Discussion note

Putting new fan-charts into use

Due to uncertainty surrounding point forecasts, many central banks nowadays calculate and officially publish prediction intervals (in the form of fan-charts) for key macroeconomic variables in order to express and communicate perceived forecast risks to professionals and the general public.[1] *Gaussian prediction bands have become a workhorse in fan-chart modelling.*[2] *Apart from being intuitive and easy to calculate, Gaussian-like prediction bands suffer from several shortcomings, which make them inconsistent (from the econometric standpoint) and unreliable (from the policy standpoint). A new system of fan-charts, based on empirical quantiles, is briefly discussed in this note.*

Traditional approach

The Gaussian intervals are given by

$$\Omega(h) = [\hat{y}_T(h) - z_u\,\hat{\sigma}(h), \hat{y}_T(h) + z_u\,\hat{\sigma}(h)], \tag{1}$$

where $\hat{y}_T(h)$ denotes a point forecast made at time T with a forecast horizon h. $\hat{\sigma}(h)$ represents the h-step ahead root mean squared forecast error (RMSFE) calculated from historical forecast errors and z_u is the u-th quantile (critical value) of the standard normal distribution $(0.5 < u < 1)$.

The list of issues associated with Gaussian-like prediction bands includes:

- **Outliers**: The standard RMSFE is not robust to outlying observations (see Maronna, 2006, Ch. 2). It means that even a single aberrant observation inflates $\hat{\sigma}_T(h)$ and, thus, the width of the prediction bands, despite the fact that the probability of such an event occurring is out of the maximum probability coverage (say, 90 %) for prediction bands.

- **Dependence**: The standard RMSFE is not a consistent measure of uncertainty for (weakly) dependent forecast errors (see Hamilton, 1994, Ch. 10). $\hat{\sigma}(h)$ underestimates the true value of $\sigma(h)$ and the bias increases with the forecast horizon h (due to increasing persistence in forecast errors).

- **Distribution**: The underlying assumption of normality of historical forecast errors seems to be at odds with empirical evidence (see Vavra, 2018). This implies that the use of the standard normal critical value z_u in (1) is questionable.

Empirical evidence suggests that the traditional approach may systematically overestimate or underestimate the width of prediction bands, depending on the distributional properties and persistence of forecast errors (see Elder et al., 2005). It is worth remarking here that the same arguments apply to the so-called two-piece (asymmetric) Gaussian prediction bands, an alternative approach which takes into account unbalanced risks of the future development of exogenous factors.[3]

---

[1] Hammond (2012) surveys the (inflation) reports of 27 central banks out of which 20 banks provide prediction intervals officially.

[2] See, e.g., Bank of Canada, Sveriges Riksbank, Norges Bank, Czech National Bank, European Central Bank.

[3] See, e.g., Bank of England, National Bank of Slovakia, South African Reserve Bank, International Monetary Fund, World Bank.

Quantile approach

Rather than focusing on fixing the above mentioned drawbacks of Gaussian-like prediction bands, we follow Tulip and Wallace (2012) and construct prediction bands using empirical quantiles of marginal distributions of historical forecast errors. Empirical quantiles are easy to calculate, robust to $100(1-u)\%$ contamination of forecast errors with outliers and, much more importantly, are consistent estimates even for (weakly) dependent observations under mild regularity conditions (see Theorem 1 in Psaradakis and Vávra, 2015). In addition, implicitly assumed symmetry of the marginal distribution of forecast errors is a much weaker condition compared to normality (see Vávra, 2018). As a result, all the above mentioned shortcomings of the Gaussian-like prediction bands are (theoretically) solved when using empirical quantiles.

However, due to the limited number of observations (less than 30 in the case of NBS), we make one modification of Tulip and Wallace's approach. Rather than relying on (raw) empirical quantiles, we calculate expected quantiles using a sieve bootstrap.

The bootstrap procedure is summarized in the following steps:

Step 1: Let $\{X_1,\dots,X_n\}$ denote a set of forecast errors for a particular variable and a given forecast horizon h.

Step 2: Select an appropriate lag order p of an AR model using the Bayesian information criterion, where the lag order is restricted by $0 \le p < \log n$, where n denotes the sample size of forecast errors. Estimate the unknown AR(p) model parameters by the OLS method.

