

Can machine learning help forecasting fiscal crises in emerging and developing countries?*



By Raffaele De Marchi and Alessandro Moro, Banca d'Italia

Keywords: Fiscal crises, debt sustainability, emerging and developing countries, machine learning techniques.

JEL codes: C18, C52, F34, H63, H68.

Pre-existing public debt vulnerabilities have been exacerbated by the effects of the pandemic, raising the risk of fiscal crises especially in emerging and developing countries. This underscores the importance of models aimed at forecasting the occurrence of sovereign debt crises and capturing the main determinants of fiscal distress episodes. In this regard, our paper shows that machine learning techniques are able to outperform standard econometric approaches. The analysis also identifies the variables that are the most relevant predictors of fiscal crises and provides an assessment of their impact on the probability of observing a crisis episode.

*The opinions expressed in the paper are those of the authors and do not necessarily reflect the views of the Banca d'Italia.

Introduction

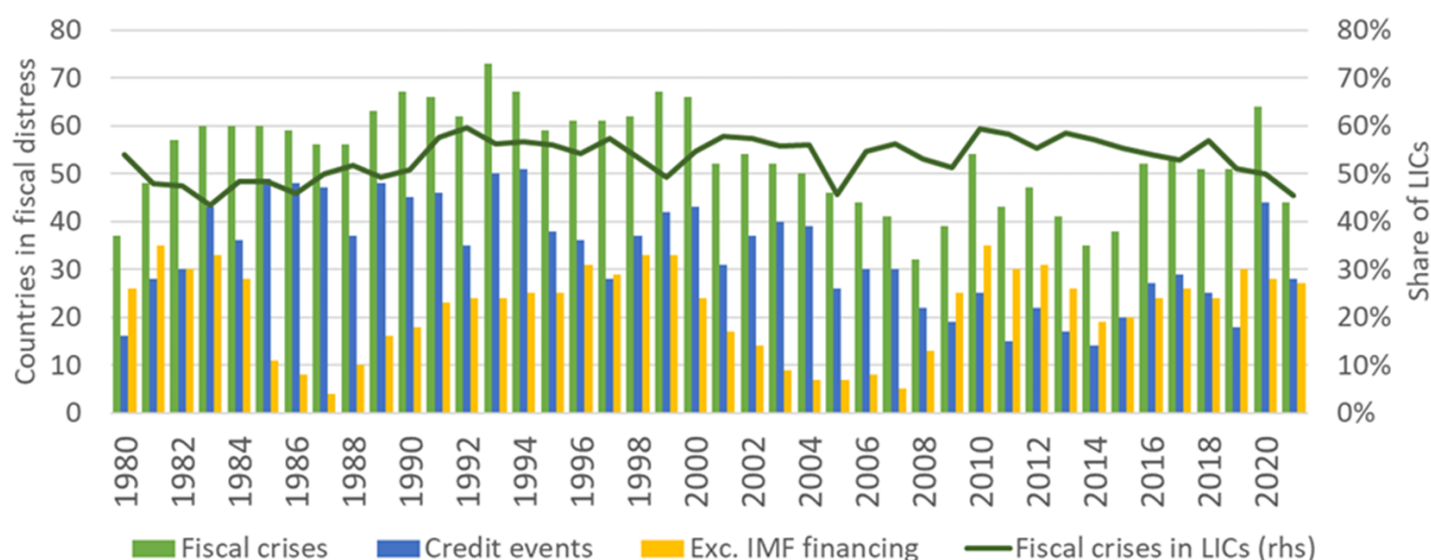
The global health crisis brought about by the outbreak of the COVID-19 pandemic and the associated economic effects have led to the strongest increase in debt recorded since World War II. While all components of debt are well above pre-pandemic levels, a significant contribution to the overall rise in global debt originated from additional borrowing by governments, as they simultaneously faced a collapse in economic activity (and therefore fiscal revenues) and a large expansion of expenditures needed to counter the health, social and economic consequences of the pandemic.

