INTERNATIONAL DEVELOPMENT ASSOCIATION
INTERNATIONAL MONETARY FUND

REVIEW OF THE DEBT SUSTAINABILITY FRAMEWORK FOR LOW INCOME COUNTRIES: PROPOSED REFORMS

August 22, 2017

Prepared by the staffs of the World Bank Group* and the International Monetary Fund

Approved by Jan Walliser (WB) and Seán Nolan (IMF)

* Prepared by a team from the IMF's Strategy, Policy, and Review Department, led by César Serra, consisting of Karina Garcia, Amr Hosny, Samuel LaRussa, Racha Moussa, Keiichi Nakatani, Joyce Saito, Nelson Sobrinho, Suchanan Tambunlertchai, Tim Willems, and Yi Xiong, under the overall guidance of Said Backhache, Mark Flanagan; and a team from the World Bank's Macroeconomic and Fiscal Management Global Practice, led by Leonardo Hernández, consisting of Jane Bogoev and Aart Kraay, under the guidance of Paloma Anos Casero, Abha Prasad, and Mark Thomas. The paper received contributions from William Battaile, Sebastian Essl, Tobias Haque, Heiko Hesse, Kareem Ismail, Henry Mooney, Anne Oeking, Zeinab Partow, Isabel Rial, David Stephan, Genevieve Verdier, and Signe Zeikate; and benefitted from discussions with an IMF inter-departmental advisory group.

EXECUTIVE SUMMARY

The Debt Sustainability Framework for Low-income Countries (LIC DSF) has been the cornerstone of assessments of risks to debt sustainability in LICs. The framework classifies countries based on their assessed debt-carrying capacity, estimates threshold levels for selected debt burden indicators, evaluates baseline projections and stress test scenarios relative to these thresholds, and then combines indicative rules and staff judgment to assign risk ratings of external debt distress.

The framework has demonstrated its operational value since the last review was conducted in 2012, but there are areas where new features can be introduced to enhance its performance in assessing risks. Against the backdrop of the evolving nature of risks facing LICs, both staff analysis and stakeholder feedback suggest gaps in the framework to be addressed. Complexity and lack of transparency have also been highlighted as causes for concern.

This paper proposes a set of reforms to enhance the value of the LIC DSF for all users. In developing these reforms, staff has been guided by two over-arching principles: a) the core architecture of the DSF—model-based results complemented by judgment—remains appropriate; and b) reforms should ensure that the DSF maintains an appropriate balance by providing countries with early warnings of potential debt distress without unnecessarily constraining their borrowing for development.

The specific proposed reforms would adapt the framework to the evolving circumstances facing LICs, seeking to make it more comprehensive, more transparent, and yet simpler and easier to use:

- The reformed DSF would move away from relying exclusively on the CPIA as the measure for assessing debt-carrying capacity; a composite measure would be used instead, based on an expanded set of economic variables that includes the CPIA.

- Adjustments to the methodology would improve the accuracy of the framework, including by better identifying debt distress episodes and enhancing the statistical accuracy of predicting debt distress.

- New tools would be introduced to shed light on the plausibility of important features of the baseline macroeconomic projections facilitating closer scrutiny of the assumptions underpinning the projections. The tools would, for example, help assess the realism of any projected fiscal adjustment and of the projected impact of public investment and fiscal adjustment on growth.

- Tailored scenario stress tests would be introduced to better evaluate specific risks of particular relevance for some countries—risks stemming from such factors as natural disasters, volatile export prices, market-financing shocks, and contingent liability exposures.

- The number of debt indicators, thresholds, and standardized stress tests would be reduced. One debt indicator—the present value of external debt to revenues would no longer be used. The number of thresholds would be reduced from 24 to 12; the number of standardized stress tests would also be halved.

- The summary assessment of risks would be expected to discuss several topics, supported by new tools. The analysis of domestic debt vulnerabilities would be sharpened; risks from shifts in market-financing conditions would be discussed; and the robustness of the debt position of countries at moderate risk of debt distress would be analyzed.

- The Staff Guidance Note that would accompany the new framework will give expanded attention to ensuring even-handed application of staff judgment in applying the framework—including in such areas as the treatment of marginal and/or transitory breaches of thresholds.

This paper also examines the appropriateness of the discount rate used in the LIC DSF. Staffs propose to keep the current five percent discount rate unchanged, revisiting this decision in future DSF reviews, and to maintain a unified discount rate for the LIC DSF, the DLP, the NCBP, and the grant element calculator.

Back-testing suggests that the proposed reforms would have a balanced impact on the framework, while significantly improving its predictive performance. Simulations of mechanical risk signals suggest these would remain unchanged for most countries, with slightly more mechanical upgrades than downgrades for remaining cases. The new framework should significantly reduce the rate of false alarms (incorrectly predicting the occurrence of debt distress) and moderately decrease the rate of missed crises (failing to predict the onset of debt distress).

The proposed reforms to the DSF would enable it to provide the basis for a richer and more informed policy discussion between Bank and Fund staffs and country authorities. Country-specific features (e.g., international reserve coverage) play an expanded role as inputs to the DSF and as drivers of the tailored scenario stress tests; the DSF's outputs would include a much richer set of information on debt developments and debt vulnerabilities, better informing the discussion of policy options and trade-offs.

The implementation of the new framework would be appropriately paced. Guidance materials, templates, and training materials would be updated, and training activities conducted ahead of implementation expected by the second half of 2018.

TABLE OF CONTENTS

BOXES

FIGURES

TABLES

APPENDIX

REFERENCES

ANNEXES

I. BACKGROUND

1. **The framework for debt sustainability analysis for LICs (LIC DSF) was introduced in 2005.** The LIC DSF uses the Country Policy and Institutional Assessment (CPIA, an index produced annually by Bank staff) to classify countries based on their debt-carrying capacity, and a set of threshold levels for selected debt burden indicators linked to elevated risk of debt distress that are statistically estimated for varying levels of debt-carrying capacity. Baseline macroeconomic projections and stress test scenarios are evaluated relative to these thresholds, and used in conjunction with staff judgment to assign risk ratings of external debt distress. The value of these risk ratings is, of course, dependent on the realism of medium- to long-term macroeconomic projections, which are critical inputs to the DSF produced by country teams (see Appendix I for an overview of the framework).

2. **The DSF has remained the cornerstone of assessments of risks to debt sustainability in LICs, with important operational implications for multilateral institutions and other creditors.** The DSF risk assessment of a country plays an important role in the Fund's debt limits policy (DLP) and the Bank's non-concessional borrowing policy (NCBP), and also influences the grant-loan funding mix of IDA support for the country. Many MDBs (including the AfDB and IaDB) have also linked their lending policies to the DSF risk assessment. Debtors often use it to inform borrowing decisions, albeit with many expressing concerns that the framework is limiting their opportunities to borrow for development. CSOs have shown increased interest in the LIC DSF and have advocated its use as a tool to help determine how to finance the large development needs of LICs in the context of the efforts to achieve the Sustainable Development Goals (SDGs).

3. **The DSF has been reviewed on three occasions, most recently in 2012 (see Box 1 for a summary of the Board discussion).** New features were added to the DSF in the 2012 review, to incorporate further country-specific information into the determination of risk ratings and to pay greater attention to domestic debt vulnerabilities in the risk assessment. Some other potential additions—such as embedding macro-linkages in stress tests and enhancing the modeling of the investment-growth nexus—were explored but ultimately left for future work.

4. **Since 2005, the economic environment in which many LICs operate has changed significantly, resulting in potentially important gaps in the DSF.** The evolving financing landscape and the challenges emerging from a weaker environment have reshaped the nature of risks facing LICs. Financing sources that have increased in importance include borrowing from non-Paris Club creditors, from domestic markets, and from international bond markets—most notably for "frontier", LICs that have attracted foreign portfolio investors. As a result, LICs are increasingly exposed to a wider set of vulnerabilities, including from market volatility.[1] The DSF, in its current form, lacks tools to assess these market-related risks.

[1]See Public Debt Vulnerabilities in LICs: The Evolving Landscape (IMF-WB, 2015).

Box 1. Board Discussion on the 2012 Review of the Debt Sustainability Framework for Low-Income Countries

Executive Directors of WB and IMF Boards agreed to introduce new features into the LIC DSF to improve its assessment of the risk of debt distress. These features were added to incorporate further country-specific information in the determination of risk ratings. The probability approach was introduced to complement the assessment of the risk of external debt distress in "borderline" cases, where risk ratings arguably fell in between categories. The introduction of remittances-augmented debt thresholds was meant to capture better the enhanced repayment capacity of countries receiving large remittances. The review also proposed strengthening the analysis of total public debt by introducing public debt benchmarks to support the assessment of the overall risk of debt distress. It also called for an in-depth analysis to determine the extent of vulnerabilities emanating from private external debt.

Directors saw merit in developing stress tests that included macro-linkages and looking more closely into the links between investment and growth:

- **Stress tests**. There was strong support for exploring stress tests that include macro-linkages. Many Directors wanted the inclusion of such stress tests to be experimental and only optional for country teams. One Director thought such stress tests should be the norm where data allows.

- **Investment-growth nexus.** There was broad support for exploring further the links between investment and growth. Several Directors raised the concern that since too little is known about the effect of investment on growth, it is best to be prudent and keep the framework conservative.

Directors unanimously supported simplifying the DSF template. Some Directors cautioned though that it should not come at the expense of adequately capturing risks.

5. **In the external consultation process that has supported the review, stakeholders have called for gaps and complexities in the framework to be addressed.**[2] In particular, external stakeholders have emphasized the need to: (i) ensure the DSF remains balanced in its treatment of risks versus borrowing opportunities; (ii) strengthen the ability of the DSF to provide adequate early warning of potential stress, including by better incorporating relevant country-specific information into the framework; (iii) ensure that the framework is appropriately aligned with the evolving nature of risks facing LICs, including by enhancing the assessment of domestic debt, better accounting for liquidity risks, and improving the assessment of contingent liabilities; (iv) introduce tools to illuminate the realism of macro projections, particularly the investment-growth nexus; (v) expand the stress testing framework to assess key risk scenarios, such as the impact of natural disasters; and (vi) provide more differentiation in characterizing the extent of vulnerabilities for countries assessed to be at moderate risk of debt distress. While calling for these

[2] The technical work of the review has been informed by a broad external consultation process. This has included dialogue with authorities from developing countries and staffs of multilateral development banks (including at the 2016 (Lusaka) and 2017 (Vienna) DMF Stakeholders' Forums, the 2016 African Caucus in Cotonou, at the 2016 and 2017 Spring Meetings – including bilateral discussions with Pacific Islands Governors and Ministers – and the 2017 Multilateral Development Bank Meeting on Debt Issues in Washington, DC), and with members of the Paris Club. Staff have also sought feedback from civil society organizations (including through an open web-based consultation and different events during the 2016 Annual Meetings).

issues to be tackled, stakeholders also emphasized the need to simplify where possible what is already a complex framework.

6. **The review has identified a package of reforms that would strengthen the DSF while providing for some simplification of the toolkit.** Key reforms include: (i) strengthening the assessment of countries' debt-carrying capacity by drawing on an expanded set of country-specific and global factors, instead of relying exclusively on the CPIA score; (ii) introducing tools to facilitate closer scrutiny of baseline macroeconomic projections; (iii) recalibrating standardized stress tests to better reflect the actual scale of shocks, while adding tailored scenario stress tests to better evaluate risks stemming from natural disasters, volatile export prices, market-financing shocks, and contingent liability exposures; (iv) providing a richer characterization of debt vulnerabilities (including an enhanced assessment of domestic debt vulnerabilities, and new tools for assessing vulnerabilities to shifts in market financing conditions and for better discriminating across countries within the moderate risk category); and (v) enhancing the guidance for a more even-handed application of judgment. At the same time, the framework would be simplified by substantially reducing the number of debt indicators, thresholds, and standardized stress tests.

7. **The remainder of this paper is organized as follows.** The next section identifies key areas for reform, informed by an assessment of the LIC DSF performance (subsection A) and its methodology (subsection B). The third section proposes reforms to the framework to align it to the evolving nature of risks facing LICs and to address the identified weaknesses, with the aim to make it more comprehensive, more transparent, and simpler. The fourth section assesses the effects of the proposed changes by presenting back-testing results. The fifth section summarizes how the process of engagement around producing a DSA would change and how the Bank and the Fund would support the implementation process. The sixth section examines the appropriate discount rate to use in the DSF. The final part of the paper concludes, summarizing the main takeaways and recommendations from the review, and raises issues for Board discussion.

II. IDENTIFYING KEY AREAS FOR REFORM

A review of the performance of the DSF and its methodology points to areas where there is scope for improving the framework. On performance, the quality of medium-term debt projections, the calibration of stress test shocks and thresholds, and the integration of country-specific information into the framework could all be enhanced. On methodology, the identification of external debt distress episodes, the specification of the statistical model for predicting debt distress, and the procedure to derive debt thresholds are all areas for improvements. Further, the design of stress tests could be strengthened (by incorporating macroeconomic linkages among key variables) while tools to assess market-related risks could usefully be added.

A. Assessing the LIC DSF Performance

8. **To help map out potentially useful reforms to the DSF, it is important to understand first where it is performing well and not so well**. The evaluation follows the overall structure of the framework; that is how it utilizes inputs and filters them to arrive at a risk rating. Since the framework starts by comparing baseline projections and stress test scenarios to thresholds, two immediate questions concern how good a job baseline projections and stress tests are doing in characterizing the possible evolution of debt; and how well the thresholds have been set. The evaluation then covers the performance of rules being used to assign risk ratings, including the use of judgment. It concludes by stepping back to bring in case studies and the additional perspective that they can bring.

Performance in Characterizing the Likely Evolution of Debt: *some weaknesses*

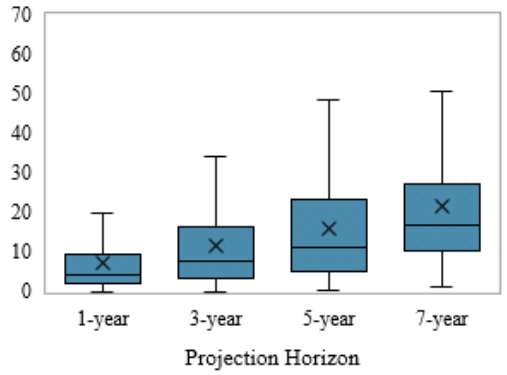
9. **While debt projections have been accurate in the near-term, they have tended to underestimate outcomes in the medium-term**. Over one and two-year projection horizons, the median (absolute) unexpected change was about 5 percent of GDP for both external and total public debt, with the median deviation more than doubling for both external and total public debt over a five-year horizon (Figure 1.a–b). About 40 percent of DSAs produced during 2007– 10 contained unexpected changes in debt over a five-year horizon in excess of 15 percentage points of GDP (Figure 1.c);[3] for small states, this occurred in more than half of the DSAs. Deviations on this scale are particularly common in DSAs for countries that are currently at high risk of external debt distress.

