

## References

- Abiad, A., D. Furceri, and P. Topalova. 2016. "The Macroeconomic Effects of Public Investment: Evidence from Advanced Economies." *Journal of Macroeconomics* 50: 224–40.
- Auerbach, A. J., and Y. Gorodnichenko. 2012. "Measuring the Output Responses of Fiscal Policy." *American Economic Journal: Economic Policy* 4 (2): 1–27.
- Blanchard, O., and R. Perotti. 2002. "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output." *Quarterly Journal of Economics* 126: 51–102.
- Fuceri, D., and B. G. Li. 2017. "The Macroeconomic (and Distributional) Effects of Public Investment in Developing Countries" *IMF Working Paper 17/217*, International Monetary Fund, Washington DC.
- International Monetary Fund, 2018. Chapter 4 "Fiscal Multipliers: How Will Consolidation Affect Latin America and Caribbean?" *Regional Economic Outlook: Western Hemisphere Region April 2018* Ch.4.
- International Monetary Fund. FAD Expenditure Assessment Tool. (EAT).
- Jorda, O. 2005. "Estimation and Inference of Impulse-Response by Local Projections." *American Economic Review* 95(1): 161-82.
- Ramey, V. A, and Zubairy S. 2018. "Government Spending Multipliers in Good Times and in Bad: Evidence from US Historical Data" *Journal of Political Economy*: 126:2, 850-901.

## A MACHINE LEARNING APPROACH TO FORECASTING GDP<sup>1</sup>

- GDP is a critical indicator of the health of the economy but is often lagged.** In addition, GDP data is subject to revisions which make it difficult to assess the current state of the economy. Suriname's GDP is released after 3 quarters but is subject to revisions for 3 years, some of which can be large. Many policymakers turn to high frequency data to make an assessment, but in many countries, such data does not exist.
- The Central Bank (CBvS) estimates economic activity for policy making using a monthly economic activity index (MEAI).** This is done using high frequency data, some of which, is not publicly available. The publicly available high frequency data is sparse and is often still subject to lags. The current lag of the MEAI is around 5 months.
- We propose a method of estimating GDP with publicly available high frequency data using the machine learning (ML) approaches.** ML is a very powerful tool but its use in macroeconomics has been somewhat limited because it requires very large datasets. We innovate a method to expand the available dataset for Suriname. We identify cross-country structural characteristics using ML, which help expand the dataset available for each individual country. We assume that countries that are structurally similar to the country of interest will be subject to the same external shocks and they will propagate through the economy in a similar way. This is done in 2 stages:

### Stage 1: Identify the countries that have structural similarities.

- Using big data on the structure of the economy and the categories of exports from the CIA Factbook, we group countries by structural similarities using two ML methods.** We use principal component analysis (PCA) for dimensionality reduction to encode countries into their latent factors and then use encoded latent factors to group similar countries using Gaussian mixture model (GMM). Our second approach is to use SimRank to find similar countries to Suriname based on their major shared industries.

Suriname's Structurally Similar Countries	
Gaussian Mixture Model	SimRank
Bolivia	Belize
Chile	Brazil
Guyana	Colombia
Haiti	Guyana
Honduras	Peru
Panama	
Paraguay	

Source: Fund staff calculations.

### Stage 2: Employ elastic net regression method to forecast the variable of interest.

- Elastic net regression is similar to an ordinary least squares regression (OLS) with two penalty terms.** The first is called the ridge penalty that compresses the estimates towards

<sup>1</sup> Prepared by Thomas Dowling (WHD), Yang Liu, and Mamoon Saeed (both ITD).

zero. The second is called the LASSO penalty that allows the coefficients to be zero when they are very small, resulting in a parsimonious model. The elastic net approach chooses to tradeoff variance for bias in order to maximize the accuracy of forecasting out of sample. We augment the naive elastic net regression model to accommodate the addition of the GDP growth rates of the countries identified in the previous exercise:

$$loss = \sum_i^N \sum_{j=1}^{N_i} (\|y_{i,j} - \mathbf{X}_{i,j}\beta - b\|^2) + \lambda_2 \|\beta\|^2 + \lambda_1 \|\beta\|_1 \quad (1)$$

$i$  from 1 to  $N$  which represents  $i^{\text{th}}$  country and  $j$  from 1 to  $N_i$  which represents the  $j^{\text{th}}$  observation in the  $i^{\text{th}}$  country samples such that,

$$\beta = \underset{\beta}{\operatorname{argmin}} (\sum_i^N \sum_{j=1}^{N_i} (\|y_{i,j} - \mathbf{X}_{i,j}\beta - b\|^2) + \lambda_2 \|\beta\|^2 + \lambda_1 \|\beta\|_1) \quad (2)$$

$$b = \underset{b}{\operatorname{argmin}} (\sum_i^N \sum_{j=1}^{N_i} (\|y_{i,j} - \mathbf{X}_{i,j}\beta - b\|^2)) \quad (3)$$

The parameters  $\beta$  and  $b$  are optimized by minimizing the loss function where  $y$  are the GDP growth rates in time  $t$  and  $\mathbf{X}$  are the predictors which include the SWIFT data in time  $t$  and the GDP growth rates in  $t-1$ . Then the nowcast equation is:

