# Policy Research Working Paper

# Debt Vulnerability Analysis

9929

# A Multi-Angle Approach

Doerte Doemeland Marcello Estevão Charl Jooste James Sampi Vasileios Tsiropoulos



WORLD BANK GROUP

Macroeconomics, Trade and Investment Global Practice February 2022

#### Abstract

Countries with high debt exposure are vulnerable to economic and financial shocks that could lead to sovereign defaults. This paper develops a methodology to identify countries that are at risk of debt default based on four elements of debt vulnerability. These elements capture the different ways in which risks associated with high debt are assessed, namely: (i) the *fundamental*, (ii) the *subjective*, (iii) the *judgmental*, and (iv) the *theoretical*. The fundamental element considers the liquidity, solvency, and institutional risk elements of debt vulnerability. The subjective element captures the investors' perceptions of debt default, while the judgmental element is based on the debt thresholds as defined by Debt Sustainability Frameworks. Finally, the theoretical element is normative and captures what ought to be. The methodology constructs an index for each of these four elements and uses them as predictors in a model of public debt default. The methodology flags countries that are at risk of default by means of machine learning techniques and delivers outputs that point to underlying causes of vulnerability. The methodology complements existing monitoring tools for assessing debt sustainability.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

This paper is a product of the Macroeconomics, Trade and Investment Global Practice. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The authors may be contacted at ddoemeland@worldbank.org; mestevao@worldbank.org; cjooste@worldbank.org; jsampibravo@worldbank.org; and vtsiropoulos@worldbank.org.

# Debt Vulnerability Analysis: A Multi-Angle Approach \*

Doerte Doemeland Marcello Estevão Charl Jooste James Sampi Vasileios Tsiropoulos

*Keywords*: Debt, Default, Factor Model, Prediction. *JEL classification*: H12, H63, C32

<sup>\*</sup>World Bank Group. Ideas in the paper are those of the authors and not that of the World Bank. We thank Timothy Kehoe and Juan Conesa for their help and for sharing their code to simulate our theoretical framework. The ideas in this paper benefited from remarks by Paloma Casero, Andrew Burns, Ivailo Isvorski, Lea Hakim, Jorge Araujo, Aart Kraay, Pablo Saavedra, Gallina Vincelette, Franziska Ohnsorge, Norbert Fiess, Stefano Curto, Naoko Kojo, Gregor Schwerhoff, and Alexander Haider. Author contacts: ddoemeland@worldbank.org; mestevao@worldbank.org; cjooste@worldbank.org; jsampibravo@worldbank.org; vtsiropoulos@worldbank.org

## 1 Introduction

A balloon is elastic and can expand when filled with air. Good latex balloons can contain even more air. One thing is certain however, inject more air than the capacity of that balloon and it will burst. The noise of the pop is loud and the consequences sudden.

A country's debt is very much like a balloon. There exists a threshold beyond which a country cannot accumulate more debt lest it defaults. The problem for the economist is that each country is a different balloon, and it is up to country officials and multilateral institutions to understand a country's capacity to accumulate debt in a sustainable manner. Many countries already have inflated balloons or sizable debts. Even though they do not necessarily need to default, it is important to understand their ability to repay debts in the face of unknown events.

Some of these debt sensitivities and vulnerabilities depend on economic fundamentals, political changes, and transparent debt accounting practices. Assuming that debt is transparently communicated and the political environment known, then policymakers can use, as best as possible, available fiscal and economic data to evaluate debt sustainability.

Using the analogy of the balloon, debt vulnerability analysis is further complicated by the environment, which is state and time dependent. First, one needs to understand the structure of debt (i.e. the balloon itself), which includes expected repayments, ownership, debt size, sensitivity to domestic economic environment, and maturity structure. Next, one needs to consider the operator of the balloon (e.g. institutions, policy makers, and politicians). Political statements and the credibility of institutions may change the fiscal policy landscape and hence influence debt sustainability. Finally, one has to consider the external environment in which the balloon floats (e.g. investor perceptions and exchange rates). Unforeseen exogenous shocks do occur and, depending on the debt structure, may have a material impact on debt sustainability. Careful debt vulnerability analysis considers all three aspects.

Reflecting on the structure of debt over the last couple of years, it is evident that public debt in IBRD and IDA countries has been rising significantly. Debt increases have been particularly acute for IBRD countries (Figure 1a) as a result of large issuance of international bonds by middle-income countries (Figure 1b).<sup>1</sup> External sovereign debt carrying variable interest rates has risen markedly in recent years. A few countries have significant and growing non-resident participation in debt. Shifts in debt composition towards foreign reliance leave countries more exposed to various shocks, such as a sharp increase in global interest rates or capital outflows followed by currency depreciation (Panizza et al. (2009)). The increase in debt also limits fiscal space, making it harder for country authorities to conduct counter-cyclical fiscal policy and requiring future resources to be diverted towards repayments.

Changes in debt composition are also followed by a reduction in the average time to maturity and higher average interest rates (Figure 2a). This leaves countries more vulnerable to liquidity shocks and debt rollover risks, although debt increases and composition changes do not necessarily imply an inevitable default. Several emerging market economies (EMEs) have increased their resilience to shocks thanks mainly to prudent policies, the

<sup>&</sup>lt;sup>1</sup>LICs = Low Income Countries; LMICs = Lower Middle Income Countries; UMICs = Upper Middle Income Countries; EMDEs = Emerging Market Developing Economies.

build-up of external and fiscal buffers, and the implementation of sound debt management strategies.



Figure 1: Public Debt

Source: WEO, WB International Debt Statistics; and WB staff calculations

**Figure 2:** External Public Debt Maturity and Average Interest Rates (In years and percentage points)



Source: WB International Debt Statistics. Excluding high-income countries; and WB staff calculations

This narrative paints a broad picture of a fictional balloon for a group of countries. The external environment is also changing rapidly, especially with increased risk of tightening monetary policy in developed economies, in general, and in the United States, in particular. Several countries have already defaulted, although risks of a systemic crisis seem limited now. The operator of the balloon is also changing with some country leaders (see Dovis et al. (2016) for a theory of debt on cycles of populism vs. austerity) pushing a populist narrative that could spook market participants.

While it is not surprising that countries default, it is surprising that lessons from the past are not internalized, thus building pressure for action, including better debt management and crisis planning. A relatively persistent number of defaults over time illustrates that hard

lessons have not been learned in spite of a growing policy dialogue on ways to manage debt and run fiscal policy in a sustainable manner (see Figure 3).



Figure 3: Number and size of defaults

Source: Own calculations using 2020 version of Beers and Mavalwalla (2017)

Consequently, the work program for addressing public debt vulnerabilities in emerging and low-income countries remains more relevant than ever. The objective is to minimize the economic and fiscal costs associated with shocks and to increase debt data transparency and the efficiency and sustainability of fiscal and debt policies. Four pillars are needed to achieve this: 1) improving debt analysis and early warnings systems; 2) enhancing debt transparency; 3) strengthening debt management capacity; and 4) reviewing debt policies. Developing an early-warning framework to monitor public debt vulnerabilities and fiscal risks for market-access countries, would go a long way to bolster this agenda.<sup>2</sup> That includes: i) developing a methodological framework for assessing debt vulnerabilities; ii) developing a tool to support the analysis; and iii) establishing a consultation process and internal mechanisms to discuss action plans WB teams could propose for countries with elevated debt vulnerabilities.

This paper presents the methodological outline of a framework for monitoring debt vulnerabilities in market-access economies. It seeks to complement existing World Bank debt vulnerability assessments, including the joint Bank-Fund Debt Sustainability Framework.

The paper produces four indicators of debt vulnerability, which are used for predicting debt default. The four indicators capture overlapping-but-different angles of vulnerabil-

<sup>&</sup>lt;sup>2</sup>MACs or market access countries are defined as countries that are not eligible for the Poverty Reduction and Growth Trust (PRGT). Emerging market countries (EMs) are a subset of MACs. MACs thus cover all advanced economies and the majority of emerging markets and may include LICs that have significant access to market financing. The classifications of advanced economies and emerging markets are defined by the IMF's World Economic Outlook.

ity. The indicators by themselves can be used to evaluate sources of risk and, hence, for designing strategies to minimize potential default. Empirical tests show that these indicators improve model predictions of default relative to simple rules of thumb regarding debt thresholds and standard multivariate regressions of vulnerability. The methodology can thus be used for describing sources of vulnerability and for assessing whether a country is becoming more vulnerable.

The rest of the paper is structured as follows: The rationale and a summary of related exercises are presented in Section 2. The overview of the framework with detailed exposition of the different indices are covered in Section 3. Section 4 presents the results of the default prediction. Section 5 discusses the shortcomings of the exercise and how some gaps are filled. Section 6 concludes the paper.

## 2 Rationale and related exercises

For a creditor like the World Bank, the timely monitoring of debt vulnerabilities is important for a variety of reasons:

- 1. **Provide better policy advice to countries**. The recognition and assessment of existing vulnerabilities is important for providing sound policy advise to countries and help prevent costly mistakes in the event of large shocks. Delayed government reactions to shocks often entail very costly policy adjustments (Laeven and Valencia (2008)), requiring steep corrective fiscal measures, often with adverse political consequences and a negative impact on growth and poverty reduction. Heightened vulnerabilities may also increase relative price volatility and overall economic and financial instability (Calvo et al. (2006)). Lack of preparedness often leads to delayed or inadequate policy responses (Solomon (2009)).
- 2. Help establish crisis prevention mechanisms. Crisis events can quickly erase gains in growth and poverty reduction. Debt vulnerability assessments can help inform World Bank responses to prevent possible crisis in countries.
- 3. Help countries build resilience. World Bank operations and analytical work focus on building resilience to financial, climate, and fiscal shocks to protect and deepen reductions in poverty and inequality. The importance of a proactive Word Bank response to anticipated crises is emphasized by the Independent Evaluation Group (IEG) of the World Bank (Group (2017)).

There already exists a large body of work regarding the prediction of defaults or earlywarning systems. To mention a few, Manasse et al. (2003) create an early warning system of default using several liquidity and solvency indicators in a logit and tree-based model. Their results suggest that a high external debt-to-GDP ratio makes a country more likely to enter a crisis (one of their measures of solvency), but is not a good predictor of remaining in a crisis. The World Bank produced a book in 2011 citing various sources of default (Primo Braga and Vincelette (2011)). High external debt and high credit spreads significantly indicate a rising probability of default. Interestingly, using a Bayesian Model Averaging methodology to uncover the importance of different explanatory variables, only external debt enters the model significantly. There is also an increase in the use of machine learning techniques to determine which variables are useful for predicting crisis. Savona and Vezzoli (2015) show that regression-tree based methods outperform competing methodologies such as logit. In their example, liquidity and past default behavior predict oncoming crisis. Finally, there are also attempts to use theoretical models to predict crisis. Gumus et al. (2017) show how a theoretical model can be used to derive endogenous time varying probabilities of default. They show that these model-based-probabilities outperform some other empirical methods such as a logit model to predict sovereign crises in Argentina.