In emerging market economies (EMEs) and low-income countries (LICs), the latest debt buildup follows a decade of rapid and persistent increase in indebtedness, with the result that the levels reached by public debt ratios have raised serious concerns about the associated vulnerabilities and the risk of a spike in crises and defaults, considering also the lower “debt tolerance” of these countries and their urgent spending needs to achieve development goals. The elevated public debt in many EMEs and LICs can thus be considered a “pre-existing condition”, which has been further aggravated by the effects of the pandemic and the war in Ukraine. According to the latest debt sustainability analysis by the International Monetary Fund and the World Bank, close to 60 per cent of LICs are classified “at high risk of debt distress” or “in debt distress”; at the same time, several EMEs are also going through difficult debt situations. In this context, it is not surprising that the topic of developing countries’ debt is currently an important priority in the agenda of the international community. After having provided liquidity support to LICs immediately after the outbreak of COVID-19 through the Debt Service Suspension Initiative, in November 2020 the G20 launched the Common Framework for Debt Treatments, to promote a more structural and coordinated approach aimed at addressing solvency issues and facilitating debt restructurings for countries with unsustainable debt.

These considerations underscore the importance, from theoretical and policy perspectives, of models aimed at capturing the main determinants of fiscal stress episodes, and therefore ultimately signaling (and possibly anticipating) the risk of sovereign debt crises. In this regard, focusing the analysis on a large sample of EMEs and LICs, in De Marchi and Moro (2023) we show that machine learning methods are able to outperform standard econometric approaches in terms of a higher accuracy when forecasting fiscal crises.

In our analysis, we identify fiscal crises using a combination of two criteria that are commonly employed in the literature: (1) credit events associated with sovereign debt defaults and restructurings, obtained by drawing information from the Sovereign Default Database of the Bank of Canada and the Bank of England; (2) recourse to large-scale IMF financing (i.e. with access above 100 percent of the country’s IMF quota and fiscal adjustment as a program objective). To separate fiscal distress episodes into distinct crisis events, at least two years of no fiscal distress between two different crises are required; if there is only one year of no fiscal distress between two crisis episodes, these episodes are lumped together in one single crisis event.

Following this methodology, we identified fiscal crisis episodes for a sample of 140 countries (83 EMEs, 57 LICs) over the 1980-2021 period. Figure 1 shows the time series of the number of countries in fiscal distress based on credit events (criterion 1), exceptional IMF financing episodes (criterion 2), and combining the two criteria (our definition of fiscal crisis). Fiscal stresses were particularly frequent in the late 80s and in the 90s (at the time of the Latin American and LICs’ debt crises), and declined thereafter before rising again following the global financial crisis of 2008-2009, with a spike during the COVID-19 pandemic in 2020. This pattern suggests that sovereign defaults are clustered around periods of global financial distress. Credit events (criterion 1) represent the main factor triggering fiscal crises in our database, while the number of large-scale IMF financing (criterion 2) shows wide fluctuations with the latest increase starting after the global financial crisis. Around half of crises occurred in LICs, which suggests their relatively higher tendency to experience fiscal distress (as LICs are less numerous than EMEs in our database).

Figure 1: Time series of the number of countries in fiscal distress according to the two triggering criteria

