Nowcasting South African GDP using a suite of statistical models

Byron Botha, Geordie Reid, Tim Olds, Daan Steenkamp, Rossouw van Jaarsveld

October 2020
The views expressed are those of the authors and do not necessarily represent those of the South African Reserve Bank or Reserve Bank policy.
Motivation

- Central banks have to assess current state of the economy but GDP data only available with a lag
- So we need to ‘nowcast’ current GDP using available data
- Statistical models been shown to be useful for estimating ‘starting points’ for central bank forecasts
Background

- True ‘real-time’ (i.e. data as available) forecast comparisons for GDP still fairly rare
  - Exceptions incl Aastveit et al. (2014) & Anesti et al. (2018)

- For SA, Kabundi et al. (2016)
  - ‘Pseudo’ real time (15th of each month)
  - Mixed frequency factor model (21 series) vs range of univariate & multivariate autoregressive models & consensus forecasts

- Model suites from other central banks
  - Kapetanios et al. (2008), Bjornland et al. (2012) and Richardson et al. (2019)
Our contribution

- Develop a large GDP nowcasting model suite
- Compare ‘true’ real-time nowcasts from our model suite to analyst forecasts & real-time projections from SARB
- Consider how our models perform at forecasting different vintages of data
- Demonstrate usefulness of statistical models in interpreting data surprises & quantifying forecast uncertainty
- We automate data loading, scheduled auto-running and reporting of nowcasts and publication to an internal blog
Why build model suites for nowcasting?

- Reduces risk associated with reliance on an individual models
  - Over time, different models perform differently at capturing time-varying underlying drivers of growth
  - Combining forecasts from different forecasting models help to better characterise statistical uncertainty around forecasts

- Strategy we pursue is to develop a large number of models from a wide variety of model classes

- To better capture process that governs evolution of GDP

- We assess whether combining forecasts based on forecast performance enhances accuracy of forecasts individual models

- Understanding statistical properties of data improves model specifications
We build a real time dataset

- Construct real-time dataset of GDP data (Q) and 25 monthly series (M1,M2,M3)

- Ensure that all models use exactly the same dataset for every nowcast

- GDP based on official seasonally adjusted series

- Transformations and dataset descriptions for each model in Appendix
Forecast comparison

- Sample of 1995/6Q1 to 2020Q1

- Out of sample forecast performance tested (2015Q1 to 2020Q1), providing nowcasts for 21 quarters

- Compare predictive accuracy of thousands of specifications of models from 9 classes, to AR(1), official forecasts, market analysts

- Q-on-Q annualized convention, 1-step ahead forecast horizon
Model suite

- ARIMA model
- Indicator model
  - Select top 20 OLS models based on ability to explain GDP, weighted using Bayesian model averaging
- Weighted VAR model
  - 7722 combinations of variables & different model specifications weighted by AIC score
- Exponential smoothing state space model
  - Separately model error, trend and seasonality of GDP
- Pattern sequence model
  - Univariate clustering into subset with equal means
- Bayesian Mixed Frequency VAR
  - Stationary version of Schorfheide & Song (2013)
- Dynamic Factor Model
  - 5 common factors from 18 variables
- Machine learning models
  - Elastic Net and LASSO (Plus MIDAS versions of each)
Importance of understanding statistical characteristics of GDP data

Change in mean

Change in standard deviation
Some statistical models outperform official nowcasts

**Figure:** Forecast RMSE relative to AR(1) in real-time
Comparing nowcasts across models

Figure: Forecasts vs real time GDP outcomes
Historical forecast errors (Selected models)
Which predictors contribute to specific nowcasts?

Figure: Example: Forecast decomposition (DFM, 2020Q1)
Model suite can be useful for assessing balance of risks
We automated data loading (programmatically and via Shiny front-end) and model runs.
We automated reporting (email and blog)

Sparta nowcasts 2020Q1

Sep 1, 2020

Real GDP growth (annualised quarter-on-quarter change)

This blog post presents the nowcasts of Real GDP growth for 2020Q1 produced by the models in the Sparta model suite. The first table shows the individual model nowcasts, the second table shows nowcasts produced using model combination techniques, while the third table provides benchmarks where available. Various charts of the nowcasts are provided below the tables.

Table 1: Individual models

<table>
<thead>
<tr>
<th>Model</th>
<th>M1 nowcast</th>
<th>M2 nowcast</th>
<th>M3 nowcast</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>INDICATOR</td>
<td>-0.32</td>
<td>-0.51</td>
<td>-1.78</td>
<td>1.41</td>
</tr>
<tr>
<td>MIDAS_ENET</td>
<td>-0.08</td>
<td>-0.47</td>
<td>-0.63</td>
<td>1.53</td>
</tr>
<tr>
<td>MIDAS_LASSO</td>
<td>-0.30</td>
<td>-0.43</td>
<td>-0.66</td>
<td>1.61</td>
</tr>
</tbody>
</table>
Summary

- Use statistical model suite & large # series to nowcast GDP

- Assess performance of techniques using ‘real-time’ data (i.e. data as available) in assessing current state of economy

- We show some models provide more accurate forecasts than official nowcast and market analysts

- GDP volatility has increased markedly over the last 5 years, making GDP forecasting more difficult

- We show that all models developed, as well as SARB’s official forecasts, have tended to over-estimate GDP growth

- We also demonstrate usefulness of statistical models for storytelling and risk assessment
Other take-aways

- As statistical properties of GDP evolve, relative performance of frameworks will change.

- We argue maintaining model suites important for uncovering these drivers & ongoing re-assessment of best frameworks.

- Our focus is not understanding contribution of methodological changes and/or macro fundamentals to change in statistical properties of GDP.

- But GDP over-prediction suggests either
  - Relationship between proxies of fundamentals & measured GDP weakening
  - Info set too limited (i.e. need to add political uncertainty, electricity supply-constraints, or fiscal consolidation proxies)