The methodology proposed in this paper describes several analytical advances. Our proposed framework would complement existing monitoring tools by establishing:

- 1. An established methodological framework and a set of analytical results that assesses country-specific public debt vulnerabilities. The Bank produces and monitors country risk from various angles. The purpose of this work is to supplement these analyses with a deeper focus on sovereign debt vulnerabilities. These include a proposed metric of public debt risks, default probabilities related to debt sustainability and to deteriorating perceptions, and country-specific safe debt thresholds, which are described in detail below.
- 2. An easy-to-use system that delivers key results to country teams and will serve as a quick one-stop shop to assess public debt risks.

## 3 An overview of the framework

Debt-vulnerable countries have a high probability of default either conditional or unconditional (e.g. standard debt dynamics when for example a large primary deficit is accompanied by real interest rates that exceed real GDP growth yielding higher debt levels in subsequent periods) on shocks. A country, by design, is not vulnerable to default if it can weather volatile political periods, can sustain a continuous and large exchange rate depreciation, changes to external creditor market conditions, manage fiscal policy in a sustainable manner, and face a slowdown in domestic and trade partner economic growth.

This paper posits that four measured aspects of an economy may capture and predict the likelihood of default.

The first considers fundamentals. Measures that capture liquidity, solvency, and institutional risks are combined into a single vulnerability variable called the **Fundamental Index**. There exists potentially hundreds of economic indicators that could affect a country's ability to service its debt. Combining these measures, using country-specific weights, adds the advantage of weighing the most important indicators for debt vulnerability while keeping all other variables in quantitative analysis. The Fundamental Index is a proxy for the core of what makes debt vulnerable. Increases in any of the indicators point to an increase in exposure and hence make a country more vulnerable to shocks.

The second aspect captures investor's subjective perceptions regarding risks. Credit default swaps, or interest rate differentials between a country's average bond interest rate and that of a risk-free rate (e.g., of bonds issued by the German or the U.S. government) are used to derive an implied **Probability of Default Index**. There are periods when market perceptions deviate from fundamentals, sometimes abruptly and significantly. An implied default probability measure can capture the market's views as they change, which would otherwise not have been picked up in low-frequency variables connected to fundamentals. As will be shown below, the index captures a key ingredient in a DSA by essentially inverting a "fan chart" to produce a measure of how far a country is from its maximal effort to implement discretionary fiscal policies in an attempt to reduce debt.

The third aspect is expert evaluation of risks. Using standard DSA debt indicators, we evaluate for each country the number of variables that exceed thresholds of risk as defined by the DSAs into a **Count Index**. DSAs have been a staple input for evaluating debt risks. Economists at the IMF and the WBG use judgment and history as a guiding principle to assess whether a country is in debt distress. The Count Index reflects some of this judgment.

The fourth and final aspect captures what ought to be. Using a theoretical framework that embeds endogenous default, we calculate time varying debt default thresholds for each country to construct our **Theory Index**. The theoretical model is used to find debt ratios that are indicative of risk. These theoretical thresholds are endogenous and hence different than DSA guidelines on debt. They are time varying and embed the possibility of a self-fulfilling debt crisis - i.e., when market perceptions deteriorate significantly from fundamentals.

These four indices are constructed for each country and then used as explanatory variables for debt default. Figure 4 illustrates the objective. The red circle represents the unit disk (i.e. 1 = default), while the blue area represents the space spanned by the four angles of debt.





Since a key aim is predicting default, the analysis attempts to incorporate as much information as possible (reduce omitted variables), while not trying to overfit relationships. To this end, the indices are used as explanatory variables in estimating observed or historical default episodes. The strength of each of the indices in terms of prediction is assessed empirically. Given that this part of the work is one related to a prediction (i.e., to default or not), several advances in econometrics and machine learning are employed. The methods employed for prediction include a classical logit regression, and two methods from the machine learning literature: random forests and the least absolute shrinkage and selection operator (LASSO) (Tibshirani (1996)).

The following subsections describe each of the analytical angles for analyzing the risks of debt default.

## 3.1 The Fundamental Index

As much data as possible are collected for each country and then reduced to a single indicator of vulnerability. A resilience index is also constructed to measure the attempts of government to offset default risks. We define resilience as a country's ability to mitigate adverse shocks. Some measures of resilience entail having healthy buffers such as large reserves for extraordinary financing if called upon, strong FDI flows instead of portfolio flows (hence reducing the likelihood of sudden stops), floating exchange rates (although this can be a risk too if a country accumulated large foreign currency denominated public debt), the existence of large and accessible sovereign wealth funds, long-maturity debt with moderate-to-low upcoming redemptions, low short-term interest rates, credible monetary policy, financial sector stability, stable exchange rates, extent of political and economic stability, and a sustainable current account balance.

The selection of variables to be used for the **fundamental** and **resilience** indices is based on the literature, which has identified that public debt vulnerabilities typically arise due to three types of fundamental risks:

- 1. Solvency risks are associated with the size of net present value liabilities exceeding the net present value of assets or a country is insolvent when the net present value of future primary balances is smaller than outstanding debt (Bouabdallah et al. (2017), Roubini (2001)). Thus, if obligations become so large that no future primary balances can finance the debt, a government might default.
- 2. Liquidity risk arises if the government does not have new financing or cash at hand to service upcoming redemptions (Bouabdallah et al. (2017), Roubini (2001)). This is often a short-term indicator of risk.
- 3. Commitment risk is related to the government's fiscal stance and the prudent management of fiscal policy.

The literature shows that one of the first warnings that a country might default is due to the maturity structure of debt - where short-term debt cannot be easily rolled over (Cole and Kehoe (1996), Rodrik and Velasco (1999)). Sanchez et al. (2018) show that before a default episode, the maturity of debt shortens and interest rate spreads increase as a consequence of lower economic growth. Bocola and Dovis (2016) show that interest rates are a significant source of default risk - they calibrate a model for the Italian economy and find that 12 percent of the spreads during 2008-2012 were due to rollover risks. Similarly, Arellano and Ramanarayanan (2012) show that higher interest rate spreads lead to a shortening of debt maturities and, thus, rollover risks.

A second warning is the size of foreign-currency (FX) denominated debt (Tomz and Wright (2013)), which exposes countries to currency risk. If countries only issued in domestic currency then a country can implicitly default by inflating debt away; an option not available for FX-denominated debt.

Political institutions matter for repaying debts. As an example, Biglaiser and Staats (2012) show that countries with strong courts are less likely to default. Another factor that could lead to default is the mere issuance of new debt to devalue existing debt - since new issuance increases the likelihood of default. This debt dilution can account for up to 78 percent of default risk and eliminating dilution decreases optimal duration by roughly two years (Hatchondo et al. (2015)).

"Twin-crisis" events occur when the financial and public sector default. Balteanu and Erce (2017) show that balance sheet interconnections (e.g. materializing contingent claims and guarantees), credit dynamics (non-performing loans increase due to higher interest rates that reflect both public and private sector risks) and economic growth (e.g. the link between asset prices, investments, and economic growth) are drivers of banking and public sector crises. This is due to falling capital ratios requiring public sector intervention in the case of a financial crisis. Balteanu and Erce (2017) also argue procyclical fiscal policy along with a lack in competition might make banks vulnerable to shocks. Obstfeld (2011) argues that foreign debt (both public and private), especially short-term liabilities, worsen the credit worthiness of the public sector.

The selected variables for the public debt fundamental index, which are associated with liquidity or solvency risk include: (i) public debt stock; (ii) financing needs (both public and external), (ii) exchange rates, (iii) interest rates, (iv) currency composition of debt, (v) debt service, and (vi) current account balances. Standard variables for liquidity risk are: (i) redemption schedules and (ii) amount of short term debt. Commitment risks include: (i) reserves proxying as buffers to unexpected shocks and (ii) differences between observed primary balances and primary balances required to stabilize debt. Links to the external and financial market include external financing needs, short-term external debt, non-performing loans, and interest rate spreads between loans and deposits. Data are sourced from the World Bank's WDI indicators, the Macro Poverty Outlook and International Debt Statistics, others are sourced from Bloomberg, the Bank of Canada and the IMF's WEO. Ideally, one would want to cast a wide net and capture all of these data but they are not all available for every market-access country in this study. Thus, the set of variables will vary across countries in our sample.

A summary of data and their sources used in this study include:

- Indicators of public debt: Debt maturity profiles (World Development Indicators (WDI), Bloomberg) and repayment schedules (Bloomberg) to assess liquidity risks; debt service costs (MPO, WDI, WEO) and currency composition (WDI, Bloomberg) to assess exchange rate risks; fiscal space (Macro Poverty Outlook (MPO), WDI, World Economic Outlook (WEO)), gross public financing needs, and required primary balances(MPO, WDI, WEO) to assess scope and ability of policy adjustments; and actual defaults (Bank of Canada, credit ratings (Bloomberg).
- Indicators of resilience: Foreign exchange reserves (WDI), exchange rate regime (Ilzetzki et al. 2017), monetary policy regimes, credible institutions (Country Pol-

icy and Institutional Assessment (CPIA) from the WDI), and sovereign wealth funds (country economists).

• Macroeconomic indicators: Core measures to calculate ratios, real variables and general indicators of economic health include real and nominal GDP, exports, public debt, the consumption price deflator, real and nominal exchange rates, average interest rates, credit spreads on public debt, public expenditures and revenues, the current and capital accounts of the balance of payments, credit ratings and default episodes (MPO, WDI, WEO, Bank of Canada, Bloomberg), and non-performing loans (WDI).

The methodology aggregates the above mentioned variables into a single index. The weights are estimated via a standard dimension reduction technique (e.g. dynamic factor models or principle components) using data starting from 2002. The advantage of this technique is that weights are neutral and completely dependent on the data. This methodology controls for country characteristics, where individual series might have a different weight across countries in each index (see Appendix A for more details on the methodology).

The proposed index measures whether country debt vulnerabilities are increasing or abating: (i) an increase in the vulnerability index highlights risk build-up, (ii) an increase in the vulnerability index relative to its trend highlights risk build-up with respect to its history, or (iii) an increase in the country index relative to regional trends to identify large outlier countries within a region. An increase in the resilience index implies that countries are building up buffers against possible unexpected shocks.

Principal components analysis (PCA) is used for extracting the vulnerability index from the annual data module. PCA is appealing because of its computational advantages and asymptotic properties in large data sets (see Bai (2003)).

After estimation, a fundamentals vulnerability index is generated. Figure 5 illustrates build-up of public and private debt vulnerabilities within each country using the proposed index. It is important to note that an increase in each of the variables used in the index point to an increase in vulnerability, such that the index has the same interpretation and such that the weights are positive (i.e., an increase in the index implies an increase in one of the variables - e.g., larger foreign denominated debt issuance will contribute to an increase in the index).

The black dots in Figure 5 represent the debt vulnerability indices by country while the red lines are the country trends of these indices (i.e. by construction the average is roughly zero). The figure summarizes vulnerability since 2002, while the making of each country's index takes the full sample size (which differs by country) into account.<sup>3</sup>

A positive value indicates that a country is now more vulnerable than, on average, in the past. As an example, Argentina's index has been rising steadily since 2012 and is above its trend. Using this measure, ex post, one would have inferred rising vulnerabilities faced by Argentina and using the decomposition chart in the Appendix, that most of the risk was emanating from upcoming amortization. Other countries in the sample (e.g., Brazil and Belarus) also have rising vulnerabilities, however, have not defaulted (in Brazil's case, probably because foreign debt is a very small share of its large public sector debt).

