

IMF Working Paper

African Department

Impact of COVID-19: Nowcasting and Big Data to Track Economic Activity in Sub-Saharan Africa

Prepared by Brandon Buell, Carissa Chen, Reda Cherif, Hyeon-Jae Seo, Jiawen Tang, and Nils Wendt¹

Authorized for distribution by Papa M. N'Diaye

April 2021

IMF Working Papers describe research in progress by the author(s) and are published to elicit comments and to encourage debate. The views expressed in IMF Working Papers are those of the author(s) and do not necessarily represent the views of the IMF, its Executive Board, or IMF management.

Abstract

The COVID-19 pandemic underscores the critical need for detailed, timely information on its evolving economic impacts, particularly for Sub-Saharan Africa (SSA) where data availability and lack of generalizable nowcasting methodologies limit efforts for coordinated policy responses. This paper presents a suite of high frequency and granular country-level indicator tools that can be used to nowcast GDP and track changes in economic activity for countries in SSA. We make two main contributions: (1) demonstration of the predictive power of alternative data variables such as Google search trends and mobile payments, and (2) implementation of two types of modelling methodologies, machine learning and parametric factor models, that have flexibility to incorporate mixed-frequency data variables. We present nowcast results for 2019Q4 and 2020Q1 GDP for Kenya, Nigeria, South Africa, Uganda, and Ghana, and argue that our factor model methodology can be generalized to nowcast and forecast GDP for other SSA countries with limited data availability and shorter timeframes.

JEL Classification Numbers: C53, C55, E37, F17, O11

Keywords: COVID-19; Econometric modeling; Economic activity; GDP; Google Search Trends; Mobile payments; NO2; Nowcasting; Short-term forecasting

Authors' E-Mails: acherif@imf.org; jiawen_tang@hks.harvard.edu; brandon_buell@alumni.harvard.edu; hseo@college.harvard.edu; cjchen@college.harvard.edu; nils.wendt@tum.de

¹ We are grateful to Karim Barhoumi and Massimiliano Marcellino for their valuable input and suggestions. We extend our thanks to IMF colleagues for their generous support: Diego Cerdeiro for sharing crucial shipping data, Seung Mo Choi, Fuad Hasanov, Xingwei Hu, Vikram Singh, and IMF seminar participants for their thoughtful feedback and discussion. We are also thankful to the South African National Treasury and the Bank of Zambia for their helpful comments. Thank you to Harvard professors Jeffrey Frankel and Dan Levy for their guidance and MPA/ID program students for their insights. We are grateful to Franck Ouattara for his help with data sourcing. This research collaboration was facilitated by the Harvard College Data Analytics Group and we thank Jerry Huang for his coordination support. Jiawen Tang served as the Project Lead for this research collaboration. All errors are our own.

Table of Contents

I. Introduction	1
II. Summary of Data Sources	2
III. Estimation Frameworks	7
A. Machine Learning Approach	7
B. Parametric Factor Model Approach	9
IV. Model Results	14
A. Machine Learning Model Predictions	14
B. Parametric Factor Model Predictions	17
C. Parametric Factor Model Economic Activity Latent Factors	19
D. Evaluation of Model Performance by Country	21
E. Discussion of Model Variable Importance	25
V. Conclusion	30
Appendix I. Machine Learning Model Supplemental Materials	32
Appendix II. Factor Model Supplemental Materials	38
Appendix III. Supplemental Visualizations	43
Appendix IV. Data sourcing and aggregation notes	45
Appendix V. Full Data Dictionary Variable List	46