[3] Debt service indicators have relatively much smaller unexpected changes. The share of DSAs with unexpected changes above 15 percentage points (in absolute value) ranges between 0 and 8 percent.

Figure 1. Performance of the DSF in Anticipating Debt Developments

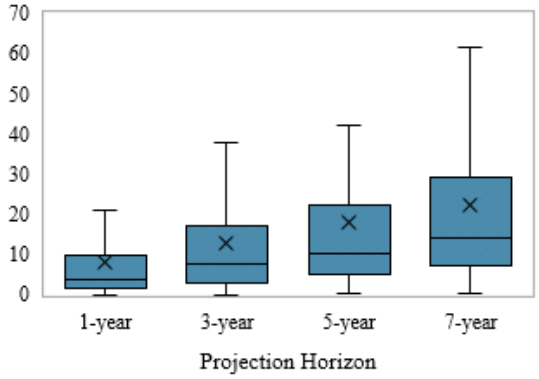
a. Public Debt Deviations 1/

(Percent of GDP; absolute value of actual – projection)



b. External Debt Deviations 1/

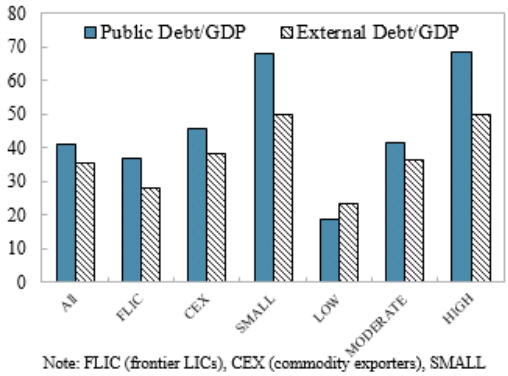
(Percent of GDP; absolute value of actual – projection)



1/ The chart shows the min/max (after removing outliers), average (marked as X), median (line within blue bars), and interquartile range (blue bars) for unexpected changes in the public debt to GDP and external debt to GDP ratios over different projection horizons.

c. Share of DSAs with Sizable Deviations

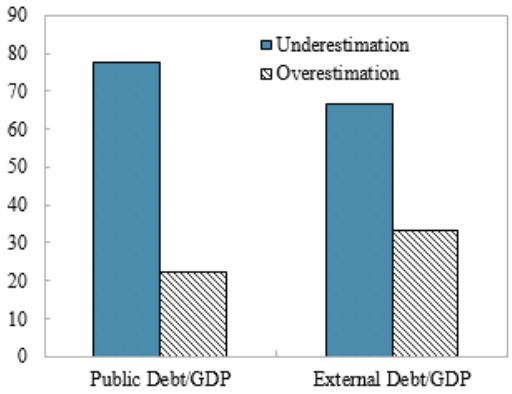
(5-year projection horizon)



Note: FLIC (frontier LICs), CEX (commodity exporters), SMALL (small states).

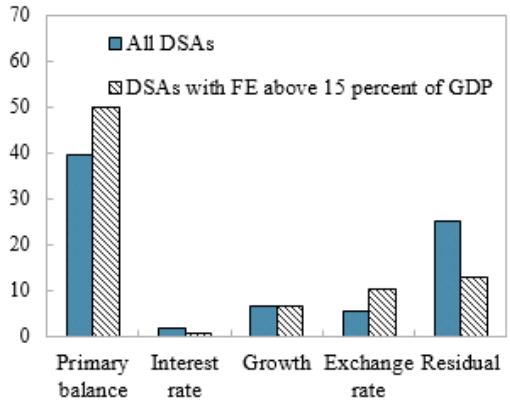
d. Direction of Bias

(Percent of DSAs with sizable deviations)



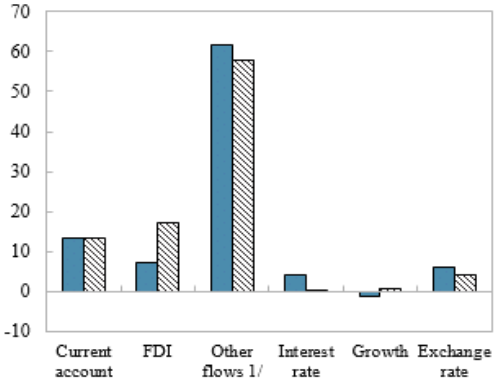
e. Decomposition of Unexpected Changes in Public Debt

(Percent, median contribution)



f. Decomposition of Unexpected Changes in External Debt

(Percent, median contribution)



1/ BOP, includes residual.

Source: Fund staff calculations.

10. **Sizable medium-term underestimation of debt outcomes are more common, and seem to be related to particular shocks:**

- **The majority of DSAs with sizable medium-term deviations involve an underestimation of debt outcomes.** Among external and public DSAs with sizable unexpected changes in debt (larger than 15 percentage points of GDP), about 70 and 80 percent, respectively, underestimate debt outcomes (Figure 1. d).[4] This point is clearly captured in the projected dynamics of the median public debt to GDP ratio across LICs for each of the last five DSA vintages (Figure 2); it is noteworthy that projections quickly revert to a downward debt path even after significant short-term upward shifts.

- **Unexpected changes in debt ratios have been primarily driven by fiscal deviations and balance of payments (BOP) shocks.** Figure 1.e–f show the decomposition of unexpected changes over a 5-year horizon for the total public debt and external debt-to-GDP ratios, for all DSAs produced during 2007–10 and for those DSAs with sizable positive deviations.

  - For total public debt, the analysis signals that, in addition to the primary deficit, unanticipated positive residuals have contributed to sizable deviations (suggesting that the materialization of contingent liabilities has played a role). Growth and exchange rate shocks have been less important.

  - For external debt, the DSA methodology does not allow to disentangle between the contribution of private and public sector elements in the balance of payments. However, the evidence points to a much larger role of financial account flows driving unexpected changes in total external debt. This suggests larger than expected access to sources of finance (a manifestation of the changes in the financing landscape).

11. **The standardized stress tests used in the DSF to characterize risk scenarios have generally not been well aligned with actual drivers of debt changes:**

- **A close look at the impact of stress tests on total public and external debt projections shows a mismatch with the impact observed in the data** (Figure 3. a). The simulated impact on debt from shocks to the primary balance is modest compared to the actual impact of primary balance deviations. One problem has been the way the shock is calibrated. Post-shock values in stress tests are set at one standard deviation below historical averages, with the key assumption being that historical averages are good predictors for the next 2–3 years. However, this assumption did not hold for the primary balance: actual primary surpluses (deficits) turned out to be consistently lower (larger) than historical averages during 2008–15 (see Annex I).[5] In contrast, the simulated impact on the debt-to-GDP ratio from shocks to the exchange rate is large relative to the actual impact of realized deviations (and with the magnitude of the

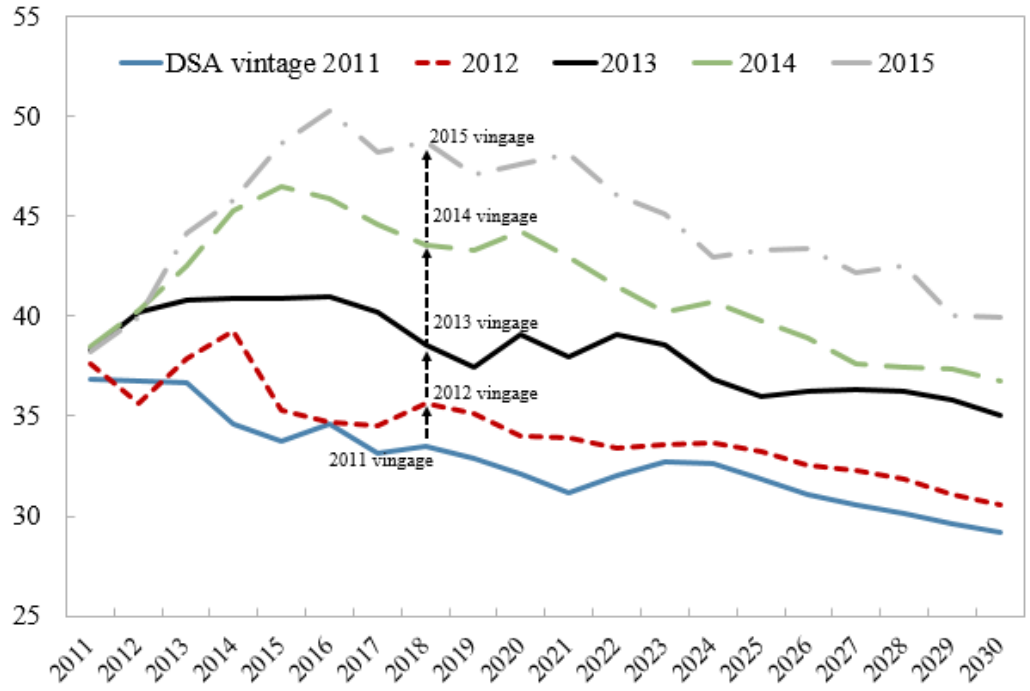
[4]Among DSAs with sizable unexpected changes in debt, the share of those underestimating debt outcomes rises from about 55 percent to more than 80 percent when moving from the 1- to the 7-year projection horizon.

[5]This relationship between historical and actual primary balances continues to hold when the sample is expanded to cover the longer period of 2000–15.

exchange rate shock correctly calibrated -see Annex I-, the larger simulated impact on the debt ratio suggests that key macro interaction factors are not captured in the stress testing framework).

- **Some stress tests have rarely played a role in the determination of the external risk rating**. Simulated shocks to exports, nominal depreciation, other flows (transfers and FDI), and a combination shock have played the main role in signaling risks of external debt distress. The external financing alternative scenario and the real GDP growth and GDP deflator stress tests are rarely identified as sources of risk (Figure 3. b–f).

Figure 2. Long-term Optimism Bias
(Public debt/GDP, median)



Source: Fund staff calculations.

12. **Some important risks facing sub-groups of LICs have been only partially covered by the standardized stress tests used in the LIC DSF, or by other means:**

- **Risks from natural disasters have been narrowly covered in LIC DSAs**. Less than 50 percent of countries severely exposed to natural disasters (as identified in IMF (2016)) used a customized scenario to assess their associated impact on debt vulnerabilities.

- **Few country teams have sought to directly assess the risks to public debt positions stemming from the realization of contingent liabilities**. Recent examples of such analysis

have included Nicaragua (2015), which examined the transfer of private external debt related to the oil collaboration with Venezuela onto the government balance sheet, and Solomon Islands (2016), which looked at potential liabilities from power purchase agreements.

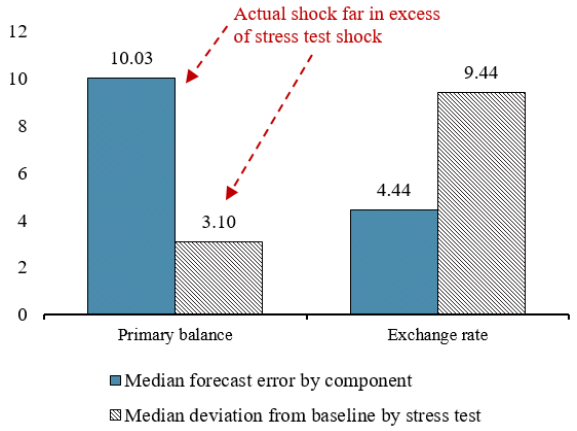
- **Commodity price shocks have also not been directly assessed in relevant cases**. Over the last two years, 8 of 11 LICs that experienced a downgrade in their risk rating were commodity exporters; the generic export shock scenario in the DSF did not capture the full impact on export and fiscal revenues from price shocks in these cases.

Figure 3. Adequacy of Stress Tests

A2: external financing; B1: real GDP growth; B2: exports; B3: GDP deflator; B4: other flows, B5: combination; and B6: depreciation stress tests.

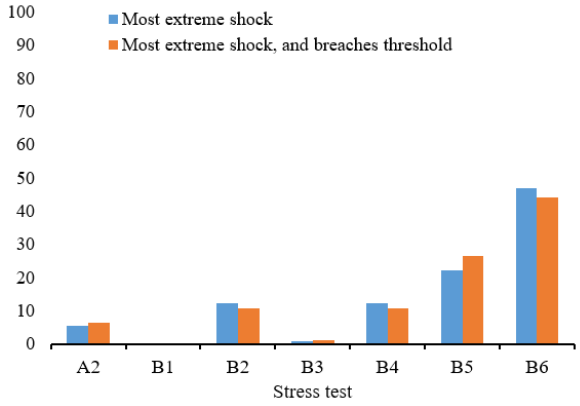
a. Sources of Unexpected Changes in Debt: Modeled vs. Actual

(Percent, Total Public Debt-to-GDP)



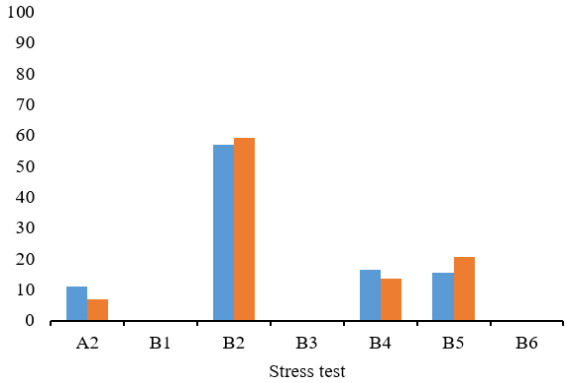
b. Freq. Most Extreme Shock: PV Debt-to-GDP

(Percent of Total DSAs)



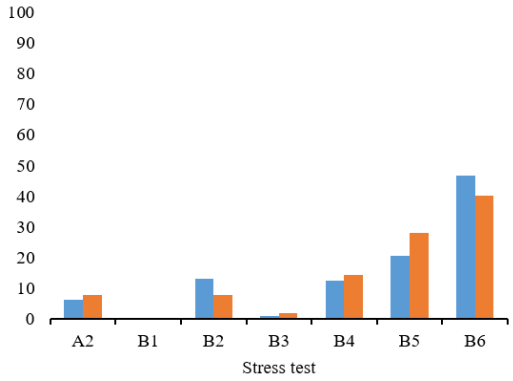
c. Freq. Most Extreme Shock: PV Debt-to-Exp.

