BANCA D'ITALIA

EUROSISTEMA

Temi di discussione

(Working Papers)

Forecasting fiscal crises in emerging markets and low-income countries with machine learning models

by Raffaele De Marchi and Alessandro Moro

March 2023

Number 1405

BANCA D'ITALIA
EUROSISTEMA

Temi di discussione

(Working Papers)

Forecasting fiscal crises in emerging markets and low-income countries with machine learning models

by Raffaele De Marchi and Alessandro Moro

Number 1405 - March 2023

ISSN 2281-3950 (online)

Designed by the Printing and Publishing Division of the Bank of Italy

FORECASTING FISCAL CRISES IN EMERGING MARKETS AND LOW-INCOME COUNTRIES WITH MACHINE LEARNING MODELS

by Raffaele De Marchi* and Alessandro Moro*

Abstract

Pre-existing public debt vulnerabilities have been exacerbated by the effects of the pandemic, raising the risk of fiscal crises in emerging markets and low-income countries. This underscores the importance of models designed to capture the main determinants of fiscal distress episodes and forecast sovereign debt crises. In this regard, our paper shows that machine learning techniques outperform standard econometric approaches, such as the probit model. Our analysis also identifies the variables that are the most relevant predictors of fiscal crises and assesses their impact on the probability of a crisis episode. Finally, the forecasts generated by the machine learning algorithms are used to derive aggregate fiscal distress indices that can signal effectively the build-up of debt-related vulnerabilities in emerging and low-income countries.

JEL Classification: C18, C52, F34, H63, H68.
Keywords: fiscal crises, debt sustainability, emerging and low-income countries, machine learning techniques.
DOI: 10.32057/0.TD.2022.1405

Contents

* Bank of Italy, International Relations and Economics Directorate.

1. Introduction[1]

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. In 2020, global debt (public and private, excluding financial corporations) increased by 29 percentage points to 257 percent of GDP. Even if the debt to GDP ratio decreased by 10 percentage points in 2021, due in large part to a sharp rise in nominal GDP in the largest economies (driven by the economic rebound from COVID-19 and the swift rise in inflation), global debt remained very high in an historical context and continued to rise in dollar terms, reaching USD 235 trillion at end-2021.[2] While all components of debt are well above pre-pandemic levels, a significant contribution to the overall rise in global debt came 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.

The surge in public debt in the aftermath of the pandemic has been very sizable in advanced countries, while it has been comparatively more limited in emerging market economies (EMEs) and low-income countries (LICs), due to a narrower fiscal space available and much tighter financing constraints. However, the levels reached by public debt ratios in developing countries have raised serious concerns about the associated vulnerabilities and the risk of a spike in crises and defaults. In fact, the latest build-up followed a decade of rapid and persistent increase in indebtedness, as many EMEs and LICs borrowed heavily, taking advantage also of low global interest rates. While being significantly lower than the average levels in advanced economies, current public debt ratios in developing countries are therefore quite worrisome, considering their lower “debt tolerance” and 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. Immediately before the COVID-19 pandemic, around 50 per cent of countries (mainly low-income) subject to the Debt Sustainability Analysis (DSA), jointly carried out by the International Monetary Fund (IMF) and the World Bank (WB), were already assessed as being “at high risk of debt distress” or “in debt distress”. This number has further increased following the pandemic, highlighting the difficult situation faced by many countries (additionally aggravated, more recently, by the repercussions of the war in Ukraine) and the existence of tangible and widespread risks in terms of debt sustainability. According to the latest DSA classification as of November 2022, the share of countries classified “at high risk of debt distress” or “in debt distress” has increased to 55 per cent.

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. Immediately after the outbreak of COVID-19, in April 2020, the G20 launched the Debt Service Suspension Initiative (DSSI), which provided a suspension of debt service payments due by eligible developing countries to official bilateral creditors.[3] Recognising that the mere liquidity support offered by the postponement of debt service payments was a temporary measure, which cannot adequately address the severe debt situation faced by many developing countries, in November 2020 the G20 endorsed also the Common Framework for Debt Treatments beyond the DSSI. In essence, the main goal of the Common Framework is to

[1] The opinions expressed in the paper are those of the authors and do not necessarily reflect the views of the Bank of Italy. The paper has been peer-reviewed by Sabina Marchetti and two Bank of Italy’s anonymous referees. We thank Sabina Marchetti, Mirko Moscatelli, Francesco Paternò, and the two anonymous referees for very helpful comments and suggestions.

[2] Source: IMF (2022), based on the 2022 Update of the IMF’s Global Debt Database.

[3] The 73 DSSI eligible countries include all IDA countries, that are current on any debt service to the IMF and the WB, and all least developed countries as defined by the United Nations, that are current on any debt service to the IMF and the WB. The DSSI eligible countries largely coincide with low-income developing countries in the IMF WEO definition.

provide a more structural and coordinated approach aimed at addressing solvency issues and facilitating debt restructurings for countries with unsustainable debt. In any case, developing countries will need a substantial amount of new financing, including from official sources, to support the recovery from the pandemic and enable meaningful progress towards their economic development goals.

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 signalling the risk of (and possibly anticipating) the occurrence of sovereign debt crises. In this regard, focusing the analysis on a sample of EMEs and LICs, our paper shows that machine learning methods are able to outperform standard econometric approaches in terms of a higher accuracy when forecasting fiscal crises.

There is an extensive literature on early warning indicators and crisis prediction, especially with reference to currency and financial crises. The literature on fiscal crises is relatively less developed, with the various studies differing on several dimensions, such as the definition of crisis events, the methodologies employed, the countries' coverage, as well as the main findings in terms of the variables most relevant in signalling the build-up of vulnerabilities.

Regarding the definition of fiscal crisis, our identification strategy is based on a combination of two criteria that are commonly adopted in this stream of literature (Hajivassiliou, 1994; Manasse et al., 2003; Ciarlone and Trebeschi, 2005; Kraay and Nehru, 2006; Fioramanti, 2008; Manasse and Roubini, 2009; Dawood et al., 2017; IMF, 2017; Liu et al., 2021). These two criteria are: credit events, such as defaults or restructurings, and recourse to large-scale IMF financing.

Focusing on methodological aspects, the current approach employed by the IMF to assess debt vulnerabilities in LICs is based on the probit model (IMF, 2017). Recently, the literature has proposed alternative approaches based on machine learning methods: decision trees (Rodriguez and Rodriguez, 2006), neural networks (Fioramanti, 2008), ensemble tree learning methods, particularly random forests (Badia et al., 2020; Jarmulska, 2022) and random forests and boosting trees (Hellwig, 2021). One of the main contributions of this paper is to compare several machine learning models, with the aim of assessing their relative usefulness and accuracy in predicting fiscal crises, and thus proposing possible enhancements to the existing literature on early warning indicators. In particular, we focus on various techniques, which have become increasingly popular in several fields of economic analysis (neural networks, support vector machines, decision trees, ensemble tree methods), and show that some of them deliver a significant improvement, in terms of out-of-sample predictive performance, compared to more traditional and standard econometric approaches, such as the probit model.

It is important to emphasise that a higher predictive performance implies also a more accurate estimation of the relationship between predictors and fiscal crises, and therefore a more plausible identification of the most relevant determinants than the one achieved using a parametric model with a poorer fit to the data. The results obtained from a more accurate model can thus provide better risk signals to be used in conjunction with the application of the analyst's judgment, which remains essential to any debt sustainability assessment as it allows to bring in country-specific considerations and other information that cannot be captured by a model.

In line with the previous considerations regarding the ongoing public debt vulnerabilities in developing countries, we focus our analysis on a large number of EMEs and LICs. This latter category, in particular, has been covered only in a limited number of studies so far (IMF, 2017; Cerovic et al., 2018; Badia et al., 2020; Hellwig, 2021), and its inclusion in our analysis provides relevant insights for the current policy debate. Moreover, while most studies have focused their attention on short-term forecasting horizons of two years, we extend the analysis by investigating the prediction accuracy of the alternative methods on several horizons, of two, five and ten years. This offers insights about the risks of crisis on a longer time span, closer to the typical maturities of the lending facilities provided by the IMF and other international financial institutions.

In addition to the goal of improving the prediction accuracy of fiscal crisis episodes, this paper aims to shed some light on their main determinants. Machine learning techniques allow to fit complex functional forms to the available data, taking into account nonlinearities and interactions not easily captured by standard econometric techniques, with the result of achieving a better out-of-sample predictive performance. However, flexibility comes at the cost of a lower interpretability of results, in the absence of the conventional tools of classic econometric models, such as the significance tests on the estimated coefficients. In order to reduce this limitation, we rely on alternative statistical procedures to identify a set of robust predictors of fiscal crises, providing information on the relative importance of variables and the shape of their relationship with the probability of fiscal crisis occurrence (as in Cascarino et al., 2022).

In this regard, we have tried to strike a balance between having a limited number of variables subjectively selected (as in Rodriguez and Rodriguez, 2006; Fioramanti, 2008; IMF, 2017; Cerovic et al., 2018; Jarmulska, 2022) and expanding the dataset to include a very large number of predictors, following a big data approach (as in Badia et al., 2020; Hellwig, 2021).[4] On one hand, a limited set of explanatory variables is likely to deliver a poor out-of-sample forecasting performance. On the other, the inclusion of an extremely large set of highly correlated variables limits the ability of machine learning procedures to correctly assess the relative importance of each predictor. We therefore solve this trade-off by constructing a dataset of 21 explanatory variables, selected from a wide range of categories (such as total and external public debt, past crisis history, fiscal variables, external accounts, institutional quality and global factors), informed by economic theory considerations as well as drawings from the most frequent variables employed in the literature.[5]

Our work finds that the random forest and, to a lesser extent, the adaboost algorithms are able to outperform traditional approaches, like the probit, as well as other machine learning techniques, including simple decision trees, neural networks and support vector machines. This result holds for all forecasting horizons and both types of countries (EMEs and LICs); the outperformance is especially strong at short forecasting horizons (2 years). 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 (henceforth, public external debt), 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-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-growth differential are likely to 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. Finally, we employ the outcomes from the random forest model to construct an aggregate index, which is able to signal effectively the evolution of fiscal risk in EMEs and LICs.

The rest of the paper is organized as follows. Section 2 specifies the definition of fiscal crisis and illustrates the data used to identify crisis occurrence in EMEs and LICs. Section 3 presents our forecasting problem, discussing the models employed in the analysis and the methodologies adopted

[4] In these studies, the idea of a very large dataset of predictors is taken to the extreme, by including a vast array of variables (the only constraint being data availability) as well as several transformations of each of them, yielding a total of 748 indicators.

[5] A valid alternative could have been to insert all the explanatory variables of interest, aggregate highly correlated variables in separate groups, and then consider the importance at group level instead of at individual variable level. The drawback of this alternative is that it is not trivial to define groups having a low level of correlation between them.

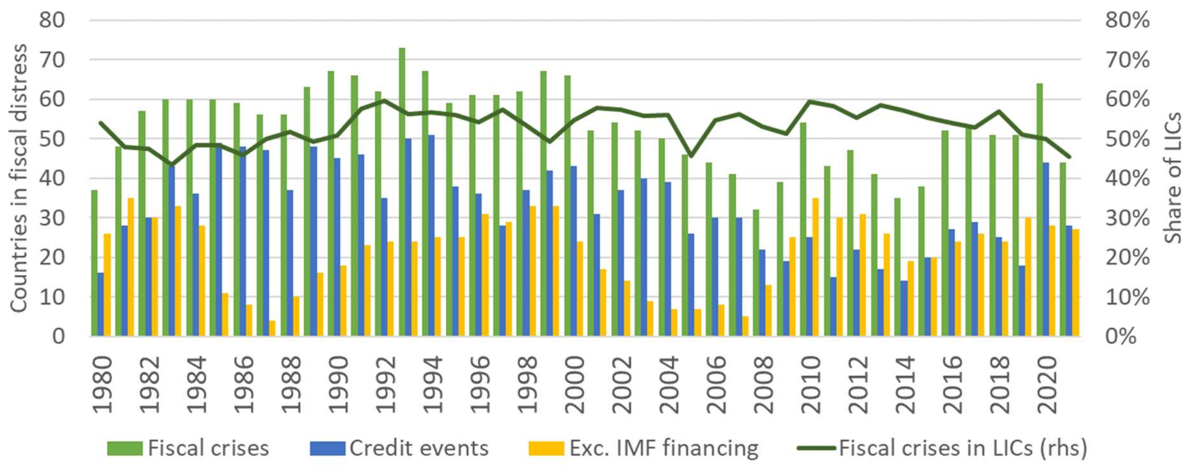
to evaluate the forecasting accuracy, to assess the importance of predictors and to estimate their effects on the probability of crisis. Section 4 summarises the main empirical results. Finally, Section 5 offers some concluding remarks.

2. Identifying fiscal crises

The empirical literature on fiscal crises has developed several approaches for the identification of fiscal distress episodes. Our identification strategy is based on a combination of two criteria that are commonly employed in this stream of literature (Hajivassiliou, 1994; Manasse et al., 2003; Ciarlone and Trebeschi, 2005; Kraay and Nehru, 2006; Fioramanti, 2008; Manasse and Roubini, 2009; Dawood et al., 2017; IMF, 2017; Liu et al., 2021). These two criteria are: (1) credit events associated with sovereign debt defaults and restructuring; (2) recourse to large-scale IMF financing.

Regarding the first criterion, and following Gerling et al. (2018), a credit event in country *i* at time *t* is identified by looking at the aggregated nominal stock of sovereign debt obligations in default. Data are taken from the Bank of Canada (BoC) - Bank of England (BoE)'s annual database of sovereign defaults. This database includes sovereign defaults to both private and official creditors (such as the Paris Club, other bilateral lenders, international financial institutions, etc.), defined as any debt operation that inflicts an economic loss on creditors (e.g., outright default, restructuring, or rescheduling). Both local currency defaults and external defaults on sovereign debt denominated in foreign currency are reported. In order to exclude small-scale technical defaults, we consider default amounts of substantial size (above 0.2 percent of GDP). Moreover, we only consider defaulted nominal amounts that grow by a substantial rate (more than 10 percent per year) to exclude cases of continued reporting of previously defaulted amounts.[6]

Fig. 1: Time series of the number of countries in fiscal distress according to the two triggering criteria and LICs' share of total fiscal crises



The second criterion captures the circumstances in which exceptionally large official financing from the IMF allows countries in fiscal distress to avoid outright default or restructuring. In fact, this support is usually provided to countries with balance of payments imbalances that have difficulties in meeting their international obligations. According to this criterion, a crisis event is identified in country *i* when the country signs an IMF financial arrangement with access above 100 percent of its quota and fiscal adjustment as a program objective, for all the years *t* within the duration of the program. The threshold is consistent with previous works in this research area, such as Manasse et al.

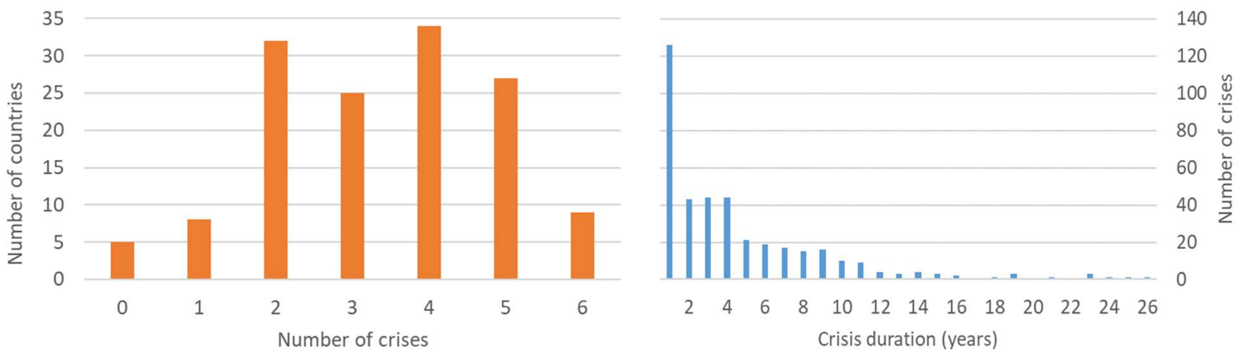
[6] The thresholds on size and growth of defaults are coherent with other works, such as Gerling et al. (2018), Badia et al. (2020), and Hellwig (2021).

(2003), Ciarlone and Trebeschi (2005), Fioramanti (2008), Manasse and Roubini (2009), Gerling et al. (2018), and Hellwig (2021).[7]

Then, for each country i and year t we construct a crisis dummy variable $c_{i,t}$ taking the value of one when at least one of the two criteria is met, and zero otherwise. 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 identify fiscal crisis episodes for a sample of 140 countries (83 EMEs, 57 LICs) over the 1980-2021 period. The complete list of countries included in the analysis is presented in Table A1 in Annex A. Fig. 1 shows the time series of the number of countries in fiscal distress considering credit events (criterion 1), exceptional IMF financing episodes (criterion 2), and combining the two criteria (our definition of fiscal crisis).[8] 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 (Kaminsky and Vega-Garcia, 2016). 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. Fig. 1 displays also the evolution of the share of fiscal crises involving LICs: around half of crises occurred in LICs, which suggests their relatively higher tendency to experience fiscal distress (LICs are less numerous than EMEs in our database).

Fig. 2: Distribution of the number of fiscal crises (left panel) and crisis duration (right panel)



In the 1980-2021 period there were 472 crisis events. Most of the countries in our sample (118 out of 140 countries) experienced from 2 to 5 crises (Fig. 2, left panel): the average number of crises per country is 3.4. Around two thirds of all fiscal distress episodes last between 1 and 4 years, while the average duration of a fiscal crisis is 4.6 years (Fig. 2, right panel).

