Joint Prediction Bands for Macroeconomic Risk Management

Farooq Akram, Andrew Binning, Junior Maih

BI NORWEGIAN
BUSINESS SCHOOL

Joint Prediction Bands for Macroeconomic Risk Management[☆]

Farooq Akram[a], Andrew Binning[b], Junior Maih[b,c]

[a]*Research Department, Norges Bank*
[b]*Monetary Policy Department, Norges Bank*
[c]*CAMP, BI Norwegian Business School*

May 6, 2016

Abstract

In this paper we address the issue of assessing and communicating the joint probabilities implied by density forecasts from multivariate time series models. We focus our attention in three areas. First, we investigate a new method of producing fan charts that better communicates the uncertainty present in forecasts from multivariate time series models. Second, we suggest a new measure for assessing the plausibility of non-central point forecasts. And third, we describe how to use the density forecasts from a multivariate time series model to assess the probability of a set of future events occurring. An additional novelty of this paper is our use of a regime-switching DSGE model with an occasionally binding zero lower bound constraint, estimated on US data, to produce the density forecasts. The tools we offer will allow practitioners to better assess and communicate joint forecast probabilities, a criticism that has been leveled at central bank communications.

1. Introduction

Inflation targeting central banks increasingly use fan charts to communicate the uncertainty associated with their forecasts for key macroeconomic variables, especially inflation and the output gap (see Franta et al., 2014, for an overview). Fan charts are typically constructed from sequences of prediction intervals calculated using marginal forecast densities. As a result the fan charts quantify the uncertainty in a forecast for a given variable at a specific point in time. They do not communicate how likely a given set of trajectories for different variables will be in different time periods, that is they do not communicate the joint probabilities implied by forecasts produced using multivariate time series models. Moreover,

---

[☆]This Working Paper should not be reported as representing the views of Norges Bank. The views expressed are those of the authors and do not necessarily reflect those of Norges Bank. We have received valuable comments from seminar participants at Norges Bank. Corresponding author: andrew.binning@norges-bank.no.

these fan charts do not convey the probabilistic trade-offs policy makers are faced with, as Leeper (2003) points out in his assessment of central bank communications.[1] This is especially relevant for central banks that pursue multiple policy objectives, such as flexible inflation targeting central banks and the US Federal Reserve with its dual policy mandate.

In this paper we investigate tools for improving the quantification and communication of the joint forecast probabilities exhibited by density forecasts produced using multivariate time series models. More specifically we concentrate our efforts in three areas. First, we investigate a new method of producing fan charts that better communicates the forecast uncertainty present in density forecasts produced using multivariate time series models. Second, we propose a new measure for assessing the plausibility of non-central point forecasts. And third, we show how to use the density forecasts to assess the probability of a set of future events occurring. We briefly discuss each of these areas in turn.

Several methods have been proposed for constructing fan charts using the joint probabilities from a set of density forecasts. Jordà & Marcellino (2010) provide one such methodology although they assume the density forecasts are asymptotically normal. Wolf & Wunderli (2012) show that if the density forecasts are not asymptotically normal then Jordà & Marcellino's (2010) method may understate the forecast uncertainty. Kolsrud (2015) departs from the assumption of asymptotic normality by employing non-parametric distance measures to construct fan charts that communicate the joint predictive probabilities in a frequentist setting. We adopt Kolsrud's (2015) methodology and extend it to a Bayesian setting.

One particular feature of fan charts produced using this approach is that they are relatively wide (see Kolsrud, 2015). As we discuss later in this paper, because we use a Bayesian approach, this can be addressed through soft conditioning as in Maih (2010), or through entropic tilting as in Robertson et al. (2005). We present an example where we show how entropic tilting can be used to narrow the density forecasts.

In addition to producing fan charts, many central banks communicate forecast uncertainty through point forecast scenarios. Leeper (2003) has also called for the assessment of the plausibility of these point forecasts. Doan et al. (1986) have proposed a plausibility index based on the size of the shocks used to construct non-central point forecasts. Leeper & Zha (2003) create a univariate modesty index based on the distance between alternative point forecasts and the mean or central forecast. They assume asymptotic normality of the predictive densities when determining if certain policy interventions are modest or not. Adolfson et al. (2005) extend this idea to a multivariate context also assuming normality. However, Wolf & Wunderli's (2012) criticism of asymptotic normality also applies here, that is, a multivariate modesty or plausibility index based on normality could understate uncertainty and hence the plausibility of some point forecasts. To circumvent this problem we use the non-parametric distances from Kolsrud's (2015) fan chart methodology to create a new measure for assessing the plausibility of point forecasts. Like Adolfson et al. (2005) our

---

[1]Leeper (2003) has called for better communication of joint predictive probabilities in his assessment of the monetary policy reports of the central banks of Sweden, New Zealand and the United Kingdom.

measure is multivariate, but has the additional advantage that it is non-parametric and free from distributional assumptions.

The methodology we follow allows us to assess the probability of sets of future events occurring. If the set of events is non-central or bounded, assessing these probabilities using fan charts is not always possible. However we can always assess these probabilities using the draws from the density forecasts used to construct the fan charts. Applying methods outlined by Garratt et al. (2002) we show how to answer questions like: "what is the probability that GDP growth is between 0 and 1% for the first year of the forecast horizon?", "what is the probability of a recession (two consecutive quarters of negative growth) over the forecast horizon?", or "what is the probability that interest rates are at zero and inflation negative for at least one year, over the forecast horizon?".

An additional novelty of our paper is the use of a regime-switching DSGE model with an occasionally binding zero lower bound constraint to produce the density forecasts. We use the simple New Keynesian model of Binning & Maih (2016), estimated on US data between 1985Q1 and 2015Q2, which allows us to handle the zero lower bound period and produce forecasts for the US. Moreover this model produces density forecasts that are not in general asymptotically normal, a property we use to justify our non-parametric approach to assessing and communicating forecast uncertainty.

The remainder of this paper is organized as follows. The next section defines multivariate and time-simultaneous fan charts and related concepts based on Kolsrud (2015) and the fan chart methodology. The model, and solution and estimation methods are briefly described in Section 3.1. Section 3.2 discusses the multivariate and time-simultaneous fan charts and compares them with typical prediction interval-based fan charts. In Section 4 we discuss techniques to rank the plausibility of different paths. These techniques may be useful for assessing the likelihood of particular paths of different variables. Section 5 discusses how to calculate the probability of a set of events occurring and Section 6 concludes. The Matlab codes used to produce the fan charts under the different methodologies are made available as part of the RISE toolbox.[2]

2. Concepts, Definitions and Methodology

In this section, we first define uncertainty measures relevant for single and multiple variables as in Kolsrud (2015). Then we describe how to produce fan charts using the conventional method, based on the marginal predictive densities, and using the multivariate and time simultaneous (MVTS) prediction bands of Kolsrud (2015).

These definitions are useful for describing the construction of fan charts from density forecasts under the different methodologies. In essence, the methods aim to ascribe uncertainty to a single entity, but each defines the entity differently. When constructing fan charts using the marginal predictive densities, as is commonly done, the single entity is a period-specific forecast of a single variable. When constructing fan charts using the MVTS

---

[2]https://github.com/jmaih/RISE_toolbox

prediction bands, the forecasts of all variables of interest over the entire forecast horizon are treated as the single entity. Graphically, instead of assigning different levels of uncertainty to a period-specific "dot", one assigns uncertainty measures to a multidimensional entity spanning the entire forecast horizon when using the MVTS prediction bands. This is also referred to as a "hyperrectangular box". We compare and contrast the different methods of constructing fan charts in Section 3.2.