(Percent of Total DSAs)



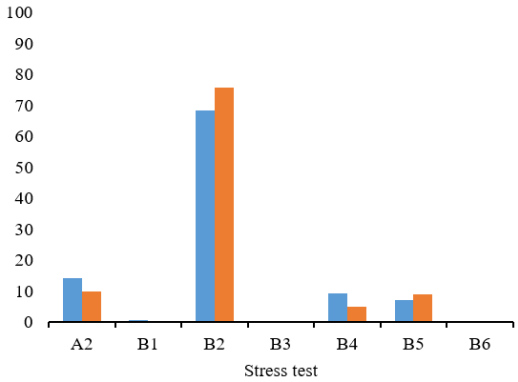
d. Freq. Most Extreme Shock: PV Debt-to-Rev.

(Percent of Total DSAs)



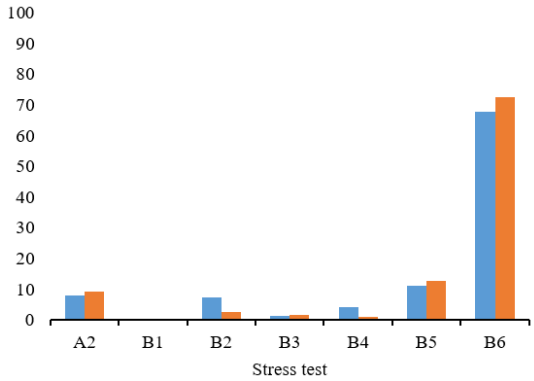
e. Freq. Most Extreme Shock: Debt Serv.-to-Exp.

(Percent of Total DSAs)



f. Freq. Most Extreme Shock: Debt Serv.-to-Rev.

(Percent of Total DSAs)



Sources: LIC DSAs during 2008–15, and Fund staff calculations.

Evaluation of debt indicators and thresholds: ***some not playing a role***

13. **Some debt burden indicators have played a minor role in signaling the risk of debt distress:**

- **Debt service indicators have rarely determined risk signals of external debt distress.** Only in 8 percent of all DSAs, high risk signals of external debt distress have been exclusively informed by breaches of debt service thresholds (Figure 4. a). Moreover, the distance of debt service indicators to their thresholds in the run up to recent debt distress events remained significant (Figure 4. b). In some cases, even after sizable fiscal deviations for several years and rising market risks, debt service thresholds were not breached. These facts suggest that either debt service measures contain little information in terms of signaling debt distress (which is counter-intuitive) or that the relevant thresholds have been poorly estimated.

- **Among debt stock indicators, the PV of external debt to fiscal revenues appears to be redundant in determining external risk signals.** Across all LIC DSAs since the inception of the framework, there was only one case in which this threshold was breached under the baseline without breaches in other debt stock indicators (Figure 4.c). In contrast, the PV of external debt to exports has signaled high risk of debt distress in about 80 percent of all cases. This could reflect the information content of these debt stock indicators and/or the appropriateness of the existing debt stock thresholds.

14. **As regards the new features added in the 2012 review (Box 1), the framework has a high bar for applying these, and the added value of making use of them has been limited:**

- **Remittances-augmented thresholds.** To inform the determination of risk ratings, 13 out of 69 countries have incorporated remittances as an integral part of the framework. Of those, five saw upgrades in their risk ratings. Extending the application of the remittances-augmented debt thresholds to all countries for which remittances data are available since 2014 (not only those that qualify under DSF rules) would not have improved any risk rating, and would have produced 14 downgrades. This outcome is the combined result of the way in which remittances-augmented thresholds were derived in 2012 and the highly-skewed distribution of remittances.[6]

[6] The highly-skewed distribution of remittances implies that only countries receiving sizable remittances would benefit from the estimation of remittances-augmented thresholds based on the full sample of countries. Not surprisingly, countries that have benefited from the use of remittances-augmented thresholds received remittances flows as percentage of GDP and exports way in excess of the eligibility cutoffs (averages of 30 percent of GDP and 150 percent of exports).

- **Probability approach** (which involves moving beyond discrete thresholds set by the CPIA-based country classification to consider all country-specific information in the model). Since the new features became operational in 2014, only nine out of 69 countries have used the probability approach to inform their risk ratings (leading to one final risk rating improvement). Had it been more broadly applied to all DSAs produced since 2012 it would have upgraded the risk ratings in 17 DSAs in the sample, while there would have been a downgrade of risk ratings in 28 of them.

Performance of rules to set risk ratings: ***too many false alarms and scope for deeper risk analysis***

15. **Overall, the mechanical framework underlying the DSF produces a high rate of false alarms.**[7] While debt burden indicators perform poorly in terms of their individual capacity to predict debt distress, the joint use of all debt burden indicators—with a breach by any one sufficing to signal high risk—significantly improves the DSF's overall predictive performance. Still, the results imply that the existing framework yields a large rate of false alarms (about 50 percent of the cases in which debt distress is not observed) and a more modest rate of missed crises (about 20 percent of the cases in which debt distress is observed).[8] Updating the sample period and re-estimating debt thresholds based on the existing methodology would imply a significant tightening of debt thresholds, leading to a much larger rate of false alarms (about 60 percent) and a very low rate of missed crises (about 8 percent).

16. **The use of staff judgment for the final determination of external risk ratings has generally been beneficial, but has not been evenly applied.** Since the inception of the DSF, judgment has been applied to override the mechanical risk signal in about 25 percent of all cases, leading to an upgrade about 80 percent of the time. Staff judgment has reduced the rate of false alarms vis-à-vis the mechanical application of the framework by about 10 percentage points. Country teams have typically used judgment to override risk signals driven by marginal and/or temporary breaches, although this does not happen in all cases. For instance, the mechanical risk signal has been overridden in 48 percent of all marginal breaches (measured as a 5 percent deviation from the threshold).

17. **Deeper attempts to disentangle risks have not been a strong feature:**

- **The moderate risk category conceals diverse debt vulnerabilities in a large number of countries**. Figure 4.d shows that the number of years in which debt indicators are projected

[7]In case of individual debt indicators, missed crises are defined as cases in which debt distress is observed but the associated debt threshold was not breached, while false alarms are defined as cases in which debt distress is not observed but the associated debt threshold was breached. Type I and II errors are missed crises and false alarms as a share of all cases in which debt distress is observed and non-observed, respectively.

[8]The rates of false alarms and missed crises are measured using historical data during the sample period (i.e., 1970–2014). It is not straightforward to assess how policy responses to high risk ratings may have affected the number of false alarms, but any such impact may be only present in the few last years of the sample period in which such ratings were produced (2008-2014).

to breach thresholds under stress tests varies widely, with some countries having only temporary breaches while others showing protracted ones. A similar result is found for the size of breaches, where some moderate risk countries have large breaches while others have minor ones.

- **The new feature added during the 2012 review—to provide for a deeper assessment of total public debt and an overall risk rating—has been only sporadically used**. Since 2014, DSAs documented 26 countries with significant vulnerabilities emanating from total public debt.[9] However, only 10 countries reported an overall risk rating, and only six of those provided an in-depth discussion of domestic public debt vulnerabilities. Most DSAs did not provide a discussion of the extent of vulnerabilities stemming from rollover risks, the increasing participation of non-residents in domestic local-currency bond markets, or from the structure of domestic public debt.

18. **Finally, the DSF timeframe for considering risks appears excessive.** Breaches of debt burden indicators are concentrated during the early years of the projection horizon (Figure 4. e-f). The fact that debt forecasts become subject to increasing underestimation errors as the projection horizon is extended is likely playing a role in this result, but even with better forecasts these would still be subject to significant uncertainty.

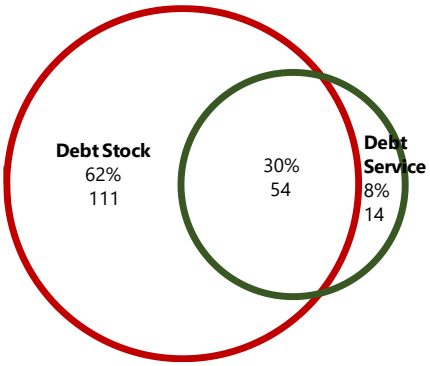
Case Studies

19. **Select case studies show that the DSF has been successful in signaling impending debt difficulties in several country cases, but its effectiveness has been limited in others due to a variety of factors.** In several cases (e.g., Chad, Mongolia) the framework signaled impending debt difficulties one year or more in advance, capturing the impact of expansionary fiscal policy and the external debt service outlook. However, other case studies highlight areas in which the framework would benefit from improvements, such as securing stronger baseline debt projections (e.g., Maldives, Djibouti); enhancing the analysis of key risks in stress tests (e.g., Sri Lanka, Samoa); and expanding the assessment of broader risks (e.g., Ghana, Central African Republic). See Annex II for more details.

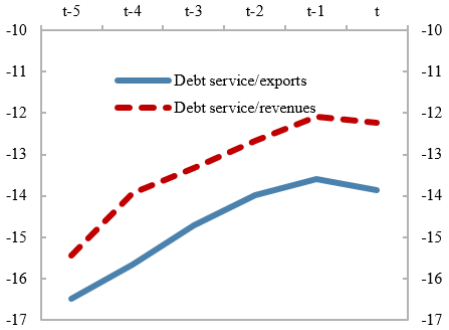
[9]Such vulnerabilities are signaled in relation to the relevant benchmark of the PV of total public debt to GDP ratio.

Figure 4. Performance of the DSF in Setting Risk Ratings

a. Breaches Under Baseline by Indicator Type
(All DSAs)

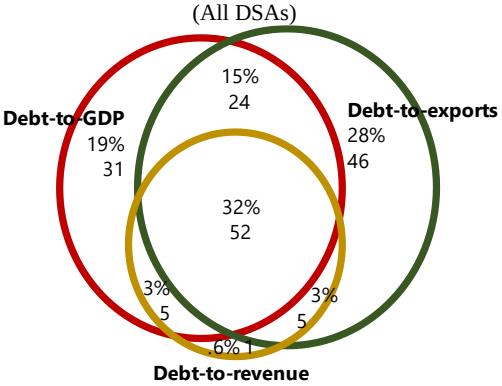


b. Distance of Debt Service Indicators to Thresholds
(Percent, Median)

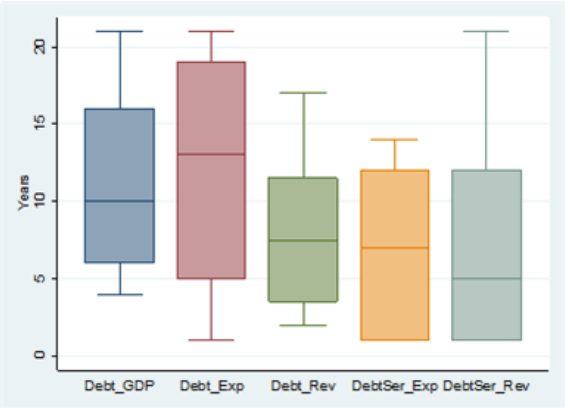


Note: t denotes the start of a debt distress event. Y-axis measures the distance of debt service indicators to thresholds.

c. Breaches Under Baseline by Stock Indicator
(All DSAs)

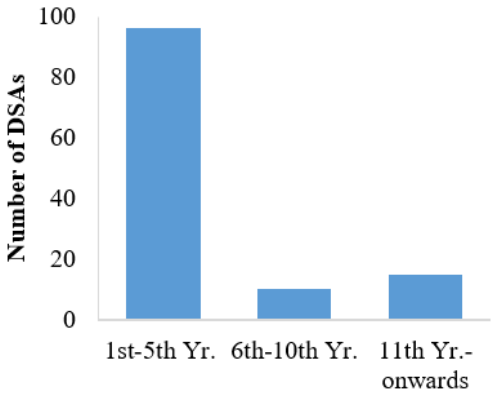


d. Dispersion within Moderate Risk Countries
(number of years breaching thresholds)

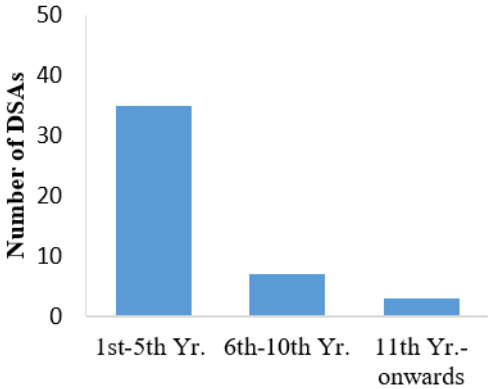


Note: The chart shows the min/max, median and interquartile range for each debt burden indicator.

e. Timing of Breaches of Debt Thresholds: PV Debt-to-Exports



f. Timing of Breaches of Debt Thresholds: Debt Service-to-Revenue



Source: Fund staff calculations.

B. Assessing the Model Specification and Methodology Underpinning the LIC DSF

20. **A closer look at the technical approach underlying the DSF framework also reveals potential areas for reform**. Underpinning the DSF is a core model of debt distress, estimated by looking at actual debt distress episodes. Thus, two immediate questions arise concerning how well debt distress episodes have been defined—an incorrect definition would reduce performance—and how well specified and accurate the model is. Since the framework then works off thresholds derived from the model (i.e., comparing baseline forecasts and stress test scenarios to them), another key question concerns the methodology used to derive these thresholds. Finally, it is also worthwhile to explore limitations in the existing methodology for stress tests.

21. **The accuracy of the estimated core statistical model depends on correctly identifying external debt distress episodes, but the existing approach has limitations** (see also Annex III):

- **The existing approach focuses on identifying severe debt distress episodes by requiring distress signals to be observed for at least three consecutive years**.[10] Distress events such as pre-emptive restructurings that take less than three years are ignored.

- **The identification of distress signals based on IMF disbursements could be improved**. The emergence of debt difficulties is captured at the time cumulative IMF disbursements exceed 50 percent of quota, providing either a false alarm—sooner or later many GRA arrangements breach such a cutoff without necessarily involving debt difficulties—or a missed or late call.

- **The identification of distress signals based on Paris Club restructurings warrants revision**. A debt restructuring deal with creditors, currently used to signal the start of the distress episode, typically comes well after the onset of debt problems. In addition, HIPC treatments are over-represented given that multi-round debt restructurings between the decision and completion points are considered as separate signals.

22. **The specification of the core statistical model, which relates the probability of external debt distress to debt burden indicators and country characteristics, could be enhanced.** While the existing model has done a reasonable job of accounting for historical patterns of external debt distress, it makes use of few variables and, when the sample period is updated (to include seven years of extra data) and the revisited identification procedure of debt distress episodes used, its performance deteriorates sharply.

[10]The DSF relies on three distress signals to identify conditions under which a country is experiencing external debt difficulties: i) cumulative IMF disbursements from the General Resource Account (GRA)—under Stand-By Arrangements (SBA) and Extended Fund Facilities (EFF)—exceeding 50 percent of the member's quota; ii) restructuring of claims held by Paris Club creditors; and iii) accumulation of arrears on external PPG debt in excess of 5 percent of the outstanding stock of external PPG debt.