<sup>&</sup>lt;sup>3</sup>See Appendix E for the same chart but for the resilience indicator.

The results suggest that these countries should be closely scrutinized and action plans to mitigate a debt disaster should be set.



Figure 5: Country indices

The results can also be summarized for a cross-section (i.e., a given year). The vulnerability and the resilience indices jointly determine the "Red-Square" measure (see Figure 6). The horizontal line depicts the vulnerability indices (centered on zero) for public and external debt, while the vertical line represents the resilience index.

A country that has rising vulnerabilities and a reduction in resilience falls into the red square. Countries with a reduction in vulnerabilities and an increase in resilience find themselves in the green square and are possibly less vulnerable to default risk. The two pink squares still indicate risk, but less than the red-square measure. The red square highlights countries such as Argentina, Lebanon, and Pakistan with heightened public debt exposure. For a country like Turkey, the predominant debt risk is related to the private sector and with falling reserves related to attempts to defend the currency.

#### Sources of vulnerability

The methodology, with its estimated weights, allows decomposing the index into its components (see Appendix A for the relative weights). These weights summarize the importance of each of the variables in the construction of the indices in terms of total explained variation. The sum of the weights add up to 1. All the indicators, as mentioned above, are constructed such that an increase in any one of them should lead to an increase in vulnerability. This is done to avoid counteracting effects. As an example, for Argentina, 1 percent of the variation in the Fundamental Index is due to larger short-term debt, while

#### Figure 6: The red-square



16 percent accounts for upcoming redemptions. The rest is split between the remaining variables.

### 3.2 The Default Index

#### Proxying government ability

The fiscal solvency constraint states that if debt is larger than the sum of the discounted future primary balances, then a country defaults. This gives rise to a debt stabilizing primary balance, or the deficit required to keep debt equal to the previous period. One can use this measure and compare it to a country's historical efforts (measured as past primary balances). If the required primary deficit is larger than the historical deficits, then the likelihood that a country will default increases. In most DSAs this information is captured through a "fan chart" that summarizes the possible paths of debt as an indication of sustainability. If debt is rising (or if its colored fans are expanding upwards) then there is a chance that debt will not be kept fixed at a certain level - in this case the risk of default increases (and policy changes may not be sufficient to cover upcoming interest repayments). The fan chart is constructed by drawing random values from a joint distribution (usually real interest rates and GDP growth from a macroeconomic model) and constructing the subsequent debt paths.

One can invert the formulas and proxy it by calculating the probability distribution function of primary balances and contrast it to the required primary balance to stabilize debt. This is not a perfect proxy, but yields an estimate of a country's default probability associated with ability (i.e., its past actions speak to its resolve to keep debt stable). A large required primary deficit to finance debt implies a larger probability of default.

The primary balance required to stabilize debt is defined by a stable or decreasing level of debt given projections for GDP growth, interest rates, and primary balances. This can be expressed as:

$$pb_t^* = \frac{r-g}{1+g}d_{t-1}$$
(1)

where r is the average implied real government interest rate on debt, g is average real GDP growth, d is the gross debt to GDP ratio and  $pb_t^*$  is the primary balance required to stabilize debt.

Assuming that historical primary balances  $pb_t$  give us an indication of the government's effort to conduct sustainable fiscal policy, the probability of default associated with ability is defined as:

$$0 \le P(pb_t \le pb_t^*) \le 1 \tag{2}$$

If  $pb_t < pb_t^* \forall t$  then a government cannot meet the debt stabilizing condition. We assume that this probability can be expressed as a cumulative normal distribution ( $\Phi$ ). The use of a normal distribution is based on convenience and simplifies the analysis (Figure 7 removes deficits larger than 20% and surpluses larger than 20% for all countries in the sample historical primarily balances are approximately normal). Future work may improve this part of the work by including distributions that have fatter tails. The probability of default can then be expressed as:

$$P(pb_t^* \le pb_t) = \Phi\left(\frac{pb_t - E(pb_t^*)}{\sigma(pb_t^*)}\right)$$
(3)

Figure 7: Distribution of of primary balances



That is, we compute rolling averages of the required primary balance  $pb^*$  using the interest rate–growth differential each period.

Figure 8 illustrates this idea. Assume that a country requires a primary balance of 6 percent of GDP to stabilize debt in the following year. According to the historical outcomes





of the primary balance, it is not in that country's ability to meet that requirement and it will very likely default.

#### **Proxying investor perceptions**

The probability of default is related to the ability of servicing future debts. This subsection shows how to calculate investors' pricing of default probability. The CDS gives us a rough approximation of default probabilities (see Hull et al. (2004) for some examples). The default probability is calculated following the approach proposed by O'Kane (2011). This approach looks at the CDS premium <sup>4</sup> for each single bond as a discounted present value of the risk neutral expected loss associated with the cash flow of that bond.<sup>5</sup> For our purpose, we approximate the term structure (e.g., the interest rate on five-year government bonds) with a flat yield curve, which we use for discounting (r).<sup>6</sup> We also use a fixed neutral loss-given-default assumption of  $L^*$  percent<sup>7</sup> and a risk neutral default hazard rate  $\pi_t^*$ .

The following condition holds:

$$x\frac{1-(1+r)^{T-1}}{r} = \pi^* L^* \frac{1-(1+\pi^*)^T (1-r)^{-T}}{r+\pi^*}$$
(4)

with x denoting the CDS premium, T the maturity date, and r the discount rate (Yield rate). The default, if it happens, is assumed to strike at the end of each period. Therefore,  $\pi_t^* = x/L^*$  for small enough values of r. However, short-run fluctuations can arise because

<sup>&</sup>lt;sup>4</sup>The premium is the amount paid by the buyer to the seller as a proportion of the nominal value of the contract (in basis points).

<sup>&</sup>lt;sup>5</sup>The present value of the CDS premium is equivalent to the present value of the expected loss,  $\pi^*L^*$ .

<sup>&</sup>lt;sup>6</sup>CDS price valuation associated with one bond requires a proper discount rate for the expected cash flow associated with that bond, for which standard literature uses the yield bond.

<sup>&</sup>lt;sup>7</sup>The loss-given default value is taken from the Unsecured Recovery Rates published by Moody's Financial Services.

of political instability or debt related issues. O'Kane (2011) proposes to account for additional volatility by incorporating a Gaussian term, where the default rate is defined as follows:

$$\pi_t = \pi_t^* + \sigma_{CDS,t} \tag{5}$$

where  $\sigma_{CDS,t}$  is the CDS conditional volatility (e.g. estimated from conditional heteroskedasticity models, ARCH or GARCH). Finally, the probability of default associated with deteriorating perceptions (SDP) is calculated as follows:

$$SDP_t = 1 - e^{-\pi_t} \tag{6}$$

Figure 9 illustrates the perceived probability of default for Argentina, Lebanon, Turkey, and the República Bolivariana de Venezuela until 2019 as an example using the methodology above. Argentina shows higher risk during the 2009-2014 period, when EMBI spreads peaked at 900 bps, mainly because of the public debt restructuring negotiations. Similarly, Turkey's probability sharply declined during the 2012-2013 period, after Fitch upgraded Turkey's credit rating to investment grade following an 18-year gap. Finally, the República Bolivariana Venezuela shows a constant risk-trend since 2014, after the Maduro election. Nonetheless, a resurrection of risk is pointed out during 2018/2019 for all of them.

#### Figure 9: Perceived probability of default



Figure 10 below summarizes the two types of default calculations. Rational expectations suggest that investor perceptions should be aligned with fundamentals and that there should be no arbitrage opportunities, implying that the two probability measures (one aligned with fundamentals and ability and one aligned with perceptions) should co-move. In most cases the two default probabilities move in the same direction. However, in some cases perceived default exceeds the probability associated with ability. We view this case as a heightened risk of default (or the probability of a panic), which may lead to self-fulfilling debt crisis. These estimates are used later in the estimation of debt thresholds.



#### Figure 10: Default probabilities

## 3.3 The Count Index

An important element of the analysis is a display of the raw data and organizing the variables by indication of vulnerability or lack thereof. This is important for several reasons, which include (i) transparency about the data used, (ii) serving as an additional filter to spot risks that might not be picked up by the indices, and (iii) serving as a benchmark against other models of vulnerability. Some of these indicators have already been used in standard DSA tables. We rely on thresholds (determined jointly by Bank-Fund staff) to highlight risk in the various indicators (Appendix D highlights the table of indicators).

In total there are 12 indicators for which thresholds are provided: Current account balance (CAB), which has a threshold of 5 percent to GDP; external debt as a share of GDP (threshold equals 60 percent); external private non-guaranteed debt as a share of GDP (threshold equals 20 percent); external gross financing needs as a share of GDP (threshold equals 15 percent); short-term external debt as a share of total external debt (threshold equals 5 percent); gross public debt as a share of GDP (threshold equals 60 percent); external public debt as a share of GDP (threshold equals 60 percent); external public debt as a share of GDP (threshold equals 60 percent); external public debt as a share of GDP (threshold equals 60 percent); external public debt as a share of GDP (threshold equals 60 percent); external public debt as a share of GDP (threshold equals 60 percent); external public debt as a share of GDP (threshold equals 60 percent); budget deficit as a share of GDP (threshold equals 4 percent); the primary deficit as a share of GDP (threshold equals 60 percent);

equals 3 percent); short-term debt as a share of GDP (threshold equals 4 percent); public debt held by non-residents as a share of gross debt (threshold equals 45 percent); and EMBI global diversified yield spread (threshold equals 600 basis points). The debt sustainability framework (DSF) uses indicative thresholds, linked by country classification, to analyze the risk of external debt distress. Thresholds are (statistically determined) bounds above which the risk of debt distress is considered elevated. Note that these thresholds were derived to maximize their joint correlation with a very specific definition of a debt distress event conditional on a very specific set of control variables. In this methodology we use the data on actual defaults, which may not correlated perfectly with the debt distress indicators used for deriving these thresholds. <sup>8</sup>

Figure 11 compares the unweighted average debt to GDP ratios for different shares of indices exceeding their thresholds over time. In the majority of cases very high debt is associated with a higher share of indicators breaching their thresholds.



Figure 11: Share of indicators exceeding thresholds

Note that DSA's are subject to revisions. This makes sense as new types of information and analytical tools help to make analysis more concise. As an example, the new MAC-DSA framework includes several new features and is being rolled out towards the end of 2021.<sup>9</sup>

### 3.4 The Theory Index

In this section we derive an angle of vulnerability by constructing an index related to a theoretical model of default. This theoretical model derives the economic conditions for when it is optimal for a country to default, and when it is strategic to avoid defaults. The derived quantities of the model produces time varying threshold debt to GDP ratios, which are codified into an index used in the DVA. The Theory Index relies on the models developed by Cole and Kehoe (1996, 2000) and Conesa and Kehoe (2017) which we extend to study the debt-thresholds of multiple economies. The underlying model is a dynamic stochastic general equilibrium model that characterizes the values of government debt and

<sup>&</sup>lt;sup>8</sup>For more detailed discussion of why the DSF employs thresholds, and on the probit methodology, please see IMF (2017), and Kraay and Nehru (2006).