2.1. Definitions and concepts

Following Kolsrud (2015) we define relevant forecast uncertainty concepts as follows:

Definition 1. *Prediction Interval:* $PI_t = [l_t, h_t]$ *is a Prediction Interval with a coverage probability $1-a$ for a single random variable y at time t:*

$$\Pr(y_t \in PI_t) = \Pr(l_t \leq y_t \leq h_t) = 1 - a, \tag{1}$$

where l_t and h_t are the lower and upper bounds of the interval respectively, and $0 \leq a < 1$. This prediction interval is expected to include the actual value of y at time t with probability $1 - a$.

Definition 2. *Prediction Band:* $PB = [\boldsymbol{l}, \boldsymbol{h}] = ([l_t, h_t])_{t=1}^T$ *is a time-simultaneous Prediction Band with coverage probability $1 - \alpha$:*

$$\Pr(\mathbf{y} \in PB) = \Pr(\mathbf{l} \leq \mathbf{y} \leq \mathbf{h}) = \Pr\left(\bigcap_{t=1}^{T}(l_t \leq y_t \leq h_t)\right) = 1 - \alpha, \tag{2}$$

where $\mathbf{y} = (y_1, y_2, ... y_T)$ is a forecast path of the variable y over T periods, a so called univariate path; $\mathbf{l} = (l_1, l_2, ... l_T)$ and $\mathbf{h} = (h_1, h_2, ... h_T)$ are the lower and upper bound paths of $\mathbf{y}$, respectively, while $0 \leq \alpha < 1$. According to (2), the prediction band PB is likely to include the actual path of $\mathbf{y}$ in all T periods (simultaneously) with probability $1 - \alpha$. The width of individual intervals $[l_t, h_t]$ and corresponding coverage probabilities defining a prediction band can differ across time periods.

Definition 3. *Uniform Prediction Band: Such a band has constant coverage probability across time periods while the width of corresponding intervals can vary:*

$$\Pr(l_t \leq y_t \leq h_t) = 1 - a \text{ for } t = 1, 2, ... T. \tag{3}$$

In the case of non-uniform prediction bands, each prediction interval constituting a prediction band has its period-specific coverage probability: $\Pr(l_t \leq y_t \leq h_t) = 1 - a_t$.

A prediction band can be seen as a sequence of prediction intervals. However, while a_t denotes the probability of a point value outside its interval at time t, α is the probability that a path breaks out of the band at any single period or in several periods. Thus, $1 - a$ is the pointwise coverage probability while $1 - \alpha$ is the simultaneous coverage probability. As $\alpha \geq a$, the simultaneous coverage probability cannot exceed the pointwise coverage

probability: $1 - \alpha \leq 1 - a$. In the case of non-uniform prediction bands, the simultaneous coverage probability has to be smaller than the smallest pointwise coverage probability, which implies $\alpha \geq \max_t a_t$.

In a multivariate setting, one would be interested in several variables and their co-dynamics. Let $\mathbf{y}_t = (y_1, ...y_G)'_t$ be a vector of G variables at time t and Y be a $G \times T$ matrix where T denotes the total time periods:

$$
Y = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ . \\ . \\ \mathbf{y}_G \end{pmatrix} = \begin{pmatrix} y_{1,1} & . & . & y_{1,T} \\ y_{2,1} & . & . & y_{2,T} \\ . & . & . & . \\ . & . & . & . \\ y_{G,1} & . & . & y_{G,T} \end{pmatrix}
$$

Definition 4. *Multivariate Prediction Box: $MBP = [L, H] = ([\mathbf{l}_g, \mathbf{h}_g])_{g=1}^G$ is a multivariate time simultaneous prediction (hyperrectangular) box with coverage probability $0 < 1 - A \leq 1$ for a random multivariate trajectory $Y = (y_{g,t})$:*

$$
\begin{aligned}
\Pr(\mathbf{Y} \in MPB) = \Pr(\mathbf{L} \leq \mathbf{Y} \leq \mathbf{H}) &= \Pr\left(\bigcap_{g=1}^G (\mathbf{l}_g \leq \mathbf{y}_g \leq \mathbf{h}_g) \right) \\
&= \Pr\left(\bigcap_{g=1}^G \bigcap_{t=1}^T (l_{g,t} \leq y_{g,t} \leq h_{g,t}) \right) = 1 - A,
\end{aligned}
$$

where $\mathbf{L} = (l_{g,t})$ and $\mathbf{H} = (h_{g,t})$ are $G \times T$ matrices containing the lower and upper paths of the G random variables over T periods. The probability of any of the G variables breaking out of the corresponding prediction band, defined over T periods, in one or more periods is denoted by A. The difference between a and A is that the former denotes the probability of a single "dot" being outside its prediction interval, while the latter denotes the probability of a single time-extensive multidimensional entity being out of its prediction space. It therefore takes only one single scalar value $y_{g,t}$ to be outside its prediction space for the multivariate trajectory not to be completely inside the box.

A multivariate prediction box is useful when a decision maker cares about attaining targets and satisfactory predictions of several variables. In such cases, an unsatisfactory misprediction, over one or several periods, of even a single variable, from a set of variables, could be considered a failure. One example of when there are arguably two equally important objectives is the Fed's pursuit of a dual objective: price stability and maximum employment.

Kolsrud (2015) utilizes the following Chebyshev concepts in defining and deriving central and non-central forecasts over multiple periods. The presumption is a sample of N forecasted paths for a variable y, contained in a matrix $\mathbf{y}_N$ of dimension $T \times N$, where T is the forecast horizon.

Chebyshev's interval and distance may be defined as follow:

Definition 5. *Chebyshev Interval: $CI(\gamma; \mathbf{y}_N) = CI(1; \mathbf{y}_{\langle M \rangle}) = [l, h]$ where the lower bound is $l = \min_m y_{\langle M \rangle}$ and the upper bound is $h = \max_m y_{\langle M \rangle}$. The interval has, by construction,*

an in-sample coverage of $0 < \gamma = M/N \leq 1$, with $1 \leq M \leq N$, for a randomly selected y_n, from the sample $\mathbf{y}_N$. $\mathbf{y}_{\langle M \rangle} = \left\{ y_{\langle m \rangle} \right\}_{m=1}^{M}$ denotes the subsample of paths containing the M most central paths, where $y_{\langle m \rangle}$ denotes the value in the sample with the mth smallest distance z, while M/N is the ratio between the M most central paths and the total number of (simulated) paths, N.

Definition 6. *Chebyshev distance: $c_n = c(y_n; \bar{y}, s) = \max_t z_{n,t}$, where $z_{n,t} = |y_{n,t} - \bar{y}_t|/s_t$ is the standardized distance, y_n is the nth simulated path of variable y, the period-specific mean of the N-paths of variable y is defined as $\bar{y} = N^{-1} \sum_{n=1}^{N} y_n$ while the standard deviation is defined as $s_t = \left[(N-1)^{-1} \sum_{n=1}^{N} (y_{n,t} - \bar{y}_t)^2 \right]^{1/2}$.*

Definition 7. *Multivariate Chebyshev distance $c_n = c(\mathbf{y}_n; \bar{\mathbf{y}}, \mathbf{s}) = \max_{g,t} z_{n,g,t}$, where $z_{n,g,t} = |y_{n,g,t} - \bar{y}_{g,t}|/s_{g,t}$; the path of a variable y over T periods is $\mathbf{y}_n = (y_{n,1}, \ldots, y_{n,T})$; a period-specific mean of variable y is $\bar{y}_t = N^{-1} \sum_{n=1}^{N} y_{n,t}$ where N is the total number of paths; while the corresponding vector of period-specific means of variable y over the forecast horizon T is $\bar{\mathbf{y}} = (\bar{y}_1, \ldots, \bar{y}_T)$. The period-specific mean and standard deviation of a variable y_g are defined as $\bar{y}_{g,t} = N^{-1} \sum_{n=1}^{N} y_{n,g,t}$ and $s_{g,t} = \left[(N-1)^{-1} \sum_{n=1}^{N} (y_{n,g,t} - \bar{y}_{g,t})^2 \right]^{1/2}$, respectively.*

2.2. Methodology

We consider density forecasts produced using a multivariate time series model estimated with Bayesian methods.[3] The density forecasts are constructed by simulating the model multiple times, taking into account both shock and parameter uncertainty. Each simulation begins from initial conditions, estimated using historical data and a Kalman filtering procedure, and involves randomly drawing a vector of parameters from the posterior parameter distribution and a sequence of shocks from the estimated historical distribution.