23. **There are limitations in the methodology used to derive debt thresholds:**

- **Thresholds are derived individually for each of the five debt indicators without regard to the information contained in other debt indicators.** This is at odds with the DSF's aggregation rule, which looks at all debt indicators and produces a risk signal if there is a threshold breach for any indicator. This may have introduced a downward bias in the estimation of debt thresholds (see Berg and others, 2014).

- **The framework would benefit from a transparent policy choice for the tolerance of missed crises and false alarms.** Instead of direct specification of weights on the rate of missed crises (type I error) and false alarms (type II error) for deriving debt thresholds, the present framework averages the results over a range of weights (from 50 to 75 percent) for type I error. This in the view of some observers is too conservative and may have introduced a downward bias in the estimation of debt thresholds (see Berg and others, 2014).

- **Country-specific information could be better integrated into the methodology for classifying countries.** Countries are classified into weak, medium or strong performers based only on the strength of their policies and institutions, measured by the CPIA score, without regard to other factors that can be shown to be of value in predicting external debt distress.

24. **The stress testing framework can be improved along two dimensions:**

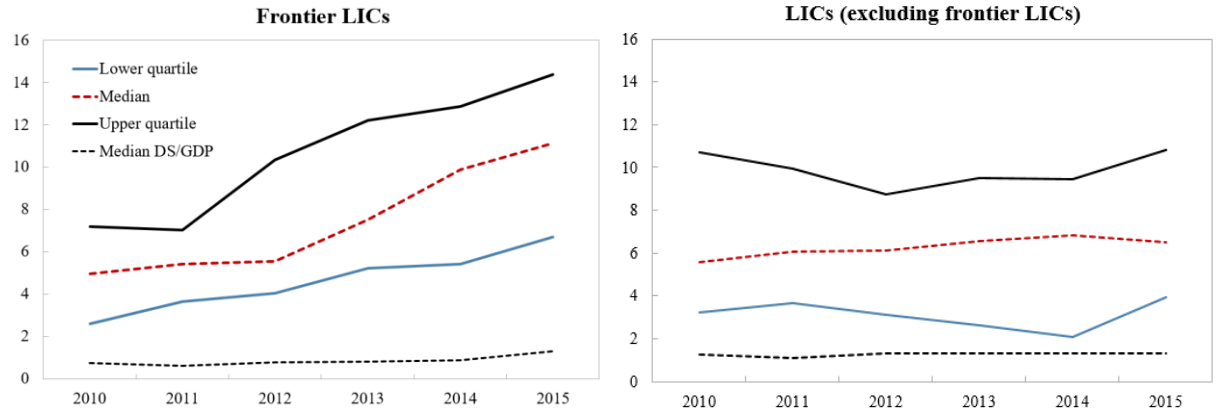
- **The shocks that are modeled could seek to capture macro-linkages among key variables.** The lack of macro-linkages seems to be a key reason for the miscalibration of the impact of the exchange rate shock (paragraph 11). In any event, lack of macro-linkages contradicts recent experiences with LICs and evidence in the literature (see Aisen and Hauner, 2008, for instance, on the primary balance and borrowing cost relationship).

- **A common set of shocks should be applied to external and total public debt**. For example, the primary balance shock has been considered only in public DSAs but not in external DSAs, although deteriorations in fiscal positions can affect external public debt and debt service through the channels of both higher public external borrowings and higher borrowing costs. Similarly, some simulated shocks to external public debt did not inform the assessment of total public debt vulnerabilities.

25. **Finally, the DSF does not provide explicit tools for assessing risks associated with market financing**. In recent years, several frontier LICs have received significant amounts of market and other forms of non-concessional financing, thereby transforming their public debt profiles and the nature of risks to which they are exposed (e.g., risks from shorter maturities and a more diverse creditor base). As a result, countries with market access are increasingly exposed to more frequent spikes in financing needs, particularly due to issuances of bonds with bullet

amortization.[11] In addition, external debt service indicators alone may no longer be sufficient to capture the extent of liquidity risks (Figure 5). Indeed, the experience in some countries (e.g., Ghana and Sri Lanka) showcases the challenges that country teams had in accounting for rollover risks in the face of growing fiscal deficits and the resulting high gross financing needs (GFNs), and often had to rely on their own analysis outside of the LIC DSF.

Figure 5. Gross Financing Needs and External Debt Service
(Percent of GDP)



Source: Fund staff calculations.

[11]Since 2005, 14 LICs have issued 29 Eurobonds worth US$20 billion. Of the 29 issuances, 25 or US$17 billion had bullet payments. Principal repayments of sovereign external debt (in percent of exports of goods and services) by frontier LICs are projected to exceed those of the 17 largest EMs over the next five years (see IMF-WB, 2015).

III. PROPOSED REFORMS

26. **The assessment of the DSF performance and methodology has pointed to areas where the framework should be strengthened.** The approach taken in refining the framework has been guided by two overarching principles. First, there is a need to preserve the core architecture of the DSF—model-based results complemented by judgment (Figure 6). A model helps ensure consistent and transparent application of the framework grounded in empirical observations. At the same time, a model cannot capture the intricacies of every country, so a role for judgment is imperative. Second, there is a need to ensure continued balance in the application of the DSF. The framework should continue providing countries early warnings of debt distress—as these episodes are very costly when they occur—but without generating multiple false alarms (which would unnecessarily advise against additional borrowing).

27. **The specific reforms proposed would correct weaknesses and gaps in the framework, adapt it to the evolving circumstances facing LICs, make it more transparent, and yet simpler to use**. Specifically, the limitations noted with the DSF methodology would be addressed by updating the approach to identify debt distress episodes and the specification of the core model. These changes would then allow the integration of key country-specific information into the classification and risk assessment of countries. This would also serve to make the framework more transparent and complete. New realism tools—to understand the key assumptions underpinning macroeconomic projections—would both support the implementation of the new approach to classify countries and help address the problems identified with debt projections. Greater simplicity would be achieved by dropping the various features that have proved, in practice, to add little value to the assessment of risks. Finally, added tools would help with a deeper assessment of risks and firmer application of judgment. Figure 6 and Table 1 summarize the proposed reforms.

28. **The rest of this section is organized as follows:** Subsection A covers adaptations to the core debt distress model; subsection B discusses revisions to the approach to classify countries; subsection C discusses new macro realism tools; subsection D covers modifications to the framework to determine mechanical risk signals; subsection E covers the analysis of other risk factors; subsection F covers enhanced guidance for the application of judgment; and subsection G summarizes the elements needed to draw conclusions on the risk of debt distress.

Figure 6. Structure of the Reformed LIC DSF

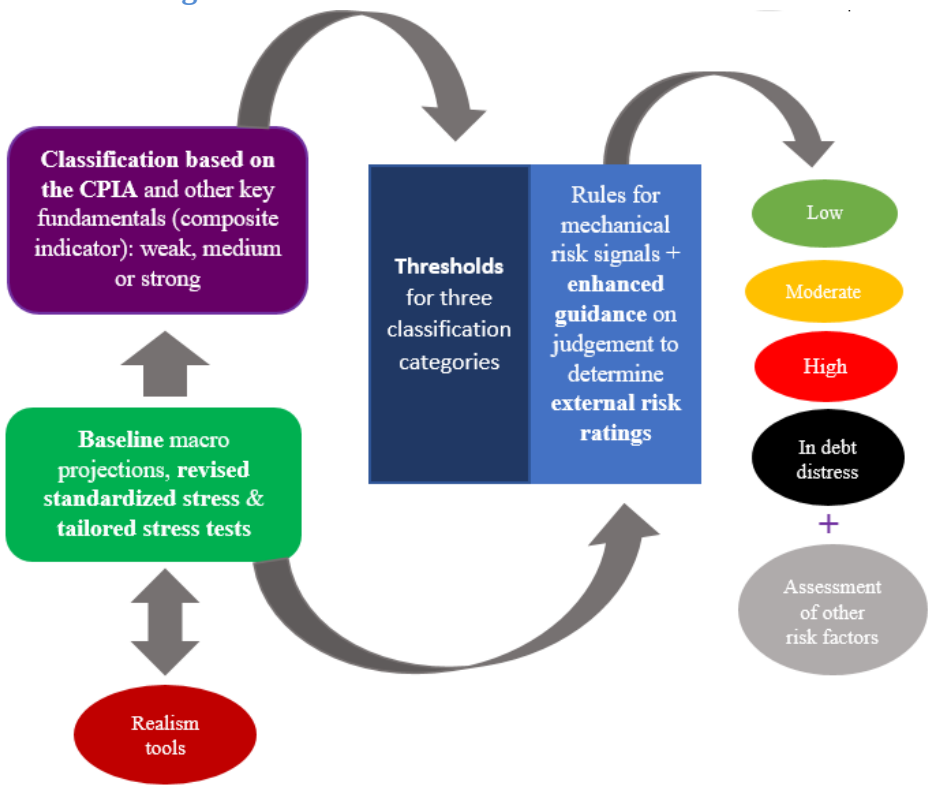


Table 1. Proposed Reforms to the LIC DSF

	Existing DSF	Reformed DSF
Core debt distress model	• Identifies *only* severe debt distress episodes. • Few country-specific explanatory variables.	• Enhanced methodology to identify *all* debt distress episodes. • Expanded specification including key country-specific fundamentals to improve predictive capacity.
Country classification	• Relies exclusively on the CPIA. • Backward-looking classification.	• Based on a composite measure covering the CPIA, growth, reserve coverage, remittances, and world growth. • Incorporate forward-looking elements (enhancing engagement with country authorities).
Realism tools		• To support stronger baseline projections and implementation of new classification (e.g., realism of projected fiscal adjustment and the investment-growth nexus).
Debt indicators and thresholds	• Complex: five debt indicators and 24 thresholds. • Thresholds are derived individually without regard to the information of other debt indicators to predict debt distress (introducing conservative bias).	• Significant simplification: four debt indicators and 12 thresholds. • Thresholds are derived jointly in line with the DSF's aggregation rule (eliminating a source of conservative bias).
Standardized stress tests	• 16 stress tests, lack of macro-linkages, and non-common testing across the external and public DSA.	• 8 common re-calibrated and re-designed stress tests across the external and public DSA, incorporating macro-linkages.
Tailored stress tests		• To better evaluate scenario risks of relevance for some countries (e.g., natural disasters).
Assessment of other potential risk factors	• Tools to assess: Domestic debt vulnerabilities.	• Tools to assess: Domestic debt vulnerabilities. Market-financing pressures. Diversity of debt vulnerabilities in countries rated as moderate risk.
Enhanced guidance for the application of judgment		• On marginal/transitory breaches. • On severe domestic-debt vulnerabilities and exposure to external market-financing pressures, among other factors.

A. Strengthening the Statistical Model for Predicting Debt Distress

29. **The core engine of the DSF framework is the statistical model to predict debt distress.** The model identifies the key factors affecting the probability of debt distress, and the DSF uses the estimated model to derive debt thresholds consistent with different carrying capacities for debt. To be effective at predicting debt distress, the model must build off of the right identification of debt distress episodes, and incorporate the right explanatory factors.

30. **The estimation of the core model in this review has used a revised procedure for identifying external debt distress episodes to improve model performance.** The aim has been to eliminate episodes that do not represent debt distress while bringing in those that clearly do. Debt distress episodes would now also include distress periods that last only one or two years (except when the shorter episode is driven solely by the occurrence of external arrears).[12] The existing criteria used to identify distress episodes would also be redefined: (i) the identification of distress signals based on IMF disbursements would be refocused on large upfront financing disbursements; (ii) information from new debt restructuring databases would be incorporated; and (iii) a better treatment of the timing of debt restructurings would be added (see Annex III for details).

31. **To strengthen predictive capacity, the model specification has been expanded to incorporate additional explanatory variables** (see Table AIII.2 in Annex III)**.** In particular, the new model includes two important proxies of capacity to repay (international reserves scaled by imports, and remittances scaled by nominal GDP), as well as a proxy for global shocks (world growth).[13] These additional controls have been used extensively in the literature on sovereign debt crises and sovereign borrowing (e.g., Manasse and others (2003), Manasse and Roubini (2005), Gelos and others (2011), Catão and Milesi-Ferretti (2014)). To further improve its accuracy and applicability, the new model has been estimated based only on data from all LICs (the previous model also included data from MICs).

32. **Alternative variables and approaches were considered, but not incorporated as they did not improve predictive performance:**

- Other macro controls considered include real GDP per capita, the current account balance/GDP and foreign direct investment/GDP, a measure of the country risk premium, global interest rates, commodity prices, the terms of trade, and measures of natural disasters and conflicts.[14] None improved the predictive performance of the baseline specification.

[12]This helps rule out episodes associated with one-off and temporary occurrence of external arrears (including for technical reasons) that do not necessarily signal debt distress.

[13]See Box AIII.2 and AIII.3 (Annex III) on data sources and issues, including a discussion of reserve measurement in currency unions.

[14] Sub-indices of the CPIA could not be considered due to insufficient data.

- The role of trade openness—which has been shown to be relevant in predicting debt distress in emerging market economies—was explored empirically. However, trade openness seems to be less relevant for LICs when it is included alongside remittances in the statistical model, which partly reflects the high correlation between the two variables. The model including only remittances performs better.

- Consideration was given to estimating different models for specific country groupings (e.g., small states)—but this was not technically feasible given the limited number of debt distress episodes within smaller country groups. In any event, a specialized model is not needed for such cases, given the scope for exercise of judgment to complement the statistical model. Moreover, new tools (i.e., tailored scenario stress tests; see more below) would allow key vulnerabilities facing small states to be incorporated into the risk assessment.

B. Classifying Countries by their Debt-carrying Capacity

33. **A second core element of the framework concerns how a country's debt-carrying capacity is assessed.** This helps determine the thresholds against which debt projections are compared to derive mechanical risk signals (as discussed below, an aggregation rule and risk-weighting scheme are also technically necessary to set the thresholds). The framework at present utilizes just one explanatory factor in the core statistical model—the country policy and institutional assessment (the CPIA, produced every year by Bank staff)—to classify a country's debt-carrying capacity.