3. Forecasting methodology

Machine learning approaches are increasingly gaining popularity in econometrics, especially when considering forecasting tasks (Mullainathan and Spiess, 2017). On one hand, these methods are characterised by highly non-linear functional forms that are able to represent complex data patterns.

[7] We notice that the current thresholds for the IMF Exceptional Access Policy are much higher (annual and cumulative limits are at 145 and 435 per cent of a country's quota, respectively); however, we prefer to keep the crisis identification threshold at 100 per cent to be coherent with the rest of the literature.

[8] It is important to stress that the number of fiscal crisis episodes in a given year is not necessarily the sum of credit events and exceptional IMF financing episodes because there might be some degree of overlapping between the two criteria.

On the other hand, such approaches are able to efficiently solve the trade-off between having highly parametrised 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 maximise the out-of-sample forecasting accuracy (Hastie et al., 2009).

Our forecasting exercise can be formalised as follows. Let $y_{i,t+1}^{h}$ denote the event equal to one if at least one crisis episode happens in country *i* between year *t*+1 and *t*+*h*, and zero otherwise:

$$y_{i,t+1}^{h} = \bigcup_{s=1}^{h} \{c_{i,t+s}\} \quad (1)$$

in which h is the forecasting horizon. In particular, for each time period *t*, we are interested in calculating the probability of a fiscal crisis in a given country within a horizon of two, five and ten years (short, medium and long-term forecasts) as a function of a set of time *t* predictors. Therefore, the aim of the analysis is to find the best (in terms of forecasting accuracy) machine learning model $m(.)$ for such probability:

$$\pi_{i,t+1}^{h} \equiv \mathbb{P}\left(y_{i,t+1}^{h} = 1 | c_{i,t} = 0, X_{i,t}\right) = m\left(X_{i,t}; \theta, \lambda\right) \quad (2)$$

where $X_{i,t}$ is a vector of country-specific and global variables, θ is a vector of parameters while λ are the model's hyper-parameters. The conditioning for $c_{i,t} = 0$ makes explicit the fact that we consider only transitions from non-crisis states to crisis or non-crisis episodes, as in the majority of empirical studies in this field (e.g., Badia et al., 2020; Hellwig, 2021); hence, we do not model the permanence in a state of crisis or the exit from a crisis.

Table 1: Model predictors

Country-specific variables			**Global variables**		
Variable	*Meaning*	*Source*	*Variable*	*Meaning*	*Source*
$D_{i,t}$	Total public debt on GDP ratio	WEO	$WGDPg_t$	World real GDP growth	Refinitiv
$ED_{i,t}$	Public external debt on GDP ratio	WDI	$\ln VIX_t$	Natural logarithm of the VIX index	Refinitiv
$DSED_{i,t}$	Debt service on public external debt relative to exports of goods, services and primary income	WDI, WEO	$r_{US,t}^{Short}$	US 3-month treasury bill rate	Refinitiv
$PB_{i,t}$	Primary balance (primary net lending/net borrowing) on GDP ratio	WEO	$r_{US,t}^{Long}$	US 10-year government bond yield	Refinitiv
$r_{i,t} - g_{i,t}$	Interest rate-growth differential (*)	WEO	$\Delta\ln P_t^{com}$	Growth of the S&P GSCI commodity price index	Refinitiv
$Rem_{i,t}$	Remittances on GDP ratio	WDI, WEO	$\Delta\ln P_t^{oil}$	Growth of the Brent oil price index	Refinitiv
$FXR_{i,t}$	FX reserves on imports ratio	IFS, WEO	TC_t	Percentage of countries in a fiscal crisis state	BoC-BoE, IMF
$\ln GDPpc_{i,t}$	Natural logarithm of GDP per capita in purchasing power parity terms	WEO			
$CPI_{i,t}$	CPI inflation	WEO			
$\Delta\ln REER_{i,t}$	Real effective exchange rate growth	IFS, Darvas (2012)			
$CA_{i,t}$	Current account balance relative to GDP	WEO			
$ka\_open_{i,t}$	Chinn-Ito index of financial account openness	Chinn and Ito (2006)			
$gee_{i,t}$	Government effectiveness indicator	WGI			
$CH_{i,t}$	Crisis history, defined in each year *t* as the historical frequency of crisis years from 1980 up to time *t*.	BoC-BoE, IMF			

Notes: (*) $r_{i,t}$ is calculated as the difference between primary balance and overall fiscal balance divided by the stock of public debt (average of debt stocks at *t* and *t-1*).

Following the theoretical and empirical literature on the determinants of fiscal crises (for an overview see IMF, 2017), we select a parsimonious set of predictors that includes 14 country-specific variables and 7 global factors (Table 1). 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-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 (Chinn and Ito, 2006), institutional quality indicators (government effectiveness index), and the crisis history of each country. Global variables are the traditional push factors that drive international capital flows (Koepke, 2019): short and long US 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 variables should capture the systemic component of sovereign debt crises and their clustering around periods of financial distress (Kaminsky and Vega-Garcia, 2016).

In line with the majority of the papers in this stream of literature, a parsimonious set of explanatory variables is preferred because it facilitates the interpretability of results and allows a better assessment of the relevance of each variable. In fact, the importance of each variable is derived more accurately when the degree of correlation among predictors is low (see the subsection on assessing variable importance). An alternative approach, followed by Badia et al. (2020) and Hellwig (2021), consists in the inclusion, as predictors, of many variables together with several transformations of them. However, since these variables are highly correlated, the increase in the forecasting performance is rather limited, with significant costs in terms of a more problematic interpretability of results.

As a robustness check, we have also included short-term debt, the structure of official bilateral debt by creditor country (United States, United Kingdom, France, Germany, Japan, China), public revenues and expenditures, a variable capturing past borrowing history from the IMF, domestic monetary policy rates, the stock of private debt, the stock of external private debt, the interaction between the oil price index and a dummy for fuel exporters, as well as additional capital flows variables (portfolio, other investments and FDI inflows and outflows). We also control for the vulnerability to climate change using the University of Notre Dame's ND-GAIN Country Index. However, these predictors do not improve the overall forecasting performance:[9] for this reason, we decide to exclude them from our specification. Results are also robust when measuring institutional quality with other indicators, such as the rule of law index (widely used in the literature but less related to public debt sustainability than government effectiveness), the regulatory quality or the political stability and absence of violence index.[10] Replacing the Chinn-Ito index with capital restriction indicators (Fernandez et al., 2016) does not alter the main results.

Predictive models

Regarding the choice of the class of models $m(.)$ to forecast fiscal crises, our analysis considers several methods as explained below.[11]

Probit model: this is a benchmark in the traditional literature on debt crises, and is also the model underlying the debt sustainability framework for LICs developed by the IMF and the World Bank (see IMF, 2017). In this case,

$$m(X_{i,t};\beta) = \Phi(X_{i,t}\beta) \quad (2.1)$$

[9] This is probably also due to a lack of data availability for all countries and time periods (especially in relation to LICs, where data issues are especially acute), which requires the imputation of many missing values.
[10] We perform a horse race among these indicators and the government effectiveness index results as the best predictor.
[11] For a detailed description of these models, see Hastie et al. (2009).

where $\Phi(.)$ is the cumulative probability distribution of a standard Gaussian distribution, β are the parameters to be estimated and there are no hyper-parameters.

Decision tree: decision trees (Breiman et al., 1984) are sequential decision rules by which a sample is recursively divided into subgroups with different levels of default risk. Each split is performed recursively according to the couple (predictor, threshold value of the predictor) that is able to reach the lowest impurity according to the Gini's index.[12] Once the space of the variables is partitioned into K regions R_k, with n_k observations in each region, the probability of default is computed as:

$$m(X_{i,t};\tau,\nu) = \sum_{k} \left(\frac{1}{n_k} \sum_{j \in R_k(\tau,\nu)} y_{j,t+1}^{h} \right) I\{X_{i,t} \in R_k(\tau,\nu)\} \quad (2.2)$$

The parameters of the model are the thresholds (τ) used to perform the splitting while the unique hyper-parameter (ν) is a cost-complexity parameter that controls the tree depth, i.e., the maximum number of splits: the higher the number of splits, the higher the goodness of in-sample fit but (potentially) the lower the out-of-sample forecasting performance (risk of over-fitting).

Random forest: the random forest algorithm for classification problems (Breiman, 2001) is based on the estimation of a high number (N_T) of trees, each of them characterised by a large depth and hence intentionally over-fitted with the aim of capturing some specific features of the data. Then, the risk of over-fitting is significantly reduced in the forest by taking the average of the predictions obtained from the different trees (m_j):

$$m(X_{i,t};\tau,\omega) = \frac{1}{N_T} \sum_{j=1}^{N_T} m_j(X_{i,t};\tau_j,\omega) \quad (2.3)$$

To diversify predictions and reduce the correlation among trees, each tree is estimated on a synthetic sample drawn randomly from the original estimation sample. Each split is performed choosing the best predictor from a limited number (ω) of randomly selected splitting variables. The coefficients of this model are the thresholds used by each tree to split the data, while the main hyper-parameter is represented by the number of variables selected for each split ω (a rule-of-thumb is to choose the number of variables equal to the squared root of the number of original variables in the dataset).

Adaboost: the adaboost algorithm (Freund and Schapire, 1997) sequentially applies trees with a limited size (characterised by one or two splits and hence by a poor fit) to modified versions of the original dataset. In particular, at each iteration a simple tree is estimated on a weighted version of the original dataset that gives more weight to the observations misclassified in the previous step. In this way, the misclassified observations have more influence on the new classifier. This new classifier is thus forced to learn from the classification errors of the previous model and in this way improves its performance. The final prediction of the algorithm is given by a weighted average (with weights w_j) of the estimated probabilities of the simple trees, in which more weight is given to the outcomes of the trees with a better fit:

$$m(X_{i,t};\tau,N_I) = \sum_{j=1}^{N_I} w_j m_j(X_{i,t};\tau_j) \quad (2.4)$$

[12] Impurity refers to the heterogeneity in the values assumed by the dependent variable within the two groups of observations created by the split. In our problem, impurity is minimised (and equal to zero) when one of the two groups includes all the crisis events and the other group includes all the non-crisis events.

The coefficients of this classifier are the thresholds used by each tree to split the data, while the hyper-parameter we consider is the number of iterations (N_I).

Neural network: neural networks (Intrator and Intrator, 1993) are complex non-linear functions of the original variables. In particular, a neural network with two layers (M, P), where M and P are the number of so-called neurons in the first and second layer, respectively, can be written as:

$$m(X_{i,t}; \alpha, \delta, M, P) = f\left(\alpha^{(3)} + \sum_{p=1}^{P} f\left(\alpha_p^{(2)} + \sum_{m=1}^{M} f\left(\alpha_m^{(1)} + X_{i,t}\delta_m^{(1)}\right)\delta_p^{(2)}\right)\delta^{(3)}\right) \quad (2.5)$$

in which $f(.)$ is the activation function (e.g., the logistic function in classification problems). Each neuron in the first layer tries to capture one particular feature of the data applying different weights to the original variables. Then, the information collected by the different neurons in the first layer are combined by the neurons in the second layer. Finally, the information processed by the neurons in the second layer is further combined to obtain a probability measure. The parameters of the model are the weights $(\alpha, \delta) = \left(\alpha_m^{(1)}, \alpha_p^{(2)}, \alpha^{(3)}, \delta_m^{(1)}, \delta_p^{(2)}, \delta^{(3)}\right)$, while the hyper-parameters are (M, P). Under certain conditions, neural networks can be seen as a generalisation of the logit model: in fact, removing the hidden neuron structure, equation (2.5) becomes a logistic regression in the original $X_{i,t}$ variables.[13]

Support Vector Machine (SVM): support vector machines (Cortes and Vapnik, 1995) try to identify the best hyperplane in the space of (transformed) variables that is able to separate crisis and non-crisis observations into two distinct regions with a predefined tolerated error C. The hyperplane is a linear combination of the transformed original variables:

$$m(X_{i,t}; \rho, C) = \varphi(X_{i,t})\rho \quad (2.6)$$

in which $\varphi(.)$ is a possibly non-linear function (in our analysis we choose a radial kernel) and ρ is the vector of parameters. The hyper-parameter C to be tuned is the maximum number of observations allowed to be on the wrong side of the hyperplane. In particular, if $C = 0$ no violation is tolerated, therefore the model obtains a perfect fit to in-sample data but it might display a poor out-of-sample classification performance. It is worth stressing that, in contrast to previously discussed methods, SVMs do not provide as output the probability of a crisis event but only discrete predictions (crisis/no crisis).

Model estimation, hyper-parameter tuning and out-of-sample forecasts

In order to evaluate the forecasting performance of the alternative methods, out-of-sample forecasts are calculated and compared using an iterative forecasting procedure with a rolling threshold. More precisely, for each year *t*, the sample is divided into two sets. Data from year *t-h-s+1* to *t-h* (in-sample observations/ training set) are used to train the model, i.e. for hyper-parameters tuning ($\hat{\lambda}$) and for the

[13] This contributes to explain the reason why we prefer to consider the probit model as the benchmark traditional approach. In any case, the logit model as well as the complementary log-log model deliver results very close to the ones of the probit, being these three approaches characterised by the absence of hyper-parameters and, hence, by the same degree of flexibility.

estimation of model parameters ($\hat{\theta}$).[14] Time t observations (test set) are employed to compute out-of-sample forecasts for the probability of a crisis during the period t+1,.., t+h. In this way, there is no overlapping between in-sample and out-of-sample information. Then, the cut-off is moved by one year and the procedure is repeated.

The parameter s defines the length of the estimation window in terms of periods used to train the model. A higher value increases the number of observations used to estimate models but makes the estimation more dependent on past and potentially obsolete observations: after a sensitivity exercise, we choose a 10-year window.

At each time t, model hyper-parameters are tuned with k-fold cross-validation (we choose k=10).[15] The training sample (data from year t-h-s+1 to t-h) is divided into k non-overlapping parts. For each hyper-parameter value (chosen on a grid of possible values), the model is estimated k-times using k-1 subsamples and employed to calculate out-of-sample forecasts on the excluded subsample. Using these forecasts it is possible to determine the hyper-parameter with the best predictive performance according to a given criterion (more on this below). Once the best hyper-parameter is selected, the model is estimated using the overall training sample.

Using this iterative procedure, we produce out-of-sample forecasts for the period 2000-2021[16] in terms of estimated probability of fiscal crisis $\hat{\pi}_{i,t+1}^{h}$ or as predicted 0-1 outcome $\hat{y}_{i,t+1}^{h} = I\left(\hat{\pi}_{i,t+1}^{h} > \phi\right)$ - which depends on the choice of the probability threshold ϕ. The forecasting performance of each method is then assessed using alternative criteria:

1) The best model is the one that maximises the area under the curve (AUC) of the receiver operating characteristic (ROC) curve. Let define false positive rates (FPR) the fraction of non-crisis periods wrongly classified as crisis events and true positive rates (TPR) the fraction of crisis periods correctly classified as crisis events. Both of them depend on the choice of the threshold ϕ. The ROC curve is the line that connects the pairs of (FPR,TPR) for different values of ϕ. The best model in the space (FPR, TPR) is the one closer to the point (0,1), i.e. the model that maximises the area under the ROC curve.

2) The best model is the one that maximises the log-likelihood calculated using the out-of-sample forecasts (N_f is the number of forecasts):

$$\ln L^{h}\left(\hat{\theta}, \hat{\lambda}\right) = \frac{1}{N_f}\sum_{i}\sum_{t}\left[y_{i,t+1}^{h}\ln\left(\hat{\pi}_{i,t+1}^{h}\right) + \left(1 - y_{i,t+1}^{h}\right)\ln\left(1 - \hat{\pi}_{i,t+1}^{h}\right)\right] \quad (3)$$

3) The best model is the one that minimises the mean squared error (MSE):

$$MSE^{h}\left(\hat{\theta}, \hat{\lambda}\right) = \frac{1}{N_f}\sum_{i}\sum_{t}\left[y_{i,t+1}^{h}\left(1 - \hat{\pi}_{i,t+1}^{h}\right)^{2} + \left(1 - y_{i,t+1}^{h}\right)\left(\hat{\pi}_{i,t+1}^{h}\right)^{2}\right] \quad (4)$$

Annex B provides further details on the algorithm employed. Denoting with $\Omega_{m}^{h}\left(\hat{\theta}, \hat{\lambda}\right)$ one of the three criteria above evaluated for method m, it is possible to test whether the difference in

[14] It is worth stressing that in models requiring the resampling of observations, such as the random forest algorithm, we take into account the temporal dimension using time as a stratification variable for the sampling, in order to generate synthetic data with the same panel structure of the original dataset.

[15] In each period, we observe on average 86 countries. Having selected a 10-year window for the estimation, the cross-validation exercise is performed using roughly 860 observations.

[16] Hence, we have 20 iterations when the forecasting horizon is h=2, 17 iterations when h=5, and 12 iterations when h=10.

performance of each machine learning method with respect to the probit model is statistically significant using the following test statistic (Demler et al., 2017):

$$z_m^h = \frac{\Omega_m^h(\hat{\theta}, \hat{\lambda}) - \Omega_{probit}^h(\hat{\theta}, \hat{\lambda})}{\sqrt{\mathrm{var}_B\left(\Omega_m^h(\hat{\theta}, \hat{\lambda}) - \Omega_{probit}^h(\hat{\theta}, \hat{\lambda})\right)}} \quad (5)$$

and comparing it with a standard Gaussian distribution;[17] $\mathrm{var}_B\left(\Omega_m^h(\hat{\theta}, \hat{\lambda}) - \Omega_{probit}^h(\hat{\theta}, \hat{\lambda})\right)$ at the denominator is the variance of the differences in performance evaluated using bootstrap resampling. A standard number of bootstrap samples (*B*=100) is used, which guarantees a precise estimation of the variance.