For a given set of density forecasts, fan charts are typically constructed using a sequence of prediction intervals (see Cogley et al., 2005). These are calculated by slicing the multivariate forecast density into a sequence of marginal distributions. The resulting sequence of prediction intervals is reconstituted into a matrix of time series data, where the cross-sectional elements represent the percentiles. This new time series is plotted to create the fan charts. This method of illustrating the uncertainty in a forecast fails to take into account the time and covariation that underlie forecasts generated by multivariate time series models.

The MVTS prediction bands method developed by Kolsrud (2015) aims to describe graphically the joint probabilities that underlie a multivariate density forecast. The method calculates the probability of each draw used to construct the density forecast and then constructs the fan charts by tracing out the envelope of all draws that fit within a given quantile. The enveloped draws form a (quantile-specific) multivariate prediction box. More precisely, the $G \times T$ matrices of forecasts (G variables over T periods) are concatenated into row vectors,

---

[3]Kolsrud (2015) demonstrates his algorithm in a frequentist environment and generates forecast uncertainty through a bootstrapping procedure.

normalized by their standard deviations (calculated from the N draws) and then the point-wise distances for each draw relative to the mean forecast (calculated from the N draws) are summed together. This creates a distance measure, known as the multivariate Chebyshev distance, for each draw. The N draws are then ranked based on their distance. The draws with a smaller distance are deemed to be more likely than those with a larger distance. The percentile bands used to construct the fan chart, also known as the multivariate prediction box, are then obtained from the ranking of the draws and by tracing out the envelope of all draws that fit within the percentile band in question.

We make the observation that the MVTS prediction bands can be interpreted as a graphical representation of a non-parametric multivariate modesty or plausibility index in the spirit of Leeper & Zha (2003) and Adolfson et al. (2005). In essence, the MVTS bands are just the envelope of the draws that fall within a certain level of modesty or plausibility as measured by the multivariate Chebyshev distance. We make use of this observation to propose a method for evaluating the plausibility of individual point forecasts. We illustrate this method in Section 4.

3. Model, Forecasts and Fan Charts

We demonstrate the empirical difference between typical interval-based fan charts with time-simultaneous fan charts using the simple New Keynesian DSGE model from Binning & Maih (2016). We use the version of the model where a combination of factors can shift to generate the ZLB steady state. This regime-switching model is chosen because it can handle contemporary US data including the ZLB period, although the procedures developed in this paper are compatible with any time series model.

We generate density forecasts that incorporate both parameter and shock uncertainty. The associated fan charts are produced using MVTS prediction bands. An advantage of this procedure is that density forecasts from the model are not required to be (asymptotically) normal, as in Jordà & Marcellino's (2010). Such a requirement is likely to be violated when a regime-switching model, which can imply multimodal density forecasts, is employed. For comparison, fan charts based on prediction intervals are constructed as in Cogley et al. (2005). Before explaining the computational aspects of the fan charts' production, we briefly describe the model, which is explained in detail in Binning & Maih (2016).

3.1. Model

The model consists of a representative household that receives utility from consumption and disutility from working. Households supply labor services to firms and receive labor income in return. They receive dividends from the firms they own and trade bonds that have net zero supply. There is a continuum of firms, normalized to unit mass, that produce intermediate goods using labor services and a common neutral technology. Firms are subject to a quadratic adjustment cost when changing prices (see Rotemberg, 1982, for example). The intermediate goods are aggregated into final goods by a perfectly competitive aggregator firm. The model is closed by the monetary authority, which sets the policy rate according

to a Taylor-type rule when interest rates are not at the ZLB. The full set of model equations is given in Appendix A.

The model is solved using a first-order perturbation method for regime-switching models (see Maih, 2015) and estimated using Bayesian methods (see Alstadheim et al., 2013) on US data between 1985Q1 and 2015Q2. The observable variables include the log change in per capita GDP, GDP deflator inflation and the Fed funds rate. A full description of the data can be found in Appendix B. The estimation details and results can be found in Appendix C.

3.2. MVTS Prediction Bands vs Interval Prediction Bands

In this section we produce fan charts using MVTS prediction bands as outlined in Kolsrud (2015) and compare their properties against fan charts produced using a sequence of prediction intervals. We have data up to 2015Q2 while forecasts are made for the five year period: 2015Q3–2020Q2. Our density forecasts consist of 200,000 draws. We simultaneously draw parameter vectors from the posterior parameter distributions and random sequences of shocks from the estimated historical shock distributions.

Before investigating the properties of the MVTS prediction bands, we highlight some of the properties of the density forecasts from our model. Because we use a regime-switching model to produce the forecasts, multimodal density forecasts are generated. That is, our density forecast for interest rates exhibits two modes, the first centered around the effective lower bound and the second centered around the interest rate that would be set under a Taylor-type rule. Figure 1 displays the multimodal interest rate density forecasts.

We produce two sets of fan charts from our forecast densities. For both sets of fan charts we calculate bands at the 10%, 30%, 50% and the 68% levels. The first set of fan charts produced using the MVTS prediction bands are plotted in the left column of Figure 2. The second set of fan charts produced using the sequence of prediction intervals are plotted in the right column of Figure 2. The shadow rate refers to the Taylor-rule interest rate under the ZLB regime. Both sets of figures are plotted on axes with the same scale to make comparisons of their coverages easier. To make further comparisons of the fan charts' coverage and to ease their interpretation, we take a single draw from the 200,000 draws used to construct the density forecasts, and plot it on top of the fan charts in Figure 2.

Comparing the left and right columns of Figure 2 we note that the fan charts are substantially wider when produced using the MVTS prediction bands. This is as expected as they embed the cumulative and multivariate uncertainty present in forecasts from multivariate time series models, as pointed out by Kolsrud (2015).

MVTS prediction bands of the same color contain draws of the same level of uncertainty over time and across variables. That is, draws used to construct a band of a particular color will always fall within those bands across time and through multivariate space. The left column of Figure 2 makes this apparent: the single draw (red line) always falls within the darkest band (10% level) at each point in time and for each variable.