<sup>&</sup>lt;sup>9</sup>See https://www.imf.org/en/Publications/Policy-Papers/Issues/2021/02/03/ Review-of-The-Debt-Sustainability-Framework-For-Market-Access-Countries-50060

the debt's maturity structure consistent with a sustainable debt dynamics in the case of financial crises brought on by a loss of confidence in the government. The model also characterizes the optimal policy response of governments to the threat of such a crisis. The results of this section yield time varying debt theoretical thresholds that we can compare against actual debt outcomes. Three thresholds are estimated - one for a safe level of debt (where the likelihood of default is extremely low), one for risky levels of debt (where a country is vulnerable to exogenous shocks and can default), and another for default levels of debt to GDP (this is where a country will default). A country that exceeds the latter for a given year is a assigned a value of 3 for that index. A country that is in the safe area is assigned a value of 0. Thus, increasing values represent a deterioration in the debt sustainable outlook - and hence an increasing chance of default.

We start by defining the theoretical model used for calibrating risky debt thresholds.

Following Conesa and Kehoe (2017) we simplify the Cole and Kehoe (1996, 2000) model by eliminating the representative household's consumption-investment choice, while extending Cole and Kehoe (1996, 2000) by introducing debt maturity. The state of the economy in every period is:  $s = (B, z_{-1}, \zeta)$ , where *B* is the level of government debt, whether default has occurred in the past  $z_{-1} = 0$  or not  $z_{-1} = 1$ , and  $\zeta$  is the value of the sunspot variable. The country's GDP is:

$$y(z) = Z^{1-z}\bar{y} \tag{7}$$

As in Conesa and Kehoe (2017) and Cole and Kehoe (1996, 2000) the drop in productivity by the factor Z is the country's default penalty. Once z = 0, it stays equal to 0 forever. Here the default penalty occurs in the same period as the crisis. We keep the structure of the income shocks deliberately simple, so that incentives to smooth consumption even in periods of vulnerability to lenders' panics are tractable. Government's tax revenue is  $\theta y(z)$ , where we assume that the tax rate  $\theta$  is fixed. Given that there is no consumptioninvestment choice as in Conesa and Kehoe (2017), the consumption of the representative household is

$$c(z) = (1 - \theta)y(z) \tag{8}$$

The benevolent government offers B' in new bonds for sale and chooses whether or not to repay the fraction  $\delta$  of the existing stock of debt that becomes due,  $\delta B$ . As in Conesa and Kehoe (2017) we follow Chatterjee and Eyigungor (2012) and Hatchondo and Martinez (2009) which use a "memory-less" debt maturity form, so it is not necessary to keep track of the entire distribution of maturities of debt. Hence, the government's budget constraint is:

$$g + z\delta B = \theta y(z) + q(B', s)(B' - (1 - \delta)B)$$
(9)

where q(B', s) is the price that international bankers pay for B', g is government expenditure, and z is a binary variable that denotes the government decision to default or repay. Notice that the problem reduces to the one-period debt case when  $\delta = 1$ , or to infinitely lived debt as  $\delta$  tends to 0.

In every period,  $\zeta$  is drawn from the uniform distribution on a support [0,1]. If  $\zeta > 1-\pi$ , international bankers expect a crisis to occur and do not lend to the government if such a crisis would be self-fulfilling. The empirical part of our methodology allows us to set the

probability of a self-fulfilling crisis at a specific value  $\pi$ ,  $0 \le \pi \le 1$ . The timing within each period is like that in Cole and Kehoe:

- 1. The shock  $\zeta$  realizes, the aggregate state is  $s = (B, z_{-1}, \zeta)$ , and the government chooses how much debt B' to sell.
- 2. Each of a continuum of measure one of international bankers chooses how much debt b' to purchase. In equilibrium, b' = B'.
- 3. The government makes its default decision z, which determines y, c, and g.

Given this timing, we can reduce the government's problem choosing c, g, B', z to solve

$$V(s) = \max \ u(c,g) + \beta EV(s')$$

$$c = (1 - \theta)y(z)$$

$$g + z\delta B = \theta y(z) + q(B',s)(B' - (1 - \delta)B)$$
(10)

International bankers are risk neutral with discount factor  $\beta$  so that the bond prices q(B', s) are determined by the probability of default in the next period. There is a continuum of measure one of bankers. For further details please refer to Conesa and Kehoe (2017). The solution for the international bankers' utility maximization problem implies that

$$q(B',s) = \begin{cases} \beta[\delta + (1-\delta)q'(\cdot)], & \text{if } B' \leq \overline{b} \\ \beta(1-\pi)[\delta + (1-\delta)q'(\cdot)], & \text{if } \overline{b} < B' \leq \overline{B} \\ 0, & \text{if } \overline{B} < B' \end{cases}$$
(11)

where  $\bar{b}$  is the lower threshold and if  $B' \leq \bar{b}$  the government does not default even if international bankers do not lend.  $\bar{B}$  is the upper threshold and if  $\bar{b} < B' \leq \bar{B}$ , the government does not default if international bankers lend. If  $B' > \bar{B}$  the government defaults even if international bankers lend.

To solve for the equilibrium, we choose the same functional form for the utility function as in Conesa and Kehoe (2017),

$$U(c,g) = \log(c) + \gamma \log(g - \bar{g})$$
(12)

For the equilibrium definition and further details please refer to Conesa and Kehoe (2017) and Cole and Kehoe (1996, 2000).

The framework identifies thresholds that distinguish three zones of debt that rely on calibrations for each country:

• In the safe zone, the government would not find it optimal to default even if there is a panic in financial markets and it could not roll over its stock of debt. The debt ratio for the countries in the safe zone is relatively small. Hence, the optimal strategy is to be solvent and sacrifice consumption to repay the stock of debt, rather than experiencing an output loss because of an insolvency.

- In the crisis zone, the economy is vulnerable to a self-fulfilling crisis and the government's optimal policy is to default only if it could not roll over its stock of debt. Countries that fall in this zone are susceptible to uncertain shocks. These countries might find themselves in a crisis without even knowing that they are at risk when an uncertain shock were to materialize.
- In the default zone, the interest payments are so large that the government would choose to default even if investors were willing to refinance the stock of debt. Countries in the default zone can be considered vulnerable since they are estimated to face an elevated risk of default in the event of a crisis.

To produce outputs, the model requires a set of calibrated parameters. The model requires an estimate of the probability of a panic (proxied by the subjective default rates identified above <sup>10</sup>); the maturity composition of public debt;<sup>11</sup> the structural levels of revenues and expenditures; output loss upon a default; and the average yield to maturity.

As an example Cole and Kehoe (1996) illustrates the usefulness of these models in the case of Mexico. Mexico was not facing imminent default right before the 1994-1995 crisis. Yet, public debt was high and the share of short-term debt was high. Mexico had debt levels that were in the crisis zone due to **short maturity bonds**. In 1994, there was political turmoil followed by the assassination of the ruling party's presidential candidate Luis Donaldo Colosia. Subsequently, Mexico defaulted. Mexico could have been better prepared to address a potential shock by increasing its maturity structure and issuing bonds in local currency. Also, the government might have pursued fiscal consolidation more forcefully if there had been more awareness about the risk of default (Cole and Kehoe, 1996). The Theory index would then have picked up if Mexico was getting close to the default zone and would have coded it as a value of 3.

Figure 12 below illustrates an output for a hypothetical country given a set of calibrated parameters. In this chart, the y-axis expresses the future debt-to-GDP ratio, while the x-axis the current debt-to-GDP ratio. The red line is a 45-degree line. Any value above the red line suggests that government's optimal policy (black line) is to increase the debt-to-GDP ratio. On the contrary, any value below the red line implies an optimal policy of reducing future debt-to-GDP ratios. In this scenario, the lower threshold is about 70 percent of GDP and it divides the safe zone from the crisis zone. The upper threshold is about 90 percent of GDP and it divides the crisis zone from the default zone. Above this level, debt drops to a value of zero given that the country's optimal decision is to default even if investors were willing to refinance the stock of debt. Hence, investors will not lend to this country and the future debt level drops to a value of zero.

Finally, once we compute the time varying probability of a panic, we can express the thresholds also as a varying function of time. In Figure 13, we illustrate this for South Africa and Argentina in the sample using a simple 1 period maturity. The upper and risky thresholds change over the sample period. Argentina breaches the threshold by 2018, while South Africa is reaching the upper threshold as it accumulates more debt. Interestingly, the

<sup>&</sup>lt;sup>10</sup>Due to computational restrictions, we have created a mapping between the theoretical model's probability of panic and the subjective probability. The probability of panic is 0.4 times the subjective probability.

<sup>&</sup>lt;sup>11</sup>We obtained maturity rates on non-concessional borrowing for 42 countries from 5 regions using Bloomberg. To increase the sample size to 159 countries, we used the average rate of the 42 countries for each of the 5 regions.

#### Figure 12: Growing indebtedness



upper threshold is declining for Argentina, which is in line with a higher probability of a panic. For South Africa, the threshold has remained fairly constant.

Figure 13: Illustrating time varying debt thresholds



## 4 Results and tests

Apart from using the DVA as a means of uncovering underlying vulnerabilities, one can also test how well it does in terms of predicting defaults. This is an essential part of an early warning system. Does it warn that a country is becoming more exposed to shocks (and hence more vulnerable) and does it indicate whether a country's default probability is rising? The empirical results are only useful if they provide reasonable signals about actual sovereign defaults.

We compare the DVA against several alternative specifications. These alternative specifications include simple explanatory variables that typically give rise to defaults, to more complex multivariate systems (see as an example Kraay and Nehru (2006)). For the simpler models we use the debt to GDP ratio as a simple indicator of default - the assumption is that higher debt makes a country more vulnerable and hence should be associated with a higher default prediction. We also use a simple interest rate spread (each country's interest rate relative to the U.S. average interest rate on debt) as an indicator of default. The assumption here is that a higher spread signals liquidity risk and should be associated with a higher default prediction. The last model for comparison includes several of the variables in Kraay and Nehru (2006) that analyse periods of debt distress. The variable list includes short-term external debt, debt service of external public debt relative to exports, foreign currency composition of debt, governance (as a proxy for the non-available CPIA variable) and we also include external financing needs, which equals short term debt plus amortization of debt minus the current account balance.

We use several methodologies to predict defaults. We start by using a simple Logit specification as a first case. There are other methods that can improve on Logit regressions. The data-science literature has made several advances to test whether additional variables in regressions aid in predicting outcomes, or whether they are creating noise or are correlated with other explanatory variables. We apply two of the these methodologies. Thus, for the second regression methodology we use random forests. Random forests build on the Classification and Regression Tree (CART) methodology for prediction. The basic idea is to partition results into smaller groups based on the predictors. Random forests reduce the correlation between trees; the objective is to reduce the variance of each tree by introducing randomness into the tree growing process (or selection of input variables) via bootstrapping (see Breiman (2001)).<sup>12</sup> The third empirical methodology applies a Lasso, equation (15), by means of zeroing out coefficients that do not increase prediction. For the bi-variate equations, Lasso makes very little sense, unless the intercept is better at prediction than the explanatory variables. In the case of the DVA and the longer set of variables for Kraay and Nehru (2006) (which we refer to Classical in the results section), Lasso regressions may help us to eliminate variables that have poor prediction properties. In one sense, instead of using PCA for the generating the Fundamental index, we could have just used a Lasso regression to rid the regression of poor predictors. This would have certainly eliminated several variables. The only reason for doing this is to generate better predictions, but this comes at the cost of losing a descriptive part of the DVA - i.e., to assess where the vulnerabilities might arise. Hence we retain the causal implication of the Fundamental index using PCA.