34. **The new framework would use the additional information in the updated core statistical model to move away from relying exclusively on the CPIA to classify countries** (Figure 7). To this end:

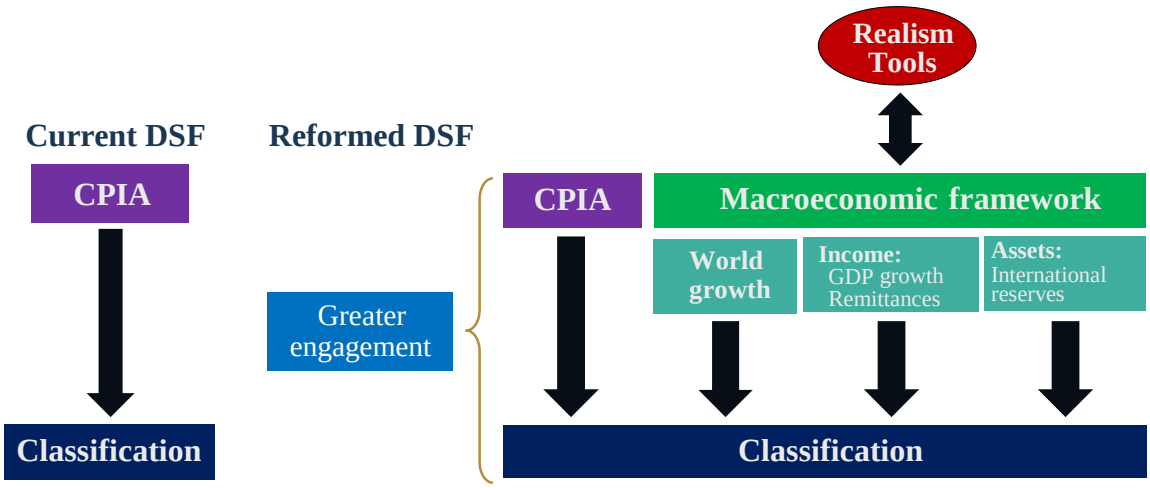
- **A composite indicator would be constructed, covering the CPIA, country growth, reserve coverage, remittances, and world growth** (with weights given by the estimated coefficients of these variables in the model). The cross-country distribution of the average composite indicator over a 10-year period (2005–14) can then be used to establish classification cutoffs. In this framework, a country's debt-carrying capacity would be assessed as *weak* if its composite indicator fell below the 25th percentile of this distribution, *medium* if it fell between the 25th and 75th percentiles, and *strong* if its composite indicator fell beyond the 75th percentile (see Annex III for details).

- **Countries would then be fit into this classification scheme based on the *country-specific* composite indicator.**[15] In this approach, forward-looking elements can be integrated into the classification of countries; it is proposed that the country-specific composite indicator be calculated based on the latest five years of historical data and the first five years of

[15] The determinants of the country-specific composite indicator would be shown in the DSA output to facilitate understanding of what is driving the classifications and where any changes have come from.

projections.[16] Mixing historical data and projections allows the framework to capture ongoing changes in the outlook for countries' fundamentals. Changes in a country's classification would continue to require two consecutive signals in which the country's composite indicator exceeds its classification cutoff, to mitigate concerns about undue volatility in applying the framework. Note that since the country-specific composite indicator would be anchored on a long-term average (10 years of data), this would attenuate concerns that countries' classifications could be influenced by cyclical considerations.

Figure 7. New Methodology for Classifying Countries Debt-carrying Capacity



35. **The proposed new classification methodology has significant advantages over relying exclusively on the CPIA score:**

- The new approach would give the overall framework greater predictive power. Ignoring the rich information brought in by the additional country-specific variables would misclassify countries and raise both missed crises and false alarms.

- Threshold effects—the fact that a small change in the CPIA can produce large changes in debt thresholds under the current framework—would be mitigated. The several components of the new composite indicator—whose co-movements are not perfectly synchronized—and the long-term averages used for its computation would make it less likely that a movement to another classification category would be due to a change in only one variable.

- The use of forward-looking information would allow country authorities to understand the relationship between their policy framework and their debt-carrying capacity, enhancing engagement. More generally, classification would be more transparent, as the additional

[16]Forecasts of the additional variables in the composite indicator are routinely produced in the WEO database and individual DSAs. Recognizing its slow-moving nature, the forecast for the CPIA rating would consist of its most recent value.

country-specific information is objective and readily available to country teams, country authorities, and other stakeholders.

36. **Staff considered an alternative to the debt threshold approach to classification, but does not see it as viable.** Conceptually, a debt threshold approach has some shortcomings compared to a probability threshold approach. The advantage of the latter is that it uses country-specific information more efficiently to predict debt distress (see Berg and others 2014), while avoiding the so-called threshold effects implied by grouping country-specific information into three classification categories.[17] On the other hand, conditional on a soundly-based classification of countries' debt-carrying capacity, a debt threshold approach has a number of operational advantages over a probability approach: (i) it rules out situations resulting from the strict application of the statistical model where debt thresholds would take extreme values in case of countries with very weak or very strong fundamentals (which would imply zero borrowing space for some LICs and implausibly-high debt limits for others); (ii) debt targets in the context of Fund-supported programs, the Fund's DLP and the Bank's NCBP can be more directly mapped to debt thresholds than to probability cutoffs; and (iii) debt thresholds are more intuitive and easier to interpret than some maximum level for the probability of debt distress.

C. Improving the Realism of Baseline Projections

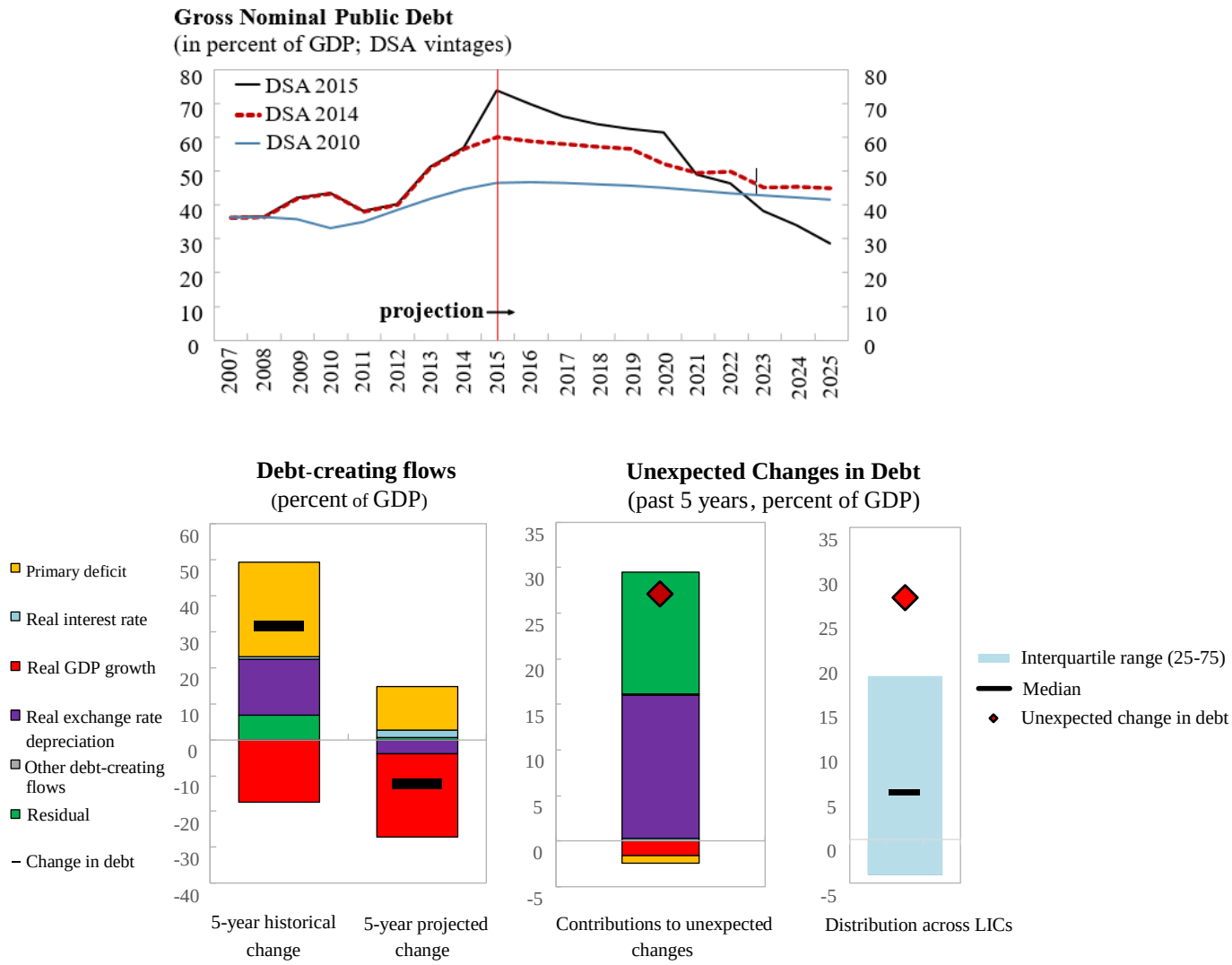
37. **To support stronger baseline debt forecasts, and implementation of the new classification methodology, the new DSF would include realism tools.** Such tools, which are already included in the DSF for market access countries, provide a point of comparison for forecasts, whether drawing on the country's own history, cross-country experience, or on relationships drawn from economic theory. They help to inform users of situations where important drivers of the macroeconomic baseline debt projections deviate markedly from experience (either with an optimistic or a pessimistic bias). They are not meant to be prescriptive (in the sense of requiring that the baseline scenario comply with comparator experience), but rather to highlight key assumptions underpinning the projections, thereby focusing discussion on these features of the macroeconomic framework and their realism.

38. **A first realism tool would present a decomposition of past and projected drivers of debt dynamics.** Specifically, DSF users would be shown the evolution of projections of external and public debt to GDP ratios over DSA vintages (one-year and five-years ago). It would provide several summary charts to help DSF users identify and scrutinize marked changes in historical and projected drivers of debt dynamics (Figure 8 shows an illustrative country). For example, a high past contribution of unexpected primary deficits would caution against projecting an excessive reliance on fiscal adjustment. This tool would also shed light on differences between

[17] As highlighted in the 2012 review, threshold effects occur when small changes in the predictors of debt distress lead to discrete jumps in debt thresholds. For instance, in the current framework, a weak-performer (say with a CPIA score of 3.24) would face a debt-to-GDP threshold of 30 percent, whereas a medium performer (say with a CPIA score of 3.26) would face a threshold of 40 percent (see IMF (2012), Appendix 1).

the historical and projected contributions of the current account and FDI flows to external debt, which may signal potential optimism or pessimism in projected reserve accumulation.

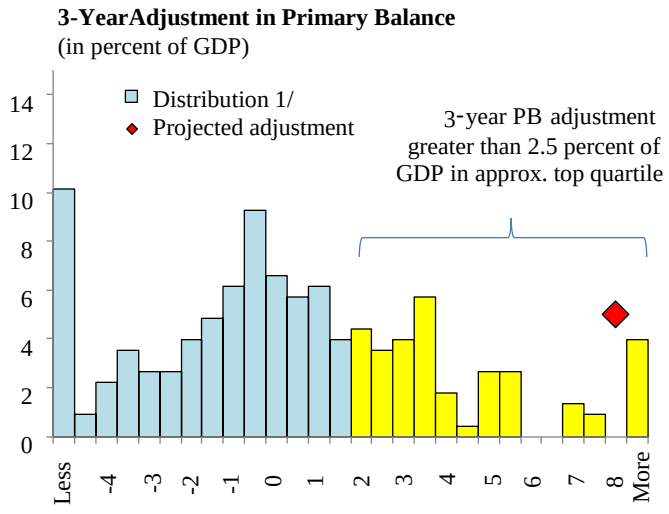
Figure 8. Drivers of Debt Dynamics–Baseline Scenario



Source: Fund staff calculations.

39. **A second tool would contrast projected fiscal adjustment with the historical experience of countries.** The comparison group focuses on a set of LICs requesting a Fund-supported program, as these countries generally have faced a need to adjust their fiscal positions.[18] The tool would present the distribution of observed headline primary fiscal adjustment over a three-year horizon, against which a country's projected primary fiscal adjustment would be compared. A fuller discussion of the credibility of the fiscal path would be called for if the projected primary fiscal adjustment over any three years during the projection horizon exceeds, say, 2½ percent of GDP, which is approximately equal to the top quartile of the distribution of observed primary fiscal adjustment (Figure 9 shows an illustrative case).[19]

Figure 9. Realism of Planned Fiscal Adjustment



Source: Fund staff calculations.
1/ Data cover Fund-supported programs in LICs (excluding emergency financing) from 1990 to 2013. Only the 3-year adjustment since the program request per country is included (horizontal axis). Percent of sample on vertical axis.

40. **A third tool would provide benchmarks for assessing growth assumptions in light of public investment dynamics or fiscal adjustment:**

[18]While the tool would ideally use cyclically-adjusted primary balances, these are difficult to compute for LICs given the high degree of uncertainty in estimating output gaps. At the same time, measuring fiscal adjustment as the change in headline primary balances across all LICs may include cases in which improved fiscal performance was driven by exogenous factors (e.g., coming on stream of natural resource projects).

[19] This issue will be addressed in the Staff Guidance Note.

- **Assessing the investment-growth nexus.** A formal assessment of the impact of public investment on growth is beyond the scope of the DSF.[20] The proposed tool would use a simple growth accounting framework to decompose projected growth rates into a: (i) contribution from the changes in the government capital stock due to public investment dynamics, and (ii) contribution from other sources (see Annex IV for details on the methodology). These two sources of growth could then be compared with historical data and past projections. Figure 10.a shows an illustrative case in which public investment was projected to be scaled-up in the 2013 DSA vintage. The tool's output would trigger a deeper discussion of the underlying assumptions underpinning growth projections in the context of changes in public investment, which is generally now absent in DSA write-ups.

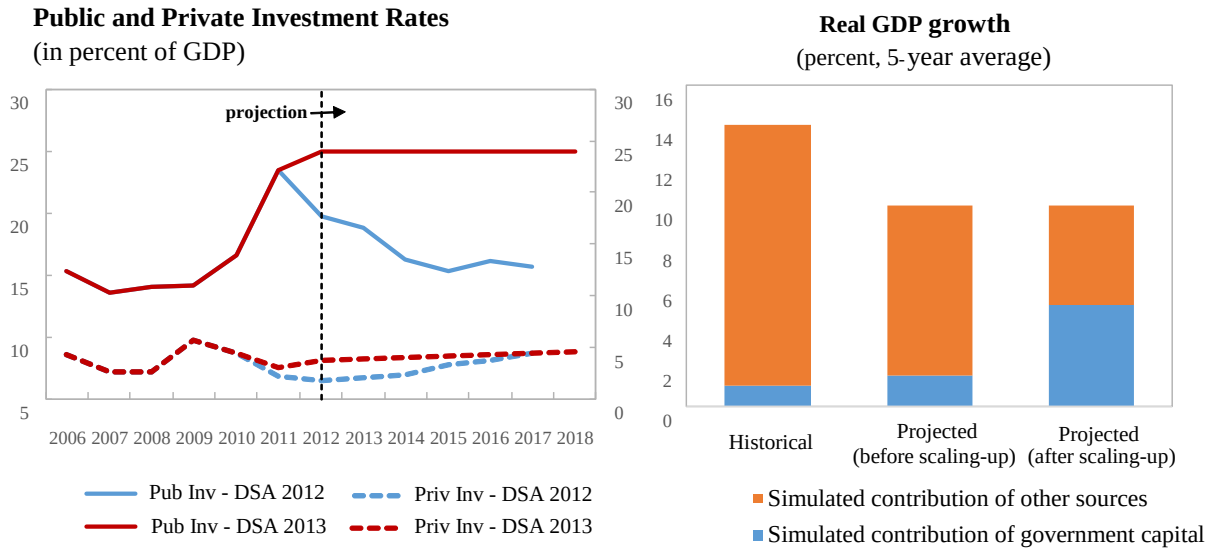
- **Assessing the impact of fiscal adjustment on growth.** Negative growth surprises have been identified as the main factor derailing fiscal consolidations (see Mauro and Villafuerte, 2013). While many factors may contribute to this, a key issue to avoid is overly optimistic assumptions about fiscal multipliers. This tool would shed light on the impact of the planned fiscal adjustment on growth projections under a range of plausible fiscal multipliers and persistence parameters, allowing a comparison with the baseline projected growth path (Figure 10.b) (see Annex IV for details on the methodology).[21]

[20]A more detailed analysis may be done outside the confines of the DSF, as a means of informing Fund-supported programs, World Bank growth diagnostics, and the policy dialogue more generally. Available tools supported by Fund and Bank staff for this purpose include the IMF's Debt-Investment-Growth model (see Buffie and others, 2012), and the World Bank's Long-Term Growth model (see Pennings, 2017).