However, the same can not be said for the coverage of the fan charts produced using the conventional method. In the right column of Figure 2 the same draw crosses all the bands and in some periods falls outside the 68% interval. If we were to assess the probability of
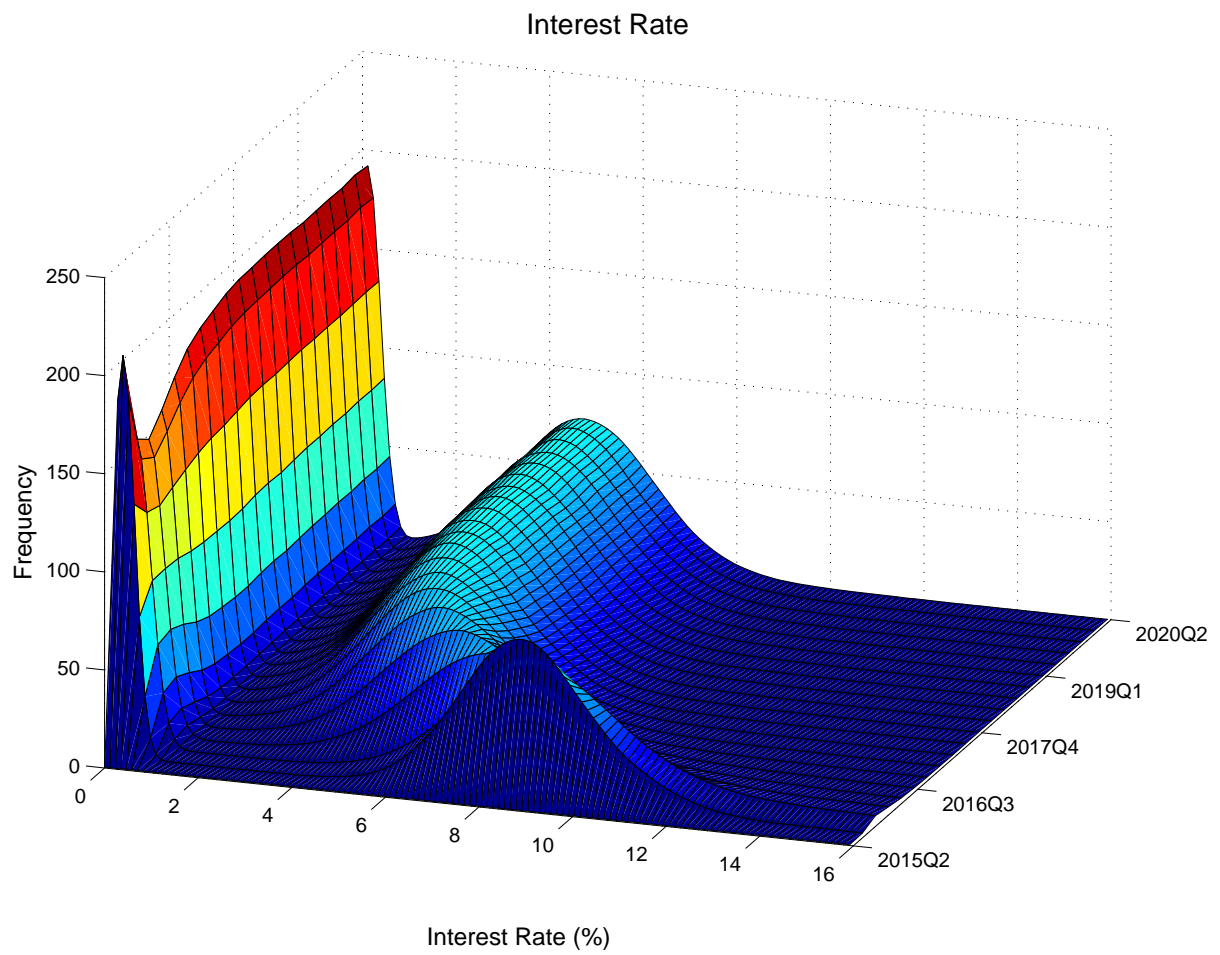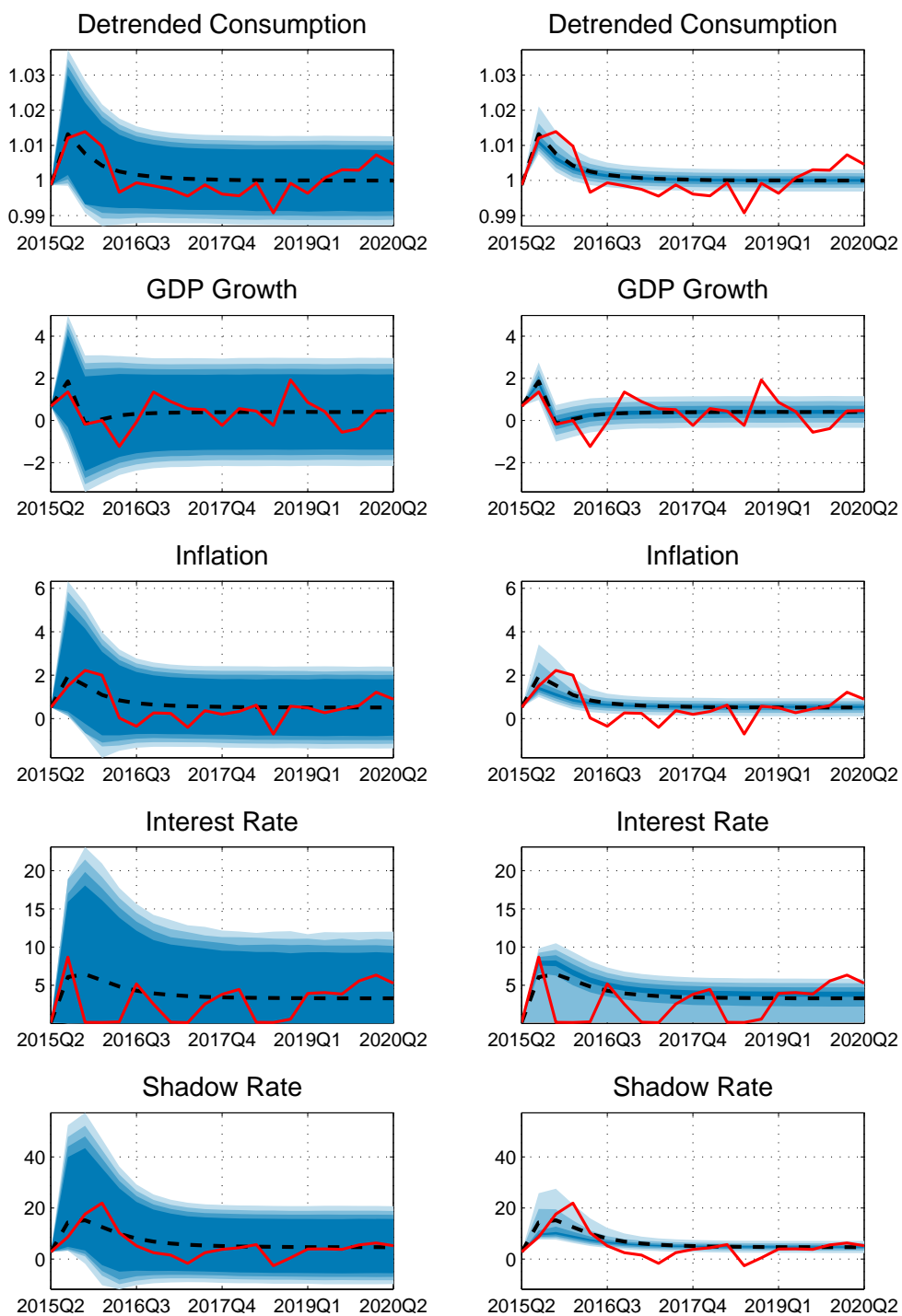
Figure 1: Multi-Modal Forecasts

Figure 2: Multivariate and Time-Simultaneous Prediction Bands vs. Prediction Interval Bands

this particular forecast using the fan charts constructed with the conventional method, we might reach the conclusion that this is an implausible forecast. However, if we assess the probability of this forecast occurring using the MVTS prediction bands, we would correctly conclude that this forecast is quite plausible.

It is well known that density forecasts produced using DSGE models are usually wider than is considered reasonable due to model misspecification (see Gerard & Nimark (2008), Wolters (2013) and Diebold et al. (2015)). Moreover, practitioners may have extraneous information making them less uncertain about the forecasts, compared with the model. To remedy the wider-than-desired density forecasts, practitioners could adjust the forecasts using soft conditioning (see Maih, 2010) or entropic tilting (see Robertson et al., 2005).