How do we test our models? We follow the data-science literature by dividing the data into a training and a test data sample. The data set is split using a random break. Our training data consists of 80% of the entire data pool, while our test data is the remaining 20%. We train each of the models using cross-validation methods. This means that we randomly select five data sets from our training data and estimate the regressions on each of those. The reason for doing this is to find the coefficients in the regressions that minimize a loss function (e.g., RMSE's, ROC, or type 1 or type 2 errors). Once our regressions are

<sup>&</sup>lt;sup>12</sup>500 trees were selected to minimize the out of bag sample error.

"trained" we then analyze how well they predict the test data set.<sup>13</sup>

Our dependent variable, historical defaults, is sourced from The Bank of Canada (Beers and Mavalwalla (2017)). A default is defined as a late interest payment, an outright payment default, or when creditors incur losses owing to an extension of maturities or a reduction in rates. Using historical defaults helps us avoid having to define a vulnerability indicator and then trying to predict that. We are interested in predicting when a default occurs in time , i.e., when the 0 flips to a 1 in the sample. We test two cases for the timing of defaults, one where default occurs in period T and data are expressed in period T - 1, and the other where both default and data are contemporaneous. The results do not change much when we vary the timing. Admittedly the variables that lead to default might be different than the variables that keep a country in default. We are not separating out these effects and instead code the default data as equal to 1 when in default and 0 otherwise, and not only 1's when the default occurs. This is partially related to the default data, where default states tend to be very persistent in the sample of countries.

Furthermore, all our data are stacked. We thus have t annual observations from 2008 until 2020 for k = 65 market access countries. We control for within country specific fixed effects via demeaning.

### 4.1 Econometric specification

The standard logistic regression (with binary response equal to default (1) and no default (0)) can be expressed as:

$$Pr(Y_t = 1 | X = x) = \Pi_t = \frac{e^{\beta_0 + x_t^T \beta}}{1 + e^{\beta_0 + x_t^T \beta}}$$
(13)

where x is the input vector (i.e., the four indices from the DVA, or the three alternative models).

The parameters are estimated via maximizing the log-likelihood:

$$ll(\beta) = \frac{1}{n} \sum_{t=1}^{n} \left( y_t log(\frac{\Pi_t}{1 - \Pi_t}) + log(1 - \Pi_t) \right) = \sum_{t=1}^{n} \left( y_t(\beta_0 + x_t^T \beta) - log(1 + e^{\beta_0 + x_t^T \beta}) \right)$$
(14)

An additional constraint is added for the Lasso. Lasso regressions are useful for model selection by shrinking coefficients to zero. As an example, it is possible that the subjective probability estimate tracks the fundamental index over time. Lasso will penalize this and shrink one of the coefficients to zero. The general goal of Lasso is to minimize the sum of squared errors by including a penalty function:

$$ll(\beta) = \sum_{t=1}^{n} \left( y_t(\beta_0 + x_t^T \beta) - \log(1 + e^{\beta_0 + x_t^T \beta}) \right) - \lambda \sum_{j=1}^{K} |\beta_j|$$
(15)

where *j* is the *j*th variable in the list of *K*, *t* is the time index, *y* is default, and *x* are the four components of debt in a matrix. The tuning parameter  $\lambda$  controls the amount of

<sup>&</sup>lt;sup>13</sup>We also run these tests over an in-sample period and a pseudo out of sample period where 2019 is used as the out of sample year.

shrinkage. When  $\lambda = 0$  then no shrinkage occurs and the minimized function boils down to a standard least squares estimate. As  $\lambda$  increases more coefficients are set to zero. Note that there is a bias-variance trade-off in parameter estimation: with high  $\lambda$  the bias increases, while a low  $\lambda$  increases the variance. The four models yield predicted probabilities of default.

In terms of statistical significance of the variables used, only the theoretical index enters significantly in the Logit estimate. The Lasso shrinks the count and fundamental index parameters to zero and indicate that the default probability calculation and theoretical index explain default. The random forest, however, suggest that the fundamental and default probability indices are relatively more important in explaining default. Using an ensemble of methods and pooling estimates may produce a more accurate reflection of default likelihoods. Figure 14 summarizes the "Gini impurity", a measure of when to split a tree. The relative values are considered here (not the scale) - the fundamental index is roughly thrice as important as the count index in this example.



Figure 14: Variable importance in random forest estimation

## 4.2 Prediction

Once the regressions are trained we can turn to prediction. A standard output is producing confusion matrices. In Appendix F we include confusion matrices for all regression variable combinations for both the training and test data sets. The column headers are labeled "no crisis" and "crisis". These correspond to actual defaults and no defaults. The row headers are labeled "no signal" and "signal" and correspond to the regression's predictions. Given that each regression will only produce estimated default values (range between 0 and 1), we need to determine cut-off values for the classification (i.e., to generate a value of 1, which equals default and a value of 0 otherwise).

To determine the cut-off values from the continuous predictions, we follow Gabriele et al. (2017) (it expands Kaminsky (1999)). This test computes a noise-to-signal ratio to obtain critical cut-offs for when an indicator points to a default. As an example, the objective here would be to find a critical value of the estimated default probabilities that minimizes the noise-to-signal ratio, which is our objective function.

The approach of Gabriele et al. (2017) attempts to minimize the noise-to-signal ratio by varying the cut-off points of an indicator (in this case the estimated probability of default from either the logit, Lasso or random forest regressions). A simple illustration is presented in Table 1. The noise-to-signal ratio is calculated as: NTSR = [B/(B+D)]/[1 - (C/(A + C))], where the capitalized letters correspond to the blocks in the table. This test thus attempts to minimize Type 2 errors. An alternative approach defined in Gabriele et al. (2017) (which is used in this paper) is to take the weighted average of both errors,  $\theta \frac{C}{C+A} + (1 - \theta) \frac{B}{B+D}$ , defined as the policymaker's loss function (*PMLF*) where  $\theta$  measures the preference between the two types of errors and is determined endogenously by maximizing  $min(\theta; (1 - \theta)) - PMLF$ . The parameter  $\theta$  is the value that minimizes average loss.

<b>TADIC I.</b> INDISC-10-Signal fall	Table	1:	Noise-	to-signal	ratio
---------------------------------------	-------	----	--------	-----------	-------

	Default	No Default
Signal	А	B (Type II)
No Signal	C (Type I)	D

The cut-offs of the Logit, Lasso and random forest regressions are summarized in Table 3. For the different cut-offs, the missed crisis and the false alarms are calculated (Appendix F). These cut-off values are then used to determine the accuracy of the test case. As an example, the accuracy of the DVA random forrest regression is equal to 72% - i.e., it correctly classified up to 72% of actual defaults. The distinction between the regressions for the different variables are not that large.

	Logit	Random Forest	Lasso
DVA	0.58	0.55	0.51
Debt	0.58	0.62	0.50
Spread	0.62	0.64	0.50
Classic	0.52	0.58	0.50

Table 2: Cut-offs

# Making a Type 2 error is not as serious in this case as making a Type 1 error. If the model predicts a crisis and no default ensues then the economic consequences are likely smaller compared to the case where a default occurs and no signal was given. Furthermore, the costs of a type 2 error (i.e., indicating a crisis when none occurs) can be a reflection of the consolidation of governments to avoid a crisis - and hence there is an endogenous anticipation response where the averted crisis due to good fiscal policy would could have resulted in a crisis without the policy intervention. This implies that type 2 errors are not really knowable when a policy action followed a signal.

	Logit	Random Forest	Lasso
DVA	0.64	0.64	0.72
Debt	0.69	0.66	0.65
Spread	0.62	0.67	0.63
Classic	0.67	0.67	0.69

Table 3: Accuracy

Figure 15: Receiver Operator Curves



We also use the receiver operator characteristic (ROC) to test the performance of the regressions. The ROC is a performance curve, and the objective is to calculate the area under the curve (AUC) to measure the degree of separability. A higher AUC means that the regression does a better job at classifying defaults and no defaults. The ROC can be represented graphically, and has as its x-axis the false positive rates and as its y-axis the true positive rate (i.e., when the estimate of the regression classifies default correctly, or area A/[A + C]). An AUC= 0.5 means there is no degree of separability, while a value less than 0.5 means a worse degree of separability - i.e., it is reciprocating defaults and no defaults. As an example, an AUS = 0.8 means that the regression has a 80% chance to distinguish between defaults and no defaults.

The ROC test attempts to find a critical cut-off that minimizes Type 1 errors. The true positive rate (TPR) is calculated when the indicator signals default when default actually occurs (TP) and expresses this as a ratio of the sum of the true positives and false negatives (FN) (TPR = TP/[TP + FN]), or area (A/[A + C]). This is then compared to the false positive rate (FPR), which is calculated as the number of false positive signals (FP) as a ratio of false positives plus true negatives (TN), or in terms of the table (FPR = B/[B+D]).

We compute the area under the curve for the various regressions. In the left hand panel of Figure 15 the AUC of the Logit regressions are displayed. The middle panel summarizes the AUC's for the random forest regressions and the far right panel summarizes the AUC's for the lasso regressions.

In all regression specifications the DVA outperforms the alternative models. The AUC for the DVA in the Logit specification is 65.5%, for the random forest specification 74.0% and for the Lasso 68.9%, respectively.

While the four indices do not predict defaults with absolute accuracy, they do better than a coin flip, rules of thumb measures, and other predictors. The methodologies can be used as a filtering mechanism to flag potentially vulnerable countries.

## 5 Discussion

The approach taken in this methodology combines several elements in the debt-default literature to not only predict possible default, but to also describe sources of vulnerabilities.

This is where the vulnerability index plays an important role. While looking at several single indicators is useful to understand potential vulnerabilities, a country-specific weighted index summarizes a large set of information. Dimension reduction techniques are well known for picking up signals in a large data group.

A key element of this analysis is a description of what is meant by default probability. There are several definitions of default probability that can be applied. However, here we use only two measures. The first relates to the ability of the government to maintain debt on a stable path. The second relates to investor perceptions regarding default. Monitoring both measures reveal interesting details regarding periods of weakened perceptions over and above the government's own ability. Once the two default estimates diverge then it points to a panic.

Our methodology attempts to incorporate expert judgment in describing defaults. DSAs provide guidance on key thresholds of risk for several indicator variables. We use these thresholds along with historical data to codify judgment. We note that this judgment is subject to revisions as more information and better analytical tools help to bridge several gaps.

A novel addition to default prediction is to incorporate more structural and theoretical models. We utilize a standard general equilibrium framework to compute debt thresholds. This is different compared to empirical relationships between the size of debt and default to define thresholds. In this example, debt thresholds are determined by the maturity structure, the probability of a panic, and the size of the government. These measures provide insights on whether countries will do well to reduce debts to reduce risks. Two thresholds provide three regions of risk. There is a debt to GDP ratio where the risk of a default is very low. At the same time, there is a debt to GDP ratio where default is very likely or imminent. Finally, there is a region between the no default and default thresholds where countries are vulnerable to exogenous shocks. These thresholds are useful from a policy dialogue perspective to identify sources of risks and policies to reduce those risks.