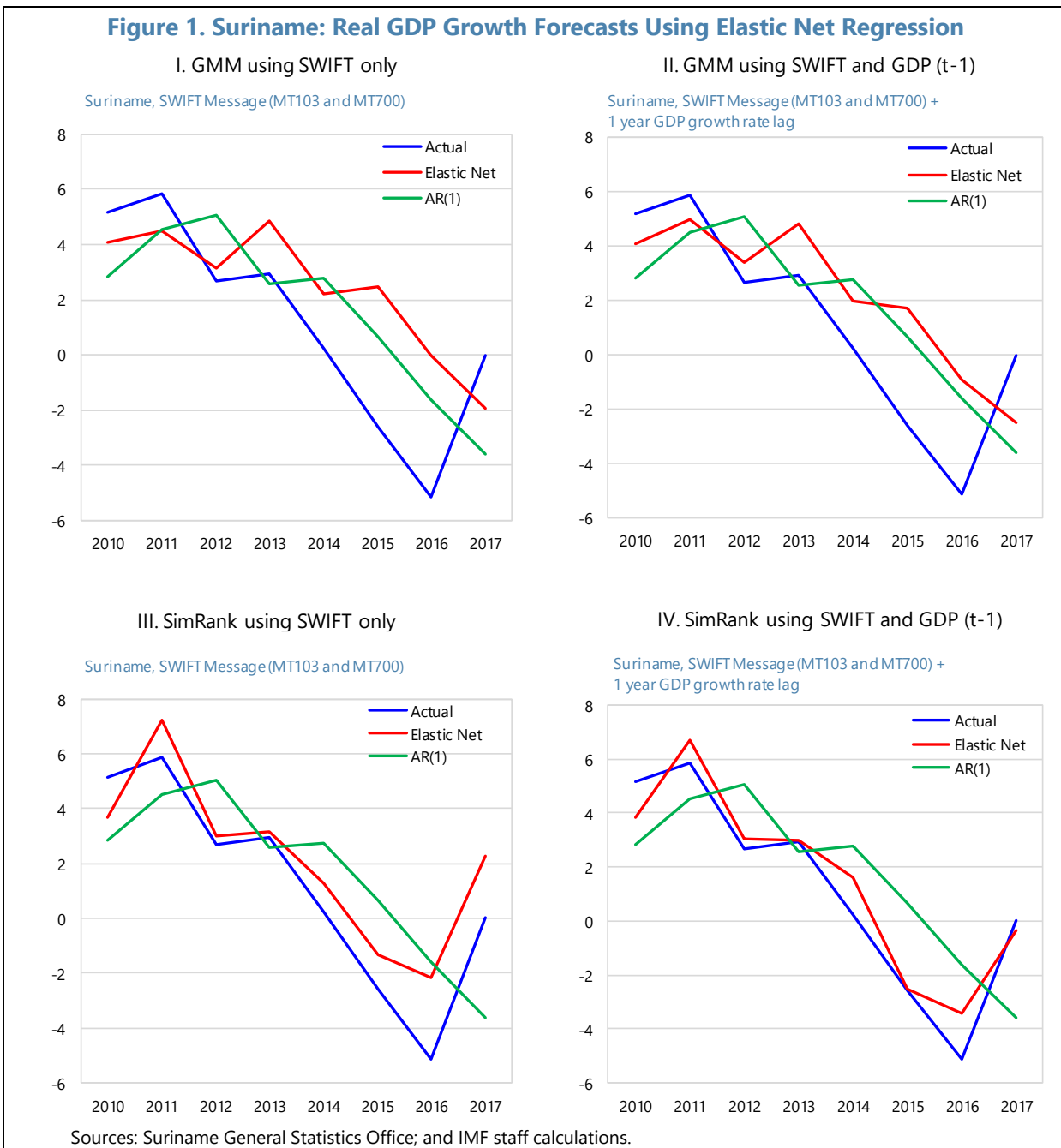
$$\widehat{y}_{SUR,t} = \hat{\beta}_1 y_{SUR,t-1} + \hat{\beta}_2 SWIFT_{103,SUR,in,t} + \hat{\beta}_3 SWIFT_{103,SUR,out,t} + \hat{\beta}_4 SWIFT_{700,SUR,in,t} + \hat{\beta}_5 SWIFT_{700,SUR,out,t} + \hat{b} \quad (4)$$

where the SWIFT messages used for inflows and outflows are MT103 (financial institutions transfers) and MT700 trade related messages.

**6. We add high frequency SWIFT data that captures financial transactions and international trade to address the lag as these data are released 9 days after the close of the period.** The forecasts are 1-step ahead and the training set used to optimize the model is from time 0 to  $t-1$ . Then we estimate the following optimizations (Figure 1):

- i. using Suriname's SWIFT data we estimate GDP growth using an AR(1) (*RMSE 2.6%*) model as a benchmarking exercise then we use the naive elastic net regression approach on the Suriname SWIFT data (*RMSE 2.9%*) and find that the AR(1) still performs better;
- ii. adding the GDP lagged by 1 period of the countries identified by GMM (*RMSE 2.5%*) we find that there is a slight improvement in the forecast;
- iii. using SWIFT data of the countries identified by SimRank and the augmented elastic net regression approach we see an improvement in the forecasting power (*RMSE 1.6%*);
- iv. adding the GDP lagged by 1 period of the SimRank countries increases the forecasting power (*RMSE 1.0%*) of the model significantly.

**Figure 1. Suriname: Real GDP Growth Forecasts Using Elastic Net Regression**



7. **The additional forecasting accuracy of the ML approaches suggest this is a useful tool for policymakers.** Additional expansion of the dataset with other big data sources such as exchange rates, financial market data, COMTRADE, APIs, or media/word count/IoT data could help increase forecasting accuracy even further. The team is developing a tool that can be easily employed for use by researchers and policymakers.