Step 3: Construct a sequence of the estimated residuals $\{\hat{\epsilon}_t\}_{\mathrm{t=p+1}}^{\mathrm{n}}$ by the recursion

$$\hat{\epsilon}_t = X_t - \hat{c} - \sum_{i=1}^{p} \widehat{\phi}_\iota X_{t-i}, \quad \text{for} \quad t \; = \; p+1,\dots,n.$$

Under the null hypothesis of marginal symmetry of forecast errors (see Tulip and Wallace, 2012), we transform the estimated residuals to ensure their marginal symmetry (around zero).

$$\widetilde{\epsilon_t} = \begin{cases} \zeta\hat{\epsilon}_t & \text{for} \quad t \; = \; p+1,\dots,n, \\ -\zeta\hat{\epsilon}_{t-n+p} & \text{for} \quad t \; = \; n+1,\dots,2n-p. \end{cases}$$

The scaling parameter $\zeta = \sqrt{(n-p)/(n-2p-1)}$ compensates for deflation of residuals due to fitting an AR(p) model (see Stine, 1987).

Step 4: Draw a sample of residuals $\{\widetilde{\epsilon}_t^*\}_{\mathrm{t=1}}^{\mathrm{n+100}}$ from the smoothed distribution function of the symmetrized residuals. The bandwidth parameter recommended in Hansen (2004) is used for smoothing the residual CDF.

Step 5: Generate bootstrap replicates $\{X_t^*\}_{\mathrm{t=1}}^{\mathrm{n+100}}$ by the recursion

$$X_t^* = \sum_{i=1}^{p} \widehat{\phi}_\iota X_{t-i}^* + \widetilde{\epsilon}_t^*,$$

where the process is initiated by a vector of sample averages: $\left(X_{-p+1}^*,\dots,X_0^*\right) = (\bar{X},\dots,\bar{X})$. The first 100 data points are then discarded in order to eliminate start-up effects and the remaining n data points are used.

Step 6: Repeat **Steps 4 - 5** independently B times to get a sample of bootstrap quantiles $\{\widehat{q_{\mathrm{u,\iota}}^*}: i = 1,\dots,B\}$. The sample quantile is calculated as $\widehat{q_u^*} = X_{(k)}^*$ using the k-th order statistic from the bootstrap sample with $k \; = \; \lceil nu \rceil$. Then, the bootstrap-based expected quantile is just a sample average calculated over all B replications $\overline{q_u^*} = B^{-1}\sum_{i=1}^{B} \widehat{q_{u,\iota}^*}$.

Finally, denoting the bootstrap-based quantiles for different forecast horizons by $\overline{q_u^*}(h)$, the proposed prediction bands are calculated by

$$\Omega^*(h) = [\hat{y}_T(h) - \overline{q_u^*}(h), \hat{y}_T(h) + \overline{q_u^*}(h)], \tag{2}$$

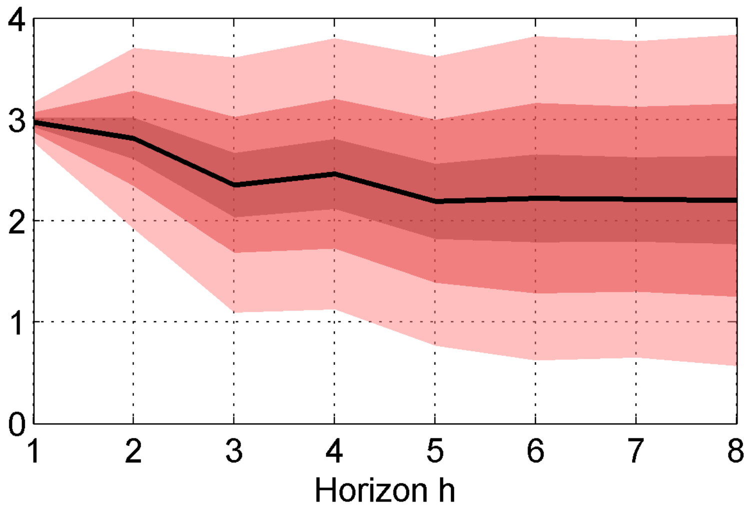
for $0.5 < u < 1$ and $h \in \{1,2,\ldots,H\}$.