All of these measures are used jointly to predict defaults. We benefit from advances in the econometric literature and test several specifications. In general, the predictions from the models do a reasonable job in predicting both in and out-of-sample defaults. Interestingly, the count index has the weakest explanatory power in predicting defaults.

There are obvious caveats to this methodology. As with any model, the outputs depend on the quality of the inputs. In the exercise described above, three significant sources of risk are not entirely captured due to lack of data: the realization of both implicit or explicit guarantees, the uncovering of hidden debts, and the risks associated with systemic shocks that spill over to countries. In cases where these data are available, the methodology can be appended to revise the results.

It should also be emphasized that the outputs of the debt vulnerability methodology are intended to supplement and support country risk assessments and not replace them. They indicate rising or abating vulnerabilities and analyze where the vulnerability build-up occurs. Consequently, the outputs in this paper tie in with expert opinion using DSAs and other assessments. As such, the DVA can be seen as a screening and filtering system to identify countries that are potentially at risk.

## 6 Conclusion

This paper describes a methodology for analyzing public debt vulnerability. Four distinct elements make up this framework: (1) A Fundamental Index identifies vulnerable countries by mapping default risk to changes in fundamentals; (2) a Default Probability Index measures the degree of weakened perceptions regarding risk; (3) a Count Index maps expert opinion to debt vulnerability; and (4) a Theory Index estimates debt thresholds that vary by the probability of a panic for each market access country in the study.

The Fundamental Index consists of a large set of variables related to debt vulnerabilities and summarizes risks by looking at available low-frequency data. The Fundamental Index shows rising/abating vulnerabilities by country.

Two default probabilities are calculated to create the Default Probability Index. One uses required primary balances and actual primary balances to produce a debt sustainability default probability related to a government's ability to finance its debts. The other one uses interest rate spreads to produce measures of defaulting expectations. The two default probabilities track each other for most of the sample, but diverge just before default. The difference between the two probability estimates is a proxy for a probability of a panic - if markets perceive a higher risk than what the fundamentals suggest, then it can lead to an increase in roll-over risk.

Thresholds from the IMF-WBG DSAs are used in the construction of the Count Index. The index measures the proportion of variables of a country that exceed a threshold. Countries that have many variables exceeding their thresholds are more likely to default.

The Theory Index of the framework consists of identifying debt thresholds that are indicative of risky, safe, and default regions. These debt-to-GDP thresholds are time varying as opposed to static thresholds typical of DSAs. They measure an angle of default related to what it ought to be. The theoretical model controls for debt maturity, spreads, and a panic probability to model endogenous default.

The four indices are used as predictors of sovereign default. Three statistical methodologies are applied. A classical logit regression, a LASSO logit regression, and random forests. The latter two methodologies apply machine learning techniques to maximize the prediction aspect of the framework.

The results suggest that the four indices do better than rules of thumb and standard estimates in identifying defaults, with a LASSO type regression yielding the highest level of precision.

This framework is distinct from traditional early-warning systems. In addition to flagging vulnerable countries, outputs help identify sources of vulnerability and hence provide a policy angle to reduce risks. The theoretical model, as an example, illustrates that debtto-GDP thresholds are unique to countries as opposed to applying a single threshold to all countries that fall within an income group. In addition, the Fundamental Index decomposes the variability of vulnerability into a set of variables.

The framework can be improved with access to more data. Some data are not publicly available for all countries - such as implicit and explicit government guarantees, which

are serious source of risk for countries. As argued in MTI-WBG (2021) debt data transparency is a fundamental agenda for improving our (debtors, creditors, researchers, and international organizations) ability for evaluating risks of debt crises.

## References

- Arellano, C. and Ramanarayanan, A. (2012). Default and the Maturity Structure in Sovereign Bonds. *Journal of Political Economy*, 120(2):187–232.
- Bai, J. (2003). Inferential Theory for Factor Models of Large Dimensions. *Econometrica*, 71:135–171.
- Bai, J. and Ng, S. (2000). Determining the Number of Primitive Shocks in Factor Models. *Journal of Business and Economic Statistic*, 25:52–60.
- Bai, J. and Ng, S. (2002). Determining the Number of Factors in Approximate Factor Models. *Econometrica*, 70:191–221.
- Balteanu, I. and Erce, A. (2017). Linking bank crisis and sovereign defaults: Evidence from emerging markets. *European Stability Mechanims*, 22. Working Paper.
- Beers, D. and Mavalwalla, J. (2017). Database of Sovereign Defaults, 2017. *Bank of Canada*, 101. Technical Report.
- Biglaiser, G. and Staats, J. (2012). Finding the democratic advantage in sovereign bond ratings: The importance of strong courts, property rights protection, and the rule of law. *International Organization*, 66:515–535.
- Bocola, L. and Dovis, A. (2016). Self-Fulfilling Debt Crises: A Quantitative Analysis. *National Bureau of Economic Research*, 22694. Working Paper.
- Bouabdallah, O., Checherita-Westphal, C., Warmedinger, T., de Stefani, R., Drudi, F., Setzer, R., and Westphal, A. (2017). Debt sustainability analysis for euro area sovereigns: a methodological framework. *ECB*, 185. Occasional Paper Series.
- Breiman, L. (2001). Rarge Forests. Machine Learning, 45(1):5–32.
- Calvo, G., Izquierdo, A., and Loo-Kung, R. (2006). Relative price volatility under sudden stops: The relevance of balance sheet effects. *Journal of International Economics*, 69(1):231–254.
- Chatterjee, S. and Eyigungor, B. (2012). Maturity, Indebtedness, and Default Risk. *American Economic Review*, 102(6):2674–2699.
- Cole, H. L. and Kehoe, T. J. (1996). A self-fulfilling model of Mexico's 1994-95 debt crisis. Staff Report 210, Federal Reserve Bank of Minneapolis.
- Cole, H. L. and Kehoe, T. J. (2000). Self-Fulfilling Debt Crises. *Review of Economic Studies*, 67(1):91–116.
- Conesa, J. C. and Kehoe, T. J. (2017). Gambling for redemption and self-fulfilling debt crises. *Economic Theory*, 64(4):707–740.

- Dovis, A., Golosov, M., and Shourideh, A. (2016). Political Economy and Sovereign Debt: A Theory of Cycles of Populism and Austerity. *NBER Working Papers*, 21948. Working Paper.
- Doz, C., Giannone, D., and Reichlin, L. (2012). A Quasi Maximum Likelihood Approach for Large Approximate Dynamic Factor Models. *Review of Economics and Statistics*, 94:1014–1024.
- Doz, C. and Reichlin, L. (2011). A Two-Step Estimator for Large Approximate Dynamic Factor Models Based on Kalman Filtering. *Journal of Econometrics*, 164:188–205.
- Durbin, J. and Koopman, S. J. (2012). *Time Series Analysis by State Space Methods*. Oxford University Press, Oxford.
- Gabriele, C., Erce, A., Athanasopoulou, M., and Rojas, J. (2017). Debt Stocks Meet Gross Financing Needs: A Flow Perspective into Sustainability. *European Stability Mechanism*. Working Paper Series 24-2017.
- Geweke, J. F. (1989). Bayesian Inference in Econometric Models Using Monte Carlo Integration. *Econometrica*, 57:1317–1339.
- Ghysels, E., Sinko, A., and Valkanov, R. (2007). Midas Regressions: Further Results and New Directions. *Econometric Reviews*, 26:53–90.
- Group, I. E. (2017). Crisis Response and Resilience to Systemic Shocks: Lessons from IEG Evaluations. World Bank.
- Gumus, I., Katayama, M., and Koeda, J. (2017). Model-based estimation of sovereign default risk. *SSRN Electronic Journal*.
- Hatchondo, J. C. and Martinez, L. (2009). Long-duration bonds and sovereign defaults. *Journal of International Economics*, 79(1):117–125.
- Hatchondo, J. C., Martinez, L., and Sosa-Padilla, C. (2015). Debt Dilution and Sovereign Default. *Center for Applied Economics and Research*, 2015(12. Working Paper.
- Hull, J., Predescu, M., and White, A. (2004). The relationship between credit defaults swap spreads, bond yields, and credit rating announcements. *Journal of Banking and Finance*, 28(11):2789–2811.
- IMF (2017). Review of the debt sustainability framework in low-income countries: Proposed reforms.
- Kaminsky, G. (1999). Currency and banking crisis: The early warning of distress. *IMF Working Paper*, 99/178.
- Kraay, A. and Nehru, V. (2006). When is external debt sustainable? *The World Bank Economic Review*, 20(3):341–365.
- Laeven, L. and Valencia, F. (2008). Systemic banking crises: A new database. *IMF Working Paper*, 08/224.

- Manasse, P., Roubini, N., and Schimmelpfenning, A. (2003). Predicting sovereign debt crises. *IMF Working Paper*, 03/221.
- MTI-WBG (2021). *Debt Transparency in Developing Economies*. The World Bank Group, Washington, DC.
- Obstfeld, M. (2011). Financial flows, financial crises, and global imbalances. *CEPR*, 8611. Discussion Paper.
- O'Kane, D. (2011). *Modelling single-name and multi-name credit derivatives*, volume 573. John Wiley & Sons.
- Onatski, A. (2012). Asymptotics of the Principal Components Estimator of Large Factor Models with Weakly Influential Factors. *Journal of Econometrics*, 168:244–258.
- Panizza, U., Sturzenegger, F., and Zettelmeyer, J. (2009). The economics and law of sovereign debt and default. *Journal of Economic Literature*, 47(3):651–698.
- Primo Braga, C. A. and Vincelette, G. A. (2011). *Sovereign debt and the financial crisis*. The World Bank, Washington DC.
- Rodrik, D. and Velasco, A. (1999). Short Term Capital Flows. NBER, 7364. Working Paper.
- Roubini, N. (2001). Debt sustainability. how to assess whether a country is insolven. *Stern School of Business, New York University*. Mimeo.
- Sanchez, J. M., Sapriza, H., and Yurdagul, E. (2018). Sovereign Default and Maturity Choice. *Journal of Monetary Economics*, 95(2018):72–85.
- Savona, R. and Vezzoli, M. (2015). Fitting and forecasting sovereign defaults using multiple risk signals. *Oxford Bulleting of Economics and Statistics*, 77(1):66–92.
- Schwaab, B., Koopman, S., and Lucas, A. (2017). Global Credit Risk: World, Country and Industry Factors. *Journal Applied Econometrics*, 32:296–317.
- Solomon, S. (2009). Regulation by the deal: The government's response to the financial crisis. *Adminstrative Law Review*, 61(3):463–541.
- Stock, J. and Watson, M. (1989). New Indexes of Coincident and Leading Economic Indicators. In Blanchard, O. and Fischer, S., editors, *NBER Macroeconomics Annual*. MIT Press, Cambridge.
- Tibshirani, R. (1996). Regression shrinkage and the selection via Lasso. *Journal of the Royal Statistical Society*, 58(1):267–288.
- Tomz, M. and Wright, M. L. (2013). Empirical Research on Sovereign Debt and Default. *NBER*, 18855. Working Paper.

