EUI Working Papers

ECO 2010/18
DEPARTMENT OF ECONOMICS

EMPIRICAL SIMULTANEOUS CONFIDENCE REGIONS FOR PATH-FORECASTS

Òscar Jordá, Malte Knüppel and Massimiliano Marcellino

EUROPEAN UNIVERSITY INSTITUTE, FLORENCE
DEPARTMENT OF ECONOMICS

Empirical Simultaneous Confidence Regions for Path-Forecasts

ÒSCAR JORDÁ,

MALTE KNÜPPEL

and

MASSIMILIANO MARCELLINO

EUI Working Paper **ECO** 2010/18

ISSN 1725-6704

Printed in Italy
European University Institute
Badia Fiesolana
I – 50014 San Domenico di Fiesole (FI)
Italy
www.eui.eu
cadmus.eui.eu

March 2010

Empirical Simultaneous Confidence Regions for Path-Forecasts*

Abstract

Measuring and displaying uncertainty around path-forecasts, i.e. forecasts made in period T about the expected trajectory of a random variable in periods T+1 to T+H is a key ingredient for decision making under uncertainty. The probabilistic assessment about the set of possible trajectories that the variable may follow over time is summarized by the simultaneous confidence region generated from its forecast generating distribution. However, if the null model is only approximative or altogether unavailable, one cannot derive analytic expressions for this confidence region, and its non-parametric estimation is impractical given commonly available predictive sample sizes. Instead, this paper derives the approximate rectangular confidence regions that control false discovery rate error, which are a function of the predictive sample covariance matrix and the empirical distribution of the Mahalanobis distance of the path-forecast errors. These rectangular regions are simple to construct and appear to work well in a variety of cases explored empirically and by simulation. The proposed techniques are applied to provide confidence bands around the Fed and Bank of England real-time path-forecasts of growth and inflation.

JEL Classification Codes: C32, C52, C53

Keywords: path forecast, forecast uncertainty, simultaneous confidence region, Scheffé's S-method, Mahalanobis distance, false discovery rate.

Òscar Jordà
Department of Economics
University of California, Davis
One Shields Ave.
Davis, CA 95616
e-mail: ojorda@ucdavis.edu

Malte Knüppel
Deutsche Bundesbank
Wilhelm-Epstein-Str. 14
D-60431 Frankfurt am Main, Germany
e-mail: malte.knueppel@bundesbank.de

Massimiliano Marcellino
Department of Economics
European University Institute
Via della Piazzuola 43
50133 Firenze, Italy
e-mail: massimiliano.marcellino@eui.eu

*This paper represents the authors' personal opinions and does not necessarily reflect the views of the Deutsche Bundesbank. We thank seminar participants at the Bundesbank for helpful comments and suggestions. Jordà acknowledges the hospitality of the Federal Reserve Bank of San Francisco during the preparation of this manuscript, and partial financial support of the Spanish Ministerio de Ciencia e Innovación, grant SEJ-2007-63098.

1 Introduction

Policy-makers are frequently confronted with a decision problem that requires assessing the set of possible trajectories that a state variable will follow over time. The prediction of such trajectories is often embedded in a control problem so that certain deviations of the state variable from pre-determined targets will trigger particular policy responses. For example, many central banks produce forecasts of inflation over two year-ahead horizons to determine the best monetary policy response (see e.g. Svensson, 2009 and references therein for a discussion on optimal policy projections. For example, inflation path-forecasts and policy path-projections are now regularly reported by the Swedish Riksbank at `www.riksbank.com`). Interest is not about the point forecast for a specific horizon but on the sequence of forecasts over the entire trajectory.

Such situations present at least two important statistical challenges: What is the proper measure of uncertainty associated with the trajectory of a random variable over time?; and How should one compute such a measure when the null distribution and the actual distribution of the forecasts may differ, or when the former may not even be available? Socioeconomic models lack the precision and stability of models of physical phenomena so that their approximative nature makes them more vulnerable to traditional assumptions commonly used to derive closed-form expressions. In the extreme, the decision-maker may in fact be ignorant of the model used to generate the forecasts themselves. This is often the case when the forecasts are not produced under the direct supervision of the decision maker. For example, several forecasting firms (Blue Chip Forecasting, Macroeconomic Advisors, Standard and Poor, etc.) release U.S. inflation forecasts but not sufficient details about the forecasting null model.

This paper addresses these challenges. Conceptually, interest about the trajectory of a variable over time reflects an interest in its path-forecast (see Jordà and Marcellino, 2009): the collection of forecasts 1- to H-steps ahead. For example, pricing many path-dependent "exotic" options (see, e.g. Goldman, Sosin and Gatto, 1979 for an early treatment and Conze, 1991 and Kwok and Lau, 2001 for more recent reviews) such as lookback, Asian and cumulative Parisian options to cite a few, requires determining the behavior of the price of the underlying asset throughout the maturity of the contract and not just the likelihood of such a price exceeding the strike price at any given time before maturity.

Providing appropriate confidence bands for these path-forecasts is complicated for two reasons: confidence bands for the path-forecast are the inverse of a multiple testing problem that results in multidimensional elliptical confidence regions; moreover, the elements of the path-forecast are correlated over time. Since communicating uncertainty about the path-forecast is as important as the forecast itself, Jordà and Marcellino (2009) have suggested how to construct rectangular approximations to the optimal confidence regions based on Scheffé's (1953) S-method.

However, the derivations needed to obtain appropriate Scheffé bands (as Jordà and Marcellino, 2009 denominate these confidence regions) require fairly strict assumptions about the Gaussianity of the error process and the correct specification of the null model. A violation of either or both of these assumptions likely reflects the behavior of macroeconomic data. Furthermore, when the forecasts themselves are available but not the model and parameter estimates used to generate them, it becomes infeasible to derive analytic formulas.

For these reasons, we investigate a simple method to construct empirical simultaneous confidence regions based on the methods by Williams and Goodman (1971), whose ideas precede but are related to the resampling methods in Efron's (1979) well-known bootstrap procedures, and subsequent literature (see, e.g. Politis, Romano and Wolf 1999).

The procedures we propose use the sequence of observed forecast errors at different horizons to construct empirical estimates of the sampling covariance matrix of the forecast errors across horizons. This approach is similar to the sample-splitting empirical approach used, e.g. by Rubin, Dudoit and van der Laan (2006) in the context of cross-sectional multiple testing environments. Because we are considering the path-forecast and simultaneous confidence regions, it is not practical, given typical predictive sample sizes, to estimate non-parametrically the joint empirical distribution of the forecast errors and the relevant quantile contours, which would be the object one would need to formally construct optimal simultaneous confidence regions (allowing the predictive sample go to infinity). Instead, we obtain rectangular confidence regions by constructing simultaneous confidence regions based on the Mahalanobis (1936) distance induced by the empirical covariance matrix of the forecast errors, and then derive rectangular regions with Scheffé's (1953) S-method. Although we rely on the appropriate quantiles of the χ^2 distribution in our examples because of the particulars of our analysis, we remark that one could use tail-probability estimates of the empirical distribution of the Mahalanobis distance when predictive samples are sufficiently large.

Standard methods used to compute forecast intervals assume the forecasting model describes the series adequately in the future. In contrast, our procedures rely on the empirical distribution of the forecast errors remaining the same in the future, which is a weaker assumption. We note that our procedures differ from common applications of bootstrap procedures in econometric models because our interest is on the joint distribution of the forecast errors for which no assumption is made about whether or not the null forecasting model is correct. Typical applications of the bootstrap in forecasting with ARIMA models (e.g. Masarotto, 1990; Thombs and Schucany, 1990; Kim, 1999; Pascual, Romo and Ruiz, 2000; and Clements and Taylor, 2001) are usually based on residual resampling given the null model, but our applications entertain the possibility that the null model is incorrect or not available to the decision-maker.

Therefore, the type of situation that we consider not only takes that data generating distribution as unknown, but possibly the null model used to produce the forecasts as well, and for this reason we cannot produce closed-form analytic expressions using large-sample arguments. Instead, we provide ample simulation evidence about how our procedures may work in practice. In addition, we provide a realistic application of our procedures by assessing the growth and inflation forecasts generated by the Bank of England and the U.S. Federal Reserve, whose null models are not revealed to the public.

Our main findings can be summarized as follows. Perhaps not surprisingly, the empirical simultaneous confidence regions that we propose provide more accurate coverage the smaller the estimation sample (if available), the larger the forecast sample (because the empirical estimates become more accurate) and in the presence of null model misspecification. These improvements are most dramatic as the length of the forecast interval increases. In contrast, the coverage rates of traditional confidence bands based on the marginal distributions of each point forecast are very poor.

When the model is correctly specified (in our case a vector autoregression or VAR), we find that coverage based on analytic formulas applied to VAR estimates is superior to coverage provided by direct forecasts using local projections. However, such differences are quickly reversed when the model is misspecified. On the question of providing accurate simultaneous confidence regions, we find that whether more traditional Bonferroni bounds or Scheffé rectangular regions are more accurate, the answer depends on the type of error control metric considered. We argue that family wise error (FWE) control is not appropriate for path-forecasts, suggesting instead that Benjamini and Hochberg's (1995) false discovery rate control (FDR) is superior. We investigate simulations using both metrics however, and show that our methods are far superior at controlling FDR while being relatively good at controlling FWE. In contrast, Bonferroni bounds control FWE but usually offer poor FDR control. Finally, we find that in both simulations and the empirical applications, empirical Scheffé bands perform best in the vast majority of cases and therefore constitute the approach that we recommend.

The paper is organized as follows. Section 2 discusses statistical issues related with measuring path-forecast uncertainty. Section 3 provides simulation evidence on the coverage rates of alternative path confidence bands. Section 4 computes and assesses competing uncertainty measures for the real-time Fed and Bank of England path-forecasts of growth and inflation. Section 5 summarizes our main findings and concludes.

2 Statistical Discussion

Suppose at time T we are interested in predicting the value of a random variable one to H periods into the future and hence define the vector of forecasts and actual realizations of these random variables respectively as:

$$\underset{H\times 1}{\widehat{Y}_T(H)} = \begin{bmatrix} \widehat{y}_T(1) \\ \vdots \\ \widehat{y}_T(H) \end{bmatrix}; \underset{H\times 1}{Y_{T,H}} = \begin{bmatrix} y_{T+1} \\ \vdots \\ y_{T+H} \end{bmatrix}.$$

We call the vector $\widehat{Y}_T(H)$ a path-forecast using the nomenclature in Jordà and Marcellino (2009). Associated to these vectors, we can define the vector of forecast errors $\widehat{U}_T(H) \equiv \widehat{Y}_T(H) - Y_{T,H}$ with a distribution that is assumed to be centered at zero when the null model is correctly specified and with $H \times H$ covariance matrix Ω_H. We call this the forecast generating distribution.

This distribution reflects a variety of sources of uncertainty associated with the predictions in $\widehat{Y}_T(H)$, among these the most common are: (1) uncertainty about the data generating distribution of $Y_{T,H}$ from which innovations are drawn; (2) uncertainty about the null model used to describe the data generating distribution; and (3) uncertainty from the parameter estimates of the null model in finite samples with respect to their population values. The reader is referred to Clements and Hendry (1998) for a more exhaustive list. For our purposes, it is unnecessary to be explicit about this breakdown. Our framework also extends to forecasting the paths of a vector of random variables and we do this in the simulations but here we prefer to keep the discussion simple. It will be important for our purposes that the forecast generating distribution be stable over time. This assumption is less restrictive than assuming that the data generating distribution is itself stable. In the next sections we discuss some of the particular statistical issues on forecasting uncertainty with particular effort in distinguishing the results that do not require distributional assumptions from those that do.

2.1 One Period-ahead Forecasting

Consider the problem of constructing a confidence interval for a one period-ahead forecast (i.e. $H = 1$). In general, the confidence region can be described as:

$$S = \{y_{T+1} : \underline{c} \leq T_1(\widehat{y}_T(1)) \leq \overline{c}\} \tag{1}$$

where the function $T_1(.) : \mathbb{R} \to \mathbb{R}$ and an example of such a function is the well-known t-ratio,

$$T_1(\widehat{y}_T(1)) = \frac{\widehat{y}_T(1) - y_{T+1}}{\sigma_1} \tag{2}$$

and where the lower and upper bounds $\underline{c}$ and $\overline{c}$ are chosen respectively so that

$$P(y_{T+1} \in S) = 1 - \alpha. \tag{3}$$

When $-\underline{c} = \overline{c} = c$ Chebyshev's inequality provides a bound for this probability regardless of the forecast generating distribution,

$$P\left(|\widehat{y}_T(1) - y_{T+1}| < c\sigma_1\right) \geq 1 - \frac{1}{c^2}$$

and hence $c = \sqrt{1/\alpha}$. For conventional 68% and 95% coverage (the usual one- and two-standard deviation limits under Gaussianity), one obtains $c = 1.77$ and $c = 4.47$ respectively, which are clearly much larger values than if it were known that $T_1 \sim N(0, 1)$, in which case, of course, $c = 1$ and $c = 2$ would do the trick.

This confidence interval can also be interpreted as the inverse of the decision problem associated with the null hypothesis

$$H_0 : \widehat{u}_T(1) = 0 \qquad vs. \qquad H_1 : \widehat{u}_T(1) \neq 0. \tag{4}$$

This notation may look a little awkward from what is conventional but is meant to convey that we are not interested in a test of the hypothesis

$$H_0 : E\left[u_T(1)|\mathcal{I}_{T-1}\right] = 0 \qquad vs. \qquad H_1 : E\left[u_T(1)|\mathcal{I}_{T-1}\right] \neq 0$$

where $\mathcal{I}_{T-1}$ refers to the conditioning information set. The difference lies in the interest to provide uncertainty about the possible realizations of $u_T(1)$ rather than its conditional mean.

Let T_1 denote a statistic associated with the null in expression (4) and denote a *false positive,* fp, as:

$$fp_1 = I\left(|T_1| > c|H_0 \text{ is true}\right)$$

where $I(.)$ is the indicator function and for convenience from hereon we proceed with the convention $-\underline{c} = \overline{c} = c$. Then, the choice of c such that

$$P\left(fp_1 = 1\right) = \alpha \tag{5}$$

for a pre-specified level α is meant to control the probability of a false positive or Type I error and hence generates a confidence interval with $1-\alpha$ coverage since

$$P\left(fp_1 = 0\right) = P\left(I\left(|T_1| \leq c\right)|H_0 \text{ is true}\right) = 1 - \alpha. \tag{6}$$

In this case, the probability of Type II error is the probability of a false negative, that is $P\left(fn_1 = 1\right) = \beta$ where

$$fn_1 = I\left(|T_1| < c|H_1 \text{ is true}\right).$$

If one is unwilling to make assumptions about the distribution of T_1, one could use a predictive sample of forecast errors $\widehat{u}_{T+j}(1)$ for $j = 1, ...N$ from which one could calculate

$$\widehat{\sigma}_1^2 = \frac{1}{N}\sum_{j=1}^{N}\widehat{u}_{T+j}(1)^2$$

and hence obtain an empirical estimate of T_1 from which the desired empirical quantile could be used to obtain a value of c that would meet condition (6) and hence define a confidence region (1). For example, the confidence region associated when one defines T_1 as in (2) would be the usual rectangular region but where the value of c is determined from the appropriate quantiles of $\widehat{T}_1$.

2.2 Path-forecasting

The one-to-one correspondence between control of Type I error and the coverage of the confidence interval breaks down when one considers multiple hypotheses and hence construction of simultaneous confidence regions. Specifically, interest now is in obtaining regions of the form

$$S = \left\{ Y_{T,H} : \underline{\delta} \leq W(\widehat{U}_T(H)) \leq \overline{\delta} \right\} \tag{7}$$

where $W(.) : \mathbb{R}^H \to \mathbb{R}$ is some function and where $\underline{\delta}$ and $\overline{\delta}$ are chosen so that

$$P\left(Y_{T,H} \in S\right) = 1 - \alpha. \tag{8}$$

Notice that the values of $Y_{T,H}$ that meet condition (8) will generate a multi-dimensional geometric object that generally cannot be represented in two-dimensional space. We will return to this issue below.

A common expression for $W(.)$ is the well-known Mahalanobis distance

$$W_M = \sqrt{\left(\widehat{Y}_T(H) - Y_{T,H}\right)' \Omega_H^{-1} \left(\widehat{Y}_T(H) - Y_{T,H}\right)} \tag{9}$$

which together with Chebyshev's inequality in the multi-dimensional case, and assuming

$-\underline{\delta} = \overline{\delta} = \delta$, can be used to obtain the probability bound

$$P\left(W_M < \delta\right) \geq 1 - \frac{H}{\delta^2}$$

so that, for given α, it is $\delta = (H/\alpha)^{1/2}$. If we knew that $W_M^2 \sim \chi_H^2$, as is conventional in many traditional multiple testing situations with cross sectional data in large samples, then δ would be about 1.65 times as large as the critical value of a random variable with distribution χ_H^2 evaluated at the 68% probability level and more than 3 times if we chose a 95% probability level instead. As we already discussed in the previous section, knowledge of the distribution makes the bounds considerably tighter relative to what we can obtain with Chebyshev's inequality.