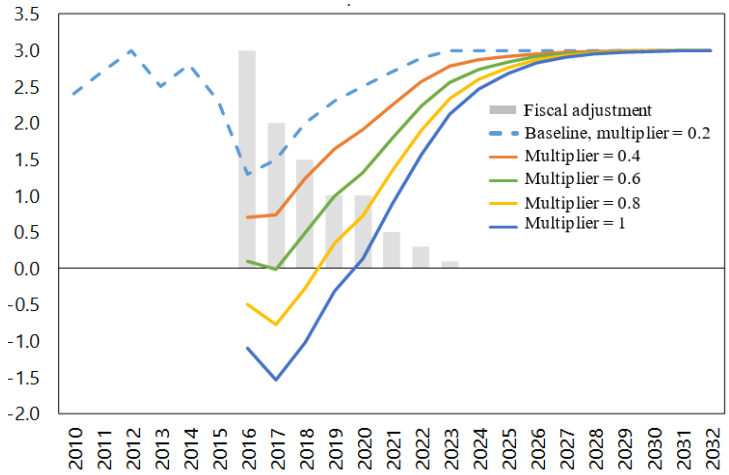
[21]While a useful check for the consistency of growth and primary balance projections is to uncover the growth path consistent with a neutral fiscal stance and assess its realism, such an approach is challenging in the case of LICs. This is because LICs' borrowing capacity is very weak in the absence of fiscal adjustment, casting doubts about the likelihood of such a counterfactual.

Figure 10. Realism of Baseline Growth Projections

(a) Public investment [1/]



(b) Fiscal adjustment [2/]



Source: Fund staff calculations.
1/ Left-hand chart shows differences in projected public and private investment rate over two DSA vintages; the right-hand chart compares the simulated contribution of government capital and other factors to real GDP growth over two DSA vintages and based on historical data.
2/ Bars refer to annual projected fiscal adjustment and lines show possible real GDP growth paths under different fiscal multipliers.

D. Determining Mechanical Risk Signals

41. **A key output of the framework is a mechanical risk signal.** This is derived by comparing debt projections in the baseline forecast and under stress tests with debt thresholds, drawing on an aggregation rule and risk-weighting scheme (i.e., how much weight is placed on the rate of missed crises versus that of false alarms). The mechanical risk signal is derived from a simple rule: a country is signaled as low risk if all debt burden indicators are below their corresponding thresholds under the baseline and stress test scenarios; moderate risk if there is any breach under the stress tests but not under the baseline; and high risk if there is at least one breach under the baseline.

42. **Two methodological improvements would be made to the technical approach to setting thresholds, to address problems noted in the previous section.** First, a transparent policy choice for the tolerance of missed crises and false alarms would be introduced. It is proposed that this be set at 2:1, recognizing the importance of keeping a strong early warning framework of the risk of debt distress (the complex averaging procedure in the existing framework produces an effective weight around this level). Second, debt thresholds for the three classification categories would be derived in a manner that is consistent with the DSF's aggregation rule, which would efficiently rebalance the role of debt burden indicators per their capacity to signal the risk of debt distress, (see Annex III).

43. **The procedure for generating mechanical risk signals would be considerably simplified in the new framework, even while a more comprehensive stress testing framework is introduced (Figure 11).** The broad approach to mechanical risk signals—the aggregation rule—would remain unchanged as it has been beneficial for predicting debt distress. However, the number of thresholds, debt burden indicators, and standardized stress tests would be considerably reduced. The reduced standardized stress tests would in turn make room for tailored scenario stress tests, better capturing key risks. The rest of this subsection considers each of these in turn.

Figure 11. Key Simplifications to the Framework for Determining External Risk Signals

	Existing DSF	Reformed DSF
1. Debt burden indicators	5	4
2. Thresholds	24	12
3. Standardized stress tests	16	8

Debt Burden Indicators and Thresholds

44. **Streamlining would address the redundancies noted in the previous section:**

- **The indicator on the present value of external debt-to-fiscal revenues would be dropped.** After incorporating all proposed changes into the framework, results confirm that this indicator is redundant (including or excluding this indicator does not affect the rate of missed crises and false alarms). This confirms the findings of the previous section (subsection A), which showed that this indicator individually has not in the past played a role in signaling high risk of debt distress.

- **The use of remittances-augmented thresholds would be discontinued.** The inclusion of remittances directly in the statistical model informs the country classification of debt-carrying capacity, removing the need for this more complex and indirect approach.

- **The projection horizon underpinning the determination of mechanical risk signals would be shortened from 20 to 10 years.** The evidence shows that threshold breaches are concentrated in the first five years of projections and that there is high uncertainty in long-term projections (see previous section). The new DSF template would continue to report 20 years of projections as is the case in the present template, while staff judgment on projected developments during the 10-20-year period could be used to override the mechanical risk signal. This approach would focus greater attention on the determinants of long-term breaches of thresholds, enhancing the transparency of DSAs.

45. **With these changes, the predictive performance of the reformed framework significantly improves on existing benchmarks** (Table 2). Under the proposed reforms, the probability of false alarms is significantly reduced (by 10 percentage points), while the model's capacity to anticipate debt distress is improved. For comparison, staff also assessed the implications of keeping the existing model specification and methodology for deriving thresholds, and re-estimated the thresholds based on the updated data (1970–2014). It was found that the existing re-estimated model explains the set of new episodes poorly, with tighter thresholds driving up the mechanical rate of false alarms to about 70 percent, without any gain in overall predictive performance (see Annex III for more details).

Table 2. Predictive Power of Existing and Reformed Framework

Statistics	Based on DSF's aggregation rule		
	Existing 2/	Existing (updated) 3/	New 4/
Type I error	0.18	0.06	0.18
Type II error	0.48	0.68	0.38
Loss function 1/	0.28	0.27	0.24

Source: Staffs calculations.
1/ Weighted sum of errors (weight on type I error equal to 67 percent).
2/ Existing framework (2012 LIC DSF review).
3/ Existing framework, after re-estimating debt thresholds using existing methodology based on the new sample (1970-2014) and proposed new definition of external debt distress.
4/ Based on all proposed changes to the existing framework in this review.

46. **The re-estimated external debt thresholds imply some additional room for countries to borrow, provided they manage their debt service well.** The new framework would maintain or increase debt stock thresholds for all countries, but countries deemed to have weak or medium debt-carrying capacity would face lower debt service thresholds (Table 3). Since the present framework had set debt service thresholds at what were typically non-binding levels (see previous section), it is not surprising that the improved methodology produces lower debt service thresholds. The impact also needs to be understood considering the changes to the classification scheme. Since the new scheme upgrades classifications after bringing in key fundamentals (see next section), effectively not all countries face lower debt service thresholds.

Table 3. Existing and Re-estimated External Debt Thresholds

Country	PV of debt-to-GDP		PV of debt-to-exports		Debt service-to-revenue		Debt service-to-exports	
Classification	Old	New	Old	New	Old	New	Old	New
Weak	30	30	100	140	18	14	15	10
Medium	40	40	150	180	20	18	20	15
Strong	50	55	200	240	22	23	25	21

Source: Fund staff calculations.

Standardized Stress Tests

47. **The new framework would apply a smaller set of standardized shocks in the external and public DSAs, to address the redundancies noted in the previous section.** The GDP deflator stress test and all permanent shocks would be dropped (although the "historical" scenario would remain as a realism check). The data show that, historically, most shocks have been of a temporary nature. The smaller common set of standardized stress tests to be applied would include shocks to: (i) real GDP growth, (ii) export growth, (iii) primary balance, (iv) other flows (transfers and FDI), (v) exchange rate depreciation, (vi) contingent liabilities, and (vii) a combination

shock.[22] These shocks would be applied to both external and public DSAs to eliminate inconsistencies in the risk assessment.

48. **The set of standardized shocks would also be re-calibrated/re-designed, to address the misalignments noted in the previous section:**

- **Interactions among the main macro variables in response to shocks would be introduced**. These include: (i) the export growth shock would negatively impact GDP growth; (ii) the exchange rate shock would positively affect inflation and net exports, which would mitigate the impact of a sharp FX depreciation on debt ratios; (iii) the real GDP growth shock would negatively affect inflation and the primary balance; and (iv) the primary balance shock would increase commercial borrowing costs. The modeling of elasticities for these interactions draws on an event study and on empirical evidence from the existing literature (see Annex I).

- **The point of application of shocks would be adjusted.** For all stress tests, post shock values would be set to the historical average minus one standard deviation—the current practice in the stress testing framework—or the baseline projection minus one standard deviation, whichever is lower for the relevant periods. This proposed change would reduce the likelihood of underestimation of primary balance shocks, but would also help improve the accuracy of all stress test shocks.

- **The contingent liability shock would be redesigned** (the existing "other debt-creating flows" stress test). The existing contingent liability shock would be re-calibrated to 5 percent of GDP (from 10 percent) aimed at capturing potential liabilities stemming from financial sector vulnerabilities.[23] A tailored shock would expand its scope where relevant, to capture potentially larger financial sector risks or those resulting from exposure to liabilities not covered under the debt concept in the DSA (see below).

Tailored Scenario Stress Tests

49. **The reformed framework would also introduce new tailored scenario stress tests where appropriate** (see Annex I for details on triggers and calibration)**.** This would fill a key gap in the framework without overburdening the DSF, and align it with the recommendation of the IMF's 2014 Surveillance Review to have better tailored scenario analysis. These stress tests would only be triggered for countries assessed to be exposed to particular risks (although they could be used on a discretionary basis in other cases). The scenarios' default settings would be based on the median shock for LICs, but in each case country teams would be encouraged to customize the scenario parameters to reflect country-specific considerations. As in the case of the standardized

[22]This stress test reflects a scenario where multiple shocks hit the economy at the same time. It applies half the magnitudes of all stand-alone shocks, and incorporates individual macro interactions as assumed under each individual shock.

[23]This is based on the average increase in debt-to-GDP ratios observed for 44 banking crisis episodes for LICs since the 1980s, as in Laeven and Valencia (2013).

stress tests, tailored scenario stress tests could move a mechanical risk signal from "low" to "moderate".

50. **The tailored stress tests would reflect risks that are common to groups of LICs.** For some countries, more than one test might apply. Annex I provides details on each test.

- **Natural disasters.** The scenario would be required for LICs identified as particularly vulnerable to natural disasters (see IMF, 2016). The scenario's default magnitudes would involve a one-off shock to public debt (capturing fiscal impacts), and real GDP and export growth decline in the year of the shock, calibrated based on data covering natural disasters during 1950–2015. Country teams would be encouraged to adjust the parameters of the scenario to country-specific circumstances, explaining the basis for the adjustments, including any assumptions about the impact of natural disasters already embedded in the baseline scenario.

- **Contingent liabilities supplement.** The re-calibrated standardized contingent liability shock would be supplemented with a tailored shock triggered when significant elements of the public sector are not covered by the public debt concept used in the DSA (e.g., other parts of general government, public private partnership exposures, or unaccounted state enterprise exposures) or when financial sector risks to the sovereign balance sheet are deemed larger than the size modeled under the standardized contingent liability shock.[24] The scenario would involve a one-time increase in the debt ratio, scaled to the size of potential exposures in those sectors not covered by the country-specific public debt concept.

- **Commodity export price shock.** The scenario would be triggered for LICs where commodity exports represent at least 80 percent of merchandise exports.[25] The proposed scenario would capture the impact of a sudden one standard deviation decline in commodity export prices (using the distribution underlying WEO forecasts), with macro interactions incorporated, based on staff event analysis and recent studies (IMF, 2015c and Aslam and others, 2016).

- **Market-financing stress test.** The shock would be triggered for LICs with access to market-based financing. A tailored combination shock would capture the impact of stressed external commercial borrowing conditions resulting from a deterioration in global risk sentiment, covering increases in the cost of new external commercial borrowing, temporary nominal depreciation, and shortening of maturities of new external commercial borrowing.

51. **The use of customized alternative scenarios would continue to provide flexibility to assess other country-specific risks not covered by standardized and tailored stress tests.**

[24] This stress test would help eliminate a disincentive to achieve full and transparent disclosure of public sector contingent liabilities.

[25] This is based on UNCTAD (2015), which is broadly in line with the definition used in IMF-WB (2015). This cutoff captures 36 countries (see Annex I).

Although standardized and tailored stress tests ensure comparability among countries, certain idiosyncratic risks may be overlooked (e.g., delays in the implementation of megaprojects, active conflict scenarios), and magnitudes or duration of certain shocks may be underestimated. The DSF template would be enhanced to better allow users to fully design customized scenarios to address such country-specific circumstances.[26] The revised Staff Guidance Note would also address good practice in handling such risks, including those related to conflicts.

E. Analyzing Other Potential Risk Factors

52. **The existing DSF calls on staffs to examine potentially significant risks not captured in the core statistical model.** In particular, staffs are called upon to conduct a deeper analysis of domestic debt vulnerabilities when public debt-GDP ratios are approaching or exceed estimated benchmark levels; to assess the risks to the balance of payments position and to government exposure to contingent liabilities of elevated levels of private external debt; and to assess the risks posed by shifts in market sentiment to debt positions in LICs that have a high share of public debt contracted on market terms with private external creditors. Where assessed to be significant, these risks can affect the overall risk of debt distress—a broader concept than the risk of experiencing external debt distress (IMF-WB, 2013b).