Consistency of the bootstrap-based prediction bands follows from Bickel and Buhlmann (1999). An important question is to what extent the AR-sieve bootstrap works for stochastic processing not stemming from the Wold representation. Bickel and Buhlmann (1997) explain that the closure of the Wold representation is fairly large. It means that for any non-linear stochastic process there exist another process in the closure of linear processes having identical sample paths with probability exceeding 0.37. This finding implies that the AR-sieve bootstrap is very likely to give satisfactory results even for stochastic processes deviating from the Wold representation.

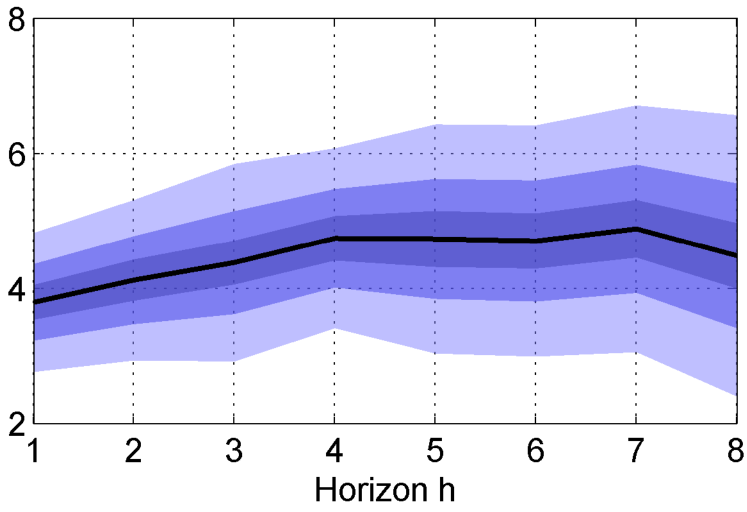
Example

In this section, we apply the above described bootstrap approach to the NBS forecast errors for HICP inflation and real GDP growth. Calculated fan-charts are depicted in Figure 1. We set the forecast horizon $h \in \{1,2,\ldots,8\}$ and $u \in \{0.65, 0.80, 0.95\}$ which corresponds to prediction intervals with 30%, 60%, and 90% probability coverage. The black lines indicate the NBS point forecasts for HICP inflation and real GDP growth for the period 2018 Q2 - 2020 Q1. The number of bootstrap replications is set to $B = 1000$. Using a larger number of bootstrap replications does not change the results significantly (see Davison and Hinkley, 1997, pp. 155-156). The lag order of AR sieve approximation is set to $p = 2$. It is important to point out that the same approach can be used to calculate (historical) data uncertainty of real GDP growth caused by statistical revisions.

Figure 1: fan-charts over the projection horizon



(a) HICP inflation



(b) Real GDP growth

References

Bickel, P. and P. Buhlmann (1997). Closure of linear processes. *Journal of Theoretical Probability* 10, 445–479.

Bickel, P. and P. Buhlmann (1999). A new mixing notion and functional central limit theorems for a sieve bootstrap in time series. *Bernoulli* 5, 413–446.

Davison, A. and D. Hinkley (1997). Bootstrap Methods and Their Application. Cambridge University Press.

Elder, R., G. Kapetanios, T. Taylor, and T. Yates (2005). Assessing the MPC's fan charts. *Bank of England Quarterly Bulletin*, Autumn 2005.

Hamilton, J. (1994). Time Series Analysis. Princeton University Press.

Hammond, G. et al. (2012). State of the art of inflation targeting. Centre for Central Banking Studies, Bank of England.

Hansen, B. (2004). Nonparametric estimation of smooth conditional distributions. University of Wisconsin, manuscript.

Maronna, R., R. Martin, and V. Yohai (2006). Robust Statistics. Wiley.

Psaradakis, Z. and M. Vávra (2015). A quantile-based test for symmetry of weakly dependent processes. *Journal of Time Series Analysis* 36, 587–597.

Stine, R. (1987). Estimating properties of autoregressive forecasts. *Journal of the American statistical association* 82, 1072–1078.

Tulip, P. and S. Wallace (2012). Estimates of uncertainty around the RBA's forecasts. *Research Discussion Paper* 2012-07.

Vávra, M. (2018). Assessing distributional properties of forecast errors. *NBS Working Paper* 3/2018.

Marián Vávra (analytici@nbs.sk)