Assessing variable importance

The most important predictors of fiscal crises can be detected using a popular model-agnostic procedure in the machine learning literature, called permutation variable importance (Breiman, 2001; Gregorutti et al., 2017; Fisher et al., 2019). The method consists in the random permutation of each variable included in a given model with the aim of evaluating how this affects the overall forecasting performance.

In our specific problem, this procedure can be applied as follows. For each variable *j* and for each period *t*, we estimate a given model using information up to time *t-h* (as before). Then, at time *t* we create a new set of predictors $X_{it}^{(j)}$ by assigning at random to the *j*-th variable in $X_{i,t}$ a value assumed by the same variable in a different country and/or a different period (before *t*). Then, we compute the out-of-sample forecasts for a crisis in *t*+1,…,*t*+*h* using $X_{i,t}$ and $X_{it}^{(j)}$, obtaining $\hat{\pi}_{i,t+1}^h = m\left(X_{i,t}; \hat{\theta}, \hat{\lambda}\right)$ and $\hat{\pi}_{i,t+1}^{h,(j)} = m\left(X_{it}^{(j)}; \hat{\theta}, \hat{\lambda}\right)$, respectively. Using one of the three criteria Ω_t cited above (AUC, log-likelihood, MSE), it is possible to compute for each year *t* a measure of the out-of-sample forecasting performance for the original set of variables $\Omega_t(\hat{\pi}^h)$ and for the permuted one $\Omega_t\left(\hat{\pi}^{h,(j)}\right)$. We can repeat this computation for each time period *t*.

By replicating the previous steps *R* times (*R*=10), it is possible to compute a measure of variable importance (VI) for variable *j* as:

$$VI_j = \frac{1}{RT} \sum_r \sum_t \left[\Omega_{r,t}(\hat{\pi}^h) - \Omega_{r,t}\left(\hat{\pi}^{h,(j)}\right)\right] \quad (6)$$

This method has the advantage that it can be applied to all models, no matter the functional form of $m(.)$. Since this procedure is based on the permutation of one variable at a time, the lower the correlation among predictors, the higher the ability of the procedure to correctly estimate the importance of a given predictor.

Accumulated local effects

In order to identify the effect of each predictor on the probability of fiscal crisis, we rely on the notion of accumulated local effect (ALE), introduced by Apley and Zhu (2020).

ALE is defined for a model $m(.)$ and a variable $X(j) \in X$ as:

[17] Given that both the log-likelihood and the MSE criteria used to evaluate the performance of alternative models are averages over countries and time periods, we can rely on the central limit theorem to approximate the test distribution with the Gaussian. Moreover, Demler et al. (2017) show that also the difference in the AUC has an asymptotic normal distribution. In addition to these theoretical results, we use bootstrapped samples to check the goodness of the Gaussian approximation in our exercise.

$$ALE_j(w) = \int_{w_0}^{w} \mathbb{E}_{X(-j)|X(j)=x} \left[\frac{\partial m(X; \hat{\theta}, \hat{\lambda})}{\partial X(j)} \right] dx + c \quad (7)$$

where w_0 is a value close to the lower bound of the effective support of $X(j)$, $X(-j)$ is the vector of predictors excluding variable j and c is a constant, usually selected so that the expected value of ALE is equal to zero. ALE measures how a change in the j-th variable affects the probability of default averaging over the other variables included in the model. The marginal effects are then accumulated (with the integral) up to value w. In intuitive terms, $ALE_j(w)$ measures the change in the probability of default when the j-th variable assumes a value equal to w with respect to the average probability of default.

The innovative aspect of ALE with respect to other tools (e.g., the partial dependence plots) is that the expectation is calculated considering only the values of $X(-j)$ that are plausible, in probabilistic terms, given the values assumed by $X(j)$. This implies that with ALE it is possible to estimate the effect of predictor j taking into account the correlation with the other variables included in the model. In other words, ALE are closer to the notion of *ceteris paribus* effects of standard econometrics, better capturing the effect of a given predictor while keeping fixed the other variables included in the specification.

4. Results

This section illustrates the main findings of our analysis, discussing them in relation to the following three aspects: i) forecasting performance of the alternative models: ii) identification of the most important predictors; iii) relationship of the relevant predictors with the probability of crisis occurrence. The results are presented for the different forecasting horizons (*h=2,5,10*), and for both the full sample of countries and for EMEs and LICs, separately. The section finally discusses a forward-looking aggregate indicator of fiscal crisis risk in EMEs and LICs, developed by using the outcomes of our best performing model.

Forecasting performance

The forecasting performance of the alternative machine learning methods, with regard to different forecasting horizons and subsamples, is presented in Table 2. It is worth to stress that the models are estimated and the forecasts are performed considering only two types of observations: transitions from non-crisis states in *t* to crises in *t+1,…,t+h,* and stays in a non-crisis situation between time *t* and *t+h*. Results in Table 2 are obtained by estimating the alternative models on our full sample of EMEs and LICs, to deliver out-of-sample forecasts for the whole sample and for the two subgroups, separately.

When we consider a forecasting horizon of two years, in the period 1980-2021 the dependent variable assumes the value one (signalling a transition from non-crisis states to crises in the following two years) in 846 cases and the value of zero (meaning stays in a non-crisis state in the following two years) in 2,612 cases. 458 of these transitions happened in EMEs (over a total of 2,320 observations for EMEs) while 388 occurred in LICs (over a total 1,138 observations for LICs). Looking at the results for the full sample, the probit model is outperformed by the random forest and the adaboost methods according to all the three criteria employed (AUC, MSE e log-likelihood). The improvements obtained with the random forest and the adaboost are statistically significant according to the AUC. The SVM and the neural network have a performance broadly similar to the probit model according to the AUC criterion,[18] while the tree algorithm is the worst classifier considering all the three criteria. Restricting the analysis to the subsamples of EMEs, the random forest and the adaboost algorithms are still the best classifiers with statistically significant improvements in terms of the AUC. In the case of LICs, the tree ensemble methods outperform the probit but the differences in performance are only marginally significant. However, it must be noted also that all methods, when forecasting crises in LICs, exhibit a worst absolute performance compared to the predictions made on the full sample and on the EMEs subsample.

Table 2: Forecasting performance of the alternative machine learning methods considering different forecasting horizons and subsamples. Models are estimated on the entire sample (EMEs and LICs)

h=2	Full sample			EMEs			LICs		
Model	*AUC*	*MSE*	*LogLik*	*AUC*	*MSE*	*LogLik*	*AUC*	*MSE*	*LogLik*
Probit	0.617	0.190	-0.594	0.581	0.165	-0.527	0.589	0.241	-0.727
Tree	0.566	0.227	-0.767	0.532	0.190	-0.649	0.536	0.301	-1.003
Random Forest	0.684***	0.175	-0.529*	0.666***	0.150	-0.473	0.621	0.224	-0.641*
AdaBoost	0.684***	0.176	-0.533	0.664***	0.148	-0.469	0.614	0.232	-0.659
Neural Network	0.590	0.194	-0.581	0.545	0.166	-0.515	0.577	0.250	-0.714
SVM	0.580	.	.	0.572	.	.	0.551	.	.
h=5	Full sample			EMEs			LICs		
Model	*AUC*	*MSE*	*LogLik*	*AUC*	*MSE*	*LogLik*	*AUC*	*MSE*	*LogLik*
Probit	0.680	0.227	-0.653	0.631	0.227	-0.653	0.658	0.226	-0.653
Tree	0.628	0.303	-1.052	0.615	0.300	-1.035	0.554	0.310	-1.086
Random Forest	0.724**	0.213*	-0.620	0.704**	0.210*	-0.610	0.645	0.220	-0.640
AdaBoost	0.712*	0.220	-0.639	0.679*	0.218	-0.632	0.654	0.225	-0.654
Neural Network	0.652	0.257	-0.748	0.624	0.260	-0.753	0.606	0.252	-0.740
SVM	0.591	.	.	0.598	.	.	0.543	.	.
h=10	Full sample			EMEs			LICs		
Model	*AUC*	*MSE*	*LogLik*	*AUC*	*MSE*	*LogLik*	*AUC*	*MSE*	*LogLik*
Probit	0.615	0.280	-0.843	0.576	0.312	-0.924	0.623	0.215	-0.681
Tree	0.547	0.338	-1.238	0.512	0.406	-1.476	0.555	0.200	-0.761
Random Forest	0.656	0.240**	-0.680	0.602	0.275*	-0.754	0.619	0.171*	-0.531
AdaBoost	0.673**	0.260	-0.764	0.619	0.303	-0.864	0.665	0.175*	-0.563
Neural Network	0.585	0.297	-1.085	0.565	0.337	-1.188	0.562	0.218	-0.879
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 a given criterion. Significance levels: *** p-value < 1%, ** p-value < 5%, * p-value < 10%.

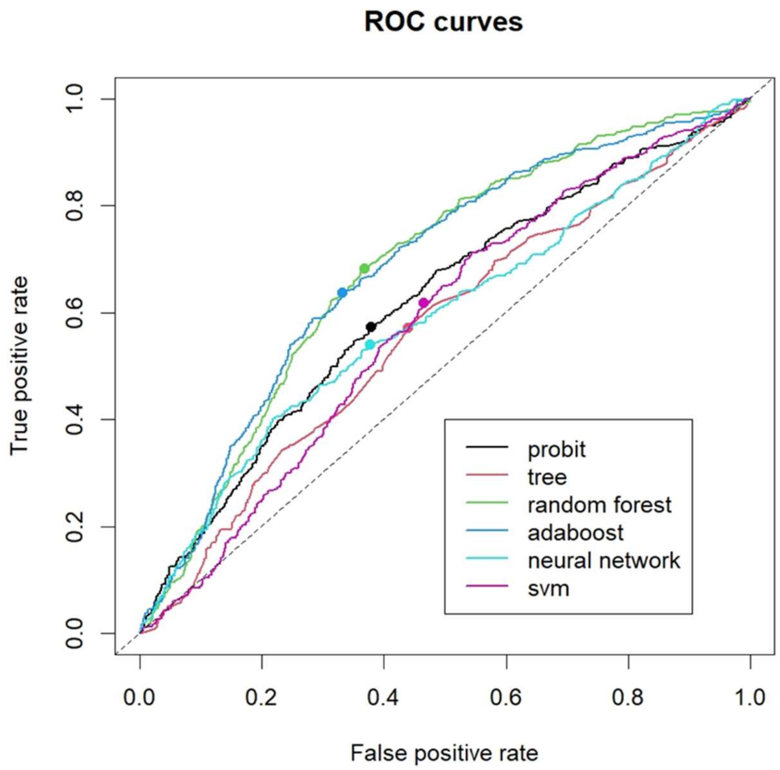
Considering a forecasting horizon of five years, the number of transitions from non-crisis states to crises increases to 1,454, while the number of stays in a non-crisis state declines to a value equal to 1,739. The increase in the number of crisis events is due to the longer time-period on which the logical union is computed in the definition of the dependent variable (see equation 1). The transitions from non-crisis states to crises that happened in EMEs were 825 (over a total of 2,143 observations for EMEs) while 629 occurred in LICs (over a total of 1,050 observations for LICs). Hence, with a forecasting horizon of 5 years, the sample is more balanced between crisis and non-crisis events, which explains the general improvement in the forecasting accuracy of all methods, with only few exceptions, in comparison to the two-year horizon forecasting exercise. The random forest and the

[18] The SVM algorithm does not allow to calculate a probabilistic outcome (see Section 3 on methodological issues). Hence, for this method it is not possible to compute the log-likelihood and the MSE.

adaboost algorithms are again the best classifiers for the overall sample of countries, as well as for the two subsamples (EMEs and LICs). In particular, these two methods display a forecasting performance that, looking at the AUC and MSE in the full as well as in the EMEs sample, remains significantly superior to the probit, even though the difference in forecasting accuracy is lower than in the case of the two-year horizon. This is due to the fact that the short-term forecasting exercise allows for a more timely update of the hyper-parameters of the machine learning models.

When the forecasting horizon is set to ten years, the number of transitions from a non-crisis to a crisis exceeds the number of stays in non-crisis states (1,714 against 992). During the period 1980-2021, crisis transitions in EMEs were 1,034 over a total of 1,822 observations, while 680 episodes were observed in LICs against 204 stays. Therefore, the sample is more unbalanced compared to the five year horizon exercise and this explains the deteriorating performance of all models. Considering the full sample, the adaboost and the random forest algorithms significantly outperform the probit considering the AUC and MSE, respectively. They remain the best classifiers when the forecasts are restricted to the subsamples of EMEs and LICs, although the differences in performance are marginally significant.

Fig. 3: ROC curves of the different algorithms based on short-term forecasts (*h=2*)



Overall, results in Table 2 show that the random forest and, to a lesser extent, the adaboost algorithms outperform traditional approaches, like the probit, in terms of forecasting accuracy. Moreover, such ensemble methods generally deliver a better performance also when compared to other machine learning approaches. Our findings on the absolute and relative performance of the different methods are largely in line with those of other recent papers, which obtain AUC in the range of 0.5-0.7 with the random forest as the best performing approach, employing a similar (Jarmulska, 2022) or a much larger set of predictors (Badia et al., 2020; Hellwig, 2021): this also suggests that models with a very large number of highly correlated variables are able to attain limited additional gains in forecasting accuracy. The better performance of the random forest and the adaboost algorithms can be appreciated graphically by comparing the ROC curves of the different models, based on the short-term forecasts (*h=2*), as depicted in Fig. 3. In fact, the ROC curves of the random forest and the adaboost are always closer to the (0,1) point – the perfect classifier case - than those of the other methods. This result can be generalised to every forecasting horizon, as shown in Fig. A1 and A2 of Annex A, even though the distance between the ROC curve of the random forest and those of the other methods narrows compared to the two-year forecasting exercise.

In Table A2 in Annex A, we also report the forecasting performance of the models estimated separately on the two subsamples of countries (EMEs and LICs) with the aim of better capturing possible distinct features of the two groups. The comparison of results in Table 2 with those in Table A2 shows that the estimation performed on the whole sample generally delivers more accurate predictions. This indicates that the relationships between the predictors and the probability of default are not too heterogeneous across the two groups of countries, with the result that the pooling of observations allows a better identification of the common complex patterns in the data.

Variable importance

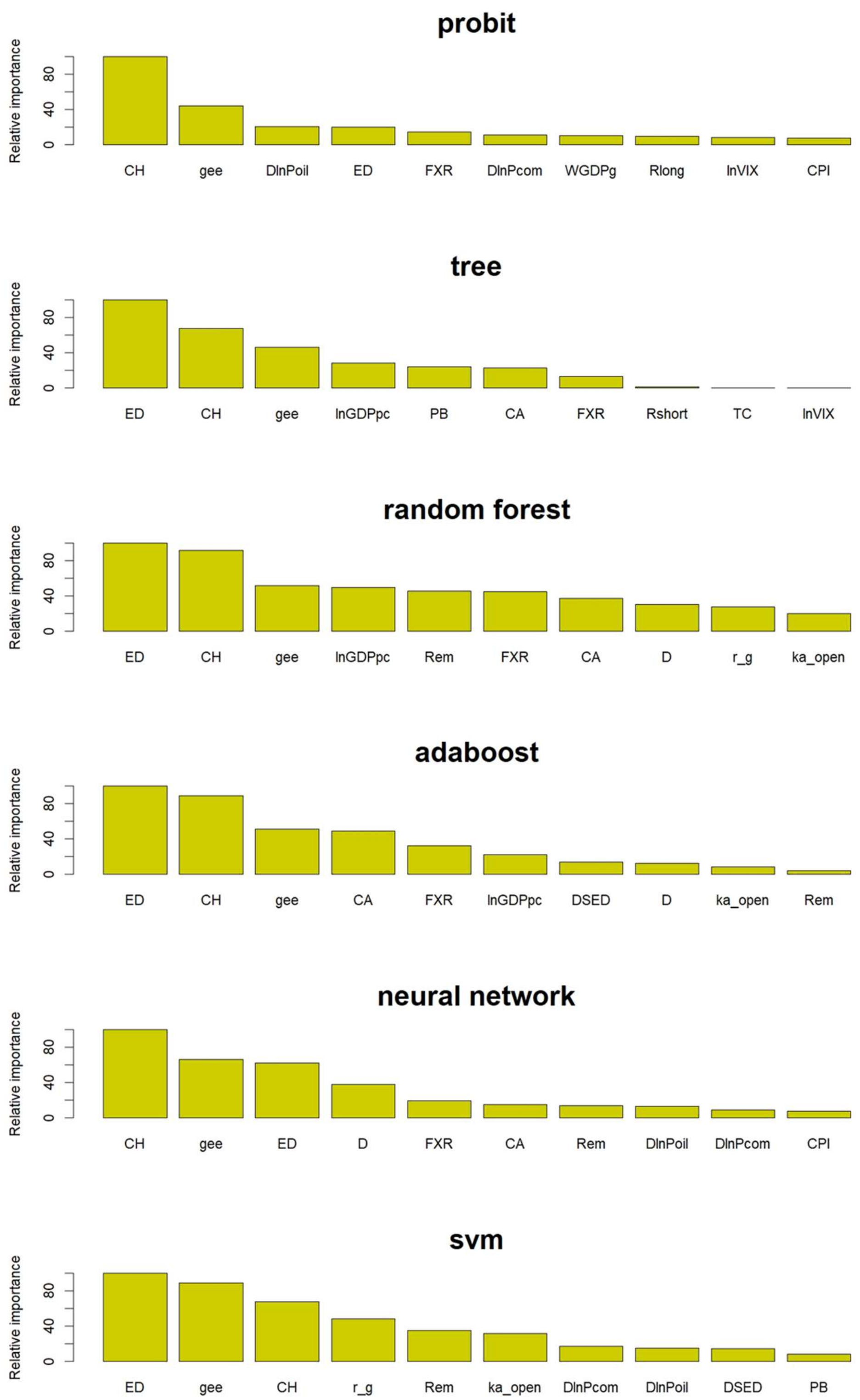
Using the permutation algorithm described in the previous section, we assess the relative importance of the selected predictors in terms of their contribution to the improvement of forecasting accuracy. In particular, the importance of each predictor is measured as the decrease in the AUC (calculated on the out-of-sample forecasts) caused by the random permutation of the same variable.