In the discussion that follows we find it convenient to note that Ω_H is symmetric and positive-definite and thus admits unique Cholesky decomposition $\Omega_H = PD_HP'$ where P is a lower triangular matrix with ones in the diagonal and D_H is a diagonal matrix. Thus, the Mahalanobis distance in (9) can be expressed as:

$$W_M = \sqrt{\widehat{U}_T(H)'\Omega_H^{-1}\widehat{U}_T(H)} = \sqrt{\widehat{V}_T(H)'D_H^{-1}\widehat{V}_T(H)}$$

where $\widehat{V}_T(H) = P^{-1}\widehat{U}_T(H)$. Furthermore,

$$W_M^2 = \widehat{V}_T(H)'D_H^{-1}\widehat{V}_T(H) = \sum_{h=1}^{H} \frac{\widehat{v}_T(h)^2}{d_{hh}} = \sum_{h=1}^{H} (T_h^*)^2 .$$

One way to think of the $\widehat{v}_T(h)$ is as the orthogonalized versions of the $\widehat{u}_T(h)$ where one projects $\widehat{u}_T(h)$ onto $\widehat{u}_T(h-1), ..., \widehat{u}_T(1)$. The d_{hh} are the elements in the diagonal of D_H and therefore the variances of the $\widehat{v}_T(h)$. The usefulness of this transformation will become clear momentarily.

The simultaneous confidence region in expression (7) is now the inverse of a multiple testing procedure that involves the intersection of the family of hypotheses

$$H_{0h} : \widehat{v}_T(h) = 0 \qquad vs. \qquad H_{1h} : \widehat{v}_T(h) \neq 0.$$

The joint null is $H_0 = \cap_{h=1}^{H} H_{0h}$ and $H_1 = \cup_{h=1}^{H} H_{1h}$. Paralleling the discussion in the previous section, let the total number of false positives among the H hypotheses be

$$FP = \sum_{j \in J} I\left(\left|T_j^*\right| > c_j^*\right)$$

where J denotes the index set containing all true null hypotheses and where the c_j^* are as of yet, to be determined. Of course, because we are talking about the uncertainty of the path-forecast then for practical purposes $J = H$, but we retain the notation J for the time being to be more precise in the discussion that follows.

Unlike the familiar single hypothesis scenario described previously, there are several error rates one may wish to control for. Specifically, the closest equivalent to expression (5) is control of the generalized family wise error ($gFWE$), defined as:

$$gFWE = P(FP \geq k) = \alpha \text{ for } 1 \leq k < H$$

from which corresponding values of c_j^* could be chosen and where it is traditional to choose $k = 1$ in which case we can simply write $FWE = P(FP \geq 1)$.

We pause our discussion of error control to remark that it has been standard practice to construct confidence intervals for path-forecasts using the cut-off values associated with control of the error rate for each individual hypothesis H_{0h} so that the c_j are chosen to meet the condition

$$P(fp_j = 1) \; \forall j \in J.$$

However, this causes Type I error for the intersection of nulls to approach 1 as the number of hypothesis considered grows, and therefore generates severe distortions on the desired probability coverage of the path-forecast.

Instead notice that

$$P(FP \geq 1) = \alpha = \sum_{J} P(fp_j = 1)$$

and hence Bonferroni's inequality suggests choosing instead

$$P(fp_j = 1) = \frac{\alpha}{J}$$

so that

$$P(FP \geq 1) = \alpha \leq \sum_{J} P(fp_j = 1) = \sum_{J} \frac{\alpha}{J} = \alpha.$$

Therefore, traditional confidence bands constructed as

$$\widehat{Y}_T(H) \pm c\, diag(\Omega_H)^{1/2} \tag{10}$$

where $diag(\Omega_H)$ is the $H \times 1$ vector of diagonal elements of Ω_h and c is chosen to ensure that condition (3) is met, can be replaced with

$$\widehat{Y}_T(H) \pm c_B\, diag(\Omega_H)^{1/2} \tag{11}$$

where c_B is chosen to meet the condition

$$P(y_{T+j} \in S_j) = 1 - \frac{\alpha}{H}$$

and

$$S_j = \{y_{T+j} : |T_j| \leq c_j\}$$

for $c_j = c_B$ for $j = 1, ..., H$.

Control of FWE has been found to be too stringent resulting in very low power (see Dudoit and van der Laan, 2008). For example, in a prediction of the path of monthly inflation over the next two years, control of FWE would result in rejection of such paths as when the trajectory of inflation is correctly predicted for 23 periods but the prediction of the last month is particularly poor.

For this reason, a number of alternative error control procedures have been proposed. Perhaps the best known is Benjamini and Hochberg's (1995) *false discovery rate* (FDR), which can be described as the proportion of incorrectly rejected hypothesis, specifically

$$FDR = E\left[\frac{\sum_J I(T_j^* > c_j^*)}{\max\left\{\left(\sum_{h=1}^H I(T_h^* > c_h^*)\right), 1\right\}}\right].$$

In our case and because we are interested in confidence regions for all H periods then

$$FDR = E\left[\frac{1}{H}\sum_{h=1}^{H} I(T_h^* > c_h^*)\right] = PCER$$

where $PCER$ is the per-comparison error rate (see Dudoit and van der Laan, 2008). In general, control of $PCER$ and FDR is less conservative than control of FWE although when all hypotheses are true, then $FDR = FWE$.

Control of FDR is more appealing for path-forecasting where interest is in constructing confidence bands that preserve the set of trajectories that mimic the overall shape of $Y_{T,H}$ rather than focusing on individual deviations of the elements of the path, as in our previous inflation example. Therefore notice that control of FDR at level α implies that

$$FDR = \frac{1}{H}\sum_{h=1}^{H} P(T_h^* > c_h^*) = \alpha$$

or

$$\frac{1}{H}\sum_{h=1}^{H} P(T_h^* \leq c_h^*) = 1 - \alpha.$$

Recall that a simultaneous confidence region based on the Mahalanobis distance with coverage $1-\alpha$ can be constructed by choosing δ such that $P(W_M \leq \delta) = 1-\alpha$. We now use Scheffé's (1953) S-method in the derivations that follow to tie this result with a confidence region that controls FDR and can be represented in two-dimensional space.

We have shown previously that $W_M^2 = \sum_{h=1}^{H} (T_h^*)^2$ and hence

$$P(W_M^2 \leq \delta^2) = P\left(\sum_{h=1}^{H} (T_h^*)^2 \leq \delta^2\right).$$

Dividing by $1/H$ on both sides of the inequality and noticing that $1/H = \sum_{h=1}^{H} 1/H^2$, we have that

$$P\left(\frac{1}{H}\sum_{h=1}^{H} (T_h^*)^2 \leq \frac{\delta^2}{H}\right) = P(W_M^2 \leq \delta^2) = 1-\alpha$$

and from the Cauchy-Schwarz inequality

$$\sum_{h=1}^{H} \frac{1}{H^2} \sum_{h=1}^{H} (T_h^*)^2 \geq \left(\frac{1}{H}\sum_{h=1}^{H} T_h^*\right)^2$$

so that

$$P\left(\left(\frac{1}{H}\sum_{h=1}^{H} T_h^*\right)^2 \leq \frac{\delta^2}{H}\right) \geq P\left(W_M^2 \leq \delta^2\right) = 1-\alpha$$

and therefore

$$P\left(\left|\frac{1}{H}\sum_{h=1}^{H} T_h^*\right| \leq \sqrt{\frac{\delta^2}{H}}\right) \geq 1-\alpha$$

but since the T_h^* have been orthogonalized, at least for elliptically contoured distributions (such as the multivariate Gaussian), we have then that

$$\frac{1}{H}\sum_{h=1}^{H} P\left(|T_h^*| \leq c_h^*\right) \simeq P\left(\left|\frac{1}{H}\sum_{h=1}^{H} T_h^*\right| \leq \sqrt{\frac{\delta^2}{H}}\right)$$

so that an approximate value of c_h^* is $\sqrt{\delta^2/H}$. If the $\widehat{U}_T(H)$ were multivariate Gaussian, then

$W_M^2 \sim \chi_H^2$ and hence one could then choose

$$c_h^* = \sqrt{\frac{\chi_H^2(\alpha)}{H}}$$

where $\chi_H^2(\alpha)$ is the α-highest quantile of a χ_H^2 distributed random variable, and therefore approximate simultaneous confidence intervals with joint $1-\alpha$ coverage (and with approximate $FDR = \alpha$) as

$$\widehat{Y}_T(H) \pm PD^{1/2}\sqrt{\frac{\chi_H^2(\alpha)}{H}}\mathbf{i}_H \tag{12}$$

where $\mathbf{i}_H$ is an $H \times 1$ vector of ones. As we will show in the simulations this approximation to the c_h^* is very accurate when the data are Gaussian.

The critical value therefore depends on the length of the path-forecast considered. As a result, the width of the bands can vary for the earlier periods in the path when the researcher chooses to evaluate a longer or shorter total horizon H. This is a natural consequence of the simultaneous nature of the evaluation problem and perfectly natural. However, if one is willing to allow the c_h^* to be possibly different from each other (since there is no such theoretical restriction) then one can use a step-down procedure similar to Holm's (1979) that would get around this feature to some extent. We have experimented with multiple values of H and found to be the case that $\sqrt{\chi_h^2(\alpha)/h} \geq \sqrt{\chi_H^2(\alpha)/H}$ for any H and $\alpha < 0.05$ (i.e. the traditional scientific standard of 95% confidence). For $\alpha = 0.32$ (one standard deviation coverage) the approximation is still rather good. For $\alpha = 0.5$, the sequential procedure tends to undercover but not by a large amount. Therefore, for practical purposes one may consider modifying the bands in (12) with the following alternative:

$$\widehat{Y}_T(H) \pm PD^{1/2}\left[\sqrt{\frac{\chi_h^2(\alpha)}{h}}\right]_{h=1}^{H} \tag{13}$$

where the vector in brackets is of dimension $H \times 1$ with typical element $\sqrt{\chi^2_h(\alpha)/h}$. In the Monte Carlo simulations and empirical application that we provide, we decided to follow this last approach as we feel researchers may feel more comfortable with this choice and in any case handicaps our methods with respect to the alternatives that we consider. Even with this handicap, we will show that Scheffé bands are a superior alternative.

Finally, all along we have tried to maintain the discussion by making minimal assumptions on the forecast generating distribution. Two conditions we have discussed is that this distribution be stable over time (even if the data generating distribution is not) and the assumption that a local projection (obtained here with the Cholesky decomposition) is sufficient to orthogonalize the forecast path. Under these conditions, we can use the empirical predictive sample to construct

$$\widehat{\Omega}_H = \frac{1}{N}\sum_{j=1}^{N} \widehat{U}_{T+j}(H)\widehat{U}_{T+j}(H)'$$

when the null model is correctly specified. Given the estimate $\widehat{\Omega}_H$, then one can construct the set of Mahalanobis distances for each of the $1, ..., N$ path-forecasts in the predictive sample, namely

$$\widehat{W}_j^2 = \widehat{U}_{T+j}(H)'\widehat{\Omega}_H^{-1}\widehat{U}_{T+j}(H) \qquad j = 1, ...N$$

which can then be ranked in ascending order to obtain the α% highest ranked value of $\widehat{W}_j$ as the natural value of $\widehat{\delta}$. However, getting an accurate estimate for a small value of α in small samples is difficult (such as $\alpha = 0.05$ for 95% coverage). We have found that in such cases, $\chi^2_H(\alpha)$ often provides a reasonable approximation, as we illustrate below.

3 Probability Coverage of Alternative Confidence Bands: Simulation Evidence

We now present extensive simulation evidence on the probability coverage of different types of confidence bands for a path-forecast. The next section describes the experimental design, the following section reports simulations when the null model coincides with the data generating process (DGP), and the last section considers simulations in which the null model is misspecified with respect to the DGP.

3.1 Experimental design

The experimental design is based on Stock and Watson's (2001) well-cited review article on vector autoregressions (VARs). In that article, Stock and Watson examine a three-variable system, specifically: P, inflation measured by the chain-weighted GDP price index; UN, civilian unemployment rate; and FF, the average federal funds rate. We estimate this VAR over the sample beginning the first quarter of 1960 and ending the first quarter of 2007 (189 observations) for the purpose of using the coefficient estimates and residual covariance matrix in our simulations.

With these estimates as parameter choices, we simulate data from this VAR as follows. Let T denote the estimation sample size, N denote the predictive sample size, and H denote the length of the path-forecast considered. Then we allow $T = 100$, 400; $N = 40, 80$ and 200; and $H = 1, 4, 8$ and 12. For each combination T, N and H we generate 10,000 Monte Carlo samples.

At each replication, we use that sample to generate path-forecasts of length H over a predictive sample N using estimates from either a VAR or local projections fitted on T observations. We consider local projections for the misspecification examples following the results in Marcellino, Stock and Watson (2006). Then, we construct three types of confidence bands: (1) marginal error bands (using expression 10); (2) Bonferroni bounds (using expression 11); and (3) Scheffé bands (using expression 13). We consider both the analytic formulas derived in Jorda and Marcellino (2009), based on the null model being the DGP, as well as the empirical formulas described in the previous section. Moreover, we examine bands for 68% and 95% probability coverage rates. We also entertain misspecification of the null model in ways that will be made explicit momentarily. Forecasts are computed for each of the three variables in the VAR (P, UN, and FF) and reported separately.

Coverage is evaluated in terms of FWE and FDR control. In FWE control, any path that has one or more elements outside the bands is considered to fall outside the bands, regardless of whether its Mahalanobis distance meets the FDR criterion. In FDR control, we compute the Mahalanobis distance for all paths and record as being outside the bands all those paths with Mahalanobis distance higher than $\chi^2_H(\alpha)$. We do not use the empirical critical value because then the Scheffé bands would score perfect FDR by construction. Furthermore, in empirical situations, a predictive sample with less than 100 observation probably does not allow for a very accurate estimate of its α-quantile. Therefore, we felt it would be best to handicap our preferred procedures to make the results even more convincing.

Two final comments are worth making. First, initial conditions to generate the data are chosen at random from the unconditional distribution. Second, in simulations when the null model is misspecified with respect to the DGP, bands are computed around the path-forecast and are not re-centered for misspecification, which is more realistic.

3.2 Simulation Results: Null Model is Correctly Specified

Tables 1-6 summarize the results of our experiments when the null model is correctly specified with respect to the DGP described in the previous section. The top panel of each of these tables contains the results for path-forecasts generated when the null model is a VAR, while the bottom panel contains results when the null model is estimated with local projections (which are less efficient but correctly specified nevertheless). Each panel computes the three types of band we consider using both analytic and empirical methods (indicated with "emp." in parenthesis) under both FWE and FDR control. Hence, Tables 1-2 summarize results for a predictive sample $N = 40$ at 68% (Table 1) and 95% (Table 2) coverage; Tables 3-4 summarize the results for $N = 80$; and Tables 5-6 for $N = 200$.

In Table 1, $H = 1$ is a good benchmark case because then FWE and FDR control are the same and marginal, Bonferroni and Scheffé bands are identical by construction. Interestingly, we find that bands calculated via the empirical method are more accurate than with the analytic method. However, as the length of the path-forecast increases and initial correspondence across methods and error control measures vanishes, the coverage of the marginal bands deteriorates very rapidly, specially for inflation (labeled P), which is a very persistent variable.

As H grows to 8 or 12, the analytic approach provides more accurate coverage than the empirical approach. This is not surprising since for $H = 8$, $\widehat{\Omega}_H$ has 36 potentially distinct entries that need to be estimated, sometimes from a sample $N = 40$. Estimation with small sample sizes ($T = 100$) however, generates its own distortions on the analytic formulas that improve markedly when $T = 400$.

Direct forecasts from local projections are systematically worse than those from a VAR, as the bottom panel of Table 1 shows. This is to be expected because VAR estimates are known to be more efficient as has been previously documented in Marcellino, Stock and Watson (2006). This is specially true when the estimation sample is small ($T = 100$). The results in Table 2 for 95% coverage essentially support the same conclusions as Table 1. We also note that Bonferroni bounds appear to control FWE but fail considerably in controlling FDR; while Scheffé bands control FDR best but still maintain respectable FWE control.

The remaining tables (Tables 3-6) display the effects of higher predictive samples ($N = 80$, 200). Since N only affects the empirical estimates (the analytic estimates are based on T), we remark that coverage rates for all analytic procedures are virtually identical to those in Tables 1-2. However, as N is allowed to grow to 200, then the empirical approach appears to be most accurate in almost every case (and even when $T = 400$).

3.3 Simulation Results: Misspecified Null Model

In this section we report simulation results in which the null model estimated at each replication and then used to generate path-forecasts and confidence bands, is misspecified. Because we consider several types of misspecification, we will focus exclusively on 95% coverage and $N = 80$, which according to Tables 1-6 appears to be a sufficiently representative intermediate case on which to experiment with misspecification. Results with 68% coverage where sufficiently similar to those with 95% coverage that are omitted here for brevity but available upon request.

We explore five different types of misspecification: Table 7 examines what happens when the null model is specified as a VAR(1) instead of a VAR(4), the DGP. Table 8 examines what happens when the null model omits the variable UN from the system. Table 9 allows for a structural break in the conditional means of the DGP that is ignored in the null model. Table 10 allows for a structural break in the residual covariance matrix of the DGP but not of the null model. Table 11 allows both types of break simultaneously. Tables 7-11 are configured like Tables 1-6 for clarity.

When the forecasts are unbiased, as in the previous section, there is no need to recenter the errors prior to estimating their variance covariance matrix (Ω_H). However, misspecification can induce bias, and therefore centering the errors could be important. Yet, it turns out that is not: the simulation results indicate that the coverage rates resulting from estimating Ω_H with centered forecast errors are in general very similar to those with uncentered errors, the gains are very limited. The reason is that in most of the cases considered the biases are small, and taking their cross products makes them even smaller. A similar feature also emerges in the empirical applications of the following section. Since, in addition, testing for bias is often inconclusive in empirical applications due to the small sample size, we report results based on estimating Ω_H with uncentered forecast errors (results with centered errors are available upon request).

We begin with a direct comparison between Tables 4 and 7, where the null model is specified as VAR(1) in the latter case. There are few substantial differences, mostly because a VAR(1) captures most of the persistence in the data anyway so that path-forecasts are not substantially biased. There is, however, a slight improvement of the direct forecasts. Omitting UN from the system also has little effect, as Table 8 reveals.