## A Estimation Methodology

#### PCA in the analysis

The formal structure of this component is summarized as:

Let  $y_{it}$  be the observed data for the *i*th variable at time *t*. In total we have *N* variables indexed by i = 1, ..., N. Also, we have *T* time periods and t = 1, ..., T. The approximate factor model decomposes *N* dimensional vectors  $y_t = (y_{1t}, ..., y_{Nt})'$ , for t = 1, ..., T, as follows

$$y_t = \Lambda f_t + \varepsilon_t \tag{16}$$

where  $\Lambda = (\lambda_1, \ldots, \lambda_N)'$  is the  $N \times r$  matrix of factor loadings with r as the number of factors,  $f_t = (f_{1t}, \ldots, f_{rt})'$  is the  $r \times 1$  vector of factors and  $\varepsilon_t$  is the  $N \times 1$  idiosyncratic disturbance term.

Equation (16) can be written by stacking the observations over t = 1, ..., T as follows

$$y = \Lambda f + \varepsilon \tag{17}$$

where  $y = (y_1, \ldots, y_N)'$  is the  $N \times T$  matrix of the observed data,  $f = (f_1, \ldots, f_r)'$  is the  $r \times T$  matrix of factors, and  $\varepsilon$  is the  $N \times T$  disturbances matrix. This alternative representation is convenient for our exposition along this paper.

Typically, the *exact factor model* assumes that idiosyncratic disturbances are mutually uncorrelated for all i = 1, ..., N and t = 1, ..., T. Unlike the exact factor model, we assume an *approximate factor structure* by allowing some serial and cross-sectional correlation among the idiosyncratic components, see Bai and Ng (2002). In approximate factor settings, the consistency and asymptotic normality of the estimators when both N and T go to infinity have been recently shown by Bai (2003), Bai and Ng (2002) and Doz et al. (2012).

The common factor,  $\Lambda f_t$ , in large panels is estimated by using the method of principal components. Consider an arbitrary number of factors r ( $r < min\{N, T\}$ ), such that rankf = r and rank $\Lambda = r$ , where  $f = (f_1, \ldots, f_r)'$  is the  $r \times T$  matrix of factors. The PCA estimates of  $\Lambda$  and  $f_t$  are obtained by solving the optimization problem

$$V = \min_{\Lambda, f} (NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} (y_{it} - \lambda_i f_t)^2$$
(18)

subject to the normalization of either  $\Lambda'\Lambda/N = I_r$  with r(r+1)/2 restrictions or  $ff'/T = I_r$ with r(r-1)/2 restrictions. We used the notation  $\lambda_i$  as the  $i^{th}$  row of  $\Lambda$  for  $i = 1, \ldots, N$ . The optimization problem is identical to maximizing tr(f(y'y)f') where  $y = (y_1, \ldots, y_N)'$ is the  $N \times T$  matrix of the observed data. Here tr() denotes the trace operator. Let Qthe  $r \times r$  diagonal matrix containing the r largest eigenvalues of sample covariance matrix  $S = \frac{1}{T} \sum_{t=1}^{T} y_t y'_t$ . The estimated factor matrix denoted by  $\hat{f}^{PCA}$ , is  $\sqrt{T}$  times the eigenvector, P, associated with Q. Hence,  $\hat{\Lambda}^{PCA} = y \hat{f}^{PCA'}/T$  is the corresponding matrix of factor loadings estimates.

The solution to the above minimization problem is not unique, even though the sum of squared residuals V is unique, see Bai and Ng (2002). An equivalent solution is explored

by Doz and Reichlin (2011) as follows

$$\hat{f}_{t}^{PCA} = Q^{-1/2} P' y_{t}$$

$$\hat{\Lambda}^{PCA} = P Q^{1/2}$$
(19)

Recent literature has shown that the principal components estimator of the common factors provides consistent estimates under the strong factor assumptions, see Bai and Ng (2000), Bai and Ng (2002) and Onatski (2012).

## **B** Decomposing the vulnerability index



Figure 16: Decomposing risks

# C Resilience indicators



#### Figure 17: Resilience indicator

## D Indicator data for the DSA Index

Red highlighted areas signal that an indicator has breached a predetermined threshold. The country names are masked in this exercise as some of the data are not available publicly (Figure 18). To construct the Count Index based on judgment, we take the variables for which the IMF-WB have provided thresholds, and count the number of indices that breach the threshold as a share of total number of indices. Countries that breach many thresholds should have a higher likelihood of default compared to those that do not.

## **E** The high-frequency indicator and sources of risk

Data updates are an important aspect of the vulnerability analysis. The core of the analysis relies on low-frequency data for each of the indices. However, this leads to risks being missed during a year. As an example, economic perceptions are largely influenced by news. Exchange rate devaluations or large changes in interest rate spreads will reflect investor perceptions when the political landscape changes, financial market stresses appear, or when hidden debts come to light during the course of a year. Attempts to stitch together high frequency indicators to the DVA should yield fruitful results. An attempt not described here connects high frequency indicators, semi-annual forecasts, and judgments to the annual risk indices described above.

One way to do this is to use nowcasting techniques (a popular approach by Ghysels et al. (2007) that is applied here as a mixed data sampling technique (MIDAS) that is described in Appendix F). This simple approach uses high-frequency data (e.g., CDS) and maps it onto low-frequency data (e.g., default probability calculations). The high-frequency data used for the nowcast include monthly country-specific variables such as EMBI spreads, exchange rates, CDS on five-year government bonds, real money supply (M3), consumer price indices, and international reserves. Those variables are easy to update every month for most EM countries.

Although, high frequency indicators are good predictors of risk, they do not differentiate between risks emanating from country-specific or global factors. Specifically, CDS premia as well EMBI spreads can rise because of uncertainty related to domestic political risk (e.g. tax or pension reforms), however it could also reflect declining commodity prices that can compromise future debt repayments by commodity exporters. In addition, financial sectors often purchase government securities from different countries and bundle them to create new asset classes to be sold to investors (e.g. FTSE World Government Bond Index). This creates an interlinked system of risks which we refer to as a global risk factor (i.e. a common factor common to all countries), meaning that a defaulted country might alter investor perceptions about a whole group of countries (e.g. Emerging Markets, EMs), which will have consequences for other countries.

Next we describe the decoupling of idiosyncratic and global contributions to the subjective default probability measure. The seminal papers of Geweke (1989), Stock and Watson (1989) and, Bai and Ng (2002) have placed the dynamic factor model (DFM) as the predominant framework for research on macroeconomic common factors (e.g. business cycles). This framework allows us to study large panels of time series through few common factors, especially, when the data series are strongly collinear.

#### Dynamic factor models in analyzing hight frequency data

Consider the approximate factor model (16) for each country c = 1, ..., C, with N indicating the number of variables for each country c, then, the approximate equation can be written as follows:

$$y_t^c = \Lambda^c f_t^c + \varepsilon_t^c \tag{20}$$

where  $y_t^c$  is a  $N \times 1$  vector that stacks all the monthly variables related to country c,  $\Lambda^c$  is a  $N \times 1$  vector such that there is only one common factor by country, and  $\varepsilon_t^c \sim NID(0, I)$ .