Fig. 4 plots the 10 most important predictors, in decreasing order of relevance (normalising the importance of the most relevant predictor to 100), for the alternative methods employed and the short-term forecasting horizon (*h=2*). There is a wide consensus among the different algorithms on the identification of the key predictors. The external stock of public debt, the historical frequency of past fiscal crises and the index of government effectiveness come out clearly as the most important predictors according to all algorithms. Other variables often appearing as relevant predictors are the stock of foreign reserves, the current account balance, the total stock of public debt and the level of economic development, as measured by the (log) GDP per capita. Moreover, results show a relatively relevant role played also by remittances, the country's financial openness, and the interest rate-growth differential.

When performing medium-term predictions (*h=5*), the external stock of public debt, the past crisis history, the institutional quality index and the current account balance continue to contribute to a large fraction of the forecasting accuracy of many models, while other variables - that were relevant in the short-term exercise - lose importance (Fig. A3 in Annex A). When analysing long-term predictions (*h=10*), the results of the alternative approaches are more heterogeneous compared to shorter forecasting horizons with the institutional quality index, the stock of total and external public debt, the interest-growth differential, the current account balance and the level of development captured by GDP per capita being the most common relevant predictors (Fig. A4). However, some caution is needed in analysing this last result given the low number of time periods involved in the ten-year forecasting exercise.

Overall, these results confirm the findings on the main determinants of fiscal crisis episodes identified by the previous literature. First, public debt (especially its sub-component of public external debt) appears to be a very relevant predictor, and actually the most important one according to the best performing models. This outcome is in line with the results obtained by other recent studies (such as Badia et al., 2020), and provides an important message, as well as a warning, to inform the policy debate in the current context of high levels of public debt in EMEs and LICs. In addition to the current level of public debt, the past history of fiscal crises also matters considerably.

Fig. 4: Importance plots for the different algorithms based on short-term forecasts (*h=2*)



Moreover, we find that some structural features play a prominent role. Indeed, the analysis shows that the level of economic development (measured by GDP per capita) and the quality of governance and institutions carry a substantial weight, validating the conjecture that less developed countries with weaker institutional settings are more prone to crises. Our work also corroborates the notion that traditional indicators of external position, such as the current account balance and the stock of foreign exchange reserves, convey useful information regarding fiscal vulnerabilities and the risk of crisis.

Unlike other recent studies (such as Badia et al., 2020, and Hellwig, 2021), a few models, including in some cases those with the best predicting accuracy, also find a role played by the interest rate-growth differential; by being one of the drivers of the dynamics of the debt to GDP ratio, this differential seems to provide some signalling value regarding the probability of a crisis, especially on longer horizons.[19]

Fiscal variables, CPI inflation and exchange rate movements rarely appear among the most relevant predictors. Moreover, despite the fact that fiscal crises appear to be clustered around time of financial turmoil, global indicators have overall a limited role, especially in the best-performing methods. This can be rationalised considering that these high-frequency indicators provide a useful signal to identify crises only in the very short-run, but not at the lower frequency horizons considered in our forecasting exercise.

Effects of predictors on the probability of crisis

Once the most relevant predictors according to the different models have been detected, we study their impact on the probability of fiscal crisis using the accumulated local effects (ALE). In particular, Fig. 5 shows the ALE plots for the probit, the random forest, the adaboost and the neural network models, considering a two-year forecasting horizon.[20]

All models show a strong positive relation between the frequency of past fiscal crises and the likelihood of a new one. This result seems to suggest that history matters, and countries that experienced fiscal crises in the past are more likely to fall in crisis again in the future than countries with a strong track record of repayments and fiscal soundness, even if they have the same levels of debt and other relevant characteristics. In other words, there appears to be a sort of vicious cycle in which a weak credit history leads to a higher probability of a new fiscal crisis, other things being equal.

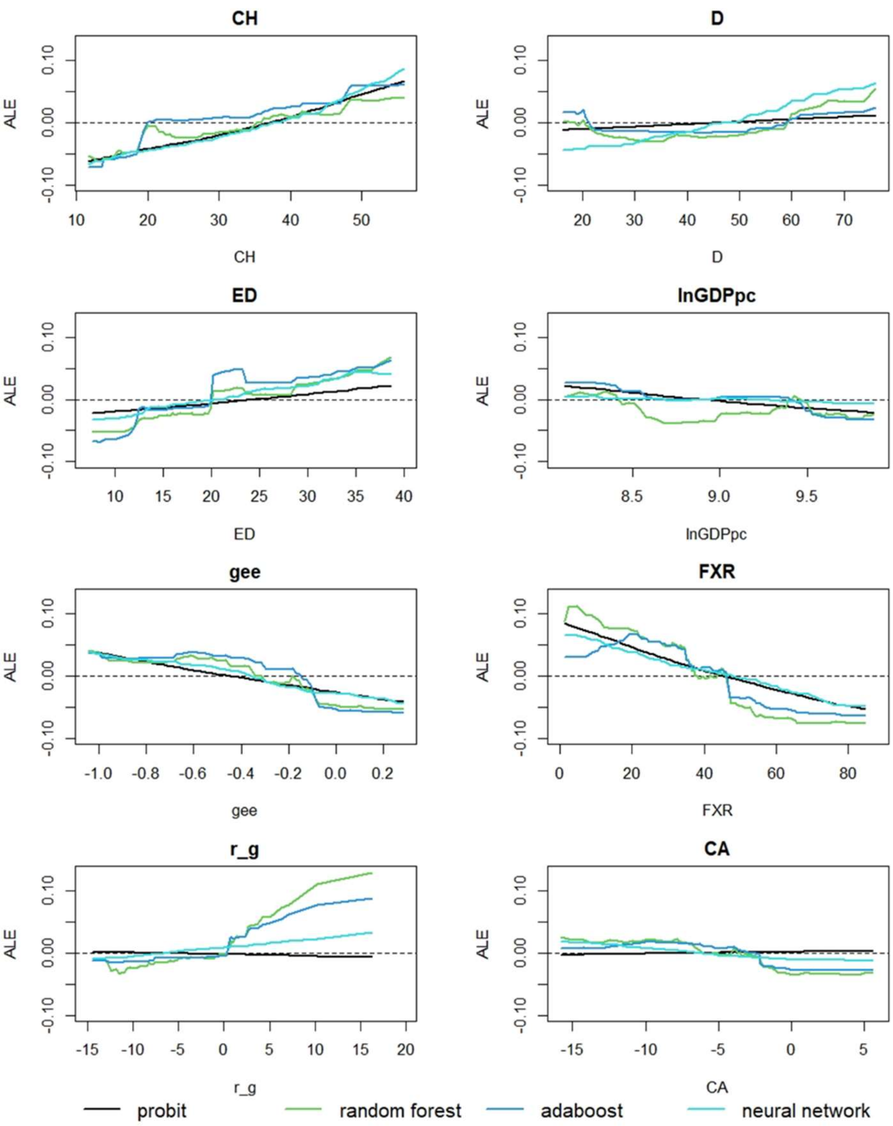
Furthermore, according to the probit model, a higher total stock of public debt increases linearly the probability of observing a fiscal crisis. The ALE plot of the neural network is also close to a line, with a steeper slope compared to that of the probit model. More complex is the relationship between the stock of public debt and crisis probability detected through the random forest and the adaboost algorithms: the crisis probability actually decreases for low values of public debt, below 40-50 per cent of GDP, but then increases rapidly when public debt exceeds this threshold. This result seems to suggest a significantly non-linear relationship, with low levels of public debt to GDP being relatively "safe", while higher values (above the identified threshold) provide strong and rapidly rising signals of mounting vulnerabilities. On the other hand, an increase of the stock of external public debt is monotonically associated with a higher probability of a fiscal crisis in all models. It must be noted, however, that the estimated effect on the probability of default is much larger according to the machine learning models compared to the probit. Moreover, also in this case the best performing machine learning models indicate the presence of strong non-linearities, with a substantial jump in

[19] At the 10-year horizon, the higher estimated relevance of the interest rate-growth differential could also reflect a decreased importance of other predictors providing very strong signalling value at shorter horizons.

[20] The SVM algorithm does not provide a probabilistic outcome (see Section 3 on methodological issues) while the ALE plots of the tree are very naive.

the crisis probabilities when the levels of the external public debt relative to GDP rise above 20 percent.[21]

Fig. 5: ALE plots of the most relevant predictors for the different algorithms calculated using short-term forecasts (h=2)



The ALE plots of the (log) GDP per capita, the government effectiveness index and the stock of foreign reserves are quite similar across the different methods. In particular, an increase in the level of GDP per capita, in the quality of government effectiveness or in the stock of foreign exchange reserves, all imply a reduction of the probability of default. These results are largely in line with what would be expected from economic theory.

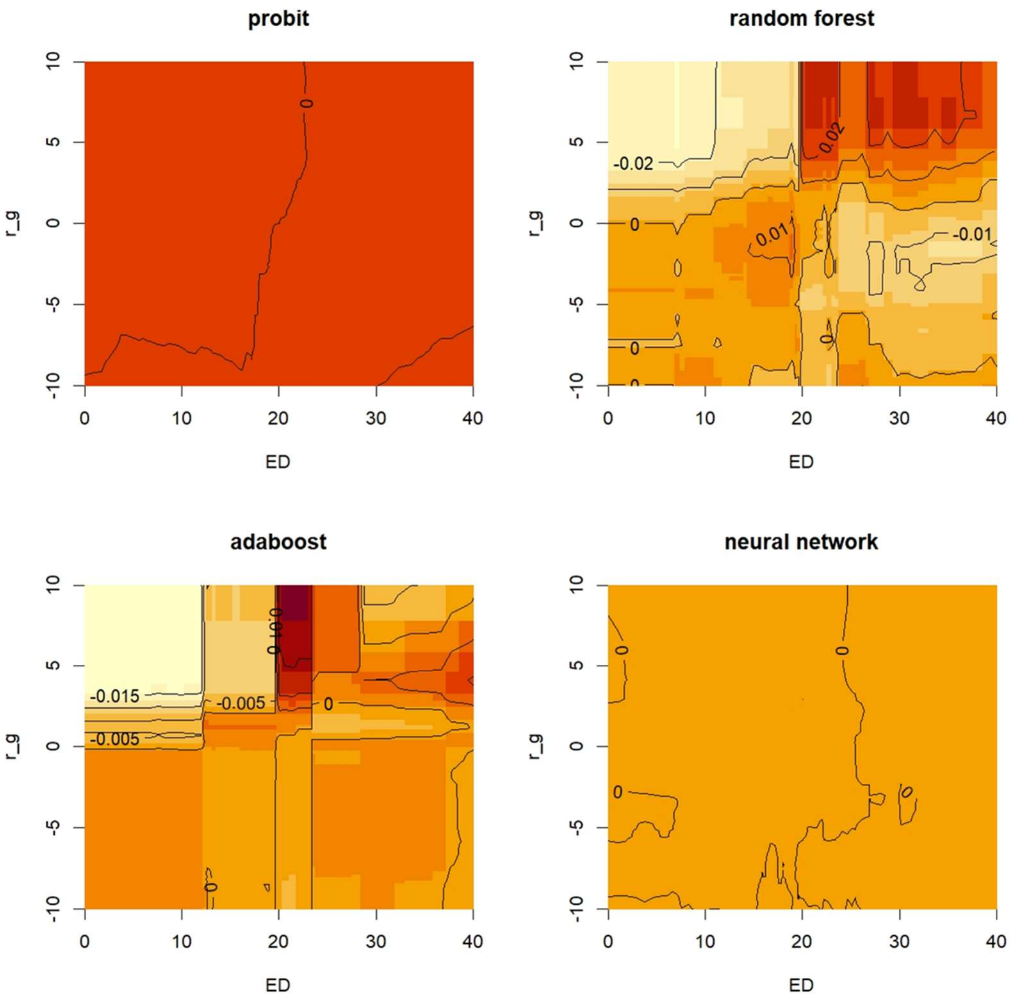
There is instead some divergence between the models on the effect of the interest rate-growth differential and the current account balance. In fact, according to the probit model, a higher differential between the cost of debt and the rate of economic growth or a more positive current account balance do not meaningfully affect the probability of default. Conversely, our best

[21] The level of the threshold over which the crisis probability accelerates is lower for external public debt compared to public debt, which is a reasonable feature as the former is a component of the latter. At the same time, the different shape of the relationships the two variables have with the crisis probability validates the fact that the two types of debt are not identical and imply a diverse degree of riskiness.

performing models (primarily the random forest and the adaboost algorithms) show a sizeable jump in the probability of fiscal crisis when the interest rate-growth differential becomes positive and large; on the other hand, these models indicate a sudden drop in the crisis probability if the current account balance is positive or moderately negative.

ALE plots calculated for longer forecasting horizons are largely in line with those depicted in Fig. 5. Even if the lower number of observations employed in the estimation makes the plots more noisy, the shapes of the different relationships are overall quite similar (see Fig. A5 and A6 in the Annex A for the five and ten-year horizon, respectively).

Fig. 6: Bivariate ALE plots showing the second-order interaction effects between the interest-growth differential and the stock of external debt in the different algorithms calculated using short-term forecasts (h=2). Darker regions are associated with a higher probability of fiscal crisis



In addition to study the relationships between single predictors and the probability of crisis, including possible non-linearities of these relationships, we employ the ALE method also to investigate the existence of interactions between variables. Among the selected predictors we find, in particular, a strong interaction effect between the interest rate-growth differential and the stock of external debt. Fig. 6 shows the bivariate ALE plot for these two variables in the context of short-term forecasts (h=2). As it emerges from the picture, the more accurate machine learning methods indicate that the estimated impact on the likelihood of fiscal crisis occurrence stemming from an elevated level of external debt is larger when the difference between the cost of debt and the GDP growth rate is higher, while the probit model and the neural network fail to uncover this additional interaction effect.[22] In other words, this suggests that even for relatively moderate levels of external debt, the probability of crisis increases considerably in case of a high interest rate-growth differential, which represents a further warning signal in the context of the global monetary tightening expected to counteract the rising inflationary pressures exacerbated by the repercussions of the war in Ukraine.

Aggregate fiscal risk index

Using the outcomes of the best performing estimated model, i.e., the random forest algorithm, we also construct a forward-looking aggregate indicator which should be able to reflect the evolution of the risk of fiscal crises in EMEs and LICs. Since the aim of this exercise is to define an aggregate index for all the countries in our sample independently on the state in which they are at time *t*, we consider the probability of fiscal crisis without conditioning on the absence of crisis in *t*: $\mu_{i,t+1}^{h} \equiv \mathbb{P}(y_{i,t+1}^{h} = 1 | X_{i,t})$. Then, we compute two versions of such an indicator, relying on two different types of output obtained through our analysis. One version relies on the expected number of countries in debt distress (over the total number of countries), and is represented by the formula:

$$I_t^{(1)} = \frac{1}{N} \sum_{i=1}^{N} I\left(\hat{\mu}_{i,t+1}^{h} > \phi_t^*\right) \qquad (8)$$

in which ϕ_t^* is the optimal threshold that identifies the point on the ROC curve (constructed using the out-of-sample forecasts prior to time *t*) closest to the perfect classification case, i.e. the point (0,1). The threshold ϕ_t^* is calibrated for the subsample of EMEs and LICs, separately. Alternatively, the fiscal risk index can be computed using directly the probabilities of fiscal crisis, and calculating their average:

$$I_t^{(2)} = \frac{1}{N} \sum_{i=1}^{N} \hat{\mu}_{i,t+1}^{h} \qquad (9)$$

This second version allows also to gain some information on the evolution of the cross-country distribution of risks, by computing the inter-quartile range of predicted probabilities ($\hat{\mu}_{i,t+1}^{h}$) in addition to the average level measured by expression (9).