Finally, we discuss the consequences of structural breaks. What we did was to revisit the parameter estimates of the Stock and Watson (2001) VAR and estimate two sets of coefficients by breaking the sample in 1984:Q4 so as to capture the well-known "great moderation" in inflation levels and output volatility (see e.g. McConnell and Pérez-Quirós, 2000). When we simulate the data for Tables 9-11 we allow for a break that splits the samples $T = 100, 400$ in such a way so as to preserve the relative sizes of the two subsamples relative to the actual data for the estimates. Table 9 uses the two sets of conditional mean parameters only; Table 10 the two sets of residual covariance estimates only; and Table 11 uses both.

While we find few differences between Tables 4 and 9 (with perhaps some visible improvements of forecasts by local projections), Tables 10 and 11 show more dramatic disparity. There is a strong upward bias in the coverage rate of the analytic Bonferroni and Scheffé bands, and the bias is even (proportionally) stronger with coverage rate 68% (results available upon request) than 95%. In this case the empirical approach provides more reliable results since estimation of the elements of Ω_H based on the predictive sample reduces substantially the bias with respect to estimation of Ω_H with the analytic formulas and VAR or local projection estimates.

Summarizing, the extensive Monte Carlo experiments we constructed in this section reveal several interesting results. First, confidence bands constructed using empirical estimates from the predictive sample are preferable to standard analytic formulas when: (1) the path-forecast is low-dimensional so that the matrix Ω_H has a small number of entries; (2) the estimation sample T is relatively small (so that the null model parameter estimates are relatively imprecise); and (3) when the residual variance of the null model is misspecified.

Second, traditional confidence bands based on the marginal distribution of the per-horizon forecast error provide very poor coverage. Since these are the bands typically reported for policy making, we think its use should be completely abandoned. Third, when the null model is correctly specified, direct forecasts based on local projections are less efficient but this disadvantage quickly disappears and turns to advantage when the null model is misspecified. Finally, Bonferroni bounds provide better control of FWE but generally poor control of FDR whereas Scheffé bands control FDR very well while maintaining reasonable FWE control.

4 Central Bank Forecasting: The Bank of England and The Federal Reserve

The Bank of England and the Federal Reserve staffs provide inflation and output forecasts to their policy-makers prior to deliberating on the future course of monetary policy. Data for the UK is available for a shorter sample and so we use the data for illustration purposes primarily. Data for the U.S. is available for much longer and hence can be used to evaluate our simultaneous confidence regions by splitting the predictive sample and saving the second subsample for evaluation. Notice that while the forecasts and the realizations are available, the null models are not and hence one has to rely entirely on the empirical approach we have proposed. Finally, we also assess the impact of using real-time data rather than only final vintage data since this is an important element when evaluating forecast uncertainty in a decision making context.

4.1 Data

The source of quarterly time series for UK output and prices and their forecasts is the Bank of England. Growth and inflation are measured as $100 * ln(y_t/y_{t-4})$, where y is either output or a price index. The predictive sample starts in 1998:Q1 and the maximum forecast horizon is $H = 9$. However, $h = 1$ is actually a nowcast, so that the path-forecasts effectively cover a period of two year-ahead forecasts. These forecasts are conditional on the market interest rates rather than conditional on a constant path for interest rates, the other guise under which forecasts are reported by the Bank of England. Inflation forecasts refer to the RPIX index (retail prices excluding mortgage interest payments) until 2003:Q4 and to the CPI thereafter. Output forecasts always refer to GDP. As final data against whose realizations we can compare the forecasts to, we use the currently available CPI and RPIX series, and chain-weighted GDP. The CPI sample ends in 2009:Q2. The GDP sample ends in 2008:Q2, since this is the last quarter for which real-time data were available as well.

Regarding the real-time data, CPI and RPIX are rarely and only marginally revised, and as far as we know they are not available in real time. For GDP, there are up to three vintages per quarter, one per month. We have used the vintages from February, May, August and November since these are available for each quarter in our sample, and we have collected the first 10 releases for each quarter (in addition to the final data, coinciding with the latest available vintage).

For the US, the data source is the Federal Reserve, with the forecasts coming from the Greenbook. Specifically, we have collected quarterly forecasts starting in 1974:Q2, for which $H = 5$ is available. As for the UK, $h = 1$ represents a nowcast, so that for the US we have path-forecasts up to 1-year ahead. The sample ends in 2003:Q4, since the Greenbook data are released with a delay of 5 years. Forecasts are for output growth (where output is measured by GNP until 1991, GDP afterwards) and inflation (measured as growth in the GNP deflator until 1991, in the GDP deflator afterwards) measured as $400 * ln(y_t/y_{t-1})$.

As final data against whose realizations we can compare the forecasts to, we use the latest available vintage of chain-weighted GNP and GDP, and of their deflators. For the real time data, we combine information from the Federal Reserve Bank of Philadelphia's real-time database and from the Alfred database maintained by the Federal Reserve Bank of St Louis.

4.2 Alternative measures of Path-forecast Uncertainty

Using forecasts and corresponding either final vintage or real time data, we have constructed sequences of path-forecast errors and tested them for normality using the Bai and Ng's (2005) statistic, obtaining non-rejection of normality in the vast majority of cases (detailed results are available upon request). We have then used the error sequences to estimate the Ω_H matrix empirically on the predictive sample. Of course, the analytic approach is unavailable since the null models that produce the forecasts are unknown.

Next, we construct marginal, Bonferroni and Scheffé bands for UK and US growth and inflation forecasts, for different nominal coverage rates. In Figure 1 we plot the (one-sided) 68% bands, centered for convenience on zero. It clearly emerges that the commonly used marginal bands can substantially under-estimate the uncertainty, in particular for large forecast horizons. The Bonferroni bands are in general wider than the Scheffé bands for low values of h, but sometimes narrower for higher values of h. A similar picture (but with larger values for each type of band) emerges for a 95% coverage rate.

Due to the extent of data revisions for output, which were large for both the UK and the US in the first part of the respective samples, the forecast errors computed with the real time growth data can be fairly different from those resulting from the final vintage data. Hence, estimates of $\mathbf{\Omega}_H$ can be also different, and as a consequence the measures of path-forecast uncertainty when computed in real time.

Figures 2 and 3 graph the UK and US path growth forecasts respectively, and the three types of bands (marginal, Bonferroni and Scheffé) when computed with each of the ten vintages of data. It turns out that for the UK the uncertainty is smaller with the real time data than with the final data. In particular, rather systematically across forecast horizons and types of band, the uncertainty increases mildly for the first 2-3 releases, remains fairly stable up to release 10, and then increases for the final (available) vintage. A similar pattern emerges for the US, where however the differences between the 10^{th} and the final releases are much smaller than for the UK, likely due to the longer sample available such that particular episodes of large revisions are averaged out.

Overall, these results indicate that the measures of path-forecast uncertainty can be fairly different in empirical applications, with the common marginal bands being the narrowest. In addition, all measures can underestimate uncertainty when computed in real time, if the forecast target is the final value of the variable of interest.

4.3 Coverage rates

The results in the previous subsection are interesting, but they do not indicate which of the three measures is the most reliable in terms of actual coverage rates. On the basis of the Monte Carlo results in Section 3, we expect the marginal bands to perform badly, but the ranking of the Bonferroni and Scheffé bands is uncertain because it also depends on what error control one is interested in. Hence, we now assess their relative performance in the application at hand.

We focus on the US, for which longer time series are available. If T denotes the full sample, then we split the sample at $T^* = 1985$:Q1. With the sample of forecast paths starting in 1974:Q2 and ending in T^*, we get 40 forecast paths and corresponding errors from which we construct an empirical estimate of Ω_H. This estimate is used to construct uncertainty bands for the forecast path from $T^* + 1$, i.e. from 1985:Q2. Then we roll the window of forecast paths for estimating Ω_H one observation at a time by deleting the first path and adding one path at the end, and we also roll the forecast path for which we construct uncertainty bands. We do so until the uncertainty bands are constructed for the final path starting in $T =$ 2003:Q4, for which the final window of forecast paths starting from 1992:Q4 to 2002:Q3 is used to estimate Ω_H. In this way, we have produced uncertainty bands for 75 forecast paths. The rolling window procedure described is repeated for every vintage of data $r = 1, ..., 10$ and for the final data.

This procedure produces a set of measures of path-forecast uncertainty that can be compared with the actual realizations in order to compute the actual coverage rate of each type of confidence band. Table 12 reports actual coverage rates under FWE and FDR control for U.S. inflation and output growth for nominal levels 50%, 68% and 95%, and for all data vintages. We consider the nominal level of 50% because of the rather small size of the evaluation sample.

As expected, the marginal bands perform very poorly, in particular for 50% and 68%

nominal coverage rates. For the same rates, the actual coverage of the Bonferroni bands is systematically higher than the nominal level; and the Scheffé bands appear to perform best throughout at FDR control. For 95% nominal coverage, Bonferroni and Scheffé bands mostly yield similar results.

For output growth and FDR control, the coverage of the Scheffé bands is close to nominal for all nominal coverage rates and all data vintages. For nominal coverage of 50% and 68% their coverage is far closer to the nominal level than that of the Bonferroni and the marginal bands.

Regarding FWE control, the Scheffé bands' coverage is smaller than the nominal level, but again far closer than the coverage of the marginal bands. The Bonferroni bands almost always cover more than the nominal level, but are mostly closer to it than the Scheffé bands.

For inflation and FDR control, the coverage of the Scheffé bands exceeds the nominal level for coverage rates of 50% and 68% and all data vintages, but only moderately so, and considerably less than the Bonferroni bands for vintages 1 to 10. Only for final release data, Bonferroni and Scheffé bands produce similar results. The marginal bands cover less than the nominal level, in some cases extremely so.

In terms of FWE control for inflation, the Scheffé bands' coverage is too small for nominal levels of 50% and 68%, but nevertheless they produce clearly the best results of all bands considered for 50% coverage. The Bonferroni bands consistently cover more than the nominal levels of 50% and 68%. For 68% coverage, the coverage of the Bonferroni bands can be closer to the nominal level than the Scheffé bands, depending on the data vintage. The marginal bands perform worst for all vintages and coverage rates.

In summary, overall the Scheffé bands yield the best results in terms of FDR control and reasonable results for FWE control, while the Bonferroni bands can be superior for FWE control, but perform clearly worse with FDR control. Moreover, the coverage of the marginal bands is in general quite far from the nominal level for both types of error control. The empirical results are in line with those found in the Monte Carlo simulations, and confirm a certain robustness in the coverage of the bands also in the likely presence of structural changes in distribution of the forecast errors.

5 Conclusions

This paper proposes a number of practical solutions to the problem of calculating and then displaying the uncertainty associated with a path-forecast. Several features make this an unusual statistical problem.

First, the forecast generating distribution is unknowable. Large sample arguments approximate the distribution of the parameters of the null model, and hence the conditional mean path-forecast. But asymptotics do not reveal the data generating distribution from which the actual realizations of the path-forecast will come. Moreover, sometimes the forecast generating distribution itself is unavailable, such as when agencies publish forecasts but not how they were generated. Even conventional use of resampling techniques is complicated or even made infeasible by such situations. In this paper we show that empirical methods based on the predictive sample provide a natural solution.

However, the empirical approach cannot be directly implemented because confidence regions for path-forecasts are a multiple comparison problem for which no unique equivalent to control of Type I error exists. Further, we argue that control of family-wise error (the closest relative to Type I error control) is inadequate for path-forecasts: should an inflation, two-year path-forecast be rejected because the prediction of one of the periods is likely to be erroneous? We argue instead in favor of evaluating the forecasts in the path jointly, and hence control the false discovery rate, a more contemporary form of error control.

Confidence regions constructed on the basis of false discovery rate control solve the problem of simultaneous evaluation of outcomes but result in multidimensional geometric objects that cannot be represented in two-dimensional space. Therefore, we show how to construct approximate rectangular regions with approximate false discovery rate control that account for the serial correlation among the elements of a path-forecast. We call the resulting confidence bands Scheffé bands because the rectangular approximation is based on Scheffé's (1953) S-method, and show how they can be calculated with the empirical approach that we propose.

Simulation evidence and our applications suggest that traditional confidence bands based on the marginal distribution of each forecast in the path provide no reliable control of either family-wise error or false discovery rate. Coverage is often off and by large amounts, prompting us to recommend that its use be discontinued. In contrast, Scheffé bands give accurate false discovery rate control and relatively good family-wise error control even when compared to Bonferroni bounds, which specifically control family-wise error but often result in very imprecise false discovery rate control.

References

Bai, Jushan and Serena Ng (2005) "Tests for Skewness, Kurtosis and Normality for Time Series Data," *Journal of Business and Economic Statistics*, 23: 49-60.

Benjamini, Yoav and Yosef Hochberg (1995) "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing," *Journal of the Royal Statistical Society, Series B* 57(1): 289-300.

Clements, Michael P. and David F. Hendry (1999) **Forecasting Economic Time Series**. Cambridge University Press: Cambridge, U.K.

Clements, Michael P. and Nick Taylor (2001) "Bootstrapping Prediction Intervals for Autoregressive Models," *International Journal of Forecasting,* 17: 247-267.

Conze and Viswanathan, Antoine (1991) "Path Dependent Options: The Case of Lookback Options," *Journal of Finance,* 46(5): 1893-1907.

Dudoit, Sandrine and Mark J. van der Laan (2008) **Multiple Testing Procedures with Applications to Genomics.** Springer Series in Statistics, Springer: New York, New York.

Efron, Bradley (1979) "Bootstrap Methods: Another Look at the Jacknife," *The Annals of Statistics*, 7(1): 1-26.

Goldman, M. Barry, Howard B. Sosin and Mary Ann Gatto (1979) "Path Dependent Options: "Buy at the Low and Sell at the High"" *Journal of Finance,* 34(5): 1111-1127.

Holm, Sture (1979) "A Simple Sequentially Rejective Multiple Test Procedure," *Scandinavian Journal of Statistics,* 6: 65-70.

Jordà, Òscar and Massimiliano G. Marcellino (2009) "Path-Forecast Evaluation," U.C. Davis working paper 08-5 (forthcoming in *Journal of Applied Econometrics*).

Kim, Jae H. (1999) "Asymptotic and Bootstrap Prediction Regions for Vector Autoregression," *International Journal of Forecasting*, 15: 393-403.

Kwok, Yue Kuen and Ka Wo Lau (2001) "Pricing Algorithms for Options with Exotic Path-Dependence," *Journal of Derivatives,* 9(1): 28-38.

Marcellino, Massimiliano G., James H. Stock and Mark W. Watson (2006) "A Comparison of Direct and Iterated Multistep AR Methods for Forecasting Macroeconomic Time Series," *Journal of Applied Econometrics,* 135(1-2): 499-526.

McConnell, Margaret M. and Gabriel Pérez Quirós (2000) "Output Fluctuations in the United States: What Has Changed Since the Early 1980's?" *American Economic Review,* 90(5): 1464-1476.

Mahalanobis, Prasanta C. (1936) "On the Generalized Distance in Statistics," *Procedures of the National Institute of Sciences of India*, 2(1): 49-55.

Masarotto, Guido (1990) “Bootstrap Prediction Intervals for Autoregressions,” *International Journal of Forecasting,* 6: 229-239.

Pascual, Lorenzo, Juan Romo and Esther Ruiz (2004) “Bootstrap Predictive Inference for ARIMA Processes,” *Journal of Time Series Analysis,* 4(7): 449-465.

Politis, Dimitris N., Joseph P. Romano and Michael Wolf (1999) **Subsampling**. Springer Series in Statistics, Springer: New York, New York.

Rubin, Daniel, Sandrine Dudoit and Mark J. can der Laan (2006) “A Method to Increase the Power of Multiple Testing Procedures Through Sample Splitting,” *Statistical Applications in Genetics and Molecular Biology,* 5(1): Art. 19.

Scheffé, Henry (1953) “A Method for Judging All Contrasts in the Analysis of Variance,” *Biometrika*, 40(1-2): 87-110.

Stock, James H. and Mark W. Watson (2001) “Vector Autoregressions,” *Journal of Economic Perspectives,* 15(4): 101-115.

Svensson, Lars E. O. (2009) “Evaluating Monetary Policy,” National Bureau of Economic Research working paper 15385.

Thombs, Lori A. and William R. Schucany (1990) “Bootstrap Prediction Intervals for Autoregression,” *Journal of the American Statistical Association*, 85: 486-492.

Williams, W. H. and M. L. Goodman (1971) “A Simple Method for the Construction of Empirical Confidence Limits for Economic Forecasts,” *Journal of the American Statistical Association,* 66(336): 752-754.

