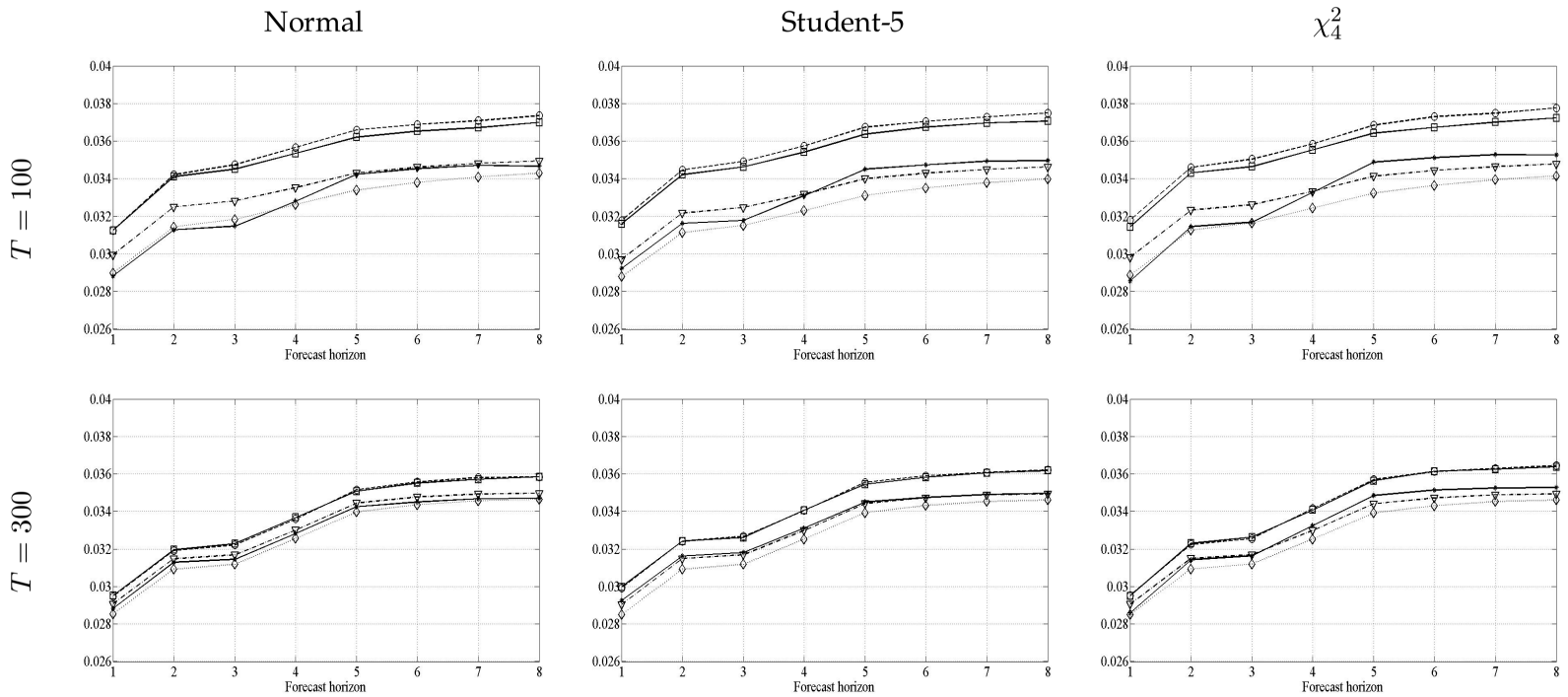


Figure C.4: Monte Carlo averages of empirical lengths of marginal forecast intervals based on : i) Gaussian (- - $\diamondsuit$), ii) Gaussian with asymptotic MSFE ($\cdot\cdot$ $\triangledown$), iii) bootstrap with BR (- $\circ$) and iv) new bootstrap (- $\cdot$ $\square$) densities, for a persistent VAR(5) model with samples sizes $T = 100$ (first row) and 300 (second row) and Gaussian (first column), Student-5 (second column) and χ^2_4 (third column) errors. Nominal coverage 95%.

Empirical coverages of marginal interval for first component ($y_{1,t}$) of a near-cointegrated VAR(10) model.