Table of Tables

Table 1. Breakdown of Variables by Frequency and Country	3
Table 2. Summary of Google Trend Search Terms Collected by Country	4
Table 3. ML Model Timeframe, Predictors and Observation Count	9
Table 4. Parametric Factor Model Timeframe, Predictors and Observation Count	9
Table 5. ML Model Train and Test periods for South Africa, Kenya, and Nigeria	14
Table 6. ML Model Predictions for YoY GDP Percent Change - Comparison with Actuals	16
Table 7. ML Model Predictions for QoQ GDP Percent Change - Comparison with Actuals	17
Table 8. Factor Model Estimation and Test periods for Uganda, Kenya, and Ghana	17
Table 9. Factor Model Predictions for YoY GDP Percent Change – Comparison with Actuals	18
Table 10. ML Model Performance Comparison – South Africa	22
Table 11. ML Model Performance Comparison – Kenya	23
Table 12. ML Model Performance Comparison – Nigeria	23
Table 13. Parametric Factor Model Performance Comparison – Uganda	24
Table 14. Parametric Factor Model Performance Comparison – Kenya	24
Table 15. Parametric Model Performance Comparison – Ghana	25
Table 16. Most important variables in top performing South Africa ML models	26
Table 17. Most important variables in top performing Nigeria ML models	26
Table 18. Most important variables in top performing Kenya ML models	27
Table 19. Factor Model Variable Importance – Uganda	28
Table 20. Factor Model Variable Importance – Kenya	29
Table 21. Factor Model Variable Importance – Ghana	29
Table 22. Machine Learning Model Variable List for South Africa	32
Table 23. Machine Learning Model Variable List for Nigeria	32
Table 24. Machine Learning Model Variable List for Kenya	33
Table 25. Data Dictionary - Full Variable List	46

Table of Figures

Figure 1. Summary of Variables by Frequency, Country, and Type	3
Figure 2. Google Trends Search Terms Time Series for South Africa, Nigeria, and Kenya	5
Figure 3. GDP YoY % Change for Top Economies in SSA Countries	6
Figure 4. Maximum-Likelihood Estimates of Factor Model Parameters	13
Figure 5. ML Model Time Series Comparison for South Africa.....	15
Figure 6. ML Model Time Series Comparison for Kenya	15
Figure 7. ML Model Time Series Comparison for Nigeria	15
Figure 8. ML Model Predictions for YoY GDP Percent Change.....	16
Figure 9. Factor Model Time Series Comparison for Uganda.....	17
Figure 10. Factor Model Time Series Comparison for Kenya	18
Figure 11. Factor Model Time Series Comparison for Ghana	18
Figure 12. Monthly Latent Factor – Uganda.....	19
Figure 13. Monthly Latent Factor – Kenya	19
Figure 14. Monthly Latent Factor – Ghana.....	19
Figure 15. Smoothed Means of the Latent Factor for South African Economic Activity	20
Figure 16. Top 20 most important variables for South Africa random forest regression	35
Figure 17. Top 20 coefficients with highest absolute value for South Africa XGBoost regression	35
Figure 18. Top 20 most important variables for Nigeria random forest regression.....	36
Figure 19. Top 20 most important variables for Nigeria XGBoost regression.....	36
Figure 20. All non-zero coefficients for Kenya XGBoost regression	37
Figure 21. Top 20 most important variables for Kenya random forest regression.....	37
Figure 22. Uganda GDP: Autocorrelation, Partial Autocorrelation, & Normal Quantile–Quantile Plots of Year-over-Year Quarterly GDP.....	38
Figure 23. Kenya GDP: Autocorrelation, Partial Autocorrelation, & Normal Quantile–Quantile Plots of Year-over-Year Quarterly GDP.....	39
Figure 24. Ghana GDP: Autocorrelation, Partial Autocorrelation, & Normal Quantile–Quantile Plots of Year-over-Year Quarterly GDP.....	40
Figure 25. Uganda Residuals: Autocorrelation and Partial Autocorrelation Comparison for Factor Model Pass 3(a) vs. Pass 3(b) Residuals.....	41
Figure 26. Kenya Residuals: Autocorrelation and Partial Autocorrelation Comparison for Factor Model Pass 3(a) vs. Pass 3(b) Residuals.....	41
Figure 27. Ghana Residuals: Autocorrelation and Partial Autocorrelation Comparison for Factor Model Pass 3(a) vs. Pass 3(b) Residuals.....	42
Figure 28. Correlation Matrix of Quarterly GDP YoY Percent Change by Country	43
Figure 29. Google Trend Search Terms Time Series for Ivory Coast, Cameroon, Ghana, and Uganda.....	43