Table 1. MC results, well specified model, nominal coverage 68%, N=40

VAR

		FWE control						FDR control					
		Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.	Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.
H:	1												
T=100	P	62.2	66.7	62.2	66.7	62.2	66.7	62.2	66.7	62.2	66.7	62.2	66.7
	UN	63.4	67.5	63.4	67.5	63.4	67.5	63.4	67.5	63.4	67.5	63.4	67.5
	FF	62.3	66.4	62.3	66.4	62.3	66.4	62.3	66.4	62.3	66.4	62.3	66.4
T=400	P	66.8	67.1	66.8	67.1	66.8	67.1	66.8	67.1	66.8	67.1	66.8	67.1
	UN	66.9	67.0	66.9	67.0	66.9	67.0	66.9	67.0	66.9	67.0	66.9	67.0
	FF	67.1	67.0	67.1	67.0	67.1	67.0	67.1	67.0	67.1	67.0	67.1	67.0
H:	4												
T=100	P	24.2	30.2	67.4	74.4	48.3	55.0	16.9	18.3	63.9	65.7	58.5	62.5
	UN	37.0	43.1	74.9	80.3	57.8	63.0	12.2	13.7	52.9	55.4	59.4	63.3
	FF	30.5	36.8	71.6	78.0	54.8	60.9	12.6	13.8	56.3	58.6	58.6	62.5
T=400	P	27.7	28.6	74.6	73.7	54.5	53.7	20.2	20.7	71.9	69.0	65.7	62.4
	UN	42.1	42.6	81.2	80.2	63.5	63.4	14.2	14.0	59.7	57.2	66.6	63.6
	FF	35.7	35.7	78.2	77.2	60.9	60.3	14.8	15.0	62.6	60.0	65.2	62.3
H:	8												
T=100	P	11.1	17.0	69.5	76.5	46.3	52.8	0.9	1.2	42.5	41.1	54.9	53.2
	UN	21.0	26.1	76.1	80.7	55.4	60.0	1.4	1.7	48.7	46.4	56.3	52.2
	FF	18.1	24.3	73.1	78.6	53.0	58.7	1.0	1.2	40.6	38.8	56.2	52.6
T=400	P	13.6	14.2	78.3	75.8	53.7	51.5	1.5	1.6	52.6	47.1	65.4	53.9
	UN	25.1	25.4	83.6	81.1	61.3	59.5	2.0	2.3	58.2	50.6	65.7	54.3
	FF	20.6	21.3	82.1	78.8	59.0	57.5	1.1	1.4	49.0	42.1	65.1	53.2
H:	12												
T=100	P	6.7	11.9	70.9	75.8	46.5	51.6	0.0	0.1	28.7	25.1	54.4	43.1
	UN	12.8	18.1	75.8	78.7	52.8	53.6	0.3	1.5	56.0	51.6	54.9	42.3
	FF	11.0	16.7	73.2	78.3	52.4	56.7	0.0	0.1	30.1	27.0	54.3	41.9
T=400	P	6.8	8.3	79.8	75.2	52.9	50.1	0.0	0.1	35.6	29.7	64.4	43.1
	UN	15.8	16.6	84.3	79.7	60.6	54.3	0.5	1.3	67.8	54.7	65.6	43.1
	FF	13.8	14.3	83.2	78.5	59.8	56.8	0.1	0.2	37.5	31.2	64.7	42.6

Local Projections

		FWE control						FDR control					
		Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.	Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.
H:	1												
T=100	P	62.4	66.7	62.4	66.7	62.4	66.7	62.4	66.7	62.4	66.7	62.4	66.7
	UN	63.7	67.5	63.7	67.5	63.7	67.5	63.7	67.5	63.7	67.5	63.7	67.5
	FF	62.6	66.4	62.6	66.4	62.6	66.4	62.6	66.4	62.6	66.4	62.6	66.4
T=400	P	66.9	67.1	66.9	67.1	66.9	67.1	66.9	67.1	66.9	67.1	66.9	67.1
	UN	66.9	67.0	66.9	67.0	66.9	67.0	66.9	67.0	66.9	67.0	66.9	67.0
	FF	67.2	67.0	67.2	67.0	67.2	67.0	67.2	67.0	67.2	67.0	67.2	67.0
H:	4												
T=100	P	21.1	30.8	62.2	74.0	44.0	55.1	17.7	18.3	65.0	65.6	55.0	62.4
	UN	33.6	43.2	70.6	79.9	56.0	63.2	11.3	13.3	51.1	55.4	56.3	63.0
	FF	26.8	37.0	66.9	78.1	52.2	60.9	12.2	14.1	55.2	58.7	55.3	62.5
T=400	P	27.1	28.5	73.5	73.6	53.6	53.6	20.5	20.6	72.2	69.0	65.1	62.6
	UN	41.0	42.4	80.7	80.2	63.2	63.4	13.9	14.0	59.5	57.1	65.7	63.3
	FF	35.0	35.9	77.4	77.3	60.5	60.2	14.8	15.3	62.5	59.8	64.7	61.9
H:	8												
T=100	P	6.7	17.1	56.6	76.6	37.3	52.8	1.2	1.2	45.1	42.1	45.4	52.7
	UN	14.9	26.8	64.4	80.6	48.8	60.0	1.6	1.9	48.8	46.3	45.6	52.0
	FF	12.3	24.7	61.2	78.9	46.9	58.8	0.8	1.3	38.9	39.5	44.8	52.2
T=400	P	12.1	14.3	76.3	75.6	52.7	51.6	1.7	1.8	53.4	47.3	63.8	53.8
	UN	23.9	25.5	82.2	81.4	60.8	59.6	2.1	2.3	58.2	50.6	64.2	54.2
	FF	19.3	21.4	80.4	78.8	58.5	57.4	1.0	1.3	48.6	42.4	63.8	53.4
H:	12												
T=100	P	2.4	12.3	51.0	76.0	33.8	51.7	0.1	0.1	32.3	26.8	35.9	42.2
	UN	7.0	18.3	59.1	79.4	34.4	51.9	1.2	2.6	62.3	52.0	36.8	41.1
	FF	5.3	17.1	54.1	78.5	40.7	56.4	0.2	0.4	32.1	28.3	35.7	40.7
T=400	P	5.9	8.3	76.9	75.1	51.8	50.0	0.0	0.1	36.2	29.8	61.2	42.9
	UN	14.4	16.7	81.8	79.7	59.0	54.1	0.5	1.4	69.1	54.6	61.8	43.1
	FF	12.0	14.0	80.6	78.8	59.3	56.4	0.1	0.1	38.0	31.2	61.7	42.5

Notes: 10,000 samples generated from VAR(4) for three variables (P, UN, FF) with stable parameters. Model coincides with DGP. Each estimated VAR or local projection (LP) on these 10,000 samples generates a forecast error variance (which includes estimation uncertainty) for the forecast path of length h, and hence the sets of bands (marginal, Bonferroni, and Scheffé) used in the analysis. Similarly, each estimated model generates N forecast paths whose associated error paths are used to generate a forecast error variance for the forecast path and hence the set of bands (marginal emp, Bonferroni emp, and Scheffé emp). Hence 10,000 actual paths are then compared with each set of 10,000 bands to determine the appropriate coverage rates. FWE control stands for "family-wise error control" and simply computes the proportion of paths strictly inside the bands. FDR control instead is the proportion of forecast paths whose Mahalanobis distance attains a value that is lower than the chi-square statistic for probability equal to nominal coverage and degrees of freedom equal to H. See text for more details.

Table 2. MC results, well specified model, nominal coverage 95%, N=40

VAR

		FWE control						FDR control					
		Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.	Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.
H:	1												
T=100	P	92.0	94.4	92.0	94.4	92.0	94.4	92.0	94.4	92.0	94.4	92.0	94.4
	UN	92.4	94.4	92.4	94.4	92.4	94.4	92.4	94.4	92.4	94.4	92.4	94.4
	FF	92.2	94.4	92.2	94.4	92.2	94.4	92.2	94.4	92.2	94.4	92.2	94.4
T=400	P	94.7	94.4	94.7	94.4	94.7	94.4	94.7	94.4	94.7	94.4	94.7	94.4
	UN	94.8	94.6	94.8	94.6	94.8	94.6	94.8	94.6	94.8	94.6	94.8	94.6
	FF	94.3	94.4	94.3	94.4	94.3	94.4	94.3	94.4	94.3	94.4	94.3	94.4
H:	4												
T=100	P	75.9	81.7	90.5	93.8	86.7	90.3	74.5	76.0	92.1	92.0	92.1	92.7
	UN	81.9	86.3	92.8	95.4	90.3	93.1	64.0	66.5	85.8	86.7	90.5	90.8
	FF	79.2	84.5	91.7	94.5	89.4	92.5	67.3	69.4	87.8	88.9	90.8	91.4
T=400	P	82.8	81.5	95.0	93.9	91.8	90.0	82.1	79.2	96.3	93.9	96.2	93.6
	UN	87.6	86.7	96.3	95.5	93.7	93.1	70.8	67.8	90.8	87.0	94.4	90.7
	FF	86.5	85.2	96.2	94.9	93.7	92.7	75.3	71.8	93.4	90.3	95.3	92.4
H:	8												
T=100	P	65.2	72.0	89.5	92.5	85.5	88.7	37.1	36.1	81.0	75.9	89.5	84.3
	UN	73.1	77.8	92.7	93.7	89.5	91.4	42.6	41.4	85.9	79.9	91.9	86.5
	FF	69.9	76.1	91.1	93.1	88.3	90.9	34.9	33.7	79.9	74.4	90.4	85.2
T=400	P	74.3	71.2	95.3	92.7	91.5	88.7	45.3	40.0	90.2	80.7	95.9	86.8
	UN	81.0	78.1	97.1	94.5	93.7	92.0	51.3	45.1	92.5	83.7	96.7	88.4
	FF	79.4	76.1	96.3	93.7	93.2	91.7	41.8	37.2	87.2	77.3	95.8	86.9
H:	12												
T=100	P	58.6	66.4	89.3	91.2	85.8	87.7	14.9	14.1	70.5	57.6	89.3	74.4
	UN	65.1	69.8	91.9	92.2	87.1	83.9	36.3	37.3	89.6	79.0	92.6	80.4
	FF	62.1	69.0	90.1	92.2	87.6	89.3	16.6	15.7	71.5	60.5	90.2	77.4
T=400	P	68.7	64.3	95.4	91.4	91.5	88.0	18.8	17.7	80.9	65.0	96.0	78.0
	UN	74.8	71.2	96.6	93.4	93.3	85.3	46.3	40.5	96.4	83.6	97.4	83.2
	FF	73.1	68.7	96.5	92.4	92.7	89.9	20.1	18.5	82.5	66.8	96.6	79.7

Local Projections

		FWE control						FDR control					
		Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.	Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.
H:	1												
T=100	P	92.2	94.4	92.2	94.4	92.2	94.4	92.2	94.4	92.2	94.4	92.2	94.4
	UN	92.5	94.4	92.5	94.4	92.5	94.4	92.5	94.4	92.5	94.4	92.5	94.4
	FF	92.3	94.4	92.3	94.4	92.3	94.4	92.3	94.4	92.3	94.4	92.3	94.4
T=400	P	94.7	94.4	94.7	94.4	94.7	94.4	94.7	94.4	94.7	94.4	94.7	94.4
	UN	94.8	94.6	94.8	94.6	94.8	94.6	94.8	94.6	94.8	94.6	94.8	94.6
	FF	94.4	94.4	94.4	94.4	94.4	94.4	94.4	94.4	94.4	94.4	94.4	94.4
H:	4												
T=100	P	71.0	81.5	86.9	93.7	82.9	90.2	75.5	76.1	92.3	92.1	90.4	92.5
	UN	77.7	86.2	89.7	95.1	89.0	93.1	63.1	66.7	83.9	86.7	88.4	90.6
	FF	74.5	84.7	88.5	94.3	87.4	92.4	66.1	69.6	86.4	88.7	88.8	91.4
T=400	P	81.9	81.3	94.6	93.9	91.6	90.0	82.3	79.0	96.4	94.0	96.1	93.6
	UN	87.3	86.7	96.1	95.4	93.7	93.1	70.7	67.8	90.8	87.0	94.3	90.6
	FF	85.8	85.0	95.9	94.9	93.6	92.6	75.1	72.1	93.3	90.3	95.1	92.6
H:	8												
T=100	P	52.7	71.9	80.0	92.6	77.2	88.5	39.5	36.9	80.8	75.9	83.0	84.0
	UN	61.2	77.6	84.5	93.6	83.9	91.2	43.7	41.5	84.6	79.6	85.9	86.3
	FF	58.4	76.4	81.9	93.3	82.7	90.8	34.8	34.0	76.4	74.4	83.6	85.0
T=400	P	72.2	71.3	94.4	92.9	91.1	88.6	45.8	40.1	90.6	81.1	95.4	86.7
	UN	79.2	78.3	96.5	94.4	93.6	91.9	51.5	45.1	92.7	83.8	96.2	88.4
	FF	77.8	76.2	95.6	93.8	93.1	91.7	41.7	36.9	87.1	77.3	95.4	86.8
H:	12												
T=100	P	39.7	66.2	73.1	91.6	71.8	87.6	18.5	15.0	68.4	58.3	74.6	73.7
	UN	47.2	69.7	78.9	92.0	64.7	81.9	44.7	38.5	89.8	79.1	80.7	79.6
	FF	43.0	69.2	74.7	92.1	74.5	88.7	19.5	17.5	69.2	61.4	76.6	75.8
T=400	P	65.5	64.8	93.9	91.7	91.2	88.0	19.5	17.9	80.8	65.4	95.0	77.8
	UN	71.6	70.8	95.7	93.4	92.3	85.0	48.7	40.9	96.5	83.8	96.6	83.1
	FF	70.2	68.7	95.3	92.6	92.4	89.8	20.0	18.6	82.4	66.6	95.6	79.4

Notes: See notes to Table 1.

Table 3. MC results, well specified model, nominal coverage 68%, N=80

VAR

		FWE control						FDR control					
		Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.	Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.
H:	1												
T=100	P	62.9	67.3	62.9	67.3	62.9	67.3	62.9	67.3	62.9	67.3	62.9	67.3
	UN	62.4	66.8	62.4	66.8	62.4	66.8	62.4	66.8	62.4	66.8	62.4	66.8
	FF	63.3	68.0	63.3	68.0	63.3	68.0	63.3	68.0	63.3	68.0	63.3	68.0
T=400	P	67.4	68.4	67.4	68.4	67.4	68.4	67.4	68.4	67.4	68.4	67.4	68.4
	UN	67.2	67.9	67.2	67.9	67.2	67.9	67.2	67.9	67.2	67.9	67.2	67.9
	FF	67.6	68.2	67.6	68.2	67.6	68.2	67.6	68.2	67.6	68.2	67.6	68.2
H:	4												
T=100	P	23.6	31.0	67.4	76.1	47.6	56.5	16.2	17.8	63.9	67.2	57.9	64.7
	UN	36.9	43.7	75.0	82.0	57.6	63.8	11.8	13.4	51.8	56.4	58.9	64.6
	FF	31.4	39.5	72.0	79.9	55.0	62.3	12.8	15.1	56.5	60.5	59.0	65.2
T=400	P	27.9	28.9	75.1	75.5	55.1	55.5	19.9	20.3	72.1	71.4	66.2	65.7
	UN	40.8	41.7	80.7	80.5	62.7	63.1	13.6	13.9	58.0	57.4	64.8	64.2
	FF	35.1	36.1	78.3	78.5	59.9	60.3	15.6	15.6	62.6	61.4	65.0	64.1
H:	8												
T=100	P	10.8	17.7	69.5	79.6	46.3	55.6	1.1	1.1	43.5	43.9	55.1	60.8
	UN	21.2	28.6	76.2	83.4	56.0	62.9	1.5	1.5	49.8	49.5	57.2	60.5
	FF	17.5	25.3	73.9	82.3	53.8	61.4	0.8	0.8	40.6	41.0	55.4	59.5
T=400	P	13.2	14.1	78.1	77.9	53.6	53.9	1.4	1.6	53.3	50.8	65.9	61.2
	UN	25.3	26.5	84.0	83.2	62.4	62.1	1.7	1.9	58.2	54.7	65.9	61.0
	FF	20.9	22.2	82.8	82.1	59.9	60.2	1.1	1.1	48.2	45.8	65.5	60.2
H:	12												
T=100	P	6.6	11.9	70.4	79.9	46.3	54.8	0.1	0.0	29.0	25.6	53.8	54.4
	UN	13.6	19.3	76.2	82.9	53.9	59.1	0.4	0.9	57.4	57.4	56.0	55.0
	FF	10.8	17.9	74.1	83.3	52.6	60.5	0.1	0.1	31.5	29.6	54.9	54.8
T=400	P	7.5	8.7	80.0	78.8	53.2	52.6	0.1	0.1	36.4	32.7	65.3	54.8
	UN	15.9	16.7	84.6	83.2	61.0	59.5	0.4	0.7	68.2	60.6	65.9	55.0
	FF	14.5	15.4	83.8	82.7	61.1	60.8	0.0	0.0	38.2	34.6	65.5	55.0

Local Projections

		FWE control						FDR control					
		Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.	Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.
H:	1												
T=100	P	63.1	67.3	63.1	67.3	63.1	67.3	63.1	67.3	63.1	67.3	63.1	67.3
	UN	62.6	66.8	62.6	66.8	62.6	66.8	62.6	66.8	62.6	66.8	62.6	66.8
	FF	63.5	68.0	63.5	68.0	63.5	68.0	63.5	68.0	63.5	68.0	63.5	68.0
T=400	P	67.5	68.4	67.5	68.4	67.5	68.4	67.5	68.4	67.5	68.4	67.5	68.4
	UN	67.3	67.9	67.3	67.9	67.3	67.9	67.3	67.9	67.3	67.9	67.3	67.9
	FF	67.6	68.2	67.6	68.2	67.6	68.2	67.6	68.2	67.6	68.2	67.6	68.2
H:	4												
T=100	P	20.6	30.8	62.6	76.3	43.4	56.6	16.9	18.1	65.2	67.7	54.7	65.3
	UN	33.2	43.4	70.9	82.0	55.5	63.8	11.5	13.7	50.5	56.7	55.8	64.8
	FF	27.8	39.6	67.6	79.7	52.6	62.2	12.9	15.1	55.2	60.5	55.5	65.2
T=400	P	27.6	28.9	74.3	75.5	54.2	55.4	20.1	20.1	72.5	71.5	65.7	65.8
	UN	40.4	41.8	80.1	80.6	62.4	63.2	13.8	13.8	57.7	57.2	64.6	64.2
	FF	34.6	36.0	77.6	78.5	59.4	60.1	15.6	16.0	62.3	61.6	64.5	64.0
H:	8												
T=100	P	6.6	18.1	57.0	79.0	38.4	55.3	1.4	1.1	45.1	44.2	45.4	59.9
	UN	15.1	28.3	65.2	83.5	50.0	62.7	1.8	1.8	50.6	49.7	47.0	60.0
	FF	11.8	25.6	62.8	82.8	47.7	61.6	1.1	1.0	39.7	41.8	45.9	59.7
T=400	P	11.9	14.3	76.0	78.0	52.6	53.8	1.4	1.6	54.0	50.8	64.2	61.1
	UN	24.2	26.8	82.5	83.0	61.8	62.3	1.7	1.9	58.1	54.5	64.2	60.9
	FF	19.7	22.6	80.9	81.9	59.3	60.3	1.1	1.2	48.2	45.7	63.7	60.4
H:	12												
T=100	P	2.3	12.4	51.0	80.2	34.6	55.3	0.2	0.1	30.9	27.9	35.7	53.7
	UN	7.4	19.4	59.8	83.0	36.2	58.1	1.2	1.6	63.4	57.5	37.7	54.6
	FF	5.4	18.4	55.0	83.0	41.0	60.2	0.2	0.2	32.7	30.6	35.7	54.1
T=400	P	6.3	8.8	77.0	78.9	51.9	52.5	0.1	0.2	37.3	33.1	62.4	54.9
	UN	14.4	16.6	82.4	83.2	59.4	59.5	0.4	0.8	69.8	60.7	62.5	55.0
	FF	12.6	15.6	81.2	82.7	60.2	60.8	0.1	0.1	38.2	34.3	61.8	54.7

Notes: See notes to Table 1.