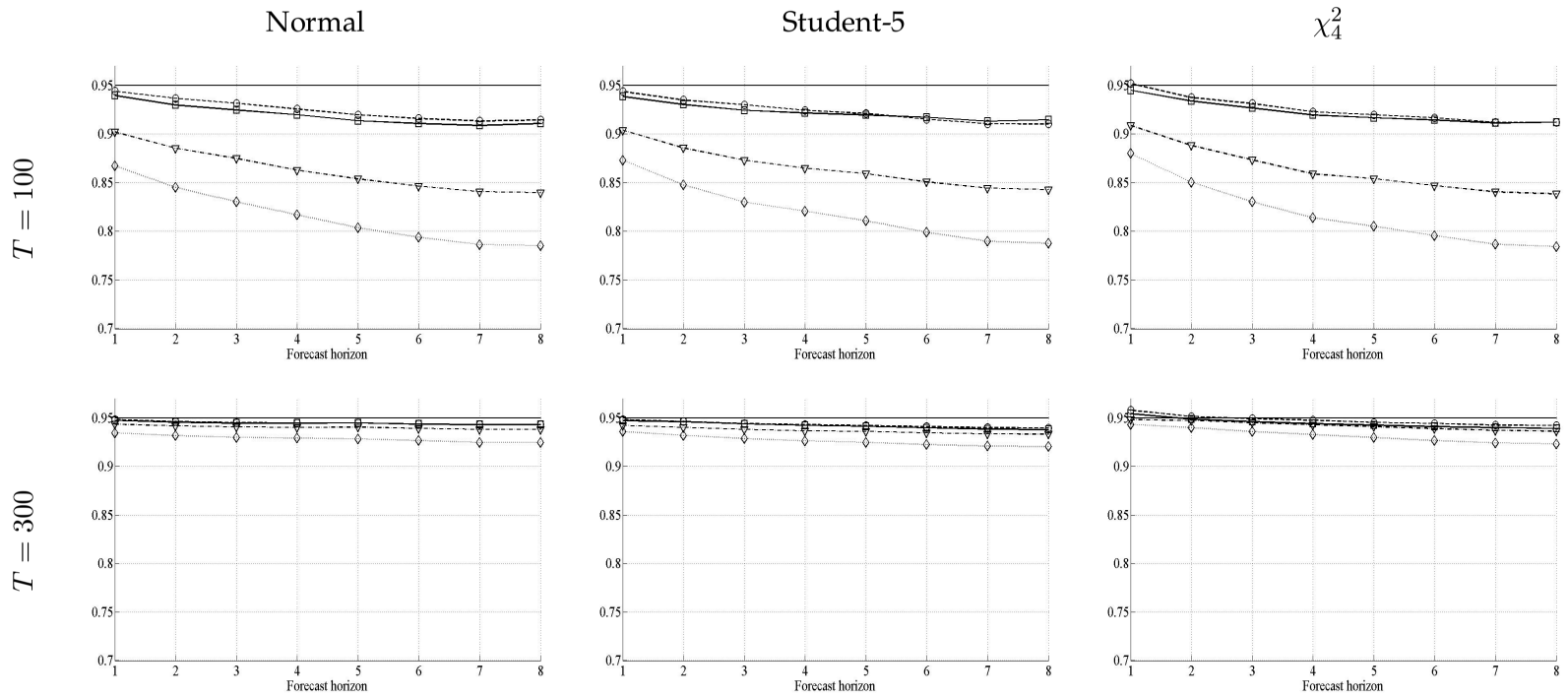


Figure C.5: Monte Carlo averages of empirical coverages of marginal forecast intervals based on : i) Gaussian (- - ♢), ii) Gaussian with asymptotic MSFE (·· ▽), iii) bootstrap with BR (- ○) and iv) new bootstrap (- · □) densities, for a near-cointegrated VAR(10) model with samples sizes $T = 100$ (first row) and 300 (second row) and Gaussian (first column), Student-5 (second column) and χ^2_4 (third column) errors. Nominal coverage 95%.

Empirical lengths of marginal interval for the first component ($y_{1,t}$) of a near-cointegrated VAR(10) model.