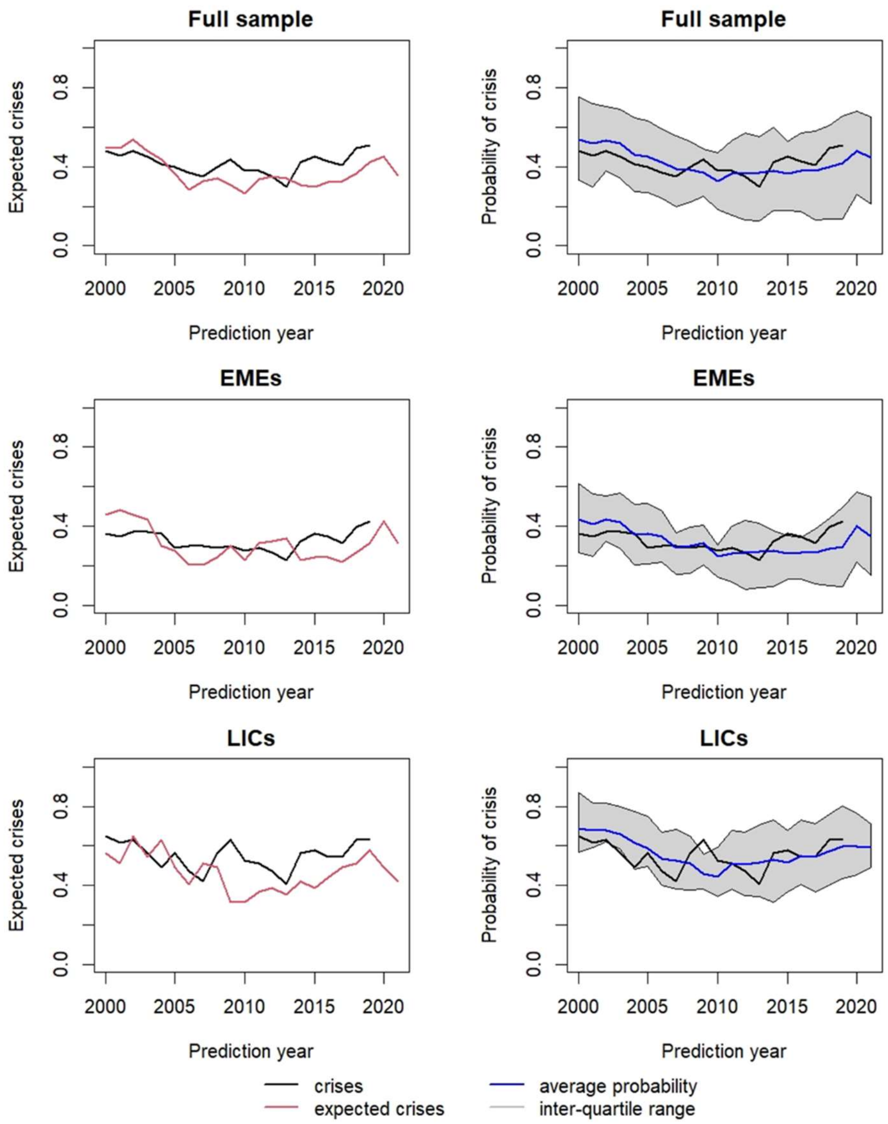
It is worth stressing that both indicators are computed at time *t* (i.e., the prediction year) considering the model-implied probabilities evaluated using predictors at time *t* but referring to forecasting crises in periods *t+1*,…,*t+h*. In this regard, the two indexes are forward-looking, as they signal at time *t* the probability of fiscal crisis occurrence (or the share of countries anticipated to enter into a crisis) in the following *h* periods.

Fig. 7 displays the time series of the two indicators considering a forecasting horizon of 2 years, for the entire set of countries and for each of the two subgroups (EMEs and LICs), separately. The first indicator is plotted in the left panels, while the second one is reported in the right panels, together with the inter-quartile range of predicted probabilities. In order to evaluate the performance of the

[22] In principle, even in the probit model the marginal effect of any predictor depends on the values assumed by the other independent variables. However, the non-linear function of the probit is not flexible enough (given the absence of hyper-parameters) to uncover important interactions among predictors: a possible solution could be to add the cross-products of all the independent variables in the probit model, introducing additional complexity and possibly making the estimation of the model numerically unfeasible. Conversely, in machine learning algorithms this is not necessary, because they are able to automatically detect the relevant interaction effects.

indexes against the occurrence of future crises, the fraction of countries actually falling in a debt crisis in the following two years is added to all the graphs. For comparability purposes, this fraction is calculated for each prediction year *t* by considering countries in fiscal crisis in periods *t+1* and *t+2*; as our database contains fiscal crisis up to the end of 2021, the last available observation is 2019, referring to crisis happened in the two-year period 2020-2021.

Fig. 7: Aggregate fiscal risk indexes calculated with the short-term forecasts (*h=2*) of the random forest algorithm



The results, for the whole set of countries as well as for the subsample of EMEs, show that the two risk indexes match quite closely the level and the dynamics of the true fraction of countries in crises. The fit of the two indexes is less accurate when considering only the subsample of LICs, a result that likely reflects the lower quality of the data for this category of countries. Comparing the two indicators, it is possible to observe that the one based on the predicted number of countries in crisis is more volatile than the index computed as a simple average of the crisis probabilities, as the former is affected also by the variability of the threshold estimate.

The two indicators signal a sharp increase of short-term risks before and during the outbreak of the COVID-19 pandemic in both EMEs and LICs. This increase is mainly driven by the worsening of the most important predictors, with regard in particular to the additional accumulation of debt stocks experienced by EMEs and LICs in the aftermath of the pandemic. The pandemic has also increased the risk of fiscal crisis in EMEs on longer horizons, while the long-run prospects of fiscal distress have remained stable at high levels in LICs, as can be appreciated by looking at the five and ten-year horizon indexes (Fig. A7 and A8 in Annex A).

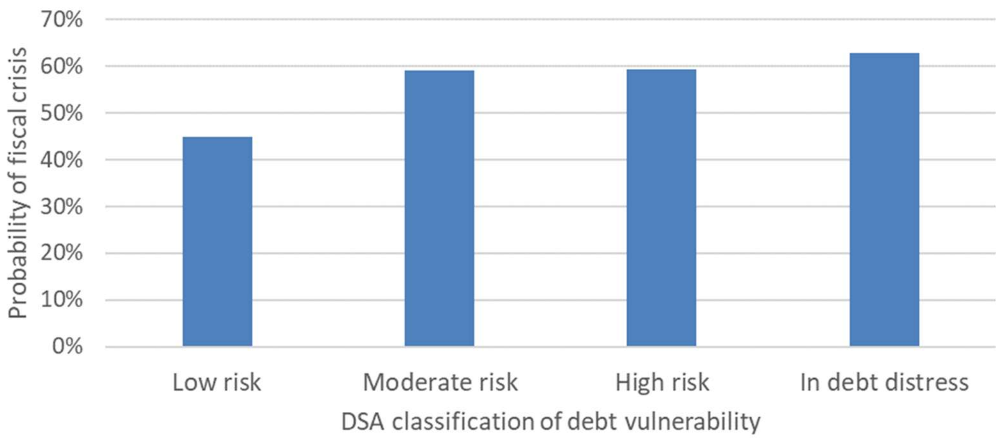
5. Concluding remarks

Our analysis supports the findings of a recent literature (Badia et al., 2020; Hellwig, 2021; Jarmulska, 2022) that highlights the merits of machine learning algorithms in improving the forecasting accuracy of fiscal crises compared to traditional approaches. In this regard, we provide additional evidence testing a wider set of machine learning models and focusing on fiscal crises in EMEs and LICs.

We find that the random forest and (to a lesser extent) the adaboost algorithms outperform both standard econometric approaches, like the probit model, as well as other machine learning techniques, including neural network and SVM. The performance of the random forest and the adaboost relative to other methods is confirmed even when considering more complex tasks, such as predictions on longer forecasting horizons, or focusing the analysis on LICs, which are characterised by a higher likelihood of fiscal distress episodes and by more noisy predictors given the lower quality and coverage of their official statistics. More accurate model-based results are important as they can better inform the analysis, complemented by the application of the analyst's judgement which is always required to reach a final debt sustainability assessment.

Our results are broadly in line with the latest assessments of debt vulnerabilities (available as of November 2022) performed by the IMF and WB in the context of their Debt Sustainability Analysis (DSA) for LICs. Fig. 8 plots the average probability of fiscal crisis in the period 2022-2023 estimated with the random forest model for each of the four DSA classes of debt distress.[23] As it is evident from the picture, the model correctly attributes a higher probability of fiscal crisis to countries classified in riskier classes of debt distress according to the DSA.

Fig. 8: Comparison between IMF LIC-DSA classification (as of November 2022) and model-implied average probabilities of fiscal crisis based on the random forest algorithm



[23] This comparison is performed on a subset of 69 countries, i.e. those subject to the DSA for LICs carried out jointly by the IMF and WB. The DSA classifies countries in four categories with ascending debt vulnerabilities: low risk of debt distress, medium risk, high risk and in debt distress.

Our analysis also detects a robust set of variables that can effectively anticipate the occurrence of fiscal crises. In this respect, the most relevant predictors are the stock of public debt (both external and total), the past crisis history, the quality of institutions, the stock of foreign exchange reserves, the degree of economic development (captured by the level of GDP per capita), the current account balance and the interest rate-growth differential. All these predictors have the expected effect on the probability of crisis occurrence.

Finally, we show how the model-implied probabilities and the expected number of countries in fiscal distress can be used to construct a reliable aggregate index of fiscal crisis risk for EMEs and LICs. This index clearly highlights the significant increase in debt vulnerabilities following the accumulation of additional debt in the aftermath of the pandemic, thus corroborating the concerns of the G20 and the international community regarding the risks to debt sustainability in developing countries.

Annex A: Additional tables and figures

Table A1: List of countries included in the analysis

EMEs		LICs	
Albania	Libya	Afghanistan	Sierra Leone
Algeria	Maldives	Bangladesh	Solomon Islands
Angola	Mauritius	Benin	Somalia
Antigua and Barbuda	Mexico	Bhutan	South Sudan
Argentina	Mongolia	Burkina Faso	Sudan
Armenia	Montenegro	Burundi	Tajikistan
Azerbaijan	Morocco	Cambodia	Tanzania
Barbados	Namibia	Cameroon	The Gambia
Belarus	Nauru	Central African Republic	Togo
Belize	North Macedonia	Chad	Uganda
Bolivia	Pakistan	Comoros	Uzbekistan
Bosnia and Herzegovina	Palau	Côte d'Ivoire	Vietnam
Botswana	Panama	Democratic Republic of the Congo	Yemen
Brazil	Paraguay	Djibouti	Zambia
Bulgaria	Peru	Eritrea	Zimbabwe
Cabo Verde	Philippines	Ethiopia	
Chile	Poland	Ghana	
Colombia	Romania	Guinea	
Costa Rica	Russia	Guinea-Bissau	
Croatia	Samoa	Haiti	
Dominica	Serbia	Honduras	
Dominican Republic	Seychelles	Kenya	
Ecuador	South Africa	Kyrgyz Republic	
Egypt	Sri Lanka	Lao P.D.R.	
El Salvador	St. Kitts and Nevis	Lesotho	
Equatorial Guinea	St. Lucia	Liberia	
Eswatini	St. Vincent and the Grenadines	Madagascar	
Fiji	Suriname	Malawi	
Gabon	Syria	Mali	
Georgia	Thailand	Mauritania	
Grenada	Tonga	Moldova	
Guatemala	Trinidad and Tobago	Mozambique	
Guyana	Tunisia	Myanmar	
Hungary	Turkey	Nepal	
India	Turkmenistan	Nicaragua	
Indonesia	Tuvalu	Niger	
Iraq	Ukraine	Nigeria	
Islamic Republic of Iran	Uruguay	Papua New Guinea	
Jamaica	Vanuatu	Republic of Congo	
Jordan	Venezuela	Rwanda	
Kazakhstan	West Bank and Gaza	São Tomé and Príncipe	
Lebanon		Senegal	

Table A2: Forecasting performance of the alternative machine learning methods considering different forecasting horizons and subsamples. The estimation is performed in the two subsamples separately

h=2	EMEs						LICs					
Model	*AUC*	*ΔAUC*	*MSE*	*ΔMSE*	*LogLik*	*ΔLogLik*	*AUC*	*ΔAUC*	*MSE*	*ΔMSE*	*LogLik*	*ΔLogLik*
Probit	0.566	-0.015	0.182	0.017	-0.586	-0.059	0.542	-0.047	0.268	0.027	-0.820	-0.093
Tree	0.593	0.061	0.188	-0.002	-0.635	0.014	0.540	0.004	0.312	0.011	-1.092	-0.089
Random Forest	0.656	-0.010	0.152	0.002	-0.478	-0.005	0.592	-0.029	0.232	0.008	-0.661	-0.020
AdaBoost	0.655	-0.009	0.152	0.004	-0.484	-0.015	0.580	-0.034	0.243	0.011	-0.687	-0.028
Neural Network	0.581	0.036	0.183	0.017	-0.579	-0.064	0.504	-0.073	0.296	0.046	-0.958	-0.244
SVM	0.577	0.005	.	.	.	.	0.488	-0.063	.	.	.	.
h=5	EMEs						LICs					
Model	*AUC*	*ΔAUC*	*MSE*	*ΔMSE*	*LogLik*	*ΔLogLik*	*AUC*	*ΔAUC*	*MSE*	*ΔMSE*	*LogLik*	*ΔLogLik*
Probit	0.611	-0.020	0.240	0.013	-0.715	-0.062	0.621	-0.037	0.267	0.041	-0.892	-0.239
Tree	0.596	-0.019	0.309	0.009	-1.118	-0.083	0.595	0.041	0.286	-0.024	-1.065	0.021
Random Forest	0.702	-0.002	0.210	0.000	-0.610	0.000	0.676	0.031	0.211	-0.009	-0.624	0.016
AdaBoost	0.670	-0.009	0.228	0.010	-0.661	-0.029	0.655	0.001	0.231	0.006	-0.733	-0.079
Neural Network	0.621	-0.003	0.272	0.012	-0.855	-0.102	0.579	-0.027	0.301	0.049	-1.163	-0.423
SVM	0.600	0.002	.	.	.	.	0.484	-0.059	.	.	.	.
h=10	EMEs						LICs					
Model	*AUC*	*ΔAUC*	*MSE*	*ΔMSE*	*LogLik*	*ΔLogLik*	*AUC*	*ΔAUC*	*MSE*	*ΔMSE*	*LogLik*	*ΔLogLik*
Probit	0.542	-0.034	0.327	0.015	-1.016	-0.092	0.480	-0.143	0.327	0.112	-1.528	-0.847
Tree	0.494	-0.018	0.431	0.025	-1.772	-0.296	0.495	-0.060	0.251	0.051	-1.195	-0.434
Random Forest	0.573	-0.029	0.275	0.000	-0.754	0.000	0.548	-0.071	0.197	0.026	-0.627	-0.096
AdaBoost	0.589	-0.030	0.308	0.005	-0.899	-0.035	0.590	-0.075	0.210	0.035	-0.794	-0.231
Neural Network	0.554	-0.011	0.357	0.020	-2.064	-0.876	0.538	-0.024	0.275	0.057	-1.281	-0.402
SVM	0.503	0.005	.	.	.	.	0.488	-0.006	.	.	.	.

Notes: the table shows both the forecasting performance in terms of AUC, MSE and log-likelihood and the difference with respect to the prediction accuracy of the corresponding models estimated on the entire sample (EMEs and LICs), i.e., Δ*AUC,* Δ*MSE and* Δ*LogLik.* For the columns showing the differences in prediction accuracy, numbers in red colour indicate a worse accuracy compared to the model estimated on the entire sample of countries.

Fig. A1: ROC curves of the different algorithms based on short-term forecasts (*h=5*)

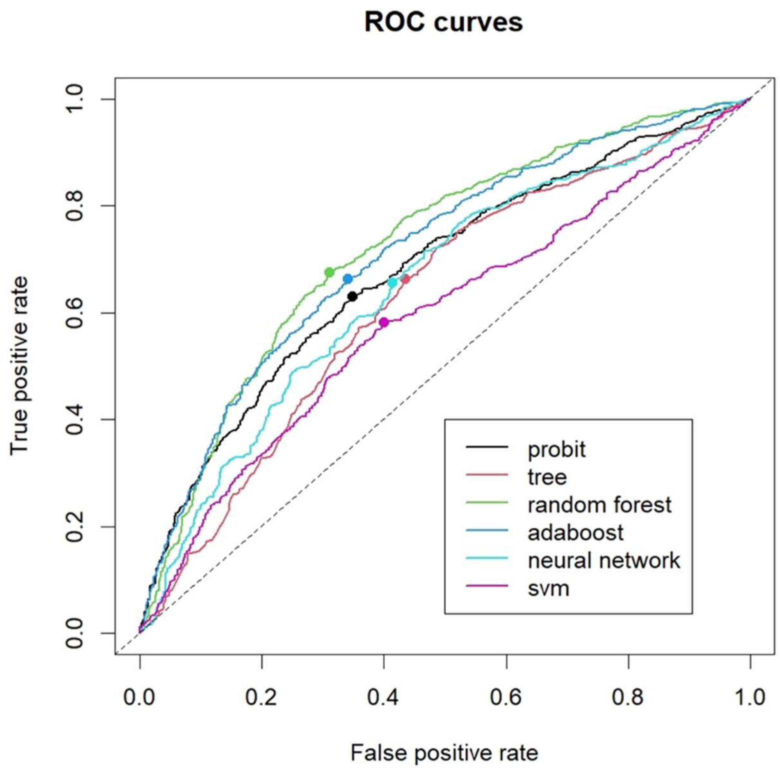


Fig. A2: ROC curves of the different algorithms based on short-term forecasts (*h=10*)