Table 4. MC results, well specified model, nominal coverage 95%, N=80

VAR

		FWE control						FDR control					
		Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.	Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.
H:	1												
T=100	P	91.8	94.6	91.8	94.6	91.8	94.6	91.8	94.6	91.8	94.6	91.8	94.6
	UN	91.8	94.6	91.8	94.6	91.8	94.6	91.8	94.6	91.8	94.6	91.8	94.6
	FF	91.6	94.5	91.6	94.5	91.6	94.5	91.6	94.5	91.6	94.5	91.6	94.5
T=400	P	94.6	94.8	94.6	94.8	94.6	94.8	94.6	94.8	94.6	94.8	94.6	94.8
	UN	94.9	95.0	94.9	95.0	94.9	95.0	94.9	95.0	94.9	95.0	94.9	95.0
	FF	94.4	94.6	94.4	94.6	94.4	94.6	94.4	94.6	94.4	94.6	94.4	94.6
H:	4												
T=100	P	75.9	83.5	90.2	94.6	86.8	91.6	74.5	77.2	92.2	93.3	92.1	94.3
	UN	81.3	87.7	92.6	95.9	89.8	93.3	64.1	69.0	86.2	88.7	90.6	92.9
	FF	78.1	85.3	91.1	95.4	88.6	93.0	66.0	70.5	88.2	90.3	91.0	93.4
T=400	P	82.8	82.9	94.8	94.6	91.9	91.6	81.5	80.3	96.1	94.8	95.9	94.8
	UN	87.5	87.5	96.2	95.9	93.7	93.6	71.0	69.8	90.8	89.6	94.5	93.2
	FF	86.2	86.3	95.8	95.9	93.7	93.6	75.6	74.4	93.3	92.1	95.4	94.6
H:	8												
T=100	P	65.2	75.5	89.9	94.4	85.6	90.9	36.6	37.2	81.8	80.6	90.0	90.4
	UN	72.1	80.2	92.2	95.8	89.5	93.1	42.9	43.5	85.8	84.7	91.8	92.1
	FF	70.2	79.6	91.0	95.5	88.4	93.2	34.6	35.2	80.1	79.2	90.2	91.5
T=400	P	74.0	73.6	95.1	94.3	91.2	90.5	45.7	43.4	89.5	86.0	95.7	92.0
	UN	80.4	79.7	96.1	95.6	93.4	93.3	50.6	48.3	92.3	87.5	96.3	92.7
	FF	79.0	78.5	96.3	95.4	93.5	93.3	42.8	40.2	88.2	83.5	96.0	92.5
H:	12												
T=100	P	58.4	69.6	89.0	94.2	84.7	90.6	15.0	13.8	69.3	64.4	88.7	86.6
	UN	65.6	74.2	92.1	94.9	87.8	90.2	37.0	39.7	89.2	86.7	92.8	91.0
	FF	62.8	73.0	90.5	94.3	88.2	92.2	16.1	15.6	71.8	67.0	90.8	88.0
T=400	P	68.7	67.9	95.5	94.2	91.4	90.7	18.7	17.9	80.8	72.8	95.6	88.7
	UN	74.7	73.1	96.4	95.2	92.9	90.5	45.5	42.5	96.2	90.5	97.5	92.1
	FF	72.9	72.2	96.0	94.6	92.9	92.4	19.6	18.7	81.9	73.2	96.3	89.9

Local Projections

		FWE control						FDR control					
		Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.	Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.
H:	1												
T=100	P	92.0	94.6	92.0	94.6	92.0	94.6	92.0	94.6	92.0	94.6	92.0	94.6
	UN	92.0	94.6	92.0	94.6	92.0	94.6	92.0	94.6	92.0	94.6	92.0	94.6
	FF	91.7	94.5	91.7	94.5	91.7	94.5	91.7	94.5	91.7	94.5	91.7	94.5
T=400	P	94.7	94.8	94.7	94.8	94.7	94.8	94.7	94.8	94.7	94.8	94.7	94.8
	UN	95.0	95.0	95.0	95.0	95.0	95.0	95.0	95.0	95.0	95.0	95.0	95.0
	FF	94.4	94.6	94.4	94.6	94.4	94.6	94.4	94.6	94.4	94.6	94.4	94.6
H:	4												
T=100	P	71.6	83.2	87.0	94.5	82.6	91.5	75.6	77.5	92.1	93.3	90.2	94.2
	UN	77.7	87.7	89.8	95.8	88.8	93.3	62.2	69.2	84.4	88.9	88.5	92.7
	FF	73.7	85.5	88.1	95.3	86.4	93.0	65.1	70.8	86.7	90.3	88.8	93.3
T=400	P	81.8	82.8	94.7	94.5	91.5	91.6	82.0	80.3	96.0	94.8	95.6	94.7
	UN	87.2	87.5	95.9	96.0	93.6	93.6	70.9	70.3	90.5	89.7	94.3	93.2
	FF	85.8	86.3	95.6	95.8	93.5	93.6	75.5	74.3	93.3	92.1	95.2	94.5
H:	8												
T=100	P	52.5	75.0	80.8	94.4	77.7	90.8	39.3	37.9	80.8	80.7	83.5	90.3
	UN	60.3	80.4	84.0	95.7	83.3	93.1	43.9	43.7	84.2	84.6	85.6	92.2
	FF	58.3	79.5	82.3	95.5	83.2	93.2	33.6	36.0	76.4	79.4	84.1	91.3
T=400	P	71.7	73.6	94.3	94.3	90.8	90.6	46.4	43.2	89.7	85.9	95.2	92.0
	UN	78.9	79.9	95.6	95.6	93.2	93.2	50.8	48.2	92.0	87.6	95.9	92.6
	FF	77.2	78.6	95.5	95.4	93.4	93.2	42.5	40.3	88.0	83.6	95.6	92.5
H:	12												
T=100	P	39.6	69.6	72.9	94.3	71.1	90.3	18.2	15.1	68.3	65.3	74.7	85.8
	UN	48.6	74.1	79.4	94.9	65.2	88.8	45.8	41.0	89.9	86.3	81.0	90.7
	FF	44.2	73.5	75.4	94.6	76.3	92.4	18.8	17.2	68.8	67.2	76.7	87.7
T=400	P	65.1	67.6	93.9	94.0	90.7	90.4	19.1	17.9	80.8	73.2	94.9	88.6
	UN	71.8	73.3	95.6	95.4	91.8	90.5	48.1	42.5	96.5	90.5	96.8	92.1
	FF	69.8	72.1	94.8	94.6	92.7	92.3	20.0	18.7	81.6	73.7	95.6	89.8

Notes: See notes to Table 1.

Table 5. MC results, well specified model, nominal coverage 68%, N=200

VAR

		FWE control						FDR control					
		Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.	Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.
H:	1												
T=100	P	62.4	67.7	62.4	67.7	62.4	67.7	62.4	67.7	62.4	67.7	62.4	67.7
	UN	63.5	68.5	63.5	68.5	63.5	68.5	63.5	68.5	63.5	68.5	63.5	68.5
	FF	61.9	67.0	61.9	67.0	61.9	67.0	61.9	67.0	61.9	67.0	61.9	67.0
T=400	P	66.5	67.3	66.5	67.3	66.5	67.3	66.5	67.3	66.5	67.3	66.5	67.3
	UN	66.9	67.7	66.9	67.7	66.9	67.7	66.9	67.7	66.9	67.7	66.9	67.7
	FF	65.9	66.9	65.9	66.9	65.9	66.9	65.9	66.9	65.9	66.9	65.9	66.9
H:	4												
T=100	P	23.3	30.8	66.5	76.6	48.1	57.5	15.9	17.2	63.6	67.9	57.8	67.0
	UN	35.6	43.7	74.4	82.2	56.9	63.8	11.1	13.0	51.7	57.0	58.9	66.7
	FF	30.1	38.0	71.0	80.3	53.7	61.5	12.4	14.4	55.1	61.1	57.8	66.3
T=400	P	27.6	28.8	74.4	75.7	55.1	56.0	19.7	19.9	71.5	71.7	65.3	66.2
	UN	41.7	43.0	82.1	83.1	63.2	64.4	14.6	14.6	60.3	60.7	67.0	67.7
	FF	35.7	37.3	79.0	80.1	60.8	61.8	16.0	16.4	63.8	64.6	66.6	67.5
H:	8												
T=100	P	11.1	18.3	70.5	81.2	47.3	57.5	1.2	1.2	44.3	45.7	56.0	65.3
	UN	21.5	28.5	75.7	84.8	55.7	63.3	1.3	1.5	49.3	52.3	56.5	65.2
	FF	17.8	25.5	74.1	84.9	53.5	61.9	1.0	1.0	41.1	43.9	56.4	65.5
T=400	P	13.4	14.3	79.0	80.1	53.3	55.0	1.4	1.5	53.7	52.4	65.9	65.4
	UN	25.2	26.6	84.2	85.1	62.0	62.7	1.8	1.9	58.0	56.7	66.0	64.8
	FF	21.6	23.3	82.6	83.6	60.5	61.5	1.1	1.1	49.8	48.9	66.3	65.6
H:	12												
T=100	P	5.9	11.7	71.0	82.9	45.6	56.5	0.0	0.0	28.4	27.3	53.1	62.3
	UN	12.7	18.7	75.8	85.3	53.4	61.8	0.3	0.6	56.5	59.6	55.9	61.9
	FF	11.0	18.9	74.2	85.5	52.2	61.6	0.1	0.0	31.1	30.3	54.8	62.6
T=400	P	7.5	8.6	80.6	81.8	53.1	54.4	0.1	0.1	36.8	35.8	65.6	63.1
	UN	15.9	16.9	84.4	85.0	61.1	61.7	0.4	0.5	68.6	66.0	66.5	63.6
	FF	13.5	15.2	84.2	85.0	59.9	60.9	0.0	0.0	37.8	37.0	65.5	62.9

Local Projections

		FWE control						FDR control					
		Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.	Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.
H:	1												
T=100	P	62.7	67.7	62.7	67.7	62.7	67.7	62.7	67.7	62.7	67.7	62.7	67.7
	UN	63.8	68.5	63.8	68.5	63.8	68.5	63.8	68.5	63.8	68.5	63.8	68.5
	FF	62.3	67.0	62.3	67.0	62.3	67.0	62.3	67.0	62.3	67.0	62.3	67.0
T=400	P	66.5	67.3	66.5	67.3	66.5	67.3	66.5	67.3	66.5	67.3	66.5	67.3
	UN	66.9	67.7	66.9	67.7	66.9	67.7	66.9	67.7	66.9	67.7	66.9	67.7
	FF	66.0	66.9	66.0	66.9	66.0	66.9	66.0	66.9	66.0	66.9	66.0	66.9
H:	4												
T=100	P	19.8	30.7	61.8	76.6	43.6	57.3	16.9	17.5	64.1	68.2	54.1	66.7
	UN	32.5	43.6	70.4	82.2	55.0	63.6	10.8	13.1	49.9	57.8	55.1	66.7
	FF	26.7	38.2	66.3	80.2	51.2	61.6	12.9	14.6	54.2	61.0	54.4	66.2
T=400	P	26.9	28.9	73.8	76.0	54.4	56.1	19.7	20.0	71.8	71.8	64.9	66.3
	UN	41.3	42.9	81.4	82.8	63.1	64.3	14.3	14.7	59.6	60.8	66.6	67.9
	FF	35.2	37.4	78.5	80.2	60.4	61.6	15.8	16.4	63.5	64.2	65.9	67.4
H:	8												
T=100	P	6.6	18.4	57.9	81.1	38.2	57.0	1.5	1.1	46.3	46.5	46.2	64.9
	UN	15.6	29.5	64.5	84.8	49.0	62.5	1.8	1.6	50.1	52.6	46.7	64.9
	FF	12.1	25.9	62.6	85.1	47.6	62.2	1.1	1.1	40.3	44.7	46.9	65.4
T=400	P	11.7	14.3	76.8	80.0	52.4	54.7	1.5	1.4	54.0	52.5	64.0	65.0
	UN	24.1	26.4	82.8	85.2	61.7	63.0	1.9	1.8	58.7	57.1	64.4	65.2
	FF	20.5	23.4	81.2	83.5	59.9	61.5	1.2	1.2	50.0	49.4	65.2	65.8
H:	12												
T=100	P	2.3	12.6	50.4	82.6	33.5	56.0	0.2	0.1	31.4	28.9	35.3	61.8
	UN	7.1	18.9	58.1	85.1	34.8	61.1	1.3	1.0	62.5	60.1	36.4	62.0
	FF	5.1	19.1	56.3	85.5	40.6	61.0	0.1	0.1	33.1	32.0	36.4	62.4
T=400	P	6.4	8.5	77.8	81.7	52.1	54.5	0.1	0.1	37.6	35.8	62.5	63.0
	UN	14.4	17.1	82.1	84.7	59.2	61.6	0.5	0.5	70.5	65.7	63.1	63.6
	FF	12.0	15.1	81.9	85.1	59.2	60.9	0.0	0.1	38.6	36.9	62.4	62.9

Notes: See notes to Table 1.

Table 6. MC results, well specified model, nominal coverage 95%, N=200

VAR

		FWE control						FDR control					
		Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.	Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.
H:	1												
T=100	P	92.0	95.1	92.0	95.1	92.0	95.1	92.0	95.1	92.0	95.1	92.0	95.1
	UN	91.9	94.8	91.9	94.8	91.9	94.8	91.9	94.8	91.9	94.8	91.9	94.8
	FF	91.8	94.7	91.8	94.7	91.8	94.7	91.8	94.7	91.8	94.7	91.8	94.7
T=400	P	94.2	94.5	94.2	94.5	94.2	94.5	94.2	94.5	94.2	94.5	94.2	94.5
	UN	94.7	95.2	94.7	95.2	94.7	95.2	94.7	95.2	94.7	95.2	94.7	95.2
	FF	94.6	94.9	94.6	94.9	94.6	94.9	94.6	94.9	94.6	94.9	94.6	94.9
H:	4												
T=100	P	76.3	84.5	90.5	95.8	86.7	92.9	74.6	78.9	92.3	94.6	92.4	95.8
	UN	81.7	88.6	92.8	96.5	90.6	94.3	64.0	69.8	85.8	90.1	90.2	93.9
	FF	78.9	86.7	91.2	96.0	88.9	93.6	67.1	73.1	87.8	91.4	90.6	94.4
T=400	P	83.0	84.1	94.9	95.4	91.7	92.6	81.9	82.1	96.2	95.9	96.0	95.9
	UN	87.2	87.9	96.0	96.2	93.1	93.5	70.9	71.4	90.6	90.7	94.2	94.3
	FF	85.5	86.4	95.4	95.9	93.3	93.6	74.7	75.3	93.0	93.1	95.3	95.1
H:	8												
T=100	P	66.0	78.2	90.3	96.3	86.6	93.0	37.4	39.0	81.7	83.7	90.0	93.8
	UN	73.0	82.2	92.1	96.9	89.3	93.8	43.1	45.5	85.9	88.4	92.0	95.2
	FF	69.4	81.4	90.8	96.5	88.4	93.7	34.4	36.6	79.5	82.8	90.4	94.3
T=400	P	74.3	75.5	95.4	95.7	91.9	92.2	46.4	45.5	90.0	89.0	95.4	94.9
	UN	79.9	81.2	96.4	96.6	93.5	93.9	50.1	48.9	91.7	90.5	96.2	95.5
	FF	78.9	80.0	96.2	96.5	93.5	94.0	42.0	41.6	87.6	86.3	95.8	95.0
H:	12												
T=100	P	59.5	72.9	89.1	96.0	85.5	92.6	15.1	13.3	70.5	69.9	89.3	92.7
	UN	65.8	76.8	91.9	96.4	88.0	93.2	37.6	40.3	89.2	91.5	92.8	95.8
	FF	63.1	76.4	90.6	96.7	88.3	93.9	15.6	15.2	72.3	72.7	90.8	94.2
T=400	P	69.0	70.4	95.9	95.7	91.9	92.4	19.1	18.6	81.2	78.6	95.5	93.9
	UN	74.7	75.5	96.5	96.6	93.3	93.4	45.7	44.9	96.4	94.8	97.5	96.5
	FF	72.8	74.5	96.2	96.7	92.8	93.5	19.6	19.0	81.3	79.2	96.1	94.7