Forecasting performance

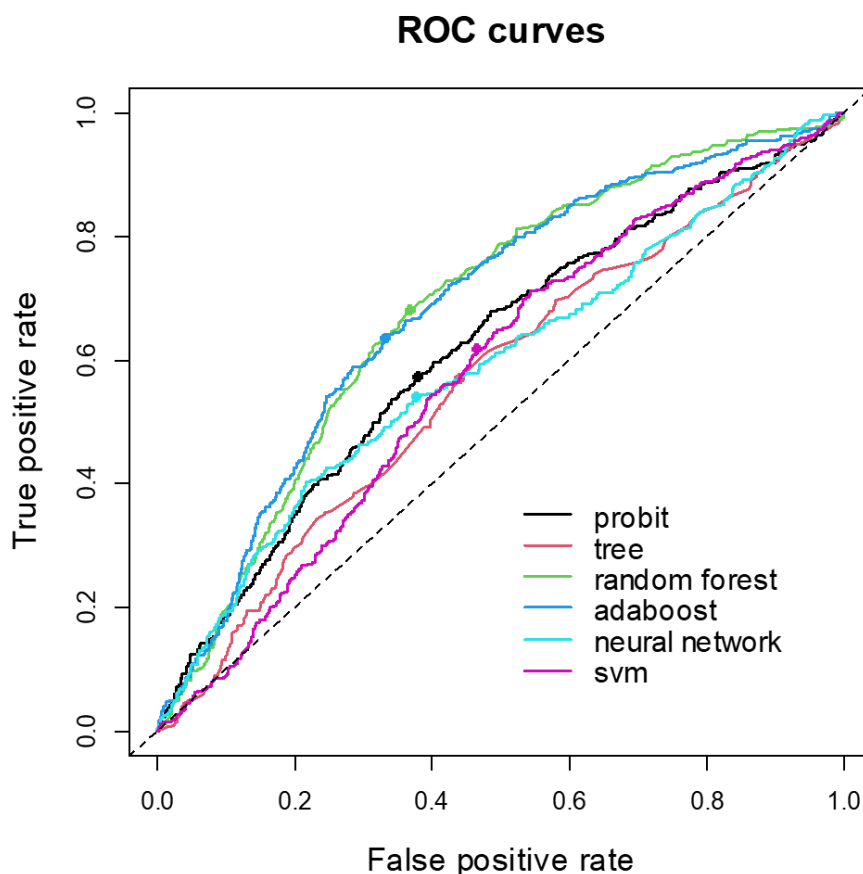
The aim of the analysis is to forecast the probability that at least one crisis episode happens in a given country within a time horizon of two (short-term forecasts), five (medium-term forecasts) and ten years (long-term forecasts), using a set of predictors and alternative predictive models.

These models include the probit, which is a standard econometric technique considered as a benchmark in the literature on debt crises (IMF, 2017), and several machine learning algorithms, including classification trees, random forests, adaboost, neural networks and support vector machines (Hastie et al., 2009).

Machine learning approaches are increasingly gaining popularity in economics, especially when considering forecasting tasks. On one hand, these methods are characterized by highly non-linear functional forms that are able to represent complex data patterns. On the other hand, such approaches are able to efficiently solve the trade-off between having highly parametrized models with a very accurate fit on observed data but volatile out-of-sample predictions (over-fitting) and simpler models with a poor fit on in-sample observations but stable out-of-sample forecasts (under-fitting). This problem is solved through the tuning of hyper-parameters, i.e. parameters that are specifically calibrated in order to maximize the out-of-sample forecasting accuracy.

Following the theoretical and empirical literature on the determinants of fiscal crises, we select a parsimonious set of predictors among both country-specific and global factors. Country-specific variables include traditional debt burden and fiscal indicators (such as total and external public debt to GDP ratios, primary balance, debt service, interest rate-GDP growth differential), macroeconomic and external sector variables (e.g., CPI inflation, variations of the real effective exchange rate, current account balance, stock of FX reserves and remittances), measures of economic development (GDP per capita), degree of financial openness, institutional quality indicators, and the previous crisis history of each country. Global variables are the traditional push factors that drive international capital flows: short and long term US interest rates, VIX index, real world growth, oil and commodity prices. We add to these variables the fraction of countries in debt distress in each year, with the aim of capturing possible contagion effects. These global variables should capture the systemic component of sovereign debt crises and their clustering around periods of financial distress.