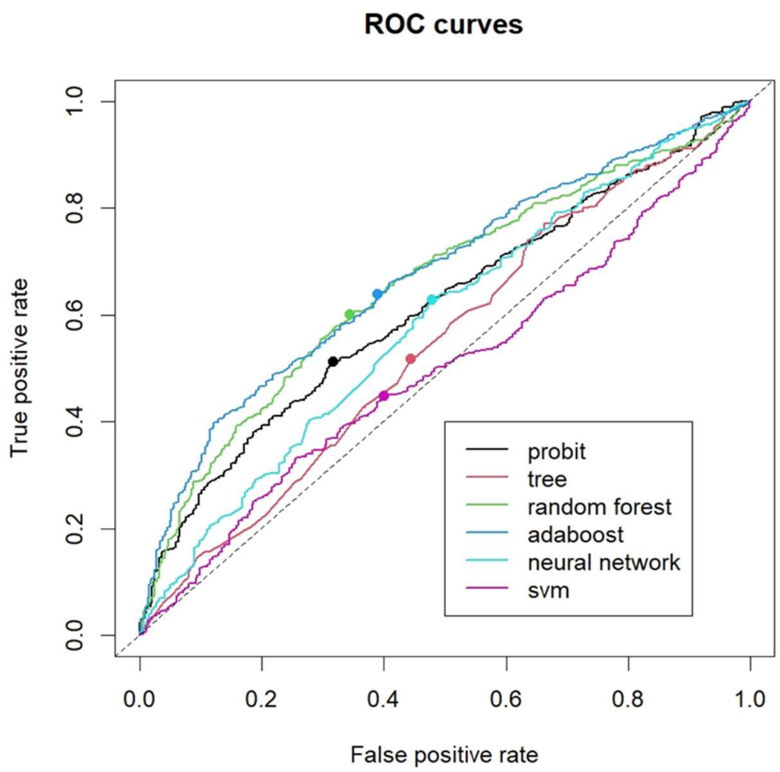


Fig. A3: Importance plots for the different algorithms based on medium-term forecasts (*h=5*)

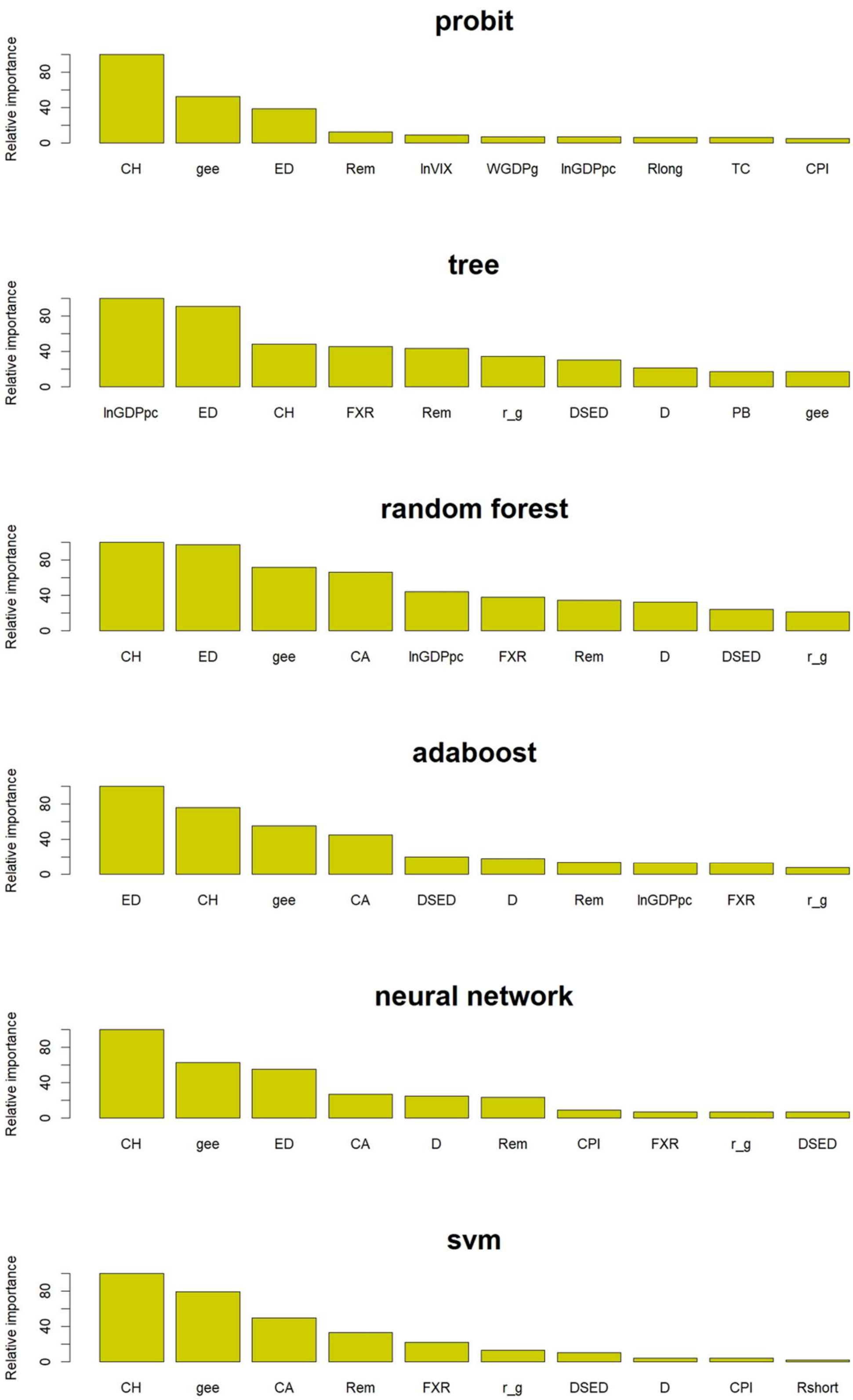


Fig. A4: Importance plots for the different algorithms based on long-term forecasts (*h=10*)

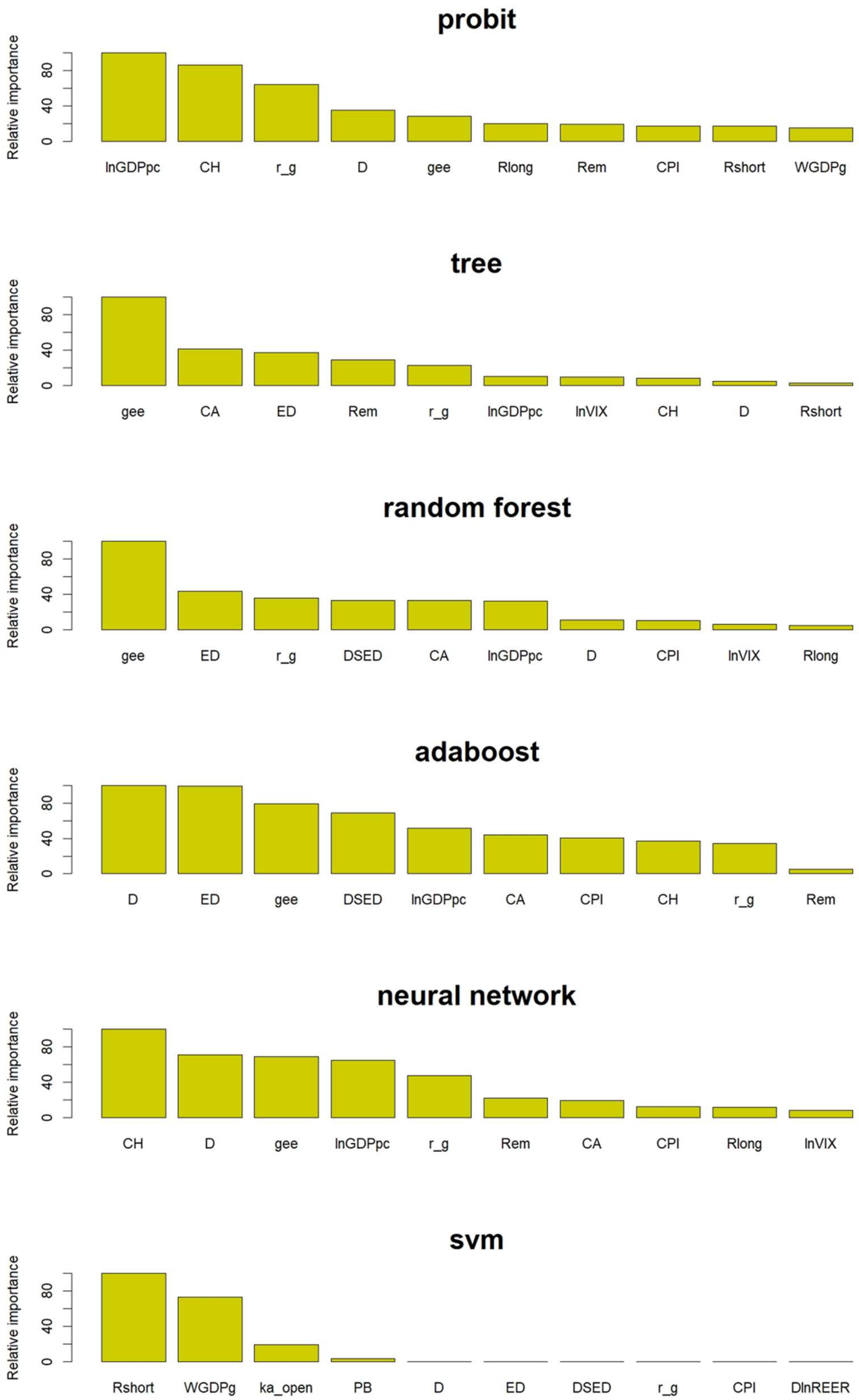


Fig. A5: ALE plots of the most relevant predictors for the different algorithms calculated using medium-term forecasts (h=5)

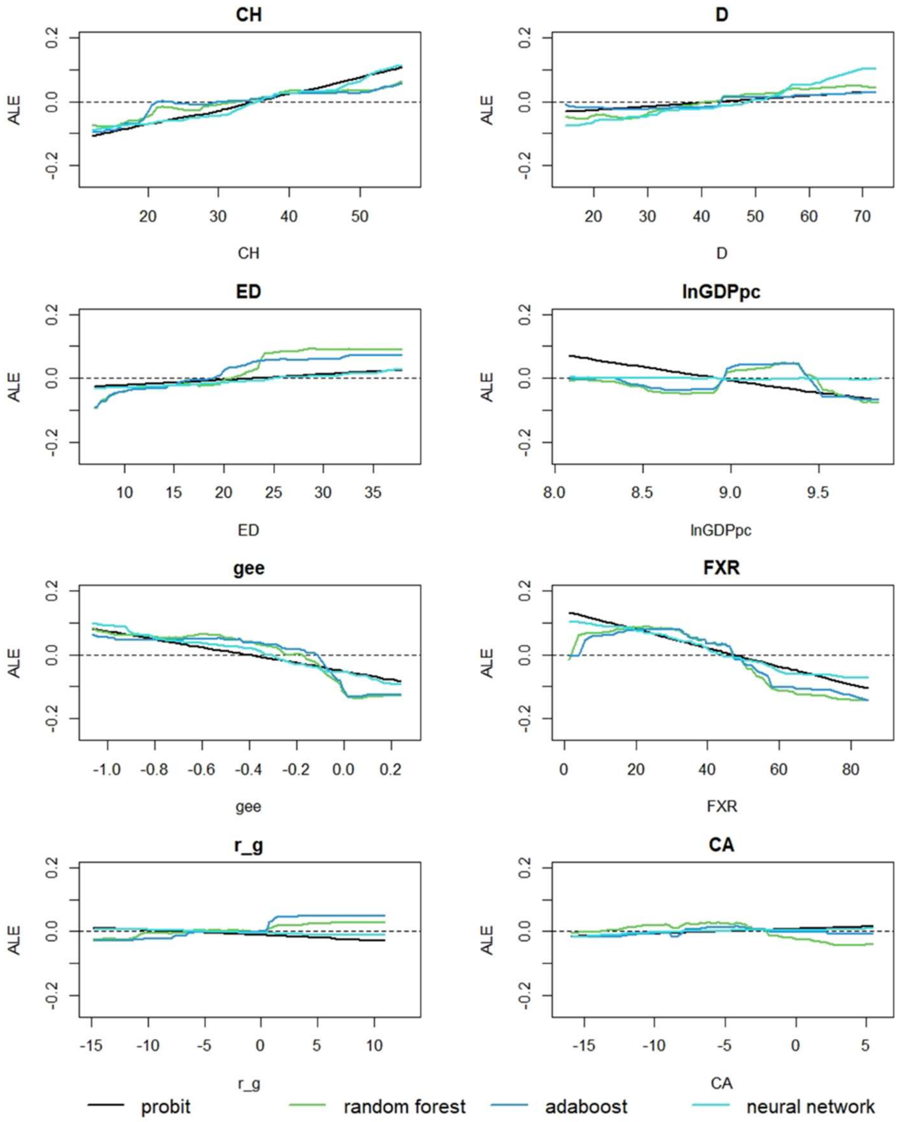


Fig. A6: ALE plots of the most relevant predictors for the different algorithms calculated using long-term forecasts (*h=10*)

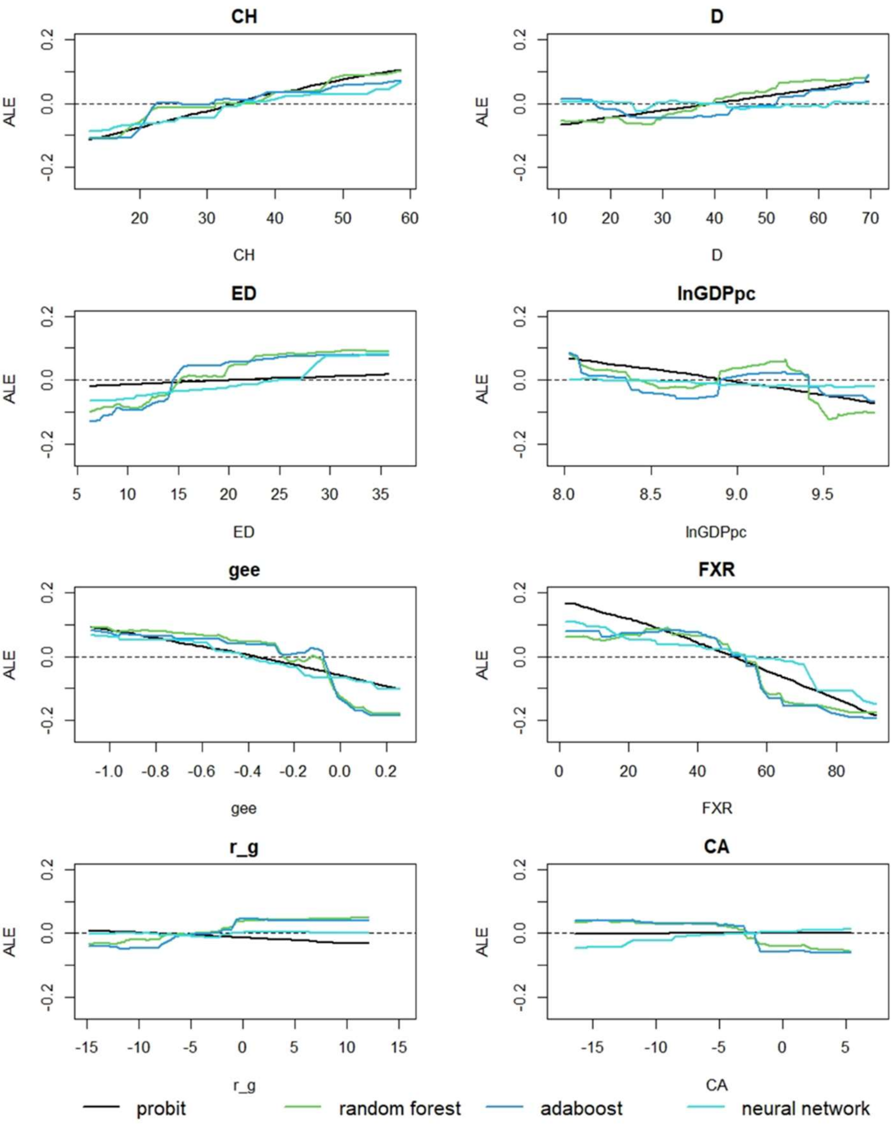


Fig. A7: Aggregate fiscal risk indexes calculated with the medium-term forecasts (h=5) of the random forest algorithm