In this paper we opt for entropic tilting and demonstrate how it can be used to narrow the density forecasts and fan charts. Entropic tilting involves reweighting the densities so that they match a chosen set of moment conditions. We briefly describe the procedure below.

Let f represent the (sample) distribution of N forecast draws so that

$$f = \{\mathbf{y}_n\}_{n=1}^{N}$$

where $\mathbf{y}_n = y_{n,t}, \ldots, y_{n,T}$ is the nth draw. Note that we drop the g subscripts. We wish to impose the following moment conditions on our density forecasts

$$E\left\{\mathfrak{g}\left(\mathbf{y}\right)\right\} = \bar{\mathfrak{g}}.$$

$\mathbf{y}$ contains the N forecast draws, E is the expectations operator, $\mathfrak{g}(\bullet)$ is a set of moment generating functions and $\bar{\mathfrak{g}}$ is a set of moments that characterize the target distribution. We can impose the moment conditions by minimizing the Kullback-Leibler distance subject to our moment conditions

$$\min_{\tilde{f} \in F} \mathrm{KLIC}\left(\tilde{f}, f\right) \text{ subject to } E_{\tilde{f}}\left\{\mathfrak{g}\left(\mathbf{y}\right)\right\} - \bar{\mathfrak{g}}.$$

The Kullback-Leibler distance between the target distribution $\tilde{f}$ and our forecast distribution f is defined as follows

$$\mathrm{KLIC}\left(\tilde{f}, f\right) = \sum_{n=1}^{N} \tilde{\pi}_n \log\left(N\tilde{\pi}_n\right)$$

$$= \log\left(N\right) + \sum_{n=1}^{N} \tilde{\pi}_n \log\left(\tilde{\pi}_n\right)$$

where $\tilde{\pi}_n$ is the new set of weights, that sum to 1, and $\pi = 1/N$ is the old set of weights. The expectation for the reweighted forecast distribution is given by

$$E_{\tilde{f}}\left\{\mathfrak{g}\left(\mathbf{y}\right)\right\} = \sum_{n=1}^{N} \tilde{\pi}_n \mathfrak{g}\left(\mathbf{y}_n\right).$$

11

The new set of weights are determined by the following set of equations

$$\pi_n^* = \frac{\exp\left(\gamma^{*\prime}\mathfrak{g}\left(\mathbf{y}_n\right)\right)}{\sum_{n=1}^{N}\exp\left(\gamma^{*\prime}\mathfrak{g}\left(\mathbf{y}_n\right)\right)}$$

$$\gamma^* = \arg\min \gamma \sum_{n=1}^{N}\exp\left(\gamma^{\prime}\left(\mathfrak{g}\left(\mathbf{y}_n\right)-\bar{\mathfrak{g}}\right)\right)$$

For illustration, we adjust the forecast densities so that the variance of interest rates and inflation are 1/4 of those implied by the model simulations. The fan charts for the MVTS prediction bands reweighted using entropic tilting are plotted in the right column of Figure 3 while the unadjusted MVTS prediction bands are plotted in the left column for comparison.[4]

---

[4]Entropic tilting is one of the techniques that may be used to align forecast densities from MVTS prediction bands with model users' beliefs about forecast uncertainty. This may be a more appropriate approach than using fan charts constructed with the conventional approach, using prediction intervals, merely for the sake of obtaining narrower fan charts.

Figure 3: MVTS Prediction Bands Before and After Entropic Tilting
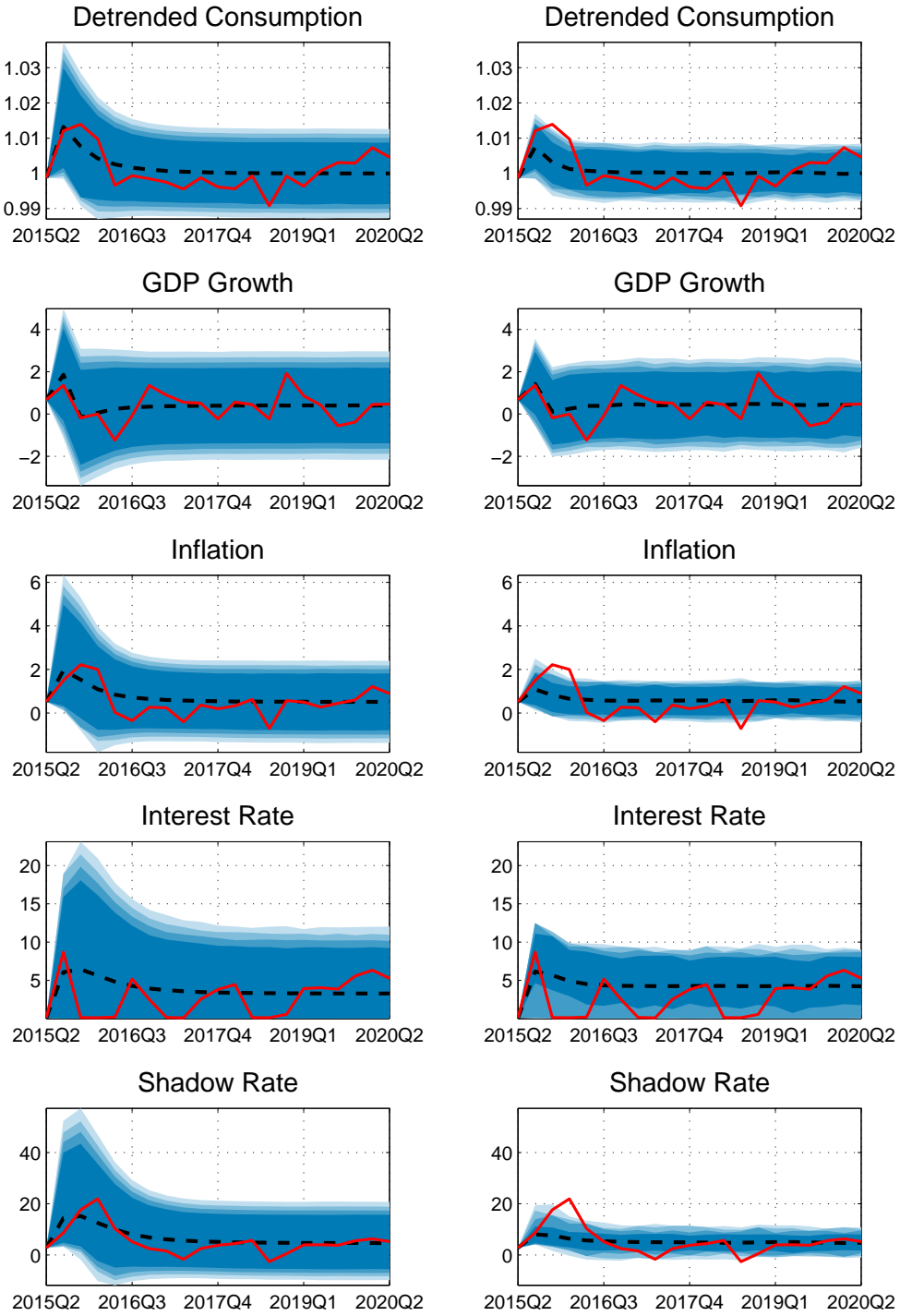
4. Assessing Point Forecast Probabilities using a Multivariate Non-parametric Modesty Index

The MVTS prediction bands provide a useful tool for graphically assessing the plausibility of a set of point forecasts. However, it may be difficult to assess the plausibility of competing point forecasts if they fall within the same colored bands. Moreover, a finer gradation between the bands may be required for assessing a given point forecast's plausibility. In this section we demonstrate how the Chebyshev distances used to construct the fan charts can also be applied to evaluate the plausibility of non-central point forecasts i.e. alternative scenarios. This can be used to evaluate the plausibility of a given point forecast relative to a set of density forecasts, or to rank multiple point forecasts (alternative scenarios) in terms of their plausibility.