Figure 2: ROC curves of the alternative models based on short-term forecasts



In order to evaluate the forecasting performance of the alternative methods, we compare the accuracy of the out-of-sample forecasts calculated for the period 2000-2021 using an iterative forecasting procedure with a rolling threshold. The outcomes of the different methods are expressed in terms of probability of observing a fiscal crisis. If this probability exceeds a given threshold, the model predicts a crisis event. The best model is the one that maximizes the area under the curve (AUC) of the receiver operating characteristic (ROC) curve. The ROC curve is the line that connects the pairs of false positive rates - the fraction of non-crisis periods wrongly classified as crisis events - and true positive rates - the fraction of crisis periods correctly classified as crisis events - changing the value of the probability threshold. A perfect classifier model is able to reach an AUC equal to one.

Figure 2 compares the ROC curves of the alternative models, calculated on the short-term forecasts. From the picture, the over-performance of the random forest and adaboost algorithms emerges sharply, as their AUCs are greater than those of the probit and the other machine learning methods. Table 1 reports the AUC for the other forecasting horizons and splitting the sample into EMEs and LICs. Overall, results show that the better forecasting accuracy of the random forest and adaboost can be generalised to the other forecasting horizons and to the two subsamples of countries, even though the over-performance with respect to the probit model is not always statistically significant.

Table 1: Forecasting performance of the alternative models measured according to the area under the ROC curve, considering different forecasting horizons and subsamples (EMEs and LICs)

Two-year horizon			
<i>Model</i>	<i>Full sample</i>	<i>EMEs</i>	<i>LICs</i>
Probit	0.617	0.581	0.589
Tree	0.566	0.532	0.536
Random Forest	0.684***	0.666***	0.621
AdaBoost	0.684***	0.664***	0.614
Neural Network	0.590	0.545	0.577
SVM	0.580	0.572	0.551
Five-year horizon			
<i>Model</i>	<i>Full sample</i>	<i>EMEs</i>	<i>LICs</i>
Probit	0.680	0.631	0.658
Tree	0.628	0.615	0.554
Random Forest	0.724**	0.704**	0.645
AdaBoost	0.712*	0.679*	0.654
Neural Network	0.652	0.624	0.606
SVM	0.591	0.598	0.543
Ten-year horizon			
<i>Model</i>	<i>Full sample</i>	<i>EMEs</i>	<i>LICs</i>
Probit	0.615	0.576	0.623
Tree	0.547	0.512	0.555
Random Forest	0.656	0.602	0.619
AdaBoost	0.673**	0.619	0.665
Neural Network	0.585	0.565	0.562
SVM	0.500	0.498	0.494

Notes: stars (*) refer to the significance levels of the test employed to assess whether each ML algorithm significantly outperforms the probit model according to the area under the ROC curve. Significance levels: *** p-value < 1%, ** p-value < 5%, * p-value < 10%.

Assessing the most important predictors and their effect on the probability of fiscal crises

The most important predictors can be detected using a popular model-agnostic procedure in the machine learning literature, called permutation variable importance (Gregorutti et al., 2017). The method consists in the random permutation of each variable included in a given model with the aim of evaluating how this reduction in the informative content of the same variable affects the overall forecasting performance of the model.

Once the most relevant predictors have been detected, it is possible to estimate their impact on the probability of fiscal crises using the notion of accumulated local effect (ALE), introduced by Apley and Zhu (2020). ALE measures how a change in a given variable affects the probability of default averaging over the other variables included in the model. This implies that with ALE it is possible to estimate the effect of each predictor taking into account the correlation with the other variables included in the model. In this respect, ALE is closer to the notion of *ceteris paribus* effects of standard econometrics.

Table 2 shows the list of the most important predictors, defined as the variables most frequently identified as the main relevant predictors by the best performing machine learning model (on the various forecasting horizons), using the permutation variable importance method. Moreover, the table reports the sign of the effect of the most relevant predictors on the probability of fiscal crises, estimated using the ALE approach.

