Council for Budget Responsibility

The presented paper offers purely statistical approach to public debt forecast and important first step for further debt sustainability analysis. It is a good complement to the deterministic approach used in the official forecast. If properly communicated, it has the potential to enrich the debt analysis framework used in Slovakia. The paper clearly identifies strengths and drawbacks of the stochastic approach and its outcomes should always be interpreted with these in mind.

We would like to stress that we find the paper well written and straightforward. The main idea is clearly explained and the methodology is properly selected. The variables used in the debt equation and in the VAR model are properly selected, and the advantage is that the set of variables includes also foreign variables and NEER.

Authors should also consider extending the approach by using more structural models where for example macro-financial linkages, government bond portfolio, share of foreign debt or tail risk are taken into account. Information about government plans could also help the anchor forecast through better estimate of fiscal reaction function and potential size of non-deficit financing of debt (SFA).

On one hand, we welcome that the paper is straightforward, on the other hand the price for simplicity is that we miss some information. Therefore, we would like to propose several amendments that could, in our view, even more increase the credibility of presented results:

1. Presentation of the results in this paper lacks, in our opinion, some key information. Figure 1 compares the official debt forecast with the stochastic one, including the estimation of uncertainty surrounding the forecast. In order to increase the transparency of the stochastic approach, the reasons behind the divergence of the two forecasts should be clearly explained. In particular, it would be useful to show how the underlying variables (variables included in the VAR model, primary deficit, non-deficit financing) differ from the official macroeconomic forecast or fiscal targets. Explanation of differences between the two debt forecasts should be used as a standard procedure also in other official documents containing both forecasts.

   author: In this work we solely focus on the forecast of the public debt hence we do not discuss the central projections of the debt equation components. Moreover, it is not straightforward to make a direct link between the central projections of the components and the central projection of the debt. The problem is that the central projections of the components do not sum up to the central projection of debt in debt equation (1) due to aggregation process. Similarly to Figure (4) we decompose both the official and alternative forecasts into the contributions of the components, see Figure 11. In Figure 12 we decompose the deviation between the forecasted levels of debt into the underlying factors. There also appears a residual caused by aggregation, although it is negligible. See explanation on page 10.

2. Regarding the VAR model, we would appreciate if the author could present the estimates of reduced form VAR. Also the number of lags is missing.

author: Based on the Schwarz information criterion the order of the model was set at three. This piece of information was added on page 6. In the VAR literature it is a common practice to report impulse response functions instead of parameter estimates. We do not identify the model hence we do not present too many details. Our ultimate goal is a forecast of the public debt and performance of the overall model is presented in Figure 7.

3. The author simply drops those simulated forecast with negative interest rates. However, we are not confident whether this is the best way to tackle this issue, especially given the fact that negative interest rates are common these days. Could you please elaborate more on this issue, or at least could you present how many simulations are affected?

   author: Recently the interest rates on some government bonds have indeed decreased into negative figures. The shorter end of the Slovak government bond term structure has been slightly negative for about six months while the longer end have been in the above-zero territory. The average of 1-year and 10-year bond yields have been positive during that period. More importantly we suspect that the model that properly captures the positive-negative yields situation may be nonlinear. Nevertheless in the appendix we added a fan-chart that includes also negative interest rate paths, see Figure 9. The central projections and the probability mass shift slightly downwards. The central forecast shifts by less than 0.5 p.p. Naturally, the wider range of interest rates forecasts translates into a slightly wider probability intervals, although the entire area becomes only marginally wider.

4. How is the average maturity (denoted as am_t) defined? It looks like there is a typo in equation (3) because maturity as such is greater than one.

   author: Inverse of average maturity rather than average maturity enters the formula in (3). The equation was corrected.

5. The parameters of the fiscal reaction function are calibrated according to the study based on the panel data. We believe that this approach is appropriate, however it will be very interesting to see how it fits the Slovak data. We would appreciate if the author could present a comparison of primary surplus forecast with actual data.

   author: See Figure 10 which depicts the actual and fitted values of the primary surplus. For additional details on the model and refinements that we made see page 8.

6. We find Figure 4 which shows the contributions to growth of public debt very interesting. We would also welcome the same type of figure to be produced for mean forecasts of debt. This can help with the interpretation of the main results presented in the paper.

   author: See point 1 above, page 10 and Appendix 4.

7. Figure 7 presents very interesting comparison of in sample forecast with actual data on public debt. However, we believe that the comparison could be even more

interesting if the author includes the official forecast of public debt. Our second point here is the way of producing forecasts. Is it possible to draw similar picture for out of sample forecasts?

author: Comparison of the in-sample forecasts with the actual official forecasts is more than difficult. The presented model forecasts were done on the current data vintage while the corresponding actual forecasts were based on different data vintages. Data revisions (for example the recent methodology change from ESA95 to ESA2010) make comparison very difficult. Out-of-sample forecast is presented in Figure 1.