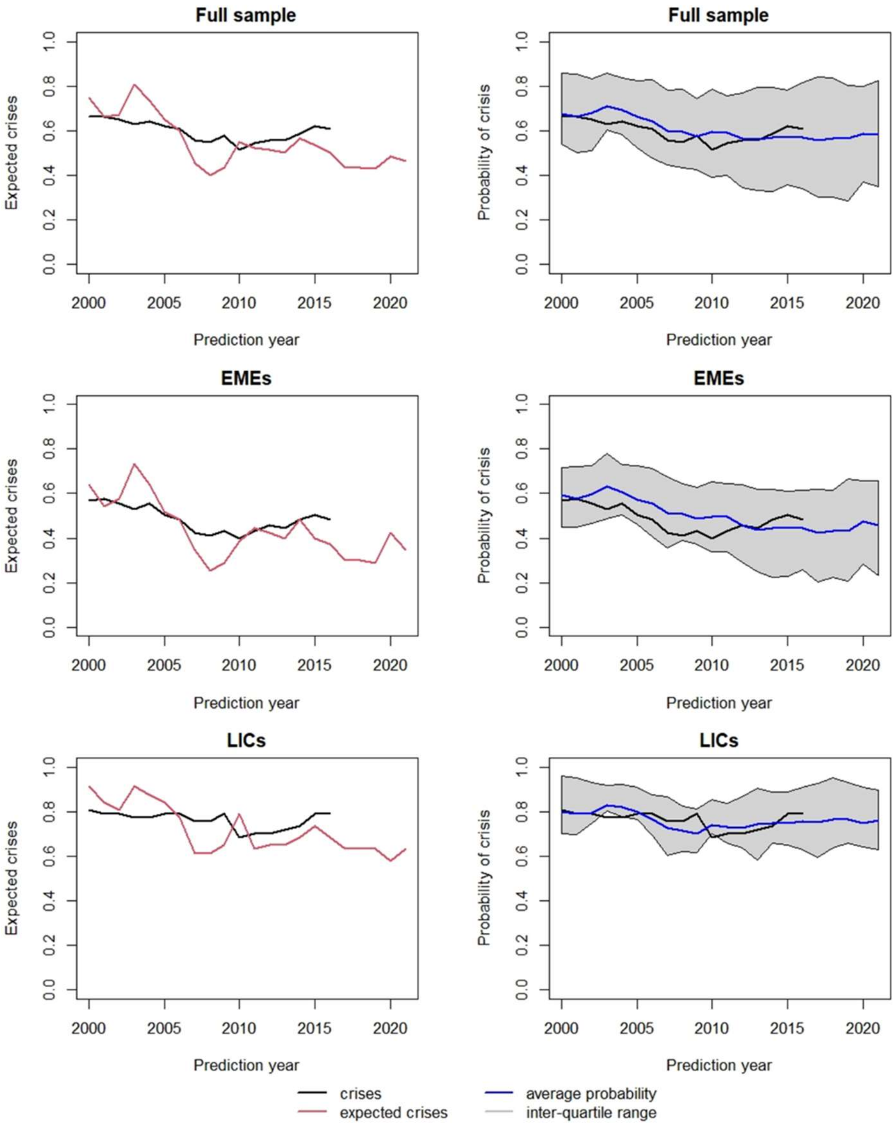
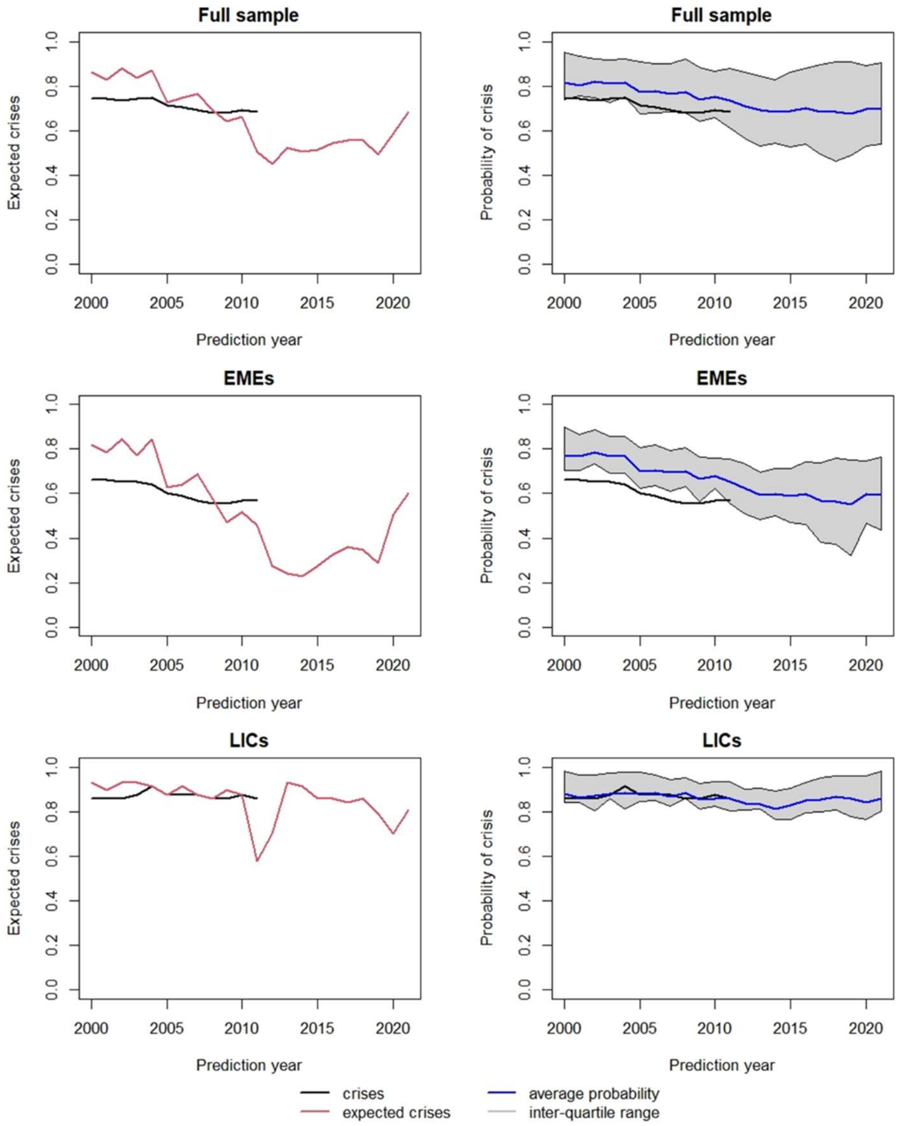


Fig. A8: Aggregate fiscal risk indexes calculated with the long-term forecasts (*h=10*) of the random forest algorithm



Annex B: Algorithm

Given a forecasting horizon $h \in \{2,5,10\}$, for each machine learning model $m(X_{it};\theta,\lambda)$ and for each year $t \in \{2000,\ldots,2021-h\}$, the sample is divided into two sets:

1) *training sample*: data from year *t-h-s+1* to *t-h* are used for both hyper-parameter tuning and for the estimation of model parameters:
   a. *hyper-parameter tuning* ($\hat{\lambda}$): the training sample is divided into *k=10* non-overlapping parts considering both the cross-sectional and the time-series dimension of the dataset. For each hyper-parameter value (chosen on a grid of possible values frequently selected in the literature, see Table A3), the model is estimated *k*-times using *k-1* subsamples and employed to calculate out-of-sample forecasts on the excluded subsample. Using these forecasts it is possible to obtain *k* measures of forecasting accuracy according to a given criterion (AUC, maximum likelihood, MSE). A simple average of these *k* measures is computed. Finally, the hyper-parameter with the best average predictive performance is selected;
   b. *model estimation* ($\hat{\theta}$): once the best hyper-parameter is selected, the model is estimated using the overall training sample.
2) *test sample*: time *t* observations X_{it} and the trained model $m(X_{it};\hat{\theta},\hat{\lambda})$ are employed to compute out-of-sample forecasts for the probability of a crisis $\hat{\pi}_{i,t+1}^{h}$ during the period *t*+1,.., *t*+*h*.

Using the out-of-sample forecasts for the probability of a crisis $\hat{\pi}_{i,t+1}^{h}$ computed for each country *i* and year $t \in \{2000,\ldots,2021-h\}$, we can compute an overall measure of forecasting accuracy according to a given criterion (AUC, maximum likelihood, MSE) for the whole sample of countries and the entire period $\{2000,\ldots,2021-h\}$.

Table A3: Tuning parameter grids

Model	Hyper-parameters	Grid
Probit	None	None
Tree	Tree depth (ν)	0.005, 0.010, 0.015
Random Forest	Number of variables (ω)	2, 5, 10
AdaBoost	Number of iterations (N_I)	1,2,…,1000
Neural Network	Neurons in the first and second layer (M,P)	(1,1), (2,1), (2,2)
SVM	Predefined tolerated error (C)	0.1, 0.5, 1, 5, 10

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.

Badia, M. M. M., Medas, P., Gupta, P., and Xiang, Y. (2020). Debt is not free. *IMF Working Papers*, 1.

Breiman, L., Friedman, J.H., Olshen, R.A., and Stone, C.I. (1984). *Classification and regression trees*. Wadsworth, New York

Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5-32.

Cascarino, G., Moscatelli, M., and Parlapiano, F. (2022). Explainable Artificial Intelligence: interpreting default forecasting models based on Machine Learning. *Bank of Italy Occasional Papers*, No. 674.

Cerovic, M. S., Gerling, M. K., Hodge, A., and Medas, P. (2018). Predicting fiscal crises. *IMF Working Papers*, 181.

Chinn, M. D., and Ito, H. (2006). What matters for financial development? Capital controls, institutions, and interactions. *Journal of development economics*, 81(1), 163-192.

Ciarlone, A., and Trebeschi, G. (2005). Designing an early warning system for debt crises. *Emerging Markets Review*, *6*(4), 376-395.

Cortes, C., and Vapnik, V. (1995). Support-vector networks. *Machine learning*, *20*(3), 273-297.

Darvas, Z. (2012). Real effective exchange rates for 178 countries: a new database. *Working Paper 2012/06, Bruegel*, 15 March 2012.

Dawood, M., Horsewood, N., and Strobel, F. (2017). Predicting sovereign debt crises: an early warning system approach. *Journal of Financial Stability*, *28*, 16-28.

Demler, O. V., Pencina, M. J., Cook, N. R., and D'Agostino Sr, R. B. (2017). Asymptotic distribution of ΔAUC, NRIs, and IDI based on theory of U-statistics. *Statistics in medicine*, 36(21), 3334-3360.

Fernández, A., Klein, M. W., Rebucci, A., Schindler, M., and Uribe, M. (2016). Capital control measures: A new dataset. *IMF Economic Review*, *64*(3), 548-574.

Fioramanti, M. (2008). Predicting sovereign debt crises using artificial neural networks: A comparative approach. *Journal of Financial Stability*, *4*(2), 149-164.

Fisher, A., Rudin, C., and Dominici, F. (2019). All Models are Wrong, but Many are Useful: Learning a Variable's Importance by Studying an Entire Class of Prediction Models Simultaneously. *Journal of Machine Learning Research*, *20*(177), 1-81.

Freund, Y., and Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of computer and system sciences*, *55*(1), 119-139.

Gerling, M. K., Medas, M. P. A., Poghosyan, M. T., Farah-Yacoub, J., and Xu, Y. (2018). Fiscal crises. *Journal of International Money and Finance*, *88*, 191-207.

Gregorutti, B., Michel, B., and Saint-Pierre, P. (2017). Correlation and variable importance in random forests. *Statistics and Computing*, 27(3), 659-678.

Hajivassiliou, V. A. (1994). A simulation estimation analysis of the external debt crises of developing countries. *Journal of applied econometrics*, *9*(2), 109-131.

Hastie T., Tibshirani R., Friedman J. (2009). *The Elements of Statistical Learning*. Springer Series in Statistics. Springer, New York

Hellwig, K. P. (2021). Predicting Fiscal Crises: A Machine Learning Approach. *IMF Working Papers*, 150.

IMF (2017). Review of the Debt Sustainability Framework for Low Income Countries: Proposed Reforms. *IMF Policy Papers*.

IMF (2022). 2022 Global Debt Monitor. *IMF Fiscal Affairs Department*.

Intrator O. and Intrator N. (1993) Using Neural Nets for Interpretation of Nonlinear Models. *Proceedings of the Statistical Computing Section*, 244-249 San Francisco: American Statistical Society.

Jarmulska, B. (2022). Random forest versus logit models: Which offers better early warning of fiscal stress?. *Journal of Forecasting*, *41*(3), 455-490.

Kaminsky, G. L., and Vega-Garcia, P. (2016). Systemic and idiosyncratic sovereign debt crises. *Journal of the European Economic Association*, *14*(1), 80-114.

Koepke, R. (2019). What drives capital flows to emerging markets? A survey of the empirical literature. *Journal of Economic Surveys*, 33(2), 516-540.

Kraay, A., and Nehru, V. (2006). When is external debt sustainable?. *The World Bank Economic Review*, *20*(3), 341-365.

Liu, L., Chen, C., and Wang, B. (2021). Predicting financial crises with machine learning methods. *Journal of Forecasting*.

Manasse, P., Roubini, N., and Schimmelpfennig, A. (2003). Predicting sovereign debt crises (No. 2003/221). *IMF Working Papers*, 221.

Manasse, P., and Roubini, N. (2009). "Rules of thumb" for sovereign debt crises. *Journal of International Economics*, *78*(2), 192-205.

Mullainathan, S., and Spiess, J. (2017). Machine learning: an applied econometric approach. *Journal of Economic Perspectives*, *31*(2), 87-106.

Rodriguez, A., and Rodriguez, P. N. (2006). Understanding and predicting sovereign debt rescheduling: A comparison of the areas under receiver operating characteristic curves. *Journal of Forecasting*, 25(7), 459-479.

RECENTLY PUBLISHED “TEMI” (*)

N. 1383 – *Bank lending to small firms: metamorphosis of a financing model*, by Paolo Finaldi Russo, Valentina Nigro and Sabrina Pastorelli (October 2022).

N. 1384 – *Macroeconomic effects of growth-enhancing measures in the euro area*, by Alessandro Cantelmo, Alessandro Notarpietro and Massimiliano Pisani (October 2022).

N. 1385 – *Revisiting the real exchange rate misalignment-economic growth nexus via the across-sector misallocation channel*, by Claire Giordano (October 2022).

N. 1386 – *Labor market spillovers of a large plant opening. Evidence from the oil industry*, by Matteo Alpino, Irene Di Marzio, Maurizio Lozzi and Vincenzo Mariani (October 2022).

N. 1387 – *Unburdening regulation: the impact of regulatory simplification on photovoltaic adoption in Italy*, by Federica Daniele, Alessandra Pasquini, Stefano Clò and Enza Maltese (October 2022).

N. 1388 – *Liberalizing the opening of new pharmacies and hospitalizations*, by Andrea Cintolesi and Andrea Riganti (November 2022).

N. 1389 – *Mediation Analysis Synthetic Control*, by Giovanni Mellace and Alessandra Pasquini (November 2022).

N. 1390 – *The effects of partial employment protection reforms: evidence from Italy*, by Diego Daruich, Sabrina Di Addario and Raffaele Saggio (November 2022).

N. 1391 – *The role of majority status in close election studies*, by Matteo Alpino and Marta Crispino (November 2022).

N. 1392 – *Issuing bonds during the Covid-19 pandemic: is there an ESG premium?*, by Fabrizio Ferriani (November 2022).

N. 1393 – *Economic fundamentals and stock market valuation: a CAPE-based approach*, by Maria Ludovica Drudi and Federico Calogero Nucera (November 2022).

N. 1394 – *The impact of “Metro C” in Rome on the housing market*, by Federica Daniele and Elena Romito (December 2022).

N. 1395 – *Gender quotas, board diversity and spillover effects. Evidence from Italian banks*, by Silvia Del Prete, Giulio Papini and Marco Tonello (December 2022).

N. 1396 – *Real-time inequalities and policies during the pandemic in the US*, by Luisa Corrado, Daniela Fantozzi and Simona Giglioli (December 2022).

N. 1397 – *The impact of sovereign tensions on bank lending: identifying the channels at work*, by Fabiana Sabatini (December 2022).

N. 1398 – *Make-up strategies and exchange rate pass-through in a low-interest-rate environment*, by Alessandro Cantelmo, Pietro Cova, Alessandro Notarpietro and Massimiliano Pisani (December 2022).

N. 1399 – *Measuring peer effects in parental leaves: evidence from a reform*, by Davide Dottori, Francesca Modena and Giulia Martina Tanzi (February 2023).

N. 1400 – *Firms' innovation and university cooperation. New evidence from a survey of Italian firms*, by Daniela Bragoli, Flavia Cortelezzi and Massimiliano Rigon (February 2023).

N. 1401 – *Schools and the transmission of Sars-Cov-2: evidence from Italy*, by Salvatore Lattanzio (February 2023).

N. 1402 – *Remittances in times of crisis: evidence from Italian corridors*, by Alessio Ciarlone (February 2023).

N. 1403 – *EU banks' dividend policies: main determinants and the role of capital ratios*, by Salvatore Cardillo and Jacopo Raponi (February 2023).

(*) Requests for copies should be sent to:
Banca d'Italia – Servizio Studi di struttura economica e finanziaria – Divisione Biblioteca e Archivio storico – Via Nazionale, 91 – 00184 Rome – (fax 0039 06 47922059). They are available on the Internet www.bancaditalia.it.