Table 2: Most important predictors and their effect on the probability of fiscal crises

Predictor	Effect on the probability of crisis
Public external debt to GDP ratio	(+)
Previous crisis history (frequency of crisis episodes)	(+)
Institutional quality	(-)
Log GDP per capita	(-)
Stock of FX reserves to imports ratio	(-)
Interest rate-GDP growth differential	(+)
Current account balance relative to GDP	(-)

Machine learning models broadly confirm the validity of the set of variables employed to predict fiscal crises in traditional approaches: the most relevant predictors are the stock of public debt, in particular the public and publicly guaranteed debt held by foreign investors, the previous crisis history, the level of economic development, the quality of institutions, the stock of foreign exchange reserves, the current account balance and the interest rate-GDP growth differential. All these predictors have the expected effect on the probability of crisis occurrence. In detail, higher levels of debt stocks, more frequent crises in the past, or a more positive interest rate-GDP growth differential increase the probability of fiscal crises. Conversely, a higher degree of economic development, a superior institutional quality, a current account surplus and a larger stock of international reserves are associated with a lower probability of fiscal stress episodes. ■

References

- Apley, D. W., and Zhu, J. (2020). Visualizing the effects of predictor variables in black box supervised learning models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 82(4), 1059-1086.
- De Marchi, R., and Moro, A. (2023). *Forecasting fiscal crises in emerging markets and low-income countries with machine learning models* (No. 1405). Banca d'Italia Working Papers.
- Gregorutti, B., Michel, B., and Saint-Pierre, P. (2017). Correlation and variable importance in random forests. *Statistics and Computing*, 27(3), 659-678.
- Hastie T., Tibshirani R., Friedman J. (2009). *The Elements of Statistical Learning*. Springer Series in Statistics. Springer, New York
- IMF (2017). Review of the Debt Sustainability Framework for Low Income Countries: Proposed Reforms. *IMF Policy Papers*.

About the authors

Raffaele De Marchi is senior adviser in the International Relations and Economics Directorate at Banca d'Italia. He is involved in G20 issues and working groups, focusing on Multilateral Development Banks, IMF issues and financial support to low-income countries. His current research interests focus on development finance and debt in low-income countries. He holds a MSc in Economics and Business from Università degli Studi di Ancona, and a Master in Money, Banking and Finance from the Université Paris 1 Panthéon-Sorbonne.

Alessandro Moro works as an economist in the International Relations and Economics Directorate at Banca d'Italia. His research focuses on international macroeconomics, the implications of climate change and digital assets for international capital flows, the role of non-bank financial intermediaries in the propagation of global shocks to emerging markets. He holds a MSc in statistical sciences from University of Padua and a Ph.D. in economics from Ca' Foscari University of Venice.

SUERF Publications

Find more **SUERF Policy Briefs** and **Policy Notes** at www.suerf.org/policynotes



SUERF is a network association of central bankers and regulators, academics, and practitioners in the financial sector. The focus of the association is on the analysis, discussion and understanding of financial markets and institutions, the monetary economy, the conduct of regulation, supervision and monetary policy.

SUERF's events and publications provide a unique European network for the analysis and discussion of these and related issues.

SUERF Policy Briefs (SPBs) serve to promote SUERF Members' economic views and research findings as well as economic policy-oriented analyses. They address topical issues and propose solutions to current economic and financial challenges. SPBs serve to increase the international visibility of SUERF Members' analyses and research.

The views expressed are those of the author(s) and not necessarily those of the institution(s) the author(s) is/are affiliated with.

All rights reserved.

Editorial Board

Ernest Gnan
Frank Lierman
David T. Llewellyn
Donato Masciandaro
Natacha Valla

SUERF Secretariat
c/o OeNB
Otto-Wagner-Platz 3
A-1090 Vienna, Austria
Phone: +43-1-40420-7206
www.suerf.org • suerf@oenb.at