National Bank of Slovakia

Technical comments:

1. How many lags are used in estimating equation (2)? Is it a VAR(2) or VAR(p) model? How is the lag order in (2) identified (AIC/BIC)?

   author: Based on the Schwarz information criterion the order of the model was set at three. This piece of information was added on page 6.

2. Residuals of equation (2) should be tested whether they are white noise. If they are not, e.g. due to neglected autocorrelation, then using the (residual) i.i.d. bootstrap may produce inconsistent results (Lahini, 2003). Moreover, if there is a residual autocorrelation (which might be suspected), then the estimation of the variance-covariance matrix Sigma should be carried out by the Newey-West procedure, as the ordinary estimate would not have been consistent either (Newey and West, 1987, 1994). It would be appropriate to test residuals of equations (2) and (4) for normality, heteroscedasticity and autocorrelation and include the diagnostic test results in the appendix.

   author: We tested the VAR residuals for presence of autocorrelation, heteroskedasticity and if they are normally distributed. The residuals did not pass the test of normality, however, we cannot reject the hypotheses of no serial correlation and homoskedasticity. For these reasons we rely on bootstrapping methods. These arguments can be found on page 7. The discussion of quality of residuals from equation (4) can be found further below. We want to avoid reporting too many technical details therefore we do not include results of the formal tests. The results are, however, available upon request from the author. This applies to both equations (2) and (4).

3. Parameters of equation (4) can hardly be estimated reliably due to insufficient data but the author should check the goodness of fit of the equation.

   author: In order to improve the goodness of fit we slightly changed the model. We keep the structural parameters of the model (i.e. sensitivity of primary balance to debt and business cycle) unchanged. We did a minor change to the constant term such that the average value of residual is zero. More importantly we re-estimate the autoregressive equation for the residuals. The coefficient on the lagged term changes from 0.7 to 0.43 while the standard deviation of the error term e_t is 2.20. See page 8. See Figure 10 which depicts the actual and fitted values of the primary surplus.

4. It is argued that residuals of equation (2) might not be normally distributed and the bootstrap method is used, yet equation (4) is simulated by normal distribution without any justification.

   author: Our sample is rather short (consists of 17 observations) and formal statistical tests do not have power to deliver conclusive results. A good example is the test for significance of the autoregressive parameter in the residual equation. Its point estimate is 0.43, which is rather far from zero, yet is statistically insignificant at 5% level. This suggests that the term $\mathcal{E}_t$ may be treated as a

serially not correlated component. Nevertheless, we allow this component to follow an AR(1) process which treats the autocorrelation issue in the error term e_t. Furthermore, in our short sample we cannot reject the null hypothesis of homoskedastic error term. Finally and not surprisingly, we do not have enough evidence that the errors are normally distributed and for this reason we opt for bootstrap approach in generating the forecasts of the primary balance. See page 9.

5. In equation (3), average maturity is used as a weight although its values are around 5 (see figure 3).

   author: Inverse of average maturity rather than average maturity enters the formula in (3). The equation was corrected.

6. Which standard deviation of non-deficit financing of debt variable (calculated over the whole sample or just over the second half) has been used after all?

   author: Standard deviation based on the entire sample was used in the simulation exercise.

7. The key problem of the debt forecast model rests on the construction of prediction intervals around the mean forecasts and, consequently, on the construction of the fan-charts. Loosley speaking, forecast errors, as well as quantiles of marginal distribution, are mutually correlated across forecast horizons. Consequently, the coverage probability of the merged marginal prediction intervals is not equal to the nominal (desired) coverage probability (e.g. 90 %). For example, consider a sequence of independent random variables over H forecast periods and calculate (standard) marginal prediction intervals (your case) with the coverage probability 0<p<1 for each horizon. Then, the probability of the mean forecast lying within the prediction bands over all forecast horizons is 0<p^H<p!!! For more details, see Wolf and Wunderli (2014). As a result, “probability” claims about the future debt development might be inaccurate and misleading.

   author: This is a commonly ignored and insufficiently understood problem in calculating h-step ahead probability intervals. We appreciate raising this point.

Formal comments:

8. The level of debt relative to GDP (instead of just the level of debt) should be appended in the text after equation (1).

   author: The level of debt relative to GDP can be found in Appendix 1.

9. The legend of figure 4 should write primary deficit instead of primary surplus.

   author: Changed

10. Figure 3 in the appendix 3 should be named figure 5.

author: Corrected

11. In order to compare uncertainty intervals at different forecast horizons in figure 5, it would be more appropriate to use the same range for both 1 year and 4 years forecast horizons on the x-axis.

    author: Accepted