Local Projections

		FWE control						FDR control					
		Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.	Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.
H:	1												
T=100	P	92.2	95.1	92.2	95.1	92.2	95.1	92.2	95.1	92.2	95.1	92.2	95.1
	UN	92.1	94.8	92.1	94.8	92.1	94.8	92.1	94.8	92.1	94.8	92.1	94.8
	FF	92.0	94.7	92.0	94.7	92.0	94.7	92.0	94.7	92.0	94.7	92.0	94.7
T=400	P	94.3	94.5	94.3	94.5	94.3	94.5	94.3	94.5	94.3	94.5	94.3	94.5
	UN	94.8	95.2	94.8	95.2	94.8	95.2	94.8	95.2	94.8	95.2	94.8	95.2
	FF	94.7	94.9	94.7	94.9	94.7	94.9	94.7	94.9	94.7	94.9	94.7	94.9
H:	4												
T=100	P	71.8	84.9	87.4	96.0	82.8	93.0	75.0	79.0	92.1	94.8	90.2	95.9
	UN	77.5	88.6	89.6	96.6	88.9	94.2	62.2	69.8	83.9	89.9	88.2	94.0
	FF	74.6	86.8	88.1	96.1	86.9	93.6	65.8	73.0	86.6	91.6	88.7	94.4
T=400	P	82.3	84.1	94.5	95.3	91.5	92.6	82.2	82.3	96.4	96.0	96.1	96.1
	UN	86.7	87.8	95.6	96.2	92.9	93.5	70.8	71.5	90.6	90.7	93.9	94.4
	FF	84.9	86.5	95.3	96.0	93.2	93.7	75.0	75.5	93.1	93.2	95.1	95.3
H:	8												
T=100	P	53.0	78.0	80.7	96.1	78.0	92.8	39.1	39.4	81.2	84.2	83.5	93.4
	UN	61.2	82.6	84.5	96.9	83.7	94.0	44.1	45.2	84.4	88.6	85.9	95.2
	FF	57.2	81.2	81.7	96.4	83.0	93.7	33.4	37.1	76.5	82.6	83.9	94.2
T=400	P	72.0	75.5	94.4	95.6	91.4	92.1	46.9	46.3	90.2	89.1	95.0	94.9
	UN	78.4	81.1	95.8	96.6	93.4	93.9	50.3	48.6	91.6	90.7	95.7	95.6
	FF	76.9	79.9	95.6	96.4	93.4	93.9	42.1	41.8	87.4	86.4	95.4	95.0
H:	12												
T=100	P	39.5	72.6	73.4	95.6	71.9	92.4	18.9	14.9	69.0	70.9	75.0	92.4
	UN	48.6	76.9	79.5	96.4	65.4	92.3	46.4	40.9	89.0	90.5	81.1	95.6
	FF	44.0	76.4	75.5	97.0	76.2	94.0	18.7	15.9	69.6	73.1	77.0	93.9
T=400	P	65.4	70.3	94.4	95.7	91.3	92.5	19.6	18.7	81.5	78.8	94.8	93.8
	UN	71.9	75.5	95.5	96.7	92.4	93.4	47.8	44.7	96.7	94.7	96.9	96.4
	FF	69.5	74.1	95.1	96.8	92.5	93.4	20.4	19.0	81.1	79.1	95.0	94.7

Notes: See notes to Table 1.

Table 7. MC results, model with wrong dynamics, nominal coverage 95%, N=80

VAR

		FWE control						FDR control					
		Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.	Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.
H:	1												
T=100	P	92.5	95.0	92.5	95.0	92.5	95.0	92.5	95.0	92.5	95.0	92.5	95.0
	UN	93.3	94.9	93.3	94.9	93.3	94.9	93.3	94.9	93.3	94.9	93.3	94.9
	FF	92.8	94.6	92.8	94.6	92.8	94.6	92.8	94.6	92.8	94.6	92.8	94.6
T=400	P	94.8	95.1	94.8	95.1	94.8	95.1	94.8	95.1	94.8	95.1	94.8	95.1
	UN	94.9	94.5	94.9	94.5	94.9	94.5	94.9	94.5	94.9	94.5	94.9	94.5
	FF	94.7	94.5	94.7	94.5	94.7	94.5	94.7	94.5	94.7	94.5	94.7	94.5
H:	4												
T=100	P	79.1	84.7	91.6	95.3	88.9	92.7	70.4	72.8	89.8	91.7	91.6	93.7
	UN	74.7	87.6	87.9	95.8	88.5	93.2	76.1	67.2	90.7	87.4	93.1	92.3
	FF	78.7	86.4	91.1	95.4	90.5	93.1	69.2	69.6	88.8	89.6	92.1	93.1
T=400	P	87.0	83.1	96.4	94.5	93.1	91.7	74.4	75.8	92.6	93.0	94.7	94.1
	UN	75.0	87.0	88.1	95.7	89.7	92.9	77.9	67.7	92.0	88.8	94.0	93.0
	FF	82.5	85.7	93.8	95.2	92.4	93.0	73.1	70.9	91.6	90.2	94.6	93.6
H:	8												
T=100	P	69.5	77.5	89.5	94.4	86.7	91.6	41.9	30.4	82.7	74.0	90.8	88.3
	UN	64.0	81.3	87.0	95.9	87.7	92.9	57.0	43.8	88.6	84.6	94.6	92.4
	FF	67.7	80.3	89.6	95.9	89.4	93.2	34.1	33.5	77.6	77.3	91.1	91.0
T=400	P	81.4	76.1	96.9	94.4	93.7	91.9	45.1	36.6	87.6	80.9	95.1	90.6
	UN	62.8	80.9	87.0	95.9	89.3	93.0	60.0	47.2	91.1	87.7	95.6	93.5
	FF	73.6	79.9	94.1	95.8	92.5	93.1	38.7	36.9	83.4	81.5	95.1	92.1
H:	12												
T=100	P	63.6	72.5	89.1	94.1	86.2	90.8	22.7	9.6	77.4	55.2	90.7	83.7
	UN	57.8	74.2	87.2	95.2	87.1	91.3	39.6	37.0	85.6	85.0	94.7	91.5
	FF	60.4	74.4	88.3	94.9	88.7	92.4	14.2	13.0	66.0	62.8	90.7	88.0
T=400	P	75.7	69.8	96.5	94.1	93.6	91.5	22.5	12.3	82.7	63.1	95.8	86.9
	UN	55.6	74.0	87.5	95.6	88.9	91.9	42.4	39.6	89.3	87.7	96.0	91.5
	FF	66.9	73.5	93.8	95.2	92.3	92.5	14.8	15.2	71.4	69.1	94.6	89.1

Local Projections

		FWE control						FDR control					
		Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.	Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.
H:	1												
T=100	P	92.6	95.0	92.6	95.0	92.6	95.0	92.6	95.0	92.6	95.0	92.6	95.0
	UN	93.5	94.9	93.5	94.9	93.5	94.9	93.5	94.9	93.5	94.9	93.5	94.9
	FF	93.0	94.6	93.0	94.6	93.0	94.6	93.0	94.6	93.0	94.6	93.0	94.6
T=400	P	94.9	95.1	94.9	95.1	94.9	95.1	94.9	95.1	94.9	95.1	94.9	95.1
	UN	94.9	94.5	94.9	94.5	94.9	94.5	94.9	94.5	94.9	94.5	94.9	94.5
	FF	94.7	94.5	94.7	94.5	94.7	94.5	94.7	94.5	94.7	94.5	94.7	94.5
H:	4												
T=100	P	75.9	84.7	90.1	95.2	87.0	92.8	75.0	74.2	92.3	92.3	92.3	94.0
	UN	82.6	87.2	93.1	95.7	91.4	93.1	64.3	66.0	85.4	87.2	90.6	91.9
	FF	80.3	86.5	91.8	95.3	90.8	93.2	68.7	69.7	88.8	89.4	91.8	92.7
T=400	P	82.9	83.4	94.5	94.7	91.8	91.9	78.4	76.0	94.6	93.2	95.4	94.2
	UN	87.9	87.0	96.4	95.7	93.6	93.0	68.6	67.1	89.5	87.6	94.2	92.7
	FF	86.0	85.5	95.7	95.2	93.4	92.8	72.8	70.8	91.8	90.0	95.1	93.4
H:	8												
T=100	P	60.8	77.0	85.8	94.8	83.6	91.9	37.8	32.1	80.7	76.3	87.1	89.2
	UN	68.9	81.0	89.5	95.5	88.1	92.5	49.0	45.9	87.5	85.6	90.6	92.1
	FF	66.8	79.6	87.7	95.6	88.2	93.2	36.1	33.3	79.9	76.8	89.2	90.7
T=400	P	74.8	75.2	95.1	95.1	92.7	92.3	40.1	36.6	86.4	81.1	95.2	91.1
	UN	81.1	80.8	96.4	95.8	93.9	92.8	52.0	47.3	92.5	87.9	96.6	93.1
	FF	79.5	79.4	96.2	95.5	93.7	93.0	39.5	36.4	85.9	81.0	95.6	92.0
H:	12												
T=100	P	48.3	71.0	80.8	94.3	80.6	91.4	16.6	11.1	68.5	58.7	82.2	84.7
	UN	57.6	73.7	86.5	94.8	75.7	88.7	51.6	44.0	92.5	88.4	89.0	91.2
	FF	54.3	73.6	83.5	94.6	84.6	92.3	20.4	15.1	73.6	64.7	85.9	88.2
T=400	P	68.6	69.4	94.8	94.6	92.4	91.8	14.5	12.5	75.1	64.5	94.9	87.7
	UN	74.0	73.8	96.2	95.3	92.8	90.1	50.8	43.9	96.8	89.6	97.2	91.5
	FF	72.2	72.0	95.9	94.8	93.4	92.3	18.4	16.2	80.4	69.8	95.9	89.1

Notes: 10,000 samples generated from VAR(4) for three variables (P, UN, FF) with stable parameters. Forecasting model has one lag only. Each estimated VAR or local projection (LP) on these 10,000 samples generates a forecast error variance (which includes estimation uncertainty) for the forecast path of length h, and hence the sets of bands (marginal, Bonferroni, and Scheffé) used in the analysis. Similarly, each estimated model generates q forecast paths whose associated error paths are used to generate a forecast error variance for the forecast path and hence the set of bands (marginal emp, Bonferroni emp, and Scheffé emp). Hence 10,000 actual paths are then compared with each set of 10,000 bands to determine the appropriate coverage rates. FWE control stands for "family-wise error control" and simply computes the proportion of paths strictly inside the bands. FDR control instead is the proportion of forecast paths whose Mahalanobis distance attains a value that is lower than the chi-square statistic for probability equal to nominal coverage and degrees of freedom equal to H. See text for more details.

Table 8. MC results, model with omitted variables, nominal coverage 95%, N=80

VAR

		FWE control						FDR control					
		Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.	Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.
H:	1												
T=100	P	93.4	94.8	93.4	94.8	93.4	94.8	93.4	94.8	93.4	94.8	93.4	94.8
	FF	92.7	94.6	92.7	94.6	92.7	94.6	92.7	94.6	92.7	94.6	92.7	94.6
T=400	P	94.6	94.7	94.6	94.7	94.6	94.7	94.6	94.7	94.6	94.7	94.6	94.7
	FF	95.1	95.1	95.1	95.1	95.1	95.1	95.1	95.1	95.1	95.1	95.1	95.1
H:	4												
T=100	P	80.0	83.9	92.7	95.0	89.3	92.1	76.2	78.3	92.5	93.8	92.9	94.5
	FF	82.5	86.9	93.5	95.6	91.5	93.7	68.1	69.7	88.3	89.5	91.5	93.1
T=400	P	83.7	83.5	95.2	94.8	92.2	92.0	81.8	80.3	95.7	95.0	95.9	95.1
	FF	86.8	86.4	96.2	96.0	93.9	93.7	73.4	71.9	92.8	90.9	95.4	94.3
H:	8												
T=100	P	70.1	75.9	92.4	94.7	88.8	91.5	40.0	38.5	84.2	81.4	91.6	91.2
	FF	72.5	79.1	92.2	95.2	89.7	92.8	33.2	32.8	78.3	76.9	90.7	90.6
T=400	P	76.5	75.5	95.9	94.6	92.3	91.6	46.1	42.8	89.8	84.8	95.7	92.0
	FF	79.7	79.2	96.3	95.5	93.5	93.2	38.8	36.0	85.4	80.4	95.7	92.2
H:	12												
T=100	P	63.5	69.8	91.7	94.2	88.0	91.3	16.1	14.2	72.6	65.5	90.8	87.0
	FF	66.3	74.5	91.7	94.9	89.9	93.0	14.7	13.4	69.4	64.3	91.0	88.3
T=400	P	70.1	68.3	95.8	94.1	92.2	90.6	19.4	17.0	81.4	71.7	96.2	88.5
	FF	75.4	73.2	96.5	95.1	93.0	92.4	16.8	16.9	77.9	70.8	96.4	89.7

Local Projections

		FWE control						FDR control					
		Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.	Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.
H:	1												
T=100	P	93.6	94.8	93.6	94.8	93.6	94.8	93.6	94.8	93.6	94.8	93.6	94.8
	FF	92.9	94.6	92.9	94.6	92.9	94.6	92.9	94.6	92.9	94.6	92.9	94.6
T=400	P	94.7	94.7	94.7	94.7	94.7	94.7	94.7	94.7	94.7	94.7	94.7	94.7
	FF	95.1	95.1	95.1	95.1	95.1	95.1	95.1	95.1	95.1	95.1	95.1	95.1
H:	4												
T=100	P	76.4	83.7	90.9	94.9	87.4	92.0	76.9	79.0	92.6	93.5	91.8	94.4
	FF	79.5	86.8	91.4	95.7	90.7	93.8	66.9	70.1	86.8	89.7	90.2	93.0
T=400	P	83.2	83.5	94.9	94.9	92.0	92.0	82.0	80.4	95.9	95.0	95.9	95.3
	FF	86.4	86.5	96.2	96.2	93.9	93.7	73.0	71.9	92.4	91.1	95.3	94.2
H:	8												
T=100	P	61.2	75.5	87.2	94.7	84.6	91.5	39.4	38.7	82.3	81.4	87.6	91.2
	FF	63.9	79.2	86.2	94.8	86.8	92.6	32.0	32.9	75.6	76.5	86.1	90.2
T=400	P	74.7	75.6	95.3	94.7	92.0	91.7	45.1	42.9	89.0	84.8	95.0	92.0
	FF	78.7	79.1	95.9	95.4	93.4	93.2	38.4	36.4	85.1	80.3	95.1	92.2
H:	12												
T=100	P	49.0	69.2	81.9	93.9	81.3	90.7	16.5	14.9	68.7	65.6	81.0	86.6
	FF	51.7	74.5	81.8	94.9	82.8	92.8	17.9	13.6	69.9	64.1	81.8	87.9
T=400	P	67.8	68.3	94.8	94.0	91.9	90.7	18.5	17.2	79.3	71.5	94.9	88.6
	FF	72.3	72.9	95.5	95.2	92.8	92.3	18.4	16.8	80.3	71.4	95.8	89.7

Notes: 10,000 samples generated from VAR(4) for three variables (P, UN, FF) with stable parameters. Forecasting model omits UN. Each estimated VAR or local projection (LP) on these 10,000 samples generates a forecast error variance (which includes estimation uncertainty) for the forecast path of length h, and hence the sets of bands (marginal, Bonferroni, and Scheffé) used in the analysis. Similarly, each estimated model generates q forecast paths whose associated error paths are used to generate a forecast error variance for the forecast path and hence the set of bands (marginal emp, Bonferroni emp, and Scheffé emp). Hence 10,000 actual paths are then compared with each set of 10,000 bands to determine the appropriate coverage rates. FWE control stands for "family-wise error control" and simply computes the proportion of paths strictly inside the bands. FDR control instead is the proportion of forecast paths whose Mahalanobis distance attains a value that is lower than the chi-square statistic for probability equal to nominal coverage and degrees of freedom equal to H. See text for more details.