	Country 2	£ vrtnuoD	6 (ninuo)	ς λι‡unoϽ	ց Հւքսոօշ	ζ γτέπυοΣ	8 YrfnuoD	6 ƙitunoʻ	01 (ninuo)	11 Yıtınıo)	21 VrinoD	country 13	41 Vitruo	ST (JJUNO)		δ1 γιτη 18 Υ	61 yıtnuoD	02 (ntruo)	۲۲ Country ک	22 (Vitano)	Country 23	42 (ntruo)	SZ (ntruno)		82 VrinuoD	62 (ninuo)	
Exposure	1 2	m	4	S	9	2	∞	6	8	11	12	13	14	5	9	17 18	19	20	21	2	23	24	25 2	6 27		8 29	
External debt risk		1								1										1				1			
CAB as a %GDP < - $5\%^2$	3.0 -11.	4.	5 -4.6	-1.7	-3.2	-0.5	0.9	-0.3	-6.6	-1.7	-5.4 -1	10.5 -	2.6 -{	3.7 -3	.0 -17	.9 -5.2	-1.6	2.5	-6.4	-4.1	-5.7	-2.2 1	0.9 -8		-9.6	3.5	
Net portfolio flows (%GDP)	IA NA	ö	0 -6.5	- 0-	۶N ۲A	1.1	0.6	-2.2	-0.2	٩N	÷ 0.0	0.3	8.5 -f	5.8	0. 0	7 -29.4	8 0.3	-0.1	٩V	0.5	AN	-6.4	0.0	1 9	0-	8 NA	
External debt as a %GDP (>60%)	8.5 2.4	.35	.2 37.0	91.5	3 73.4	32.5	14.0	38.4	41.1	68.8	0.0	13.5 7	1.0 #N	I/A 10.	5.1 49.	5 46.0	7/N# (	64.8	83.3	27.3	6.69	49.6 3	5.3 80	.1 53.	2 104	0 46.5	
External private non-guaranteed debt as a % GDP (> 20%) 2.	1.8 0.1	ö	0 7.2	36.	7 17.7	17.3	4.1	7.6	0.0	17.9	0.0	1.7 1	3.2 6	36 0.	1.3	4 6.3	9.1	17.1	39.0	2.4	30.3	17.5 1	1.0 2.	6 23.	0 52.	2 11.9	
External gross financing needs as a %GDP (>15%)	9.0 0.6	~	4 17.1	2.0	32.7	7.5	9.2	8.4	0.0	14.4	8.7 2	3.7 1	0.2 #N	I/A 0.	.0 171	.6 0.0	1/N#	0.0	34.9	6.3	10.6	0.0	0.0 32	.5 25.	4 0.0	0.0	
ST external debt (> 5% share of external debt)	9.9 49.(	0.0	2 37.3	8.5	24.3	7.7	60.2	1.0	6.8	9.6	1.9 2	1.9 1	1.3 21	5.9 4.	5 7.	5 12.6	5 5.0	7.6	7.6	1.1	3.5	28.3 4	1.9 20	.9 26.	0 14.	7 15.4	
	1		1				1	2		2	2		2	2	2		2	1				1	1	1		1	
Public debt risk																											
Total Public Debt as a %GDP (>60%)	1.0 27.0	64.	.1 57.6	58.5	9 40.1	74.0	36.2	45.2	108.8	71.0	52.7 4	4.0 11	4.5 N	1A 20	153	.6 0.0	٩N	50.6	46.8	67.9	62.6	53.3 4	1.9 71	.3 28.	2 72.	3 58.2	
External Public Debt as a %GDP (>30%)	6.5 1.0	ž	A 39.2	47.8	3 38.6	5.3	ΝA	1.8	ΝA	NA	NA 3	4.7 6	3.8 N	A 8.	8 49.	5 31.3	NA NA	1.7	40.0	21.4	77.0	AN	1.1 48	.0 11.	0 50.	AN 6	
Budget Balance as a %GDP (<-4%)	1.8 -13.	7 -4.	5 -7.9	-5.5	1.5	-8.7	-3.2	-7.4	-12.5	-2.8	-5.0	4.1	3.2 -ŝ	1.2 -6	.4 -9.	6 -3.5	-2.8	-3.1	-2.0	-4.5	-1.3	-3.8	2.6 -6	1-1-	-2	-6.3	
Primary Balance as a %GDP (<-3%)	.7 -13.	4 -2.	.1 -4.5	-3.6	3.5	-1.8	-2.3	-5.8	-3.4	-0.2	-2.4	2.9	.5 -	0.2 -5	.3	0.1-0	0.0	-0.8	-1.0	-0.2	2.2	-0.5	1.5 -3	0	1.1	-4.2	
ST debt as a %GDP (>4%)	9.5 NA	6	.9 12.7	2.7	ΝA	7.4	5.3	0.4	15.6	ΝA	2.8	1 6.5	AA 1	3.2 6.	4 32.	1 7.2	5.9	6.0	0.9	16.1	22.1	8.0 2	2.4 2.	1 2.3	4.2	3.1	
Public debt held by non-residents as %Debt (>45%)	3.7 NA	ž	A 21.1	ΝA	ΝA	9.4	1.4	14.3	18.1	37.4	38.7 3	4.7	A A	N N	۸ ۸	AN VA	٩N	13.7	40.0	21.4	37.8	NA NA	1.4 47	.1 19.	4 46.	1 NA	
Dollar denominated debt (%of total external debt)	IA NA	Ą.	.9 67.1	48.5	5.99.1	9.66	38.3	93.2	52.2	100.0	0.0 4	3.4 1C	0.0 8	5.0 10	3.0 98.	0 13.5	9 77.4	100.0	95.5	79.4	13.2	9.66	VA 21	.4 69.	5 74.	5 11.2	
EMBI Global Diversified Yield Spread (> 600 bpts.)	IA NA	610	2.11.5	9 305.	2 NA	263.5	145.6	621.2	426.5	NA 4	175.9 25	56.9 32	9.1 36	7.5 22	9.9 437	.4 176.	2 256.0	0 134.0	٨A	344.3	NA 2	66.7	N N	A 305	0 564	2 NA	
USD Redemptions (cumulative 2018 & 2019 & 2020 %GDP[2016]) N	IA NA	÷.	0 8.1	6.6	0.0	0.3	ΝA	1.5	0.5	3.0	0.0	A N	i.3 6	2.0	0 14.	8 NA	0.6	٨A	NA	1.1	6.5	1.5	NA 1	2	5.0	0.5	
:																											
Resilience																											
Reserve import coverage 6	.7 24.5	5	.7 6.0	5.9	1.3	20.3	18.7	2.4	3.8	3.6	2.5	3.9 6	5.2 7	.8	22.22.	3 6.2	4.9	6.2	3.9	5.3	5.5	5.7	9.2 3.	3.0	1 3.4	. 2.3	
Hirfendahl Concentration Index (2015)	.2 0.1	0	2 0.1	0.1	0.2	0.1	0.1	0.2	0.0	0.3	0.0	0.1	0.1	.1	ľ.	A 0.1	0.6	0.1	0.4	0.1	AA	0.1	0.1	2 0.0	0.0	0.1	
Government effectiveness (>0.5)	.2 -0.9	-i-	.2 0.5	-0.£	-1.4	0.5	-1.6	-0.2	-1.2	0.3	-1.0 (	<u></u> 2.2 ι	)C	.8 -1	.a	5 -0.6	-0.1	-0.5	-0.6	-0.7	0.2	0.6	1.1	ο, Ο	0.0	-1.4	
Exchange rate regime $(1 = free float)^4$ (	0.0 0.0	<u>.</u>	0 1.0	0.0	1.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	.0	0.0	0.0	1.0	1.0	NA	٩N	ΝA	1.0	0.0	0 1.(	0.0	AN NA	
Sovereign wealth fund as a %GDP	H.8 NA	4	8 NA	٩N	18.7	18.7	ΝA	٨A	ΝA	NA	NA F	A A V	A N	IA 85	N 0'	AN VA	0.6	٩N	٩N	٩N	AN	AN	N AV	A N	۸N	NA	
Tier 1: Capital to risk weighted assets (> 8)	IA NA	14.	.3 15.2	AN N	٩N	13.7	11.2	19.2	AN	ΝA	NA 1	0.5 1	A A V	N N	A 14.	8 NA	13.2	٩N	٩N	13.1	AN	14.5 1	4.5 N	A 13.	1 9.0	٩N	
Non-performing loans to gross loans (< 5%)	3.2 NA	ž	A NA	5.4	12.5	3.6	1.7	٩N	AN	1.9	NA	2.8	V V	A N	Ň V	AN VA	2.1	1.5	1.0	8.4	ΝA	2.8	3.1 N	A 2.8	54.	5 NA	
Liquidy-asset ratio 1.	3.0 NA	ž	A NA	32.5	1 27.4	14.6	22.5	٨A	٨A	22.4	NA 2	1.3 1	Z A	N N	ν ν	AN VA	32.0	22.5	21.2	48.3	AN	15.1 1	9.9 N	A 49.	6 53.	9 O.O	
Liquid assets to ST liabilities (> 100%)	IA NA	37.	.4 54.0	NA	NA	236.3	47.6	35.2	٨A	٨A	NA 3	4.0 1	2 47	A N	A 29.	7 NA	42.4	٩N	٨N	95.0	AN	30.9 3	0.7 N	A 66.	9 92.	1 NA	
0il exporter (>30 % of GDP = 1)	.0 1.0	ਜ -	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0 C	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
Real effective exchange rate (+ = appreciation_)		-20	0.1 -11.5	9 2.2	-8.0	6.0	-4.6	1.2	-2.8	-0.2	6.8	1.3	5.4 C	:0 -2	5.7 -0.	6 0.1	-11.(	-4.5	-0.8	-1.9	-2.0	- 6.9	4.4 -6	4 -2.	3.1.1	-0.3	
Systemically important countries <sup>5</sup>	0.0 0.0	0.0	0.0 0.0	0.0	0.0	1.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	1.0	0.0	0 1.0	0.0	0.0	
1/ CE estimates (Macro Poverty Outlook)																											
5/ Defined as large trade shares in global trade; Size of GDP; Econor	nic or mo	netar	y policy ;	zone; P	rice se	tters; Ru	egional	trade	inkage	; Finan	cial link,	ages (e	g. ban	cing set	ctor))												

4/ 1 = free float; 2 = crawling peg
5/ Economic or monetary policy zone; Price setters; Regional trade linkages; Financial linkages (e.g. banking sector)
http://www.imf.org/external/np/pp/eng/2013/111513.pdf

Figure 18: Heat map indicators

Therefore, our country factor is estimated following (19) and denoted as  $f_t^{c,PCA}$  for each country c.

On the other hand, consider the vector  $y_t^G$  that stacks all country data in addition to global variables in a single vector, such that  $y_t^G = \{y_t^1, \ldots, y_t^C, p_t^{oil}, i_t^{US}, VIX_t\}'$ .

Similarly, the approximate equation can be written as follows:

$$y_t^G = \Lambda^G f_t^G + \varepsilon_t^G \tag{21}$$

where  $\Lambda^G$  is a  $(NC + 3) \times 1$  matrix, such that there is only one global common factor, and its denoted by  $f_t^{G,PCA}$ .

Notice that the following condition is required:  $E(f_t^G f_t^c) = 0$  for each c = 1, ..., C. It implies that we need to clean up the collinear relationships between country and global factors. The Kalman filter is used to ensure that the factors are orthogonal Schwaab et al. (2017).

Denote  $y_t^f = \{f_t^{G,PCA}, f_t^{1,PCA}, \dots, f_t^{C,PCA}\}$  as the vector containing all normalized factors estimated previously. The state-space representation can be expressed as follows:

$$y_t^f = f_t + \varepsilon_t^f \qquad \varepsilon_t^f \sim NID(0, \Omega)$$
  

$$f_t = f_{t-1} + \eta_t \qquad \eta_t \sim IID(0, I_\eta)$$
(22)

where  $\Omega$  accounts for all the commonalities among country and global factors, and  $f_t$  represents the set of truly global and country factors. Notice that, the estimation procedure is determined by the state smoothing recursion method in Durbin and Koopman (2012). Now  $f_t^{c,MLE}$  and  $f_t^{G,MLE}$  are the new set of estimated global and country common factors in (22).

Figure 19 depicts the global and some country specific factors using this methodology. The global common factor points out two events of risk. The first peak captures the 2007-2008 financial crisis, while the second peak captures the sharp decline in commodity prices, primarily oil, during the 2015-2016 period. The global common factor reveals potentially a third build up of risk starting with a trend in 2018. An increase in the global factor does not necessarily imply that countries will default at peak - but simply illustrates the build up of global risks. Country risks might be amplified with an increase in the global factor. On the other hand, country common factors, as expected, are heterogeneous since these only capture country specific movements.

Finally, we model the impact of country and global factors on the probability of default related to perceptions. Specifically, we assume that country factors affect the premium directly, but may become more pronounced with an increase in the global factor risk. Thus, for some countries the global factor might amplify/subdue the effect of country-specific factors on default. To isolate this amplification effect, we estimate a Discrete Threshold Regression (DTR) model for each country c, as follows:

$$SDP_{c,t} = \begin{cases} \alpha_c + \beta_c f_t^{c,MLE} + \epsilon_{c,t}, & \text{if } f_t^{G,MLE} < s^c \\ \alpha_c + \gamma_c f_t^{c,MLE} + \epsilon_{c,t}, & \text{if } f_t^{G,MLE} > s^c \end{cases}$$
(23)

for each country c = 1, ..., C. The parameter  $s^c$  is the threshold that indicates the global conditioning regime for each country. Likewise,  $SDP_{c,t}$  is the default probability estimated

#### Figure 19: Common factors



for each country *c*. The threshold parameter captures the idea of a global risk effect, where weaker global perceptions might lead to an increase in risk within countries. If there is a zero effect of the global factor then  $\gamma_c$  would be equal or lower than  $\beta_c$ . Therefore, if  $\gamma_c > \beta_c$ , global conditions amplify the country effects.

## F Confusion matrix on the test sample

The confusion matrix shows that the DVA indicators signal correctly 83% (71/86) for the logit case, 91% (78/86) for the Lasso and 86% (74/86) the random forest regressions.

	DVA Logit				Debt Logi	t			Spread Log	it			Classic Log	it	
	No Crisis	Crisis			No Crisis	Crisis			No Crisis	Crisis			No Crisis	Crisis	
No Signal	11	15	26	No Signal	2	1	3	No Signal	9	16	25	No Signal	2	3	5
Signal	30	71	101	Signal	39	85	124	Signal	32	70	102	Signal	39	83	122
-	41	86		-	41	86		-	41	86		-	41	86	
	DVA Lasso	)			Debt Lasso	)			Spread Lass	so			Classic Lass	60	
	No Crisis	Crisis			No Crisis	Crisis			No Crisis	Crisis			No Crisis	Crisis	
No Signal	3	8	11	No Signal	10	12	22	No Signal	13	14	27	No Signal	5	6	11
Signal	38	78	116	Signal	31	74	105	Signal	28	72	100	Signal	36	80	116
	41	86			41	86			41	86			41	86	
DV	A random fo	orest		Deb	ot random f	orest		Spre	ad random	forest		Clas	sic random	forest	
	No Crisis	Crisis			No Crisis	Crisis		-	No Crisis	Crisis			No Crisis	Crisis	
No Signal	17	12	29	No Signal	16	19	35	No Signal	14	20	34	No Signal	11	10	21
Signal	24	74	98	Signal	25	67	92	Signal	27	66	93	Signal	30	76	106
	41	86			41	86			41	86			41	86	