Executive Summary and Implications: Predicting Inflation Crisis in 13 African Countries

By Camille Kirongozi

We live in a world where inflation is considered a necessity, but too much inflation can lead to inflation crisis, which can then lead to economic collapse. In this report, I analyzed historical economic data for thirteen African countries to predict inflation crisis using a logistic regression model. The countries included in the dataset are Algeria, Angola, Central African Republic, Ivory Coast, Egypt, Kenya, Mauritius, Morocco, Nigeria, South Africa, Tunisia, Zambia, and Zimbabwe. The research question that I address in this report is, which factors are significantly associated with inflation crisis in the 13 African countries observed? In my logistic regression model, I aimed to identify the most significant variables in predicting inflation crisis at a 0.05 significance level.

Hypothesis: Null hypothesis, there are no significant factors that are associated with inflation crisis at a .05 significance level. Alternative hypothesis, there are factors associated to inflation crisis at the .05 significance level.

The four assumptions of the logistic regression:

1. A binary outcome variable: The outcome variable was the inflation_crisis

   variable which had values 1 and 0.

2. Independent observations: This means there should be no repeated

   observations in the dataset. The dataset did not include any repeated

   observations. Each observation was a yearly economic summary for a country.

3. Absence of multicollinearity: I ran a variance inflation factor (VIF) command on

   the model I decided to use, and no variables had a value of more than ten. A

   value of ten would indicate multicollinearity.

4. A linearity between the numeric independent variables and the log odds: To

   test this assumption, I created a plot for the only numeric variable exch_usd against the log odds of the dependent variable. Linearity was observed.

The data met all four assumptions. The dataset did not have any missing values. However, it did include a perfect predictor variable (inflation_annual_cpi), and some outliers in the exch_usd and currency_crises variables. These were all removed along with other variables that were irrelevant to the analysis. After the dataset was prepared, I produced a summary statistic of the entire set along with univariate and bivariate plots for the variables that were featured in the final logistic regression mode.

The dataset was split into a training set and a test set, using an eighty to twenty

percent ratio. I then produced the initial logistic regression model using all dependent variables. The model resulted in four variables with significant P-values at the .05 significance level. The variables are, exch_usd (0.00276), sovereign_external_debt_default (0.00234), currency_crises (1.06e-15), banking_crisis (0.04055). With these results, I was able to reject the null hypothesis. I then created a second model only using the significant variables. The second model resulted in a lower Akaike information criterion (AIC), indicating that the second model is the better model.

The data provided only included economic information for thirteen African countries, from the years 1860 to 2014. One country had data starting from 1860, and only eleven out of thirteen countries had data in 2014. I was limited in making the predictions at a one hundred percent accuracy rate. However, the final logistic regression model provided predictions at an accuracy rate of ninety percent.

My proposal for the African countries is to focus on economic development by creating more jobs and implementing monetary policies. When jobs are available, the population can earn, which correlates with spending. As the population participates in the economy, by earning and spending, the countries can expect to observe less instances of currency crisis and banking crisis. I also strongly advise that the countries promote narrow banking. A narrow bank is a financial institution that issues demandable liabilities and invests in assets that have little or no nominal interest rate and credit risk (Pennacchi, 2012). This will mitigate banking crisis. The countries can address currency crises by adopting a floating exchange rate and raising interest rates. Improving in the two categories would help alleviate the sovereign external debt in default and in turn the exchange rate to USD. Thus, preventing inflation crisis.

Works Cited

"A Box and Whiskers Plot (in the Style of Tukey) - geom_boxplot." -

geom_boxplot • ggplot2, 2021, https://ggplot2.tidyverse.org/reference/ geom_boxplot.html.

Chiri. "Africa Economic, Banking and Systemic Crisis Data." *Kaggle*, 21 July 2019, https://www.kaggle.com/chirin/africa-economic-banking-and-systemic-crisis- data.

Cloud Team, IBM. "Python vs. r: What's the Difference?" IBM, 23 Mar. 2021, www.ibm.com/cloud/blog/python-vs-r#:~:text=R%20is%20an%20open%20source,elegant%20tools%20for%20data%20reporting.&text=Creating%20visualizations,learning%20and%20deep%20l earning%20algorithms.

"Currency Crisis: Causes, Signs, Impacts and Possible Solutions." *Penpoin.*, 15 Sept. 2021, https://penpoin.com/currency-crisis/.

Gelman, Andrew. "Statistical Modeling, Causal Inference, and Social Science." *Statistical Modeling Causal Inference and Social Science*, 10 Jan. 2020, https://statmodeling.stat.columbia.edu/2020/01/10/linear-or-logistic- regression-with-binary-outcomes/.

Hartman, Mitchell. "Why Is Inflation Necessary?" *Marketplace*, 12 Sept. 2019, https://www.marketplace.org/2019/09/12/why-is-inflation-necessary/.

Pennacchi, George. *Narrow Banking*. ANNUALREVIEWS; Journal: ARFE:Annual Review of Financial Economic, 2012, https://gpennacc.web.illinois.edu/ GPNarrowBankARFE.pdf.

Rulphus, Adam, and Roman Ahmed. "Logistic Regression Model: " Glm.fit: Fitted Probabilities Numerically 0 or 1 Occurred "." *RStudio Community*, 19 June 2018, https://community.rstudio.com/t/logistic-regression-model-glm- fit-fitted-probabilities-numerically-0-or-1-occurred/9828/15.

Schreiber-Gregory, Deanna. *Logistic and Linear Regression Assumptions: Violation Recognition and Control*. Henry M Jackson Foundation for the Advancement of Military Medicine, 2018, https://www.lexjansen.com/wuss/2018/130_Final_Paper_PDF.pdf.

Vettoretti, Martina, and Barbara Di Camillo. "A Variable Ranking Method for Machine Learning Models with Correlated Features: In-Silico Validation and Application for Diabetes Prediction." *Applied Sciences*, vol. 11, no. 16, 2021, p. 7740., https://doi.org/10.3390/app11167740.

Vidhya, Analytics. "Logistic Regression R: Introduction to Logistic Regression." *Analytics Vidhya*, 1 Nov. 2015, https://www.analyticsvidhya.com/ blog/2015/11/beginners-guide-on-logistic-regression-in-r/.