Many central banks publish non-central point forecasts also known as alternative scenarios to discuss economic and policy implications of non-central risks; see e.g. Norges Bank (2012, p.19). These can be thought of as conditional forecasts. However, the plausibility of such forecasts is seldom assessed, as pointed out by Leeper (2003) in his assessment of the monetary policy reports of the central banks of Sweden, New Zealand and the UK.[5]

Several methods have been proposed for evaluating the plausibility of conditional forecasts. Doan et al. (1986) propose using the size of the shocks used to construct the conditional forecast relative to the estimated historical shock standard deviations to evaluate the forecast plausibility. Their measure assumes that the shocks are normally distributed and that the central forecast has not been judgementally adjusted.

Leeper & Zha (2003) construct a modesty index to evaluate policy interventions. Their index is constructed using the distance between the conditional forecast and the mean forecast normalized by the forecast standard deviations. A policy intervention is deemed modest if the index is less than two, which implies the conditional forecast is within two standard deviations of the forecast distribution.

Adolfson et al. (2005) extend the modesty statistic to multivariate distributions. Jordà & Marcellino (2010) develop a similar measure for assessing the quality of point forecasts. The multivariate modesty index is based on the forecast distribution of the model variables, so it allows for judgementally adjusted density forecasts. However it assumes the forecast densities are normally distributed, an assumption that is easily violated.[6]

We propose using the Chebyshev distances as a metric for evaluating the plausibility of conditional forecasts. This is based on our earlier observation that the MVTS prediction bands can be interpreted as a graphical representation of a multivariate non-parametric modesty index in the spirit of Leeper & Zha (2003) and Adolfson et al. (2005). The same distances used to construct the fan charts can also be used to evaluate the plausibility

---

[5]This criticism may also apply to stress-tests usually conducted by policy institutions to examine the vulnerability of e.g. the financial system to some relatively extreme realisations of economic variables. It may be of interest to assess the plausibility of such tail-risks.

[6]The presence of non-linearities and binding constraints are among some of the reasons a forecast distribution may not be asymptotically normal.

of alternative scenarios. In the case of a single alternative forecast, we can construct its Chebyshev distance and rank it against the draws used to construct the density forecasts. Its position in the ranking determines its plausibility. Likewise in the case of evaluating competing alternative forecasts, we can calculate the Chebyshev distances for each of these forecasts and rank them, or we can determine their ranking against the draws used to construct the density forecasts.

Our measure has two properties in common with the multivariate modesty index of Adolfson et al. (2005). First, our measure is multivariate so it takes into account the entire predictive density implied by the model. And second, it is based on the forecast variable densities, so it allows for judgementally adjusted forecast distributions. Our Chebyshev distance measure has, however, the additional advantage of being non-parametric, making it free from distributional assumptions, especially the assumption of normality. Hence, there is a lower risk of understating the plausibility of particular point forecasts than when measures of plausibility based on a multivariate normal density are employed. This consideration could be important in cases like ours where the model does not in general produce asymptotically normal density forecasts.

In the following we calculate Chebyshev distances to evaluate the relative plausibility of competing alternative forecasts using the model and the density forecasts from the previous sections. We plot two competing alternative forecasts on top of the MVTS prediction bands we derived in the previous section, see Figure 4. In the red scenario, we hold interest rates at 4% for 8 quarters beginning in 2015Q3. In the orange scenario, we hold interest rates at 6% for 8 quarters. Both scenarios are produced using a conditional forecasting algorithm (see Maih, 2010) with all shocks unanticipated. We assume the monetary policy regime is characterized by a Taylor-type rule.

It is not immediately obvious from a visual inspection of Figure 4 which scenario is more plausible. For the majority of the simulations, the forecasts fall within the dark blue band. Calculating the Chebyshev distances for these forecasts we obtain a distance of 3.0564 for the red scenario, which falls in the 54th percentile of the draws used to construct the density forecasts. For the orange scenario we obtain a distance of 4.1655, which falls in the 87th percentile of the draws used to construct the density forecasts. Based on the Chebyshev distances, we would conclude that the red scenario is more plausible than the orange scenario.

Figure 4: Alternative Forecasts

Detrended Consumption

GDP Growth

Inflation

Interest Rate

Shadow Rate
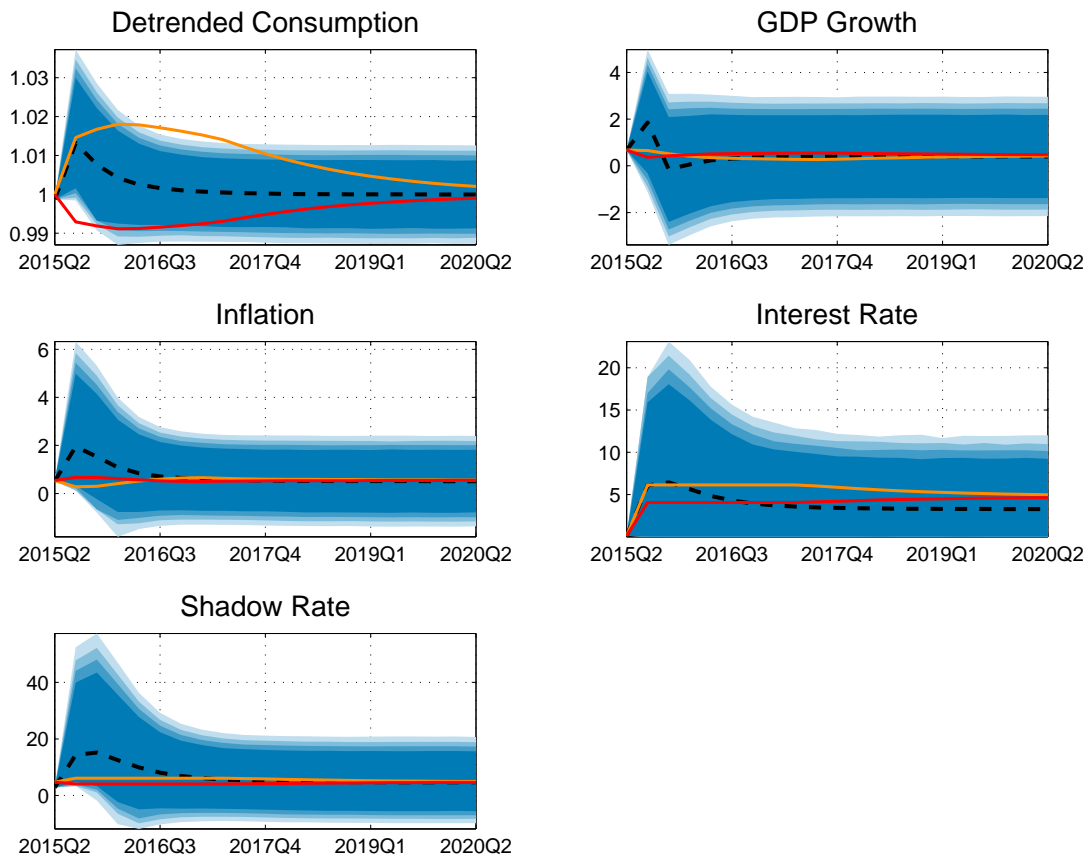
5. Scenario Set Probabilities

This section describes a more general method for constructing the joint probability of a set of events occurring. The method has been proposed by Garratt et al. (2002) in a frequentist setting. The set of events whose probability one would like to assess could be about the values of one or several variables at a single or several points in time. The set of events is usually characterized by an upper bound on the values of one or more variables, a lower bound or both. For example, one may want to infer the probability of inflation remaining within a +1/-1 band of the inflation target for a given time period over the forecast horizon. Another example could be that one wishes to infer the probability of inflation being above target in combination with negative GDP growth over the forecast horizon.