Table 9. MC results, break in DGP coefficients, nominal coverage 95%, N=80

VAR

		FWE control						FDR control					
		Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.	Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.
H:	1												
T=100	P	91.3	94.6	91.3	94.6	91.3	94.6	91.3	94.6	91.3	94.6	91.3	94.6
	UN	90.1	93.9	90.1	93.9	90.1	93.9	90.1	93.9	90.1	93.9	90.1	93.9
	FF	91.1	94.4	91.1	94.4	91.1	94.4	91.1	94.4	91.1	94.4	91.1	94.4
T=400	P	94.0	94.7	94.0	94.7	94.0	94.7	94.0	94.7	94.0	94.7	94.0	94.7
	UN	94.0	94.6	94.0	94.6	94.0	94.6	94.0	94.6	94.0	94.6	94.0	94.6
	FF	94.1	94.5	94.1	94.5	94.1	94.5	94.1	94.5	94.1	94.5	94.1	94.5
H:	4												
T=100	P	73.7	81.3	89.7	94.0	84.4	88.9	73.9	82.8	91.4	95.8	89.8	94.7
	UN	77.8	87.5	89.5	95.8	87.5	93.2	57.0	62.4	80.4	85.4	86.3	90.4
	FF	76.1	87.9	88.4	96.0	87.8	93.6	66.3	67.4	86.4	87.3	90.2	91.8
T=400	P	83.6	80.8	95.5	94.6	91.1	87.4	83.1	89.7	96.3	98.1	95.1	96.0
	UN	86.2	88.0	95.6	96.0	93.1	93.7	67.3	66.5	88.3	87.5	92.9	92.4
	FF	81.8	88.1	92.8	96.2	92.3	94.2	74.1	69.5	91.9	89.3	94.6	93.0
H:	8												
T=100	P	59.8	71.4	87.7	94.0	82.7	87.9	35.0	44.3	80.0	85.4	86.0	89.9
	UN	68.8	82.5	88.4	95.9	87.2	93.1	28.1	26.1	72.6	70.3	87.4	88.6
	FF	63.0	81.5	84.4	95.5	85.8	93.0	34.0	28.9	76.4	72.2	89.0	89.1
T=400	P	75.7	70.4	96.1	94.5	91.0	86.7	47.3	60.4	89.2	93.5	93.4	92.7
	UN	77.4	81.9	94.3	95.8	92.4	93.5	34.6	29.6	81.6	73.7	93.3	90.1
	FF	72.0	81.8	91.6	95.9	91.8	93.5	41.2	30.9	84.5	75.8	94.9	91.0
H:	12												
T=100	P	53.1	65.3	87.2	94.3	83.0	87.7	15.7	21.5	68.7	72.2	84.2	85.6
	UN	62.3	77.6	87.5	94.8	86.2	92.6	18.5	14.7	73.9	64.1	88.5	87.0
	FF	54.3	76.8	82.6	95.1	84.4	92.7	17.7	11.6	70.7	60.4	89.3	86.4
T=400	P	69.8	61.8	96.7	94.1	90.9	85.3	19.7	31.4	79.9	84.5	92.8	88.7
	UN	71.1	77.6	93.8	95.4	92.9	93.3	21.0	15.7	82.1	68.5	94.9	89.0
	FF	63.7	77.0	90.8	95.6	91.3	93.4	20.7	13.5	79.4	65.1	95.6	88.8

Local Projections

		FWE control						FDR control					
		Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.	Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.
H:	1												
T=100	P	91.5	94.6	91.5	94.6	91.5	94.6	91.5	94.6	91.5	94.6	91.5	94.6
	UN	90.2	93.9	90.2	93.9	90.2	93.9	90.2	93.9	90.2	93.9	90.2	93.9
	FF	91.3	94.4	91.3	94.4	91.3	94.4	91.3	94.4	91.3	94.4	91.3	94.4
T=400	P	94.0	94.7	94.0	94.7	94.0	94.7	94.0	94.7	94.0	94.7	94.0	94.7
	UN	94.1	94.6	94.1	94.6	94.1	94.6	94.1	94.6	94.1	94.6	94.1	94.6
	FF	94.1	94.5	94.1	94.5	94.1	94.5	94.1	94.5	94.1	94.5	94.1	94.5
H:	4												
T=100	P	68.7	81.5	85.9	94.0	78.7	88.8	76.0	82.6	92.6	95.7	88.3	94.8
	UN	73.3	87.6	85.7	96.0	85.2	93.2	54.9	63.3	77.5	85.5	83.3	90.5
	FF	71.2	87.9	84.3	96.0	85.4	93.7	63.6	67.1	83.9	87.5	87.0	91.9
T=400	P	83.2	80.8	95.2	94.5	90.8	87.7	83.3	89.4	96.5	98.1	95.2	96.0
	UN	85.3	88.1	95.1	96.0	92.9	93.8	66.7	66.2	88.1	87.4	92.8	92.3
	FF	81.0	88.2	92.5	96.1	92.0	94.0	73.6	69.3	91.5	89.2	94.5	93.2
H:	8												
T=100	P	45.3	71.4	75.9	94.2	70.2	87.9	38.5	44.5	80.5	84.9	77.5	89.5
	UN	56.5	82.2	77.9	95.8	80.8	92.9	25.4	26.2	66.8	70.1	78.0	87.9
	FF	50.2	81.0	72.8	95.1	78.1	92.8	29.8	29.3	69.0	72.4	78.6	88.5
T=400	P	73.8	70.8	95.5	94.7	90.4	86.9	47.7	60.0	89.5	93.4	92.9	92.7
	UN	75.5	82.3	93.3	95.9	92.0	93.6	34.5	29.2	81.2	73.5	92.4	90.2
	FF	69.3	82.0	90.0	96.0	91.1	93.4	39.7	30.9	83.4	76.3	94.0	90.9
H:	12												
T=100	P	31.9	66.3	67.5	94.2	62.7	87.6	19.6	22.2	68.3	71.8	67.3	84.4
	UN	43.1	77.6	71.4	94.9	71.1	92.4	19.8	16.6	67.3	63.5	71.7	85.7
	FF	36.0	76.6	64.3	95.1	69.0	92.4	17.2	13.1	62.2	61.2	71.9	85.8
T=400	P	67.0	62.4	95.5	94.0	90.3	85.7	19.8	31.2	80.5	84.1	92.1	88.8
	UN	68.0	77.6	92.2	95.5	92.2	93.3	20.5	15.4	81.3	67.7	93.7	88.8
	FF	59.5	77.0	88.5	95.6	90.5	93.4	20.4	13.5	78.8	64.8	94.2	88.6

Notes: 10,000 samples generated from VAR(4) for three variables (P, UN, FF) with break in parameters. Forecasting model has stable parameters. Each estimated VAR or local projection (LP) on these 10,000 samples generates a forecast error variance (which includes estimation uncertainty) for the forecast path of length h, and hence the sets of bands (marginal, Bonferroni, and Scheffé) used in the analysis. Similarly, each estimated model generates q forecast paths whose associated error paths are used to generate a forecast error variance for the forecast path and hence the set of bands (marginal emp, Bonferroni emp, and Scheffé emp). Hence 10,000 actual paths are then compared with each set of 10,000 bands to determine the appropriate coverage rates. FWE control stands for "family-wise error control" and simply computes the proportion of paths strictly inside the bands. FDR control instead is the proportion of forecast paths whose Mahalanobis distance attains a value that is lower than the chi-square statistic for probability equal to nominal coverage and degrees of freedom equal to H. See text for more details.

Table 10. MC results, break in DGP var-cov matrix of errors, nominal coverage 95%, N=80

VAR

		FWE control						FDR control					
		Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.	Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.
H:	1												
T=100	P	98.5	94.9	98.5	94.9	98.5	94.9	98.5	94.9	98.5	94.9	98.5	94.9
	UN	97.8	94.4	97.8	94.4	97.8	94.4	97.8	94.4	97.8	94.4	97.8	94.4
	FF	99.8	94.9	99.8	94.9	99.8	94.9	99.8	94.9	99.8	94.9	99.8	94.9
T=400	P	99.7	94.9	99.7	94.9	99.7	94.9	99.7	94.9	99.7	94.9	99.7	94.9
	UN	99.3	94.9	99.3	94.9	99.3	94.9	99.3	94.9	99.3	94.9	99.3	94.9
	FF	100.0	94.9	100.0	94.9	100.0	94.9	100.0	94.9	100.0	94.9	100.0	94.9
H:	4												
T=100	P	94.9	84.4	98.9	95.3	97.9	92.6	96.9	77.0	99.6	93.3	99.5	94.5
	UN	93.4	87.6	98.4	96.0	97.3	93.5	89.8	67.9	98.4	88.1	99.1	92.3
	FF	98.0	88.0	99.6	96.1	99.7	93.9	99.2	69.7	99.8	89.9	99.9	92.7
T=400	P	98.6	82.5	99.9	94.7	99.5	91.1	99.4	79.5	100.0	94.7	100.0	94.7
	UN	97.8	87.3	99.8	95.8	99.2	93.7	96.1	70.4	99.7	89.5	99.9	93.1
	FF	99.9	85.8	100.0	95.6	100.0	93.1	100.0	72.5	100.0	91.3	100.0	93.9
H:	8												
T=100	P	90.1	76.9	98.5	94.8	97.4	91.6	88.6	36.0	99.4	79.2	99.7	90.0
	UN	89.7	80.5	98.4	95.6	97.3	92.8	85.9	41.4	99.3	84.0	99.7	92.1
	FF	95.2	81.8	99.1	95.9	99.5	93.0	98.0	32.7	99.8	75.8	99.9	89.7
T=400	P	97.4	74.2	100.0	94.9	99.7	91.1	96.4	41.2	100.0	84.7	100.0	91.9
	UN	96.7	80.0	99.9	95.5	99.3	93.1	94.4	45.4	100.0	86.7	100.0	92.9
	FF	99.7	79.7	100.0	96.0	100.0	93.0	100.0	37.3	100.0	81.8	100.0	92.0
H:	12												
T=100	P	86.8	72.5	98.4	94.9	97.5	91.8	75.4	12.3	99.3	61.3	99.7	85.8
	UN	87.5	75.2	98.6	95.4	96.9	90.7	87.2	38.7	99.7	85.8	99.8	91.3
	FF	93.3	76.4	99.2	95.6	99.6	92.6	96.8	13.1	99.9	60.6	99.9	86.4
T=400	P	96.2	69.1	99.9	94.5	99.5	91.1	89.0	16.3	100.0	71.4	100.0	88.9
	UN	95.8	72.9	100.0	95.2	99.2	90.0	96.1	42.3	100.0	90.2	100.0	91.8
	FF	99.4	74.4	100.0	95.5	100.0	92.6	99.9	16.8	100.0	70.0	100.0	88.6

Local Projections

		FWE control						FDR control					
		Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.	Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.
H:	1												
T=100	P	98.5	94.9	98.5	94.9	98.5	94.9	98.5	94.9	98.5	94.9	98.5	94.9
	UN	97.9	94.4	97.9	94.4	97.9	94.4	97.9	94.4	97.9	94.4	97.9	94.4
	FF	99.9	94.9	99.9	94.9	99.9	94.9	99.9	94.9	99.9	94.9	99.9	94.9
T=400	P	99.7	94.9	99.7	94.9	99.7	94.9	99.7	94.9	99.7	94.9	99.7	94.9
	UN	99.3	94.9	99.3	94.9	99.3	94.9	99.3	94.9	99.3	94.9	99.3	94.9
	FF	100.0	94.9	100.0	94.9	100.0	94.9	100.0	94.9	100.0	94.9	100.0	94.9
H:	4												
T=100	P	92.3	84.1	97.9	95.2	96.7	92.5	96.1	77.0	99.3	93.3	99.0	94.5
	UN	91.4	87.6	97.3	96.0	97.0	93.5	87.1	67.9	97.3	88.3	98.3	92.3
	FF	96.6	87.9	98.9	96.3	99.3	93.9	98.2	70.2	99.6	89.7	99.6	92.9
T=400	P	98.5	82.4	99.9	94.6	99.5	91.2	99.4	79.8	100.0	94.5	100.0	94.7
	UN	97.7	87.3	99.8	95.7	99.2	93.7	95.7	70.5	99.7	89.3	99.9	93.1
	FF	99.9	85.8	100.0	95.6	100.0	93.1	100.0	72.8	100.0	91.2	100.0	93.8
H:	8												
T=100	P	80.7	77.0	94.6	94.8	94.5	91.6	82.0	36.7	97.9	79.6	97.9	89.9
	UN	81.6	80.5	95.1	95.5	94.8	93.0	81.0	41.4	98.0	83.4	98.1	91.7
	FF	88.0	81.9	96.6	96.0	98.0	93.2	92.8	34.3	98.9	76.5	99.0	89.0
T=400	P	96.8	74.3	100.0	95.0	99.7	91.2	95.9	41.4	100.0	84.7	100.0	92.0
	UN	96.3	80.1	99.9	95.5	99.3	93.2	93.8	45.5	100.0	86.7	100.0	93.0
	FF	99.6	79.6	100.0	95.9	100.0	93.0	99.9	37.5	100.0	82.0	100.0	92.1
H:	12												
T=100	P	69.2	72.1	90.9	94.8	91.2	91.7	63.2	14.3	95.4	62.5	96.0	85.2
	UN	74.5	75.1	93.3	95.3	79.8	89.0	84.0	40.8	98.8	85.6	97.1	91.0
	FF	78.8	76.7	93.7	95.4	94.3	92.6	86.5	16.5	98.3	63.3	98.2	85.7
T=400	P	95.0	69.0	99.9	94.5	99.5	91.2	87.4	16.5	99.9	71.8	100.0	88.6
	UN	94.9	72.9	99.9	95.0	99.1	89.8	96.1	42.6	100.0	89.9	100.0	91.7
	FF	98.9	74.2	100.0	95.3	100.0	92.6	99.8	17.0	100.0	70.1	100.0	88.9

Notes: 10,000 samples generated from VAR(4) for three variables (P, UN, FF) with break in var-cov matrix of errors. Forecasting model has stable var-cov matrix of errors. Each estimated VAR or local projection (LP) on these 10,000 samples generates a forecast error variance (which includes estimation uncertainty) for the forecast path of length h, and hence the sets of bands (marginal, Bonferroni, and Scheffé) used in the analysis. Similarly, each estimated model generates q forecast paths whose associated error paths are used to generate a forecast error variance for the forecast path and hence the set of bands (marginal emp, Bonferroni emp, and Scheffé emp). Hence 10,000 actual paths are then compared with each set of 10,000 bands to determine the appropriate coverage rates. FWE control stands for "family-wise error control" and simply computes the proportion of paths strictly inside the bands. FDR control instead is the proportion of forecast paths whose Mahalanobis distance attains a value that is lower than the chi-square statistic for probability equal to nominal coverage and degrees of freedom equal to H.
See text for more details.

Table 11. MC results, break in DGP coefficients and var-cov matrix of errors, nominal coverage 95%, N=80

VAR

		FWE control						FDR control					
		Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.	Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.
H:	1												
T=100	P	97.7	94.5	97.7	94.5	97.7	94.5	97.7	94.5	97.7	94.5	97.7	94.5
	UN	96.1	94.3	96.1	94.3	96.1	94.3	96.1	94.3	96.1	94.3	96.1	94.3
	FF	99.6	94.1	99.6	94.1	99.6	94.1	99.6	94.1	99.6	94.1	99.6	94.1
T=400	P	99.4	95.0	99.4	95.0	99.4	95.0	99.4	95.0	99.4	95.0	99.4	95.0
	UN	98.7	94.5	98.7	94.5	98.7	94.5	98.7	94.5	98.7	94.5	98.7	94.5
	FF	100.0	94.6	100.0	94.6	100.0	94.6	100.0	94.6	100.0	94.6	100.0	94.6
H:	4												
T=100	P	92.2	83.2	98.1	95.4	96.5	91.0	94.4	78.9	99.2	94.2	99.1	94.7
	UN	89.0	87.7	95.7	96.0	95.4	93.5	83.0	61.9	95.0	85.2	96.8	90.2
	FF	95.2	87.9	98.6	95.9	99.1	93.5	97.4	68.1	99.5	88.6	99.6	92.0
T=400	P	98.7	81.6	99.9	94.4	99.3	89.6	98.7	85.7	99.9	96.8	99.9	95.4
	UN	96.2	87.7	99.3	95.9	98.6	93.5	92.1	64.3	98.9	86.3	99.5	91.2
	FF	99.4	88.0	99.9	96.1	100.0	93.6	99.8	69.3	100.0	89.6	100.0	93.1
H:	8												
T=100	P	85.1	75.6	97.6	95.5	96.4	90.5	80.4	38.1	98.3	81.2	99.0	88.8
	UN	81.2	81.8	94.5	95.7	94.3	93.0	72.4	25.4	96.2	69.3	97.8	88.0
	FF	88.5	82.7	97.3	96.0	99.0	93.0	94.2	31.6	99.4	74.9	99.6	88.8
T=400	P	97.9	72.1	100.0	94.5	99.2	89.2	92.4	49.4	99.8	88.9	100.0	91.4
	UN	92.0	82.0	98.8	95.7	98.6	93.5	85.4	29.3	99.4	73.1	99.9	90.1
	FF	97.6	82.4	99.9	96.1	100.0	93.4	99.4	32.5	100.0	77.1	100.0	90.8
H:	12												
T=100	P	79.9	71.0	97.0	95.6	96.3	90.4	64.0	15.2	98.0	63.9	99.3	83.0
	UN	78.4	76.9	94.5	95.0	93.7	92.4	70.8	15.2	98.3	62.1	98.6	87.2
	FF	83.7	78.7	96.5	96.4	98.6	92.9	92.0	15.4	99.5	64.3	99.7	87.5
T=400	P	97.5	65.8	100.0	94.9	99.3	88.8	80.6	22.6	99.7	77.4	100.0	88.2
	UN	89.4	76.6	99.2	95.4	98.7	93.1	84.5	13.9	99.9	65.1	100.0	88.2
	FF	95.7	77.9	99.8	95.8	100.0	92.8	99.0	15.3	100.0	68.2	100.0	88.8