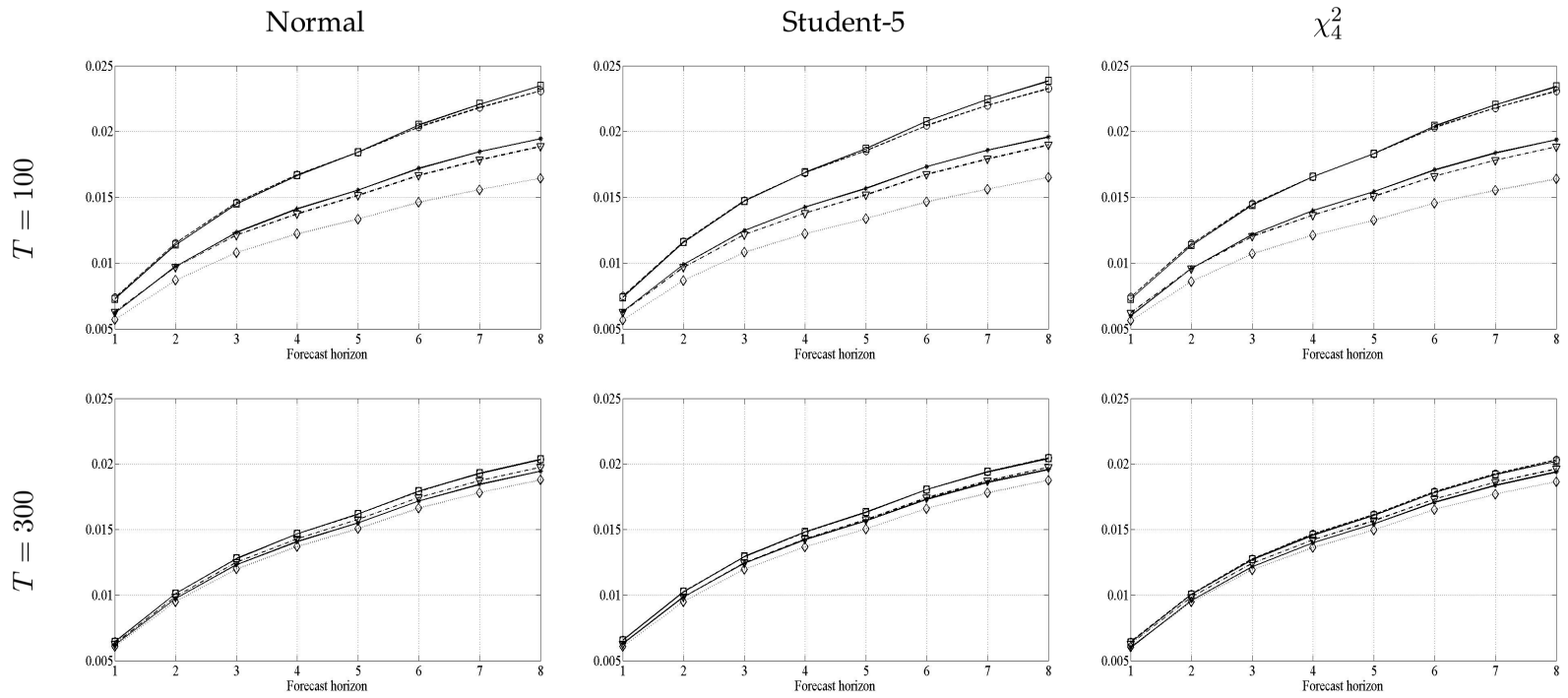


Figure C.6: Monte Carlo averages of empirical lengths of marginal forecast intervals based on : i) Gaussian (- - $\diamondsuit$), ii) Gaussian with asymptotic MSFE ($\cdot\cdot$ $\triangledown$), iii) bootstrap with BR (- $\circ$) and iv) new bootstrap (- $\cdot$ $\square$) densities, for a near-cointegrated VAR(10) model with samples sizes $T = 100$ (first row) and 300 (second row) and Gaussian (first column), Student-5 (second column) and χ^2_4 (third column) errors. Nominal coverage 95%.