2020

ALESSANDRI P. and M. BOTTERO, *Bank lending in uncertain times,* R European Economic Review, V. 128, **WP 1109 (April 2017).**

ANTUNES A. and V. ERCOLANI, *Public debt expansions and the dynamics of the household borrowing constraint,* Review of Economic Dynamics, v. 37, pp. 1-32, **WP 1268 (March 2020).**

ARDUINI T., E. PATACCHINI and E. RAINONE, *Treatment effects with heterogeneous externalities,* Journal of Business & Economic Statistics, , v. 38, 4, pp. 826-838, **WP 974 (October 2014).**

BALTRUNAITE A., C. GIORGIANTONIO, S. MOCETTI and T. ORLANDO, Discretion and supplier selection in public procurement, Journal of Law, Economics, and Organization, v. 37, 1, pp. 134-166, **WP 1178 (June 2018)**

BOLOGNA P., A. MIGLIETTA and A. SEGURA, *Contagion in the CoCos market? A case study of two stress events,* International Journal of Central Banking, v. 16, 6, pp. 137-184, **WP 1201 (November 2018).**

BOTTERO M., F. MEZZANOTTI and S. LENZU, *Sovereign debt exposure and the Bank Lending Channel: impact on credit supply and the real economy,* Journal of International Economics, v. 126, article 103328, **WP 1032 (October 2015).**

BRIPI F., D. LOSCHIAVO and D. REVELLI, *Services trade and credit frictions: evidence with matched bank – firm data,* The World Economy, v. 43, 5, pp. 1216-1252, **WP 1110 (April 2017).**

BRONZINI R., G. CARAMELLINO and S. MAGRI, *Venture capitalists at work: a Diff-in-Diff approach at late-stages of the screening process,* Journal of Business Venturing, v. 35, 3, **WP 1131 (September 2017).**

BRONZINI R., S. MOCETTI and M. MONGARDINI, *The economic effects of big events: evidence from the Great Jubilee 2000 in Rome,* Journal of Regional Science, v. 60, 4, pp. 801-822, **WP 1208 (February 2019).**

COIBION O., Y. GORODNICHENKO and T. ROPELE, *Inflation expectations and firms' decisions: new causal evidence,* Quarterly Journal of Economics, v. 135, 1, pp. 165-219, **WP 1219 (April 2019).**

CORSELLO F. and V. NISPI LANDI, *Labor market and financial shocks: a time-varying analysis,* Journal of Money, Credit and Banking, v. 52, 4, pp. 777-801, **WP 1179 (June 2018).**

COVA P. and F. NATOLI, *The risk-taking channel of international financial flows,* Journal of International Money and Finance, v. 102, **WP 1152 (December 2017).**

D'ALESSIO G., *Measurement errors in survey data and the estimation of poverty and inequality indices,* Statistica Applicata - Italian Journal of Applied Statistics, v. 32, 3, **WP 1116 (June 2017).**

DE BLASIO G., I. DE ANGELIS and L. RIZZICA, *Lost in corruption. Evidence from EU funding to Southern Italy,* Italian Economic Journal / Rivista italiana degli economisti, v. 6, 3, pp. 355–377, **WP 1180 (December 2016).**

DEL PRETE S. and S. FEDERICO, *Do links between banks matter for bilateral trade? Evidence from financial crises,* Review of World Economic, v. 156, 4, pp. 859 - 885, **WP 1217 (April 2019).**

D'IGNAZIO A. and C. MENON, *The causal effect of credit Guarantees for SMEs: evidence from Italy,* The Scandinavian Journal of Economics, v. 122, 1, pp. 191-218, **WP 900 (February 2013).**

ERCOLANI V. and F. NATOLI, *Forecasting US recessions: the role of economic uncertainty,* Economics Letters, v. 193, **WP 1299 (October 2020).**

MAKINEN T., L. SARNO and G. ZINNA, *Risky bank guarantees,* Journal of Financial Economics, v. 136, 2, pp. 490-522, **WP 1232 (July 2019).**

MODENA F., E. RETTORE and G. M. TANZI, *The effect of grants on university dropout rates: evidence from the Italian case,* Journal of Human Capital, v. 14, 3, pp. 343-370, **WP 1193 (September 2018).**

NISPI LANDI V., *Capital controls spillovers,* Journal of International Money and Finance, v. 109, **WP 1184 (July 2018).**

PERICOLI M., *On risk factors of the stock–bond correlation,* International Finance, v. 23, 3, pp. 392-416, **WP 1198 (November 2018).**

PIETRUNTI M. and F. M. SIGNORETTI, *Unconventional monetary policy and household debt: the role of cash-flow effects,* Journal of Macroeconomics, v. 64, Article 103201, **WP 1142 (October 2017).**

RAINONE E., *The network nature of OTC interest rates,* Journal of Financial Markets, v.47, article 100525, **WP 1022 (July 2015).**

RAINONE E. and F. VACIRCA, *Estimating the money market microstructure with negative and zero interest rates,* Quantitative Finance, v. 20, 2, pp. 207-234, **WP 1059 (March 2016).**

RIZZICA L., *Raising aspirations and higher education. Evidence from the UK's widening participation policy,* Journal of Labor Economics, v. 38, 1, pp. 183-214, **WP 1188 (September 2018).**

RIZZICA L. and M. TONELLO, *Persuadable perceptions: the effect of media content on beliefs about corruption,* Economic Policy, v. 35, 104, pp. 679–737, **WP 1043 (November 2016).**

SANTIONI, R., F. SCHIANTARELLI and P. STRAHAN, *Internal capital markets in times of crisis: the benefit of group affiliation*, Review of Finance, v. 24, 4, pp. 773-811, **WP 1146 (October 2017).**

SCHIANTARELLI F., M. STACCHINI and P. STRAHAN, *Bank Quality, judicial efficiency and loan repayment delays in Italy,* Journal of Finance, v. 75, 4, pp. 2139-2178, **WP 1072 (July 2016).**

2021

ACCETTURO A., A. LAMORGESE, S. MOCETTI and D. PELLEGRINO, *Housing Price elasticity and growth: evidence from Italian cities,* Journal of Economic Geography, v. 21, 3, pp. 367-396, **WP 1267 (March 2020).**

AFFINITO M. and M. PIAZZA, *Always look on the bright side? Central counterparties and interbank markets during the financial crisis*, International Journal of Central Banking, v. 17, 1, pp. 231-283, **WP 1181 (July 2018).**

ALBANESE G., E. CIANI and G. DE BLASIO, *Anything new in town? The local effects of urban regeneration policies in Italy,* Regional Science and Urban Economics, v. 86, **WP 1214 (April 2019).**

ALBANESE G., G. DE BLASIO and A. LOCATELLI, *Does EU regional policy promote local TFP growth? Evidence from the Italian Mezzogiorno,* Papers in Regional Science, v. 100, 2, pp. 327-348, **WP 1253 (December 2019).**

ALBERTAZZI A., A. NOBILI and F. M. SIGNORETTI, *The bank lending channel of conventional and unconventional monetary policy,* Journal of Money, Credit and Banking, v. 53, 2-3, pp. 261-299, **WP 1094 (Jan 2017).**

ANZUINI A. and L. ROSSI, *Fiscal policy in the US: a new measure of uncertainty and its effects on the American economy*, Empirical Economics, v. 61, 6, pp. 2613-2634, **WP 1197 (November 2018).**

APRIGLIANO V. and D. LIBERATI, *Using credit variables to date business cycle and to estimate the probabilities of recession in real time*, The Manchester School, v. 89, 51, pp. 76-96, **WP 1229 (July 2019).**

AUER S., M. BERNARDINI and M. CECIONI, *Corporate leverage and monetary policy effectiveness in the euro area,* European Economic Review, v. 140, Article 103943, **WP 1258 (December 2019).**

BANERJEE R, L. GAMBACORTA and E. SETTE, *The real effects of relationship lending,* Journal of Financial Intermediation, v. 48, Article 100923, **WP 1133 (September 2017).**

BARONE G., F. DAVID, G. DE BLASIO and S. MOCETTI, *How do house prices respond to mortgage supply?,* Journal of Economic Geography, v. 21, 1, pp.127-140, **WP 1282 (June 2020).**

BARONE G. and S. MOCETTI, *Intergenerational mobility in the very long run: Florence 1427-2011,* Review of Economic Studies, v. 88, 4, pp. 1863–1891, **WP 1060 (April 2016).**

BARTOCCI A., L. BURLON, A. NOTARPIETRO and M. PISANI, *Macroeconomic effects of non-standard monetary policy measures in the Euro Area: the role of corporate bond purchases,* The Manchester School, v. 89, S1, pp. 97-130, **WP 1241 (October 2019).**

BATINI N., A. CANTELMO, G. MELINA and S. VILLA, *How loose, how tight? A measure of monetary and fiscal stance for the euro area,* Oxford Economic Papers, v. 73, 4, pp. 1536-1556, **WP 1295 (September 2020).**

BENETTON M. and D. FANTINO, *Targeted monetary policy and bank lending behavior,* Journal of Financial Economics, v. 142, 1, pp. 404-429, **WP 1187 (September 2018).**

BUSETTI F., M. CAIVANO and D. DELLE MONACHE, *Domestic and global determinants of inflation: evidence from expectile regression,* Oxford Bulletin of Economics and Statistics, v. 83, 4, pp. 982-1001, **WP 1225 (June 2019).**

BUSETTI F., M. CAIVANO, D. DELLE MONACHE and C. PACELLA, *The time-varying risk of Italian GDP,* Economic Modelling, v. 101, Article 105522, **WP 1288 (July 2020).**

BUSETTI F., S. NERI, A. NOTARPIETRO and M. PISANI, *Monetary Policy strategies in the new normal: a model-based analysis for the Euro Area,* Journal of Macroeconomics, v. 70, Article 103366, **WP 1308 (December 2020).**

CAPOLONGO A. and C. PACELLA, *Forecasting inflation in the Euro Area: countries matter*, Empirical Economics, v. 61, 4, pp. 2477-2499, **WP 1224 (June 2019).**

CARMIGNANI A., G. DE BLASIO, C. DEMMA and A. D'IGNAZIO, *Urbanization and firm access to credit,* Journal of Regional Science, v. 61, 3, pp. 597-622, **WP 1222 (June 2019).**

CORNELI F., *Financial integration without financial development,* Atlantic Economic Journal, v. 49, 2, pp. 201-220, **WP 1120 (June 2017).**

COVA P., P. PAGANO, A. NOTARPIETRO and M. PISANI, *Secular stagnation, R&D, public investment and monetary policy: a global-model perspective,* Macroeconomic Dynamics, v. 25, 5, pp. 1267-1287, **WP 1156 (December 2017).**

DE PHILIPPIS M., *Multitask agents and incentives: the case of teaching and research for university professors,* Economic Journal, v. 131, 636, pp. 1643-1681, **WP 1042 (December 2015).**

DEL PRETE S. and M. L. STEFANI, *Women as "Gold Dust": gender diversity in top boards and the performance of Italian banks,* Economic Notes, Monte dei Paschi di Siena, v. 50, 2, e12183, **WP 1014 (June 2015).**

FERRERO G., M. LOBERTO and M. MICCOLI, *The assets' pledgeability channel of unconventional monetary policy,* Economic Inquiry, v. 59, 4, pp. 1547-1568, **WP 1119 (June 2017).**

FIDORA M., C. GIORDANO and M. SCHMITZ, *Real exchange rate misalignments in the Euro Area,* Open Economies Review, v. 32, 1, pp. 71-107, **WP 1162 (January 2018).**

GAMBACORTA L., G. RICOTTI, S. SUNDARESAN and Z. WANG, *Tax effects on bank liability structure,* European Economic Review, v. 138, Article 103820, **WP 1101 (February 2017).**

HERTWECK M., V. LEWIS and S. VILLA, *Going the extra mile: effort by workers and job-seekers,* Journal of Money, Credit and Banking, v. 54, 8, pp. 2099-2127, **WP 1277 (June 2020).**

LI F., A. MERCATANTI, T. MAKINEN and A. SILVESTRINI, *A regression discontinuity design for ordinal running variables: evaluating central bank purchases of corporate bonds,* The Annals of Applied Statistics, v. 15, 1, pp. 304-322, **WP 1213 (March 2019).**

LOSCHIAVO D., *Big-city life (dis)satisfaction? The effect of urban living on subjective well-being,* Journal of Economic Behavior & Organization, vol. 192, pp. 740-764, **WP 1221 (June 2019).**

LOSCHIAVO D., *Household debt and income inequality: evidence from Italian survey data,* Review of Income and Wealth. v. 67, 1, pp. 61-103, **WP 1095 (January 2017).**

METELLI L. and F. NATOLI, *The international transmission of US tax shocks: a proxy-SVAR approach,* IMF Economic Review, v. 69, 2, pp. 325-356, **WP 1223 (June 2019).**

NISPI LANDI V. and A. SCHIAVONE, *The effectiveness of capital controls,* Open Economies Review, v. 32, 1, pp. 183-211, **WP 1200 (November 2018).**

PAPETTI A., *Demographics and the natural real interest rate: historical and projected paths for the Euro Area,* Journal of Economic Dynamics and Control, v. 132, Article 04209, **WP 1306 (November 2020).**

PEREDA FERNANDEZ S., *Copula-based random effects models for clustered data,* Journal of Business & Economic Statistics, v. 39, 2, pp. 575-588, **WP 1092 (January 2017).**

2022

ANDINI M., M. BOLDRINI, E. CIANI, G. DE BLASIO, A. D'IGNAZIO and A. PALADINI, *Machine learning in the service of policy targeting: the case of public credit guarantees,* Journal of Economic Behavior & Organization, v. 198, pp. 434-475, **WP 1206 (February 2019).**

ANGELICO C., J. MARCUCCI, M. MICCOLI and F. QUARTA, *Can we measure inflation expectations using twitter?,* Journal of Econometrics, v. 228, 2, pp. 259-277, **WP 1318 (February 2021).**

BARTOCCI A., A. NOTARPIETRO and M. PISANI, *Covid-19 shock and fiscal-monetary policy mix in a monetary union,* Economic challenges for Europe after the pandemic, Springer Proceedings in Business and Economics, Berlin-Heidelberg, Springer, **WP 1313 (December 2020).**

BOTTERO M., C. MINOIU, J. PEYDRÒ, A. POLO, A. PRESBITERO and E. SETTE, *Expansionary yet different: credit supply and real effects of negative interest rate policy,* Journal of Financial Economics, v. 146, 2, pp. 754-778, **WP 1269 (March 2020).**

BRONZINI R., A. D'IGNAZIO and D. REVELLI, *Financial structure and bank relationships of Italian multinational firms,* Journal of Multinational Financial Management , v. 66, Article 100762, **WP 1326 (March 2021).**

CANTELMO A., *Rare disasters, the natural interest rate and monetary policy,* Oxford Bulletin of Economics and Statistics, v. 84, 3, pp. 473-496, **WP 1309 (December 2020).**

CARRIERO A., F. CORSELLO and M. MARCELLINO, *The global component of inflation volatility,* Journal of Applied Econometrics, v. 37, 4, pp. 700-721, **WP 1170 (May 2018).**

CIAPANNA E. and G. ROVIGATTI, *The grocery trolley race in times of Covid-19. Evidence from Italy,* Italian Economic Journal / Rivista italiana degli economisti, v. 8, 2, pp. 471-498, **WP 1341 (June 2021).**

CONTI A. M., A. NOBILI and F. M. SIGNORETTI, *Bank capital requirement shocks: a narrative perspective,* European Economic Review, v.151, Article 104254, **WP 1199 (November 2018).**

FAIELLA I. and A. MISTRETTA, *The net zero challenge for firms' competitiveness,* Envnmental & Resource Economics, v. 83, pp. 85-113, **WP 1259 (February 2020).**

GUISO L., A. POZZI, A. TSOY, L. GAMBACORTA and P. E. MISTRULLI, *The cost of steering in financial markets: evidence from the mortgage market,* Journal of Financial Economics, v.143, 3, pp. 1209-1226, **WP 1252 (December 2019).**

LAMORGESE A. and D. PELLEGRINO, *Loss aversion in housing appraisal: evidence from Italian homeowners,* Journal of Housing Economics, v. 56, Article 101826, **WP 1248 (November 2019).**

LI F., T. MÄKINEN, A. MERCATANTI and A. SILVESTRINI, *Causal analysis of central bank holdings of corporate bonds under interference,* Economic Modelling, v.113, Article 105873, **WP 1300 (November 2020).**

LOBERTO M, A. LUCIANI and M. PANGALLO, *What do online listings tell us about the housing market?,* International Journal of Central Banking, v. 18, 4, pp. 325-377, **WP 1171 (April 2018).**

MIRENDA L., M. SAURO and L. RIZZICA, *The economic effects of mafia: firm level evidence,* American Economic Review, vol. 112, 8, pp. 2748-2773, **WP 1235 (October 2019).**

MOCETTI S., G. ROMA and E. RUBOLINO, *Knocking on parents' doors: regulation and intergenerational mobility,* Journal of Human Resources, v. 57, 2, pp. 525-554, **WP 1182 (July 2018).**

PERICOLI M. and M. TABOGA, *Nearly exact Bayesian estimation of non-linear no-arbitrage term-structure models,* Journal of Financial Econometrics, v. 20, 5, pp. 807-838, **WP 1189 (September 2018).**

ROSSI P. and D. SCALISE, *Financial development and growth in European regions,* Journal of Regional Science, v. 62, 2, pp. 389-411, **WP 1246 (November 2019).**

SCHIVARDI F., E. SETTE and G. TABELLINI, *Credit misallocation during the European financial crisis,* Economic Journal, v. 132, 641, pp. 391-423, **WP 1139 (September 2017).**

TABOGA M., *Cross-country differences in the size of venture capital financing rounds: a machine learning approach,* Empirical Economics, v. 62, 3, pp. 991-1012, **WP 1243 (November 2019).**

FORTHCOMING

APRIGLIANO V., S. EMILIOZZI, G. GUAITOLI, A. LUCIANI, J. MARCUCCI and L. MONTEFORTE, *The power of text-based indicators in forecasting Italian economic activity,* International Journal of Forecasting, **WP 1321 (March 2021).**

BARTOCCI A., A. NOTARPIETRO and M. PISANI, *Non-standard monetary policy measures in non-normal times,* International Finance, **WP 1251 (November 2019).**

CIAPANNA E, S. MOCETTI and A. NOTARPIETRO, *The macroeconomic effects of structural reforms: an empirical and model-based approach,* Economic Policy, **WP 1303 (November 2022).**

DI ADDARIO S., P. KLINE, R. SAGGIO and M. SØLVSTEN, *It ain't where you're from, it's where you're at: hiring origins, firm heterogeneity, and Wages,* Journal of Econometrics, **WP 1374 (June 2022).**

FERRARI A. and V. NISPI LANDI, *Whatever it takes to save the planet? Central banks and unconventional green policy,* Macroeconomic Dynamics, **WP 1320 (February 2021).**

FERRARI A. and V. NISPI LANDI, *Toward a green economy: the role of central bank's asset purchases,* International Journal of Central Banking, **WP 1358 (February 2022).**

LILLA F., *Volatility bursts: a discrete-time option model with multiple volatility components,* Journal of Financial Econometrics, **WP 1336 (June 2021).**

LOBERTO M, *Foreclosures and house prices,* Italian Economic Journal / Rivista italiana degli economisti, **WP 1325 (March 2021).**

LOMBARDI M. J., M. RIGGI and E. VIVIANO, *Bargaining power and the Phillips curve: a micro-macro analysis, and wages,* Journal of the European Economic Association, **WP 1302 (November 2020).**

RIZZICA L., G. ROMA and G. ROVIGATTI, *The effects of shop opening hours deregulation: evidence from Italy,* The Journal of Law and Economics, **WP 1281 (June 2020).**

TANZI G. M., *Scars of youth non-employment and labour market conditions,* Italian Economic Journal / Rivista italiana degli economisti, **WP 1312 (December 2020).**