Local Projections

		FWE control						FDR control					
		Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.	Marg.	Marg. emp.	Bonf.	Bonf. emp.	Schef.	Schef. emp.
H:	1												
T=100	P	97.8	94.5	97.8	94.5	97.8	94.5	97.8	94.5	97.8	94.5	97.8	94.5
	UN	96.2	94.3	96.2	94.3	96.2	94.3	96.2	94.3	96.2	94.3	96.2	94.3
	FF	99.6	94.1	99.6	94.1	99.6	94.1	99.6	94.1	99.6	94.1	99.6	94.1
T=400	P	99.4	95.0	99.4	95.0	99.4	95.0	99.4	95.0	99.4	95.0	99.4	95.0
	UN	98.7	94.5	98.7	94.5	98.7	94.5	98.7	94.5	98.7	94.5	98.7	94.5
	FF	100.0	94.6	100.0	94.6	100.0	94.6	100.0	94.6	100.0	94.6	100.0	94.6
H:	4												
T=100	P	88.7	82.7	96.4	95.3	94.5	90.7	92.8	79.1	98.6	94.0	97.9	94.5
	UN	85.7	88.0	93.8	96.2	94.4	93.5	79.2	62.4	93.3	84.9	95.1	90.1
	FF	92.5	87.7	97.2	95.9	98.5	93.4	95.3	68.6	98.9	88.6	99.0	92.0
T=400	P	98.6	81.5	99.9	94.4	99.3	89.7	98.6	85.6	99.9	96.6	99.9	95.6
	UN	96.2	87.7	99.3	96.0	98.6	93.6	91.9	64.5	99.0	86.2	99.5	91.3
	FF	99.4	87.8	99.9	96.1	100.0	93.5	99.8	69.6	100.0	89.8	100.0	93.4
H:	8												
T=100	P	72.3	76.5	90.9	95.6	91.2	90.7	72.5	38.9	96.0	80.7	95.6	88.3
	UN	72.4	81.9	89.2	95.8	90.5	93.0	63.7	26.3	92.2	70.0	93.8	88.0
	FF	77.9	82.4	91.5	95.9	95.9	92.8	85.4	34.0	97.1	76.5	97.7	88.1
T=400	P	97.4	72.1	100.0	94.6	99.2	89.1	91.9	49.2	99.9	88.8	100.0	91.5
	UN	92.1	82.1	98.9	95.7	98.5	93.6	85.1	29.5	99.3	73.4	99.8	90.3
	FF	97.0	82.2	99.8	96.2	100.0	93.3	99.1	33.0	100.0	77.1	100.0	90.6
H:	12												
T=100	P	56.3	71.4	83.2	95.7	84.8	90.2	50.9	16.1	91.2	63.0	91.8	81.5
	UN	64.2	76.8	86.1	95.2	83.9	92.0	59.1	18.5	93.7	64.1	92.4	86.4
	FF	65.4	78.3	86.1	96.0	90.1	92.7	77.3	20.6	95.3	68.2	95.7	86.6
T=400	P	96.5	66.1	99.9	94.8	99.3	88.8	79.5	21.7	99.7	77.0	100.0	87.6
	UN	89.7	76.9	99.2	95.4	98.6	93.2	82.6	15.0	99.8	66.7	99.9	88.6
	FF	94.4	77.3	99.8	95.6	100.0	92.8	98.6	15.7	100.0	69.1	100.0	88.7

Notes: 10,000 samples generated from VAR(4) for three variables (P, UN, FF) with break in parameters and var-cov matrix of errors. Model has stable parameters and stable var-cov matrix of errors. Each estimated VAR or local projection (LP) on these 10,000 samples generates a forecast error variance (which includes estimation uncertainty) for the forecast path of length h, and hence the sets of bands (marginal, Bonferroni, and Scheffé) used in the analysis. Similarly, each estimated model generates q forecast paths whose associated error paths are used to generate a forecast error variance for the forecast path and hence the set of bands (marginal emp, Bonferroni emp, and Scheffé emp). Hence 10,000 actual paths are then compared with each set of 10,000 bands to determine the appropriate coverage rates. FWE control stands for "family-wise error rate" and simply computes the proportion of paths strictly inside the bands. FDR control instead is the proportion of forecast paths whose Mahalanobis distance attains a value that is lower than the chi-square statistic for probability equal to nominal coverage and degrees of freedom equal to H. See text for more details.

Table 12. Coverage of alternative bands for US inflation and growth

FWE control

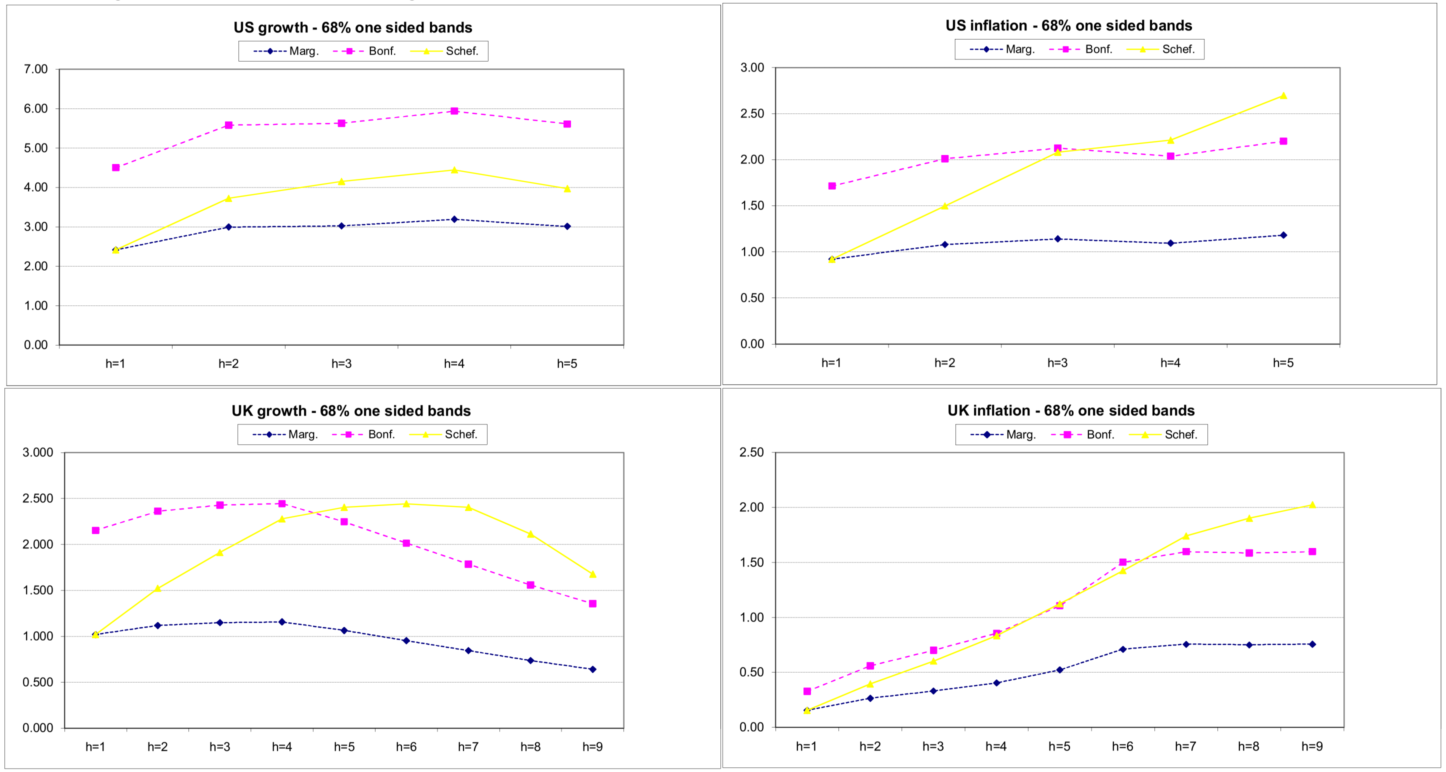
	Inflation									*Growth*								
nominal coverage	50			68			95			50			68			95		
	Marg	Bonf	Schef	Marg	Bonf	Schef	Marg	Bonf	Schef	Marg	Bonf	Schef	Marg	Bonf	Schef	Marg	Bonf	Schef
vintage 1	12	76	44	33	80	60	85	96	92	19	67	25	32	73	44	77	91	80
vintage 2	16	80	39	35	87	56	89	96	93	19	65	32	35	75	47	76	99	87
vintage 3	15	80	44	32	83	56	88	92	93	15	69	33	31	75	48	77	95	85
vintage 4	16	79	48	37	83	59	88	93	93	15	72	32	31	80	51	84	96	88
vintage 5	15	76	44	39	80	57	84	89	88	16	68	35	33	76	49	77	96	91
vintage 6	15	67	39	36	72	55	73	88	85	13	69	35	33	76	51	77	96	93
vintage 7	12	67	40	39	72	59	76	88	87	15	68	32	31	72	52	79	96	89
vintage 8	15	68	37	37	72	55	77	87	85	13	69	32	33	76	52	81	96	95
vintage 9	16	71	39	35	79	56	80	92	87	11	69	36	36	75	55	81	97	93
vintage 10	24	65	45	39	77	63	80	92	87	11	69	32	33	77	49	81	97	95
final release	16	71	47	31	79	60	88	92	93	9	64	27	27	72	47	73	100	85

FDR control

	Inflation									*Growth*								
nominal coverage	50			68			95			50			68			95		
	Marg	Bonf	Schef	Marg	Bonf	Schef	Marg	Bonf	Schef	Marg	Bonf	Schef	Marg	Bonf	Schef	Marg	Bonf	Schef
vintage 1	4	84	61	28	89	80	91	96	96	31	76	49	43	84	67	85	97	95
vintage 2	12	85	63	39	91	83	92	95	93	29	68	53	43	83	64	87	96	95
vintage 3	11	81	61	37	87	79	89	93	93	29	73	49	41	89	68	91	96	92
vintage 4	13	77	63	32	89	76	91	95	93	29	85	52	39	91	75	91	96	92
vintage 5	12	81	59	40	87	77	89	92	91	29	72	49	43	85	67	85	97	93
vintage 6	16	75	59	36	77	68	83	93	88	27	76	51	41	81	71	84	97	95
vintage 7	11	76	65	32	79	72	81	89	89	25	69	47	37	81	71	88	99	96
vintage 8	20	75	63	43	79	71	81	91	91	25	73	52	39	83	73	85	97	96
vintage 9	16	71	59	45	80	73	83	92	92	25	73	55	43	79	71	84	96	96
vintage 10	25	72	61	51	79	72	81	89	89	24	73	53	47	80	76	84	97	97
final release	3	65	67	16	73	75	75	88	91	32	81	47	40	83	67	85	99	99

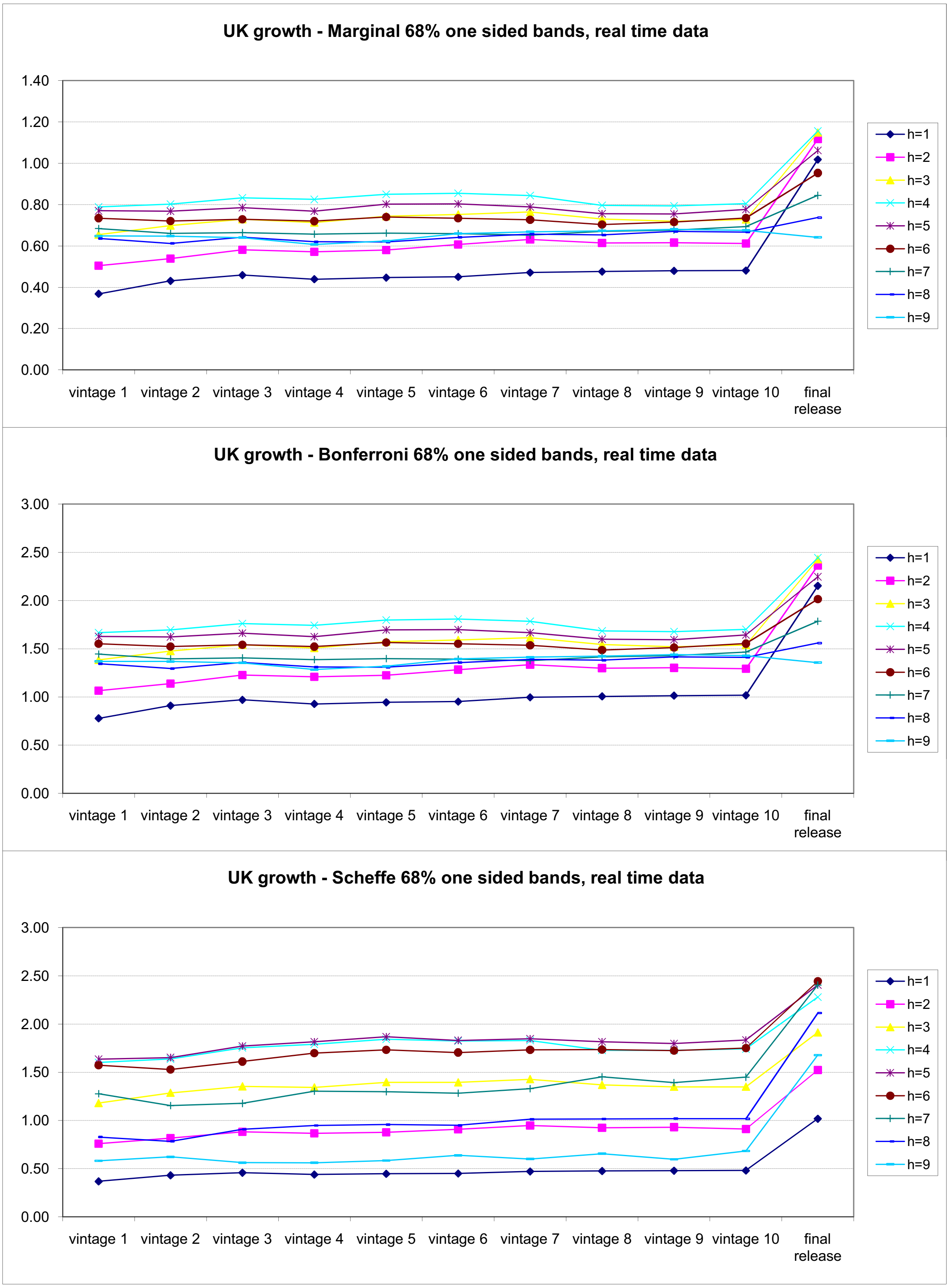
Note: The Table reports the coverage rates of alternative bands for forecast paths. Sample 1974q2-1985q1 is used for estimation of first variance covariance matrix of forecast errors. First forecast path used starts in 1985q2. Exercise is repeated in rolling fashion. Last availabe forecast path comes from 2003q4, so that the coverage of bands for 75 path forecasts is evaluated.

Figure 1. Alternative bands for US and UK growth and inflation forecasts



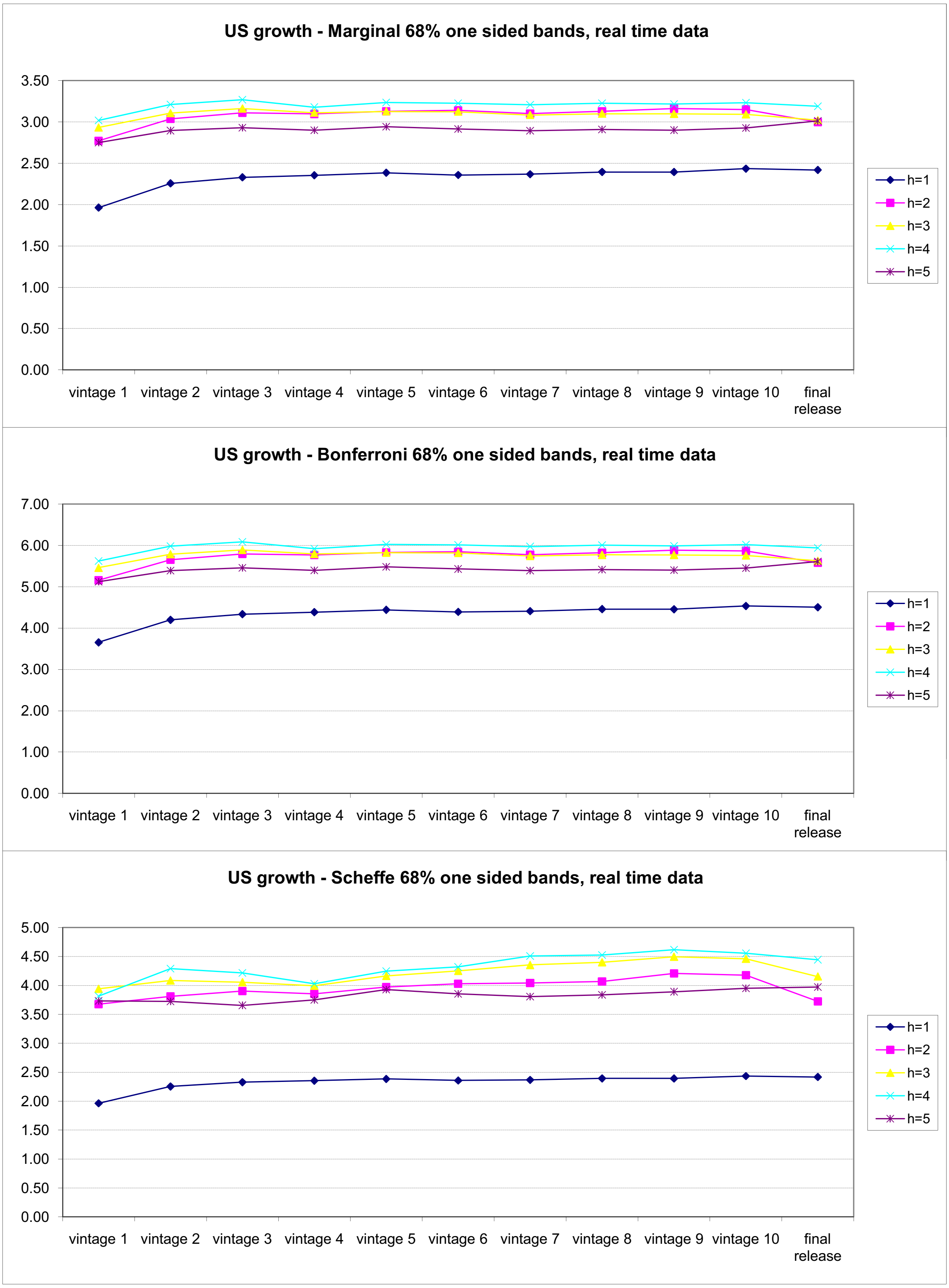
Note: US forecasts by the Federal Reserve, quarterly sample 1974q2-2003q4.
UK forecasts by the Bank of England, quarterly sample 1998q1-2008q2 for growth, 1998q1-2009q2 for inflation.
See text for details.

Figure 2. Alternative bands for UK growth forecasts, real time data



Note: UK forecasts by the Bank of England, quarterly sample 1998q1-2008q2. See text for details.

Figure 3. Alternative bands for US growth forecasts, real time data



Note: US forecasts by the Federal Reserve, quarterly sample, 1974-2003.

European University Institute
European University Press