Appendix D

Monte Carlo coverages of Bonferroni cubes based on Gaussian and several bootstrap forecast densities for a persistent VAR(5)

We plot the empirical coverages of h-steps-ahead Gaussian and several bootstrap Bonferroni cubes, $h = 1, ..., 8$, for the persistent VAR(5).

Empirical coverages of Gaussian and several bootstrap Bonferroni cubes for a persistent VAR(5).

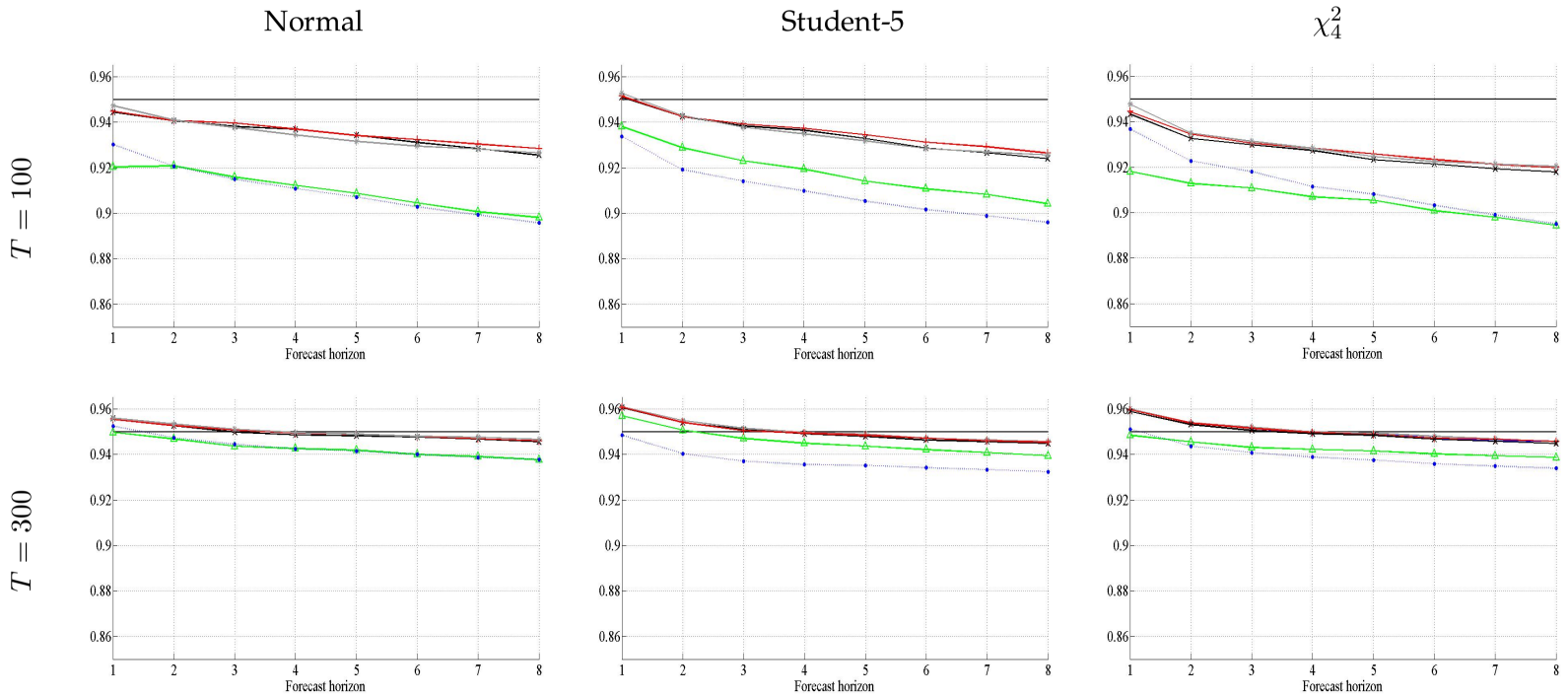


Figure D.1: Monte Carlo averages of coverages of 95% Bonferroni cubes based on STD (blue $\cdot$), DIST-B (green $\triangle$), BB (black $\times$), BCB (red $+$) and BCEB (grey $\star$), for a persistent VAR(5) model with sample sizes $T = 100$ (first row) and 300 (second row) and Gaussian (first column), Student-5 (second column) and χ^2_4 (third column) errors.