53. **The new framework would continue this approach.** It would include new tools to help users assess public debt-related risks and market financing-related risks, and a tool to help users better understand, in summary form, the diverse risk characteristics of countries at moderate risk of debt distress. The DSF template would automatically report on these broader risk assessments through a heat map.

Total Public Debt

54. **The new DSF would continue to require an assessment of risks stemming from domestic debt levels, through an analysis of total public debt.** Although external PPG debt remains the largest component of total public debt in most LICs, a formal analysis of risks from total public debt is warranted because a) domestic debt markets are an increasingly important source of financing for many LICs, and b) in some frontier LICs featuring deeper financial markets and less restrictions on capital flows, non-residents have increased their participation in local and regional debt markets, blurring the distinction between domestic and external debt.[27]

55. **The methodology to assess public debt-related risks has been strengthened.** Lack of adequate data on domestic arrears is a key impediment to identifying domestic debt distress episodes. An alternative identification procedure has been used to capture *de facto* domestic

[26]If a user yet wished to consider one of the dropped stress tests, this would also offer a channel to do that.

[27]This is particularly important in countries where external debt is defined on a currency denomination basis, which seems to be the norm in LICs, given data limitations on debt by residency (a Fund's staff survey in 2016 revealed that more than 50 percent of LICs report their external debt on currency basis).

default, based on two proxies (see Annex III). A noise-to-signal modelling approach is then used to derive benchmarks for the PV of total public debt—contrasting with the probit models deployed for this purpose in the 2012 LIC DSF review.[28] [29]

56. **The newly-estimated benchmarks for the present value of total public debt are not substantially different from the existing benchmarks** (Table 4).[30] As in the analysis of external debt distress, the new benchmarks imply a lower rate of false alarms compared to the one achieved by the existing framework at the time of the 2012 review. Estimated benchmarks are robust to variations in the definitions of the two proxies of *de facto* domestic default. The framework would signal high risk of public debt distress if any of the four external debt burden indicators or the public debt indicator breach their corresponding threshold/benchmark; moderate risk if the thresholds/benchmark were breached in stress tests; and low risk if the thresholds/benchmark were not breached in either the baseline or stress tests scenarios.

Table 4. Existing and Re-estimated Benchmarks for Total Public Debt

Variable	Thresholds			Type I	Type II	Loss
	weak	medium	strong	error	error	function
2017 review 1/	35	55	70	0.24	0.44	0.31
2012 review 2/	38	56	74	0.19	0.52	0.31

1/ Based on NTS approach, using 0.67 weight on type I error. Threshold are rounded up to the nearest 5 percent.
2/ Based on the probit model and weight on type I error ranging from 0.5 to 0.75.

Market-financing Pressures

57. **The new DSF would include a tool to signal rollover risks from external commercial borrowing exposures, where such exposure is significant**. The debt service indicators in the existing framework provide some indication of risks under the baseline scenario and the proposed market stress test (discussed in subsection D above) examines the impact of shifts in market sentiment on the four debt indicators tracked in the framework. However, while the behavior of

[28]Data on the PV of total public debt is the sum of the PV of external PPG debt and nominal domestic public debt, which is assumed to be contracted on market terms.

[29] The use of probit models for analysis of public debt vulnerabilities would produce a composite indicator different from the one derived in the external debt risk assessment block – potentially leading to a classification of countries' debt-carrying capacity different from that developed in analyzing external debt distress. This would create considerable conceptual confusion, while also increasing the complexity of the general framework.

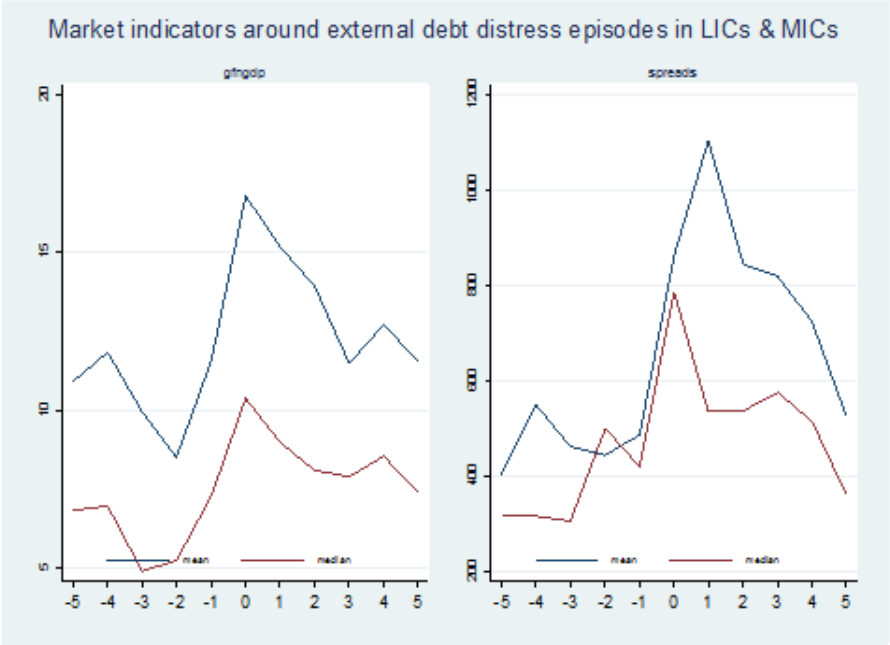
[30]Benchmarks were rounded up to align the strong classification category with the MAC DSA high-risk benchmark.

these indicators is informative, the historical experience suggests that they are not sufficient, by themselves, to fully capture the extent of liquidity risks faced by countries with market access under the baseline scenario.[31]

58. **The new tool would assess whether estimated gross public financing needs (as a share of GDP) and prevailing EMBI spreads point to risks of experiencing debt distress.** Applying the noise-to-signal approach to external debt distress episodes in a sample of frontier LICs and MICs between 1995 and 2015 produces estimated benchmark levels for public gross financing needs to GDP (a comprehensive measure of liquidity pressures) and for bond spreads (a measure of market risk perception) (see Annex V for details). The two indicators tend to move together in the run-up to debt distress (Figure 12).

59. **Market-financing vulnerabilities would be deemed significant if there were consistent early warning signals across the two indicators.** A breach of both benchmarks (GFNs above 14 percent of GDP and EMBI spreads higher than 570 bps) would signal heightened liquidity needs at times of adverse external market conditions, thereby increasing rollover risks and the likely need to seek exceptional financing. In such circumstances, country teams would be expected to provide an in-depth assessment of liquidity needs and the composition of the creditor base, as input into the final conclusions of the risk assessment.

Figure 12. GFNs and EMBI Spreads Around External Debt Distress Episodes



Source: Fund staff calculations.
Note: Data is on annual basis since the mid-1990s for which the two indicators are available.

[31] This is even more compelling in those cases where non-residents have increased their participation in local and regional debt markets, blurring the distinction between domestic and external debt.

Moderate Risk Category

60. **The new framework would provide a tool to help users better understand the nature of debt vulnerabilities in countries rated as facing a moderate risk of debt distress**. One way to illuminate the diversity of debt vulnerabilities is by measuring how far a country's debt burden indicators are from crossing the debt thresholds under the baseline scenario. This distance to breaching thresholds or "space to absorb shocks" can be assessed against the distribution of shocks that have been observed in the past to lead to a downgrade of countries to high risk of debt distress. This uses existing information in DSAs and does not require a view about the likelihood of shocks (as an "outlook" would).

61. **Using the tool, countries at moderate risk of debt distress would be characterized per their assessed "space to absorb shocks"**. Taking this approach, a country rated as facing moderate risk would be characterized as having:

- *"Limited space to absorb shocks"* if a median-size observed shock would lead to threshold breaches under the baseline, excluding cases of marginal and/or temporary breaches (i.e., if there is at least one debt burden indicator with insufficient distance to thresholds under the baseline scenario to withstand a shock equivalent to the median observed shock, estimated at 20 percent of the threshold for debt stock indicators, and 12 percent for debt service indicators).
- "*Substantial space to absorb shocks*" if threshold breaches would not occur under the baseline in all but shocks in the top quartile of the observed distribution (i.e., if all debt burden indicators have sufficient distance to thresholds under the baseline scenario to withstand all but top quartile observed shocks, estimated at 40 percent of the threshold for debt stock indicators and 35 percent for debt service indicators).

This assessment of "space to absorb shocks" would shed light on the robustness of the debt position of countries in the moderate risk category and the potential for slipping into high risk of debt distress; it would not have operational implications for Fund/Bank debt policies (see Annex VI for details on the methodology).

F. The Application of Judgment

62. **Judgment would continue to play an important role in the framework.** The mechanical risk signals from the framework would provide a first cut at determining the external risk rating, but this could be modified on the basis of staff judgment. Any application of judgment would need to be justified in the DSA write-up of the final risk assessment. Factors that could justify an adjustment in the external risk rating due to judgment would include:

- Breaches of thresholds that are assessed to be marginal and/or transitory in nature. Given the incomplete explanatory power of the probit model (as is the case for any econometric tool), the strict mechanical application of the DSF's ratings approach - where the breach of any one threshold signals high risk, could unduly penalize countries with marginal

and/or transitory breaches. Staff would be expected to avoid mechanical application of the risk assessment rules in such circumstances, while retaining the ability to do so if a convincing case can be made. A technical specification of the circumstances under which this presumption would apply will be developed for the Staff Guidance Note.

- Concerns as to the severity of domestic debt vulnerabilities could justify adjustment of the risk of external debt distress in cases where non-residents hold a sizable share of domestic government debt; such concerns would need to be flagged in the summary risk assessment even if they did not rise to a level that warranted adjustment of the external risk rating.
- Concerns about the exposure of the sovereign to external market-financing pressures could also justify adjustment of the risk of external debt distress if sufficiently severe, and would warrant mention in the summary risk assessment even if not sufficient to justify adjustment of the external risk rating.
- Other country-specific considerations (risks and mitigating factors) not captured by the statistical framework—such as the availability of sizable public financial assets (i.e., partially offsetting public debt levels), conflict risks, long-term considerations (e.g., vulnerability to climate change), availability of insurance-type arrangements, and collateralized financing arrangements. The Staff Guidance Note, to be prepared by WB and IMF teams would discuss the considerations under which these factors should be considered for the final determination of external risk ratings.

G. Drawing Conclusions on the Risk of Debt Distress

63. **Having conducted the various tests and assessments discussed above, and brought in judgment on other factors, staffs would be expected to provide:**

- A rating of the risk of external debt distress, using the established categories: low, moderate, high, or in debt distress.[32]

- A rating on the overall risk of debt distress, where vulnerabilities linked to public domestic debt levels are a serious concern: low, moderate, high, or in debt distress.

- A full discussion of the main risks to this assessment, including from factors such as data coverage, macroeconomic uncertainty, policy implementation risks, and global factors, among others.

64. **The summary risk assessment would be expected to explain how the staffs reached their conclusion and to provide perspective on the evolution of debt-related vulnerabilities over time**. This would cover the results derived from the mechanical features of the framework as well as explanations of the decisions made to deviate from the mechanical risk signals. The summary would also be expected to comment on the evolution of debt risks and vulnerabilities

[32] Actual or impending debt restructuring negotiations, the existence of arrears, or a significant or sustained breach of thresholds would generally suggest that a country is in debt distress (see IMF-WB, 2013b). The Staff Guidance Note will elaborate further.

over time, and identify key risks/mitigating factors that could shift the risk assessment going forward.

IV. BACK-TESTING RESULTS

65. **To help better illuminate the properties of the reformed DSF, staff has back-tested its mechanical features.** It is important to emphasize that mechanical back-testing results cannot be taken as a literal interpretation of how country classifications, external risk ratings, and the analysis of other risk factors would evolve. By necessity, back-testing results are based on a vintage of macro data and projections, and most DSAs (about 95 percent) used in the simulations were produced during 2015–16. Likewise, country classifications were informed by macro data and projections from the 2016 Fall WEO. Moreover, mechanical risk signals do not factor in full customization of the tailored stress tests, nor staff judgment, which as noted in the second section of the paper, has been applied in about 25 percent of DSAs since the DSF inception.

66. **On country classifications, back-testing suggests more upgrades than downgrades relative to the existing CPIA-based classification.** Out of 67 countries for which LIC DSAs are currently produced, the CI-based classification keeps the existing CPIA-based classification unchanged for the majority (40 countries, 60 percent of the sample), upgrades 22 countries (33 percent), and downgrades 5 countries (7 percent) (Table 5.a). This relative stability is in part explained by the fact that the CPIA still accounts for an important share of the CI (45 percent). Most upgrades are driven by large reserve positions and/or sizable remittances flows (Table 5.b). These results confirm that each variable in the composite indicator plays a meaningful role in classifying countries' debt carrying capacity (see Annex III).

Table 5. New Country Classification of Debt-carrying Capacity

a. Transition Matrix of Country Classification

		Composite indicator			
		Weak	**Medium**	**Strong**	**Total**
CPIA	**Weak**	20	12	0	**32**
	Medium	4	14	10	**28**
	Strong	0	1	6	**7**
	Total	**24**	**27**	**16**	**67**

Green: CI upgrades relative to CPIA	22	33%
Yellow: CI and CPIA classifications are equal	40	60%
Red: CI downgrades relative to CPIA	5	7%
Total	67	100%

b. Changes in Classification and Contribution to the CI

Country group	Sample	Average contribution to the CI					Total
		CPIA	Growth	Reserves	Remittances	World growth	
Upgraded	22	42%	4%	32%	6%	16%	100%
Same classification	40	46%	4%	28%	3%	18%	100%
Downgraded	5	50%	5%	26%	2%	18%	100%
All	67	45%	4%	29%	4%	17%	100%

67. **On mechanical risk signals, back-testing points to a balanced impact of the framework**. Out of 67 countries currently producing LIC DSAs, the reformed framework would keep the existing mechanical external risk signals unchanged for the majority of countries (47 countries, 70 percent of the sample), upgrade 12 countries (18 percent), and downgrade 8 countries (12 percent) (Table 6). Overall, back-testing results are not very different from those under the current framework at the aggregate level, with the distribution across risk signal categories remaining broadly unchanged.

Table 6. External Risk Ratings and Transition Matrix of Countries' Mechanical Risk Signals

Final risk rating (current framework)		Mechanical risk signal (current framework)		Mechanical risk signal *Transition matrix vis-à-vis mechanical risk signal based on current framework* (proposed revisions to the framework)		
Low	14	Low	12	8		Low 14
				4		
				0		
Moderate	32	Moderate	28	6		Moderate 28
				18		
				4		
High	21	High	27	0		High 25
				6		
				21		
		Upgrade		12		
		Downgrade		8		
		Unchanged		47		

68. **On the assessment of other risk factors, back-testing results give a sense of how they could be deployed:**

- **For public debt, assessments would closely parallel those given in the current DSF.** As expected, the broadly unchanged re-estimated total public debt benchmarks lead to a very few changes in mechanical signals of heightened risks from total public debt vulnerabilities (i.e., breaches of the total public debt benchmark). Most assessments under the existing framework (about 90 percent of the cases) are reaffirmed under the reformed DSF (Table 7).