Density forecasts from a multivariate time series model are useful in answering such questions. Accordingly, the probability of a set of events occurring can be calculated by counting the number of draws that are a member of the event set as a fraction of the total number of draws.

More precisely, let B denote a set of events that can occur and N the total number of draws. The probability of B occurring can be derived as:

$$p\left(Y \in B\right) = \frac{\sum_{n=1}^{N} \mathbb{1}_B\left(Y_n\right)}{N},$$

where Y_n are the forecast paths produced from the nth shock and parameter draw combination and

$$\mathbb{1}_B\left(Y_n\right) = \left\{ \begin{array}{ll} 1 & \text{if } Y_n \in B \\ 0 & \text{if } Y_n \notin B \end{array} \right.$$

We demonstrate how to calculate these probabilities through two examples. Using the density forecasts from Section 3.2, we can calculate the probability of the sets of events occurring. In the first example, we derive the probability of interest rates remaining below 1% and GDP growth between 0 and 2% over the period 2016Q2 and 2017Q1. We find that this event occurs in 1,809 draws out of 200,000, implying the probability of this event occurring to be 0.009. In the second example, the question is to assess the probability of interest rates above 1% over the period 2016Q2-2018Q1. This event occurs in 60,581 draws out of 200,000, suggesting the probability to be 0.303.

6. Conclusion

In this paper we present methods for assessing and communicating forecast uncertainty that take into account the joint probabilities implied by multivariate time series models. We focus our attention in three areas. First, we describe a method of constructing fan charts that more accurately describes the uncertainty implied by forecasts produced using multivariate time series models. Second, we create a new measure for assessing the plausibility of non-central point forecasts. And third, we describe how the information in density forecasts can be used to assess the probability of a set of future events occurring.

More specifically we take the multivariate and time simultaneous prediction bands of Kolsrud (2015) and extend the framework to density forecasts constructed using Bayesian methods, that is we produce density forecasts that take into account both shock and parameter uncertainty. We show how the Chebyshev distances used to construct the MVTS bands can also be used to assess the plausibility of individual non-central point forecasts. We then show how the draws used to construct the density forecast can also be used to assess the probability of a set of future events occurring. Our paper is also novel in that we produce all of our forecasts using a regime-switching DSGE model with an occasionally binding zero lower bound constraint, which makes the model suitable to use with contemporary US data.

Moreover, our paper provides a guide for practitioners to answer some of Leeper's (2003) criticisms of inflation reports and the communication of joint probabilities. This is especially useful for central banks facing multiple policy objectives. The codes used to produce the analysis in this paper are available in Matlab as part of the RISE toolbox.[7]

Appendix A. Model

$$\log\left(A_t\right) = \rho_A \log\left(A_{t-1}\right) + \sigma_A \varepsilon_t^A, \tag{A.1}$$

$$\mathcal{M}_{t,t+1} = \frac{1}{R_t}, \tag{A.2}$$

$$R_t = \max\left(R_t^E, R_t^*\right), \tag{A.3}$$

$$A_t^R = \rho_{A^R} A_{t-1}^R + \sigma_r \varepsilon_t^R. \tag{A.4}$$

$$R_t^E = K + \sigma_E \varepsilon_t^E, \tag{A.5}$$

$$\tilde{\lambda}_t = A_t \left(\tilde{C}_t - \chi \tilde{C}_{t-1}/\mu_t\right)^{-\sigma}, \tag{A.6}$$

$$\tilde{W}_t = \kappa N_t^\eta / \tilde{\lambda}_t, \tag{A.7}$$

$$\mathcal{M}_{t,t+1} = E_t \left\{ \beta d_{t+1} \frac{\tilde{\lambda}_{t+1}}{\tilde{\lambda}_t \Pi_{t+1} \mu_{t+1}} \right\}, \tag{A.8}$$

$$\tilde{Y}_t = N_t, \tag{A.9}$$

$$\left(\frac{\varepsilon}{\varepsilon-1}\right) \tilde{W}_t - \exp\left(\sigma_\pi \varepsilon_t^\pi\right) - \left(\frac{\phi}{\varepsilon-1}\right) \Pi_t \left[\Pi_t - \tilde{\Pi}_t\right] + \ldots$$
$$\ldots + E_t \left\{ \left(\frac{\phi}{\varepsilon-1}\right) \mathcal{M}_{t,t+1} \Pi_{t+1}^2 \frac{\tilde{Y}_{t+1}}{\tilde{Y}_t} \mu_{t+1} \left[\Pi_{t+1} - \tilde{\Pi}_{t+1}\right] \right\}, \tag{A.10}$$

---

[7]https://github.com/jmaih/RISE_toolbox

$$R_t^* = R_{t-1}^{*\rho_r} \left(\bar{R}^* \left(\frac{\Pi_t}{\bar{\bar{\Pi}}} \right)^{\kappa_\pi} \left(\frac{\tilde{Y}_t}{\tilde{Y}_{t-1}} \right)^{\kappa_y} \right)^{1-\rho_r} \exp\left(A_t^R \right), \qquad \text{(A.11)}$$

$$\tilde{Y}_t = \tilde{C}_t + \frac{\phi}{2} \tilde{Y}_t \left[\Pi_t - \tilde{\Pi}_t \right]^2, \qquad \text{(A.12)}$$

$$\Delta \log\left(Y_t \right) = \log\left(\tilde{Y}_t \right) - \log\left(\tilde{Y}_{t-1} \right) + \log\left(\mu_t \right), \qquad \text{(A.13)}$$

where $\mu_t = \exp(g_Z + \sigma_Z \varepsilon_t^Z)$ and $\tilde{\Pi}_t = \Pi_{t-1}^\xi \bar{\Pi}^{1-\chi}$. The complete set of variables in the model is defined as follows

$$\tilde{x}_t = \left[\tilde{C}_t, \, N_t, \, R_t, \, \Pi_t, \, \tilde{W}_t, \, \tilde{\lambda}_t, \, \mathcal{M}_{t,t+1}, \, \tilde{Y}_t, \, R_t^*, \, R_t^E, \, \Delta\log\left(Y_t \right), \, A_t^R, \, A_t \right]',$$

Table A.1: Variable Description

Variable	Description
$\tilde{C}_t$	Detrended consumption
N_t	Hours worked
R_t	Interest rate
Π_t	Inflation rate
$\tilde{W}_t$	Real detrended wage
$\tilde{\lambda}_t$	Marginal utility of consumption
$\mathcal{M}_{t,t+1}$	Stochastic discount factor
$\tilde{Y}_t$	Effective output
R_t^*	Shadow/Taylor rule interest rate
R_t^E	Effective lower bound interest rate
$\Delta\log\left(Y_t \right)$	Log change in per capita GDP
A_t^R	Autoregressive monetary policy shock term
A_t	Autoregressive consumption preferences shock
ε_t^A	Consumption preference shock
ε_t^R	Monetary policy shock
ε_t^E	Effective lower bound shock
ε_t^π	Stochastic subsidy shock
ε_t^Z	Productivity shock

Appendix B. Data

All data is taken from the St. Louis Federal Reserve's FRED database. We report the FRED pneumonics in brackets. The data is quarterly and spans 1985Q1 to 2015Q2. We used the log change in per capita US GDP, constructed using Real Gross Domestic Product (GDP1) and the Civilian Noninstitutional Population (CNP16OV), the percentage change in the US GDP deflator (GDPDEF), and the quarterly average of the monthly fed funds rate (FEDFUNDS). The data are not detrended.

Appendix C. Estimation

We calibrate the following parameters, $\beta = 0.9985$, $g_Z(L, N) = 0.0029, 0.0046$, $\bar{\Pi}(L, N) = 1.0046, 1.0055$, $\sigma_E = 0.0001$, $\varepsilon = 6$, where β, g_Z and $\bar{\Pi}$ are chosen based on estimates from Binning & Maih (2016). We run MCMC with 220,000 draws, burning the first 20,000 draws, to leave us with 200,000 draws. The estimation results are presented in Table C.2.

Table C.2: Parameter Estimates

Parameter	Distribution	Prior Mean	Prior Std	Post Mode	95% Posterior Probability Range
χ	beta	0.5000	0.1500	0.0772	[0.0173 0.2754]
σ	normal	2.0000	0.3000	2.2443	[1.8861 2.8357]
η	normal	2.0000	0.5000	2.7295	[1.9870 3.5228]
ϕ	normal	10.0000	2.0000	18.4246	[15.7244 22.0535]
κ_π	normal	1.5000	0.2500	2.8030	[2.2932 3.0520]
κ_y	normal	0.1200	0.0500	0.1576	[0.0430 0.2623]
ξ	beta	0.5000	0.1500	0.3778	[0.1241 0.5391]
ρ_r	beta	0.5000	0.1500	0.4637	[0.2373 0.5255]
ρ_{AR}	beta	0.5000	0.1500	0.5175	[0.3166 0.5464]
ρ_A	beta	0.5000	0.1500	0.7115	[0.4408 0.7567]
σ_R	inverse gamma	0.1000	2.0000	0.0039	[0.0038 0.0049]
σ_A	inverse gamma	0.1000	2.0000	0.0085	[0.0050 0.0117]
σ_Z	inverse gamma	0.1000	2.0000	0.0069	[0.0062 0.0089]
σ_π	inverse gamma	0.1000	2.0000	0.0055	[0.0043 0.0077]
$p_{N,L}$	beta	0.0500	0.0200	0.0809	[0.0469 0.1788]
$p_{L,N}$	beta	0.0500	0.0200	0.2697	[0.1940 0.2697]

References

Adolfson, M., Laséen, S., Lindé, J., & Villani, M. (2005). Are Constant Interest Rate Forecasts Modest Policy Interventions? Evidence from a Dynamic Open-Economy Model. *International Finance*, *8*(3), 509–544. URL https://ideas.repec.org/a/bla/intfin/v8y2005i3p509-544.html.

Alstadheim, R., Bjørnland, H. C., & Maih, J. (2013). Do Central Banks Respond to Exchange Rate Movements? A Markov-Switching Structural Investigation. Working Paper 2013/24, Norges Bank. URL https://ideas.repec.org/p/bno/worpap/2013_24.html.

Binning, A. & Maih, J. (2016). Implementing the Zero Lower Bound in an Estimated Regime-Switching DSGE Model. Working Paper 2016/3, Norges Bank. URL https://ideas.repec.org/p/bno/worpap/2016_03.html.

Cogley, T., Morozov, S., & Sargent, T. J. (2005). Bayesian Fan Charts for U.K. Inflation: Forecasting and Sources of Uncertainty in an Evolving Monetary System. *Journal of Economic Dynamics and Control*, *29*(11), 1893–1925. URL https://ideas.repec.org/a/eee/dyncon/v29y2005i11p1893-1925.html.

Diebold, F. X., Schorfheide, F., & Shin, M. (2015). Real-Time Forecast Evaluation of DSGE Models with Stochastic Volatility. Unpublished mimeo.

Doan, T., Litterman, R. B., & Sims, C. A. (1986). Forecasting and Conditional Projection Using Realistic Prior Distribution. Tech. rep.

Franta, M., Baruník, J., Horváth, R., & Smídková, K. (2014). Are Bayesian Fan Charts Useful? The Effect of Zero Lower Bound and Evaluation of Financial Stability Stress Tests. *International Journal of Central Banking*, *10*(1), 159–188. URL https://ideas.repec.org/a/ijc/ijcjou/y2014q1a5.html.

Garratt, A., Lee, K., Pesaran, M. H., & Shin, Y. (2002). Forecast Uncertainties In Macroeconometric Modelling: An Application to the UK Economy. Royal Economic Society Annual Conference 2002 82, Royal Economic Society. URL https://ideas.repec.org/p/ecj/ac2002/82.html.

Gerard, H. & Nimark, K. (2008). Combining Multivariate Density Forecasts Using Predictive Criteria. Economics Working Papers 1117, Department of Economics and Business, Universitat Pompeu Fabra. URL https://ideas.repec.org/p/upf/upfgen/1117.html.

Jordà, Ò. & Marcellino, M. (2010). Path Forecast Evaluation. *Journal of Applied Econometrics*, *25*(4), 635–662. URL http://EconPapers.repec.org/RePEc:jae:japmet:v:25:y:2010:i:4:p:635-662.

Kolsrud, D. (2015). A Time-Simultaneous Prediction Box for a Multivariate Time Series. *Journal of Forecasting*, *34*(8), 675–693. URL https://ideas.repec.org/a/wly/jforec/v34y2015i8p675-693.html.

Leeper, E. (2003). An "Inflation Reports" Report. NBER Working Papers 10089, National Bureau of Economic Research, Inc. URL https://ideas.repec.org/p/nbr/nberwo/10089.html.

Leeper, E. M. & Zha, T. (2003). Modest Policy Interventions. *Journal of Monetary Economics*, *50*(8), 1673–1700. URL https://ideas.repec.org/a/eee/moneco/v50y2003i8p1673-1700.html.

Maih, J. (2010). Conditional Forecasts in DSGE Models. Working Paper 2010/07, Norges Bank. URL http://EconPapers.repec.org/RePEc:bno:worpap:2010_07.

Maih, J. (2015). Efficient Perturbation Methods for Solving Regime-Switching DSGE Models. Working Paper 2015/01, Norges Bank. URL http://ideas.repec.org/p/bno/worpap/2015_01.html.

Norges Bank (2012). Monetary Policy Report 1—12. Tech. rep.

Robertson, J. C., Tallman, E. W., & Whiteman, C. H. (2005). Forecasting Using Relative Entropy. *Journal of Money, Credit and Banking*, *37*(3), 383–401. URL https://ideas.repec.org/a/mcb/jmoncb/v37y2005i3p383-401.html.

Rotemberg, J. J. (1982). Sticky Prices in the United States. *Journal of Political Economy*, *90*(6), 1187–1211. URL https://ideas.repec.org/a/ucp/jpolec/v90y1982i6p1187-1211.html.

Wolf, M. & Wunderli, D. (2012). Bootstrap Joint Prediction Regions. ECON - Working Papers 064, Department of Economics - University of Zurich. URL https://ideas.repec.org/p/zur/econwp/064.html.

Wolters, M. H. (2013). Evaluating Point and Density Forecasts of DSGE Models. Tech. rep.

CENTRE FOR APPLIED MACRO – AND PETROLEUM ECONOMICS (CAMP)

Centre for Applied Macro - and Petroleum economics (CAMP)
will bring together economists working on applied macroeconomic issues, with special emphasis on petroleum economics.

BI Norwegian Business School
Centre for Applied Macro - Petroleum economics (CAMP)
N-0442 Oslo

http://www.bi.no/camp