- **For market-based risks, two countries out of 17 LICs with market access would be flagged** (Table 8). Five countries do not breach any benchmark, while in the other cases either only one benchmark is breached or the assessment is inconclusive due to unavailability of EMBI spreads.

Table 7. Back-testing of Total Public Debt Benchmarks

Desk Assessment	Mechanical signal (current framework)	Mechanical signal (proposed revisions to the framework)	
Classification	CPIA		CI
Thresholds	Existing		New
Projection horizon	20 years		10 years
Heightened risks due to total public debt vulnerabilities?			
		36	
No 50	**No 39**		**No 40**
		3	
		4	
Yes 17	**Yes 28**		**Yes 27**
		24	
	Upgrade		4
	Downgrade		3
	Unchanged		60

Table 8. Back-testing of Market-Financing Risk Indicators

GFN 1/	EMBI 2/	Number of Frontier LICs
Breach benchmark?		
Y	Y	2
Y	N	1
N	Y	1
N	N	5
Inconclusive		8

1/ Max GFN over a 5-year projection horizon based on data from the latest available DSA.

2/ EMBI spreads correspond to the latest available 3-month average.

- **Finally, back-testing of the methodology to identify the extent of "space to absorb shocks" in the moderate risk category shows that it can usefully discriminate across countries**. Of the 32 countries currently rated as facing moderate risk of debt distress, 16 would be characterized as having "limited space to absorb shocks" while 6 would be characterized as having "substantial space to absorb shocks". Of note, 7 out of the 16 countries flagged as "with limited space to absorb shocks" already have a mechanical high risk signal (Table 9).

Table 9. "Space to Absorb Shocks" in the Moderate Risk Category

Risk qualification	Number of countries (currently rated as moderate risk)
Moderate with limited space	16
Moderate	10
Moderate with significant space	6

V. DISCOUNT RATE FOR THE LIC DSF

69. **Since the DSF uses indicators of debt that focus on present value, an important consideration is the discount rate used to determine this present value**. Too high a rate can underestimate the debt burden (and risks); too low a rate can over-estimate it (and over-state risks). Given drawn out and large movements in global interest rates since the global financial crisis, the issue is of particular relevance.

70. **The methodology used to set the LIC DSF discount rate has been updated twice in the last 12 years:**

- When the LIC DSF was introduced in 2005 the adopted discount rate was the then-prevailing level of the USD Commercial Interest Reference Rate (CIRR) of 5 percent. This was to be adjusted by 100 bps whenever the 6-month average USD CIRR deviated from that level by at least 100 bps for a 6-month period. Based on this rule, the discount rate was adjusted to 4 percent in 2009 and 3 percent in 2012. However, these developments were primarily driven by low interest rates in developed countries and were thus seen to have unjustifiably narrowed LICs' assessed borrowing space against the backdrop of relatively stable financing terms in LICs.

- In 2013, reflecting concerns that had arisen, staff proposed a change to the methodology by using as a reference point a long-term average (10-year average of the USD CIRR) plus a margin to reflect the term premium for loans with maturities longer than the bond yields used to derive the CIRR. The long-term average aimed to smooth out the effect of low global interest rates that was assessed to be cyclical at the time. As a result, the discount rate was set at 5 percent.[33]

71. **As requested by Executive Directors (see IMF-WB, 2013a), staff has assessed whether there is justification for an update of the discount rate.** Staff considered two methodologies:

- *The existing CIRR methodology, which links the discount rate to market rates in advanced economies.* This has several benefits: (i) the CIRR is transparent and readily available, and (ii) it yields a PV of debt that can be broadly interpreted as the amount of risk-free investment a country needs to make to be able to service its external debt.[34] Applying the existing methodology today, the discount rate would be 4.75 percent.

[33]See "Unification of Discount Rates Used in External Debt Analysis for Low-Income Countries" (IMF-WB, 2013a).

[34]The CIRR overstates the risk-free rate, thus underestimating the required investment, as the different export credit agencies (ECAs) agreed to set the CIRR as the average of market rates for long-term government debt in their own currency, plus one percent. This was designed to ensure that ECAs do not unfairly subsidize trade by setting their lending rates too low.

- *A methodology linking the discount rate with LICs' average per capita income growth*. The right discount rate for the purpose of assessing debt sustainability is arguably the one that reflects LICs' economic fundamentals and capacity to repay debt. A concept based on the nominal GDP growth rate addresses these concerns. Intuitively, higher per capita income growth implies a greater repayment capacity as future debt service obligations represent a smaller burden on future taxpayers. With the median nominal dollar GDP growth rate for LICs at 7 percent and the median population growth rate at 2.2 percent, the discount rate based on this methodology would be 4.8 percent.[35]

72. **The review proposes to keep the current 5 percent discount rate, revisiting this decision in future DSF reviews, and to maintain a unified discount rate for the LIC DSF, the debt limits policy (DLP), the non-concessional borrowing policy (NCBP), and the grant element calculator.** Both approaches above point to keeping the discount rate at 5 percent at present. Conceptually, this could differ from the rate used for the DLP, NCBP, and the grant element calculator. However, operational complications can arise from different discount rates producing different PV of debt for the same underlying loans assessed in the LIC DSF versus the DLP and NCBP (see IMF-WB, 2013a).

[35]The median dollar nominal GDP growth rate is 7 percent when based on (i) 5 years of history and 5 years of projections, (ii) 10 years of history and 10 years of projections, or (iii) 5 years of history and 10 years of projections, using data submitted in the latest LIC DSAs. The median population growth rate is 2.2 when based on (i) 5 years of history and 5 years of projections or (ii) 10 years of history and 5 years of projection, using WEO data. A longer-term average has the benefit of smoothing out cyclical components and avoiding abrupt changes in the discount rate.

VI. ISSUES IN IMPLEMENTATION

73. **The proposed features of the reformed framework would facilitate a deeper discussion of debt risks and vulnerabilities between Fund and Bank staffs and country authorities.** This expanded discussion would respond to feedback received during consultations with country authorities, particularly about the need for stronger consideration of country-specific circumstances. Thus:

- Under the new approach to classifying countries' debt-carrying capacity, there would be greater engagement with country authorities to discuss key elements of macro projections (and thus the country's policy framework) that play a role in determining a country's classification. In principle, the new classification could continue to be done annually using the country teams' projections reflected in the Fall WEO. However, there could be circumstances in which an update may be warranted based on projections submitted to WEO during the Spring or at the time of producing DSAs. The operational aspects of updating countries' classifications, including transitional rules to avoid undue volatility in classifications, would be discussed in the Staff Guidance Note.

- The customization of tailored stress tests would require discussion with country authorities about the appropriate scenario parameters given a country's experience with these shocks.

- The technical and/or country-specific considerations underlying the application of judgment and the analysis of other risk factors would be fully disclosed and justified in DSA write-ups after discussion with country authorities. The DSA write-up would incorporate the authorities' views in this regard.

- Finally, in a framework with some additional room to borrow, but where greater attention to debt service risks is required, more policy dialogue would be needed on debt management capacity and strategy.

74. **The implementation of the reformed DSF would be supported in two ways:**

- **Updated materials to support implementation** (expected to be available by end-2017):

    - The review would be followed by a thorough revision and simplification of the Staff Guidance Note, aimed at providing a user manual of the DSF template. The Guidance Note would go beyond the simple application of the framework to cover issues such as: factors that guide the application of judgment, including the appropriate use of the probability approach; use of the framework to better understand fiscal space (building on the work done to date for EMs and advanced economies); use of the framework to make judgments about debt sustainability; and issues in modeling specialized types of risk (such as conflict related risks). The Guidance Note would also address data requirements (with a view to encourage incremental improvements in debt coverage data reporting); and the applicability

of the LIC versus MAC DSA frameworks. It is expected that a Staff Guidance Note will be issued, to support implementation, by end-December 2017.

  - The DSF template would be re-programmed, simplified, and streamlined, with automated features for ease of implementation and transparency. Composite indicator calculations would be automated and transparently reported, as would be the assumptions underlying other tailored scenario shocks.

- **Interactions with country authorities, through three different channels:**

  - Joint WB-IMF high-level seminar during the 2017 Annual Meetings, focused on clarifying the conceptual underpinnings of the new framework.

  - Technical discussions with country counterparts during Fund/Bank missions.

  - Expanded training program, including through onsite and online courses and seminars. With support from donors, external training would be increased from nine missions during FY 2017 to sixteen during FY 2018, including four workshops in regional training centers during the Fall and nine training missions during the first half of 2018.[36] This expanded training would help countries to build capacity to prepare their own DSAs and, ultimately, fully own the revamped DSF template. Particular attention would be given to assisting countries with weak capacity, with training tailored to their own specific circumstances.

75. **The staffs propose that the new framework becomes effective on July 1, 2018,** predicated on the assumption that the updated staff Guidance Note and DSF template are finalized by the end of December 2017. The effectiveness date would ensure teams had adequate time to complete DSAs for all countries under the new DSF prior to July 1, 2019 – the cut-off date for determination of changes to country "traffic lights" under the IDA grant allocation framework applicable to IDA's fiscal 2020 allocations and IDA19 projections. To maintain this timetable, it will be important to disseminate supporting materials to staffs and country officials in a timely manner, complemented by internal and external training in use of the methodology. Once the new framework becomes effective, country teams would make use of the new DSF in their next cycle of engagement with Fund and/or Bank staff.[37] Such interaction, coupled with hands-on experience in tailoring the DSF template to country data and circumstances, are the key elements in building up technical capacity in assessing risks to debt sustainability.

[36]During 2016-17, 175 representatives from 40 countries participated in one-week training sessions held in several regional centers and individual countries around the world, the bulk of which were financed through the Debt Management Facility. Most of them benefited from accessing the IMF-WB online module on the LIC DSF ahead of their face-to-face training. Overall, over the same period, nearly 850 government officials familiarized themselves with the online LIC DSF materials, through the IMF Massive Open Online Course on debt sustainability and debt management.

[37] However, missions that will be conducted before the DSF enters into effect will produce DSAs based on the old framework.

76. **Bank/Fund staff would continue providing support after the rollout of the new framework.** Continuing training and other technical assistance would support capacity building in the use of the revamped LIC DSF as well as integrating debt management strategies when producing DSAs. In this connection, staff would explore options to link the DSF with debt management tools, in particular the Medium-term Debt Management Strategy (MTDS) analytical tool, for both analysis and policy formulation.

77. **The flexibility embedded in the DLP and NCBP should help smooth any implications arising from changes in final risk ratings in the wake of this review.**

- In the case of the DLP, if a country's risk rating is upgraded, the country would face looser debt conditionality (or no debt conditionality) under a Fund-supported program. If, instead, its risk rating is downgraded, it could still benefit from the flexibility embedded in the current DLP. For instance, a country whose risk rating is downgraded from moderate to high could still access non-concessional resources to finance critical development projects or to implement debt management operations that improve the overall debt profile. A country whose rating is downgraded from low to moderate would generally continue benefiting from the flexibility of the DLP as it provides room for non-concessional borrowing (provided the borrowing does not lead to a deterioration in the moderate risk rating). Where the risk rating of a member under an existing arrangement changes as a result of the new DSF, staff would seek to reach new understandings when changes in debt conditionality are called for by the DLP. Such revisions to debt conditionality would be considered by the Executive Board in the subsequent staff report.

- Similar considerations will apply under the NCBP for all IDA-only non-gap countries that are either grant recipients in the current fiscal year or MDRI recipients, irrespective of their risk rating of debt distress. All countries will continue to be provided with a choice between loan by loan exceptions to the NCBP, with countries at low or moderate risk of debt distress also provided the option to seek nominal ceilings on new non-concessional external PPG borrowing. Countries at low or moderate risk of debt distress with adequate capacity may also opt for a PV ceiling on total new external PPG borrowing. The options available to countries will remain consistent with the conclusion of the joint WB-IMF debt management capacity assessment. IDA's grant/loan allocation mix will not be affected by the changes in country risk ratings resulting from the application of the new DSF until the fiscal year 2020 (starting July 1, 2019).[38]

[38] Each year, IDA determines its grant/allocation mix for the next 12 months based on the external risk ratings available by end-June.

VII. CONCLUSION

78. **The reforms proposed in this paper imply a significant overhaul of the DSF, bringing significant advantages to all stakeholders.** The proposed reforms would address several technical criticisms of the existing framework, and reduce an excessive rate of false alarms (with improved predictive power to signal debt distress events). By incorporating additional key country-specific information into country classification and derivation of thresholds, risk assessments in the new framework would be better customized to country conditions and would provide the basis for a richer dialogue on how policy actions affect debt-related risks. Finally, new realism checks and stress test reforms, and the enhanced guidance on the application of judgment, coupled with greater attention to domestic debt vulnerabilities and market-financing risks, will allow a more informed dialogue on the risks and trade-offs that national policy-makers face.

VIII. ISSUES FOR DISCUSSION

79. **Directors may wish to offer views on the following:**

- Do Directors agree that an expanded set of country-specific information (including reserve coverage and remittances flows) should be used to underpin the assessment of countries' debt-carrying capacity in the framework?

- Do Directors see merit in introducing realism tools to promote a deeper understanding of key assumptions underlying baseline macroeconomic projections, including the assumed relationship between public investment and growth, and to support stronger debt projections?

- Do Directors agree with the need to streamline the mechanical framework, by reducing the number of debt indicators, debt thresholds and standardized stress tests, and to rebalance the debt thresholds?

- Do Directors support the proposed enhancements to the stress testing framework, including the recalibration and new features of standardized stress tests (i.e., macro-interactions) as well as the inclusion of tailored scenario stress tests to assess key risks facing LICs?

- Do Directors see a need for a better assessment of broader risks stemming from high domestic debt levels and market-financing pressures, and a tool for characterizing countries in the moderate risk category in terms of its "space to absorb shocks"?

- Do Directors see merit in enhancing the Staff Guidance Note to ensure a more uniform and transparent application of judgment?

- Do Directors support the proposed timeline for the implementation of the framework?

Appendix I. Overview of the LIC DSF

1. **A model explaining the probability of external debt distress is at the core of the methodology underpinning the DSF.** Given an approach to identify external debt distress episodes, the probability of experiencing such events is explained through probit models controlling for debt burden indicators, the strength of institutions and policies—measured by the CPIA—, and country growth.

2. **Given an estimated model, countries are classified per their borrowing capacity in three categories: weak, medium, and strong performers.** This categorization is exclusively based on the CPIA. Once countries are classified, debt thresholds derived for each category from the probit models are assigned.

3. **As a first step for determining a risk rating, risk signals are obtained from the mechanical application of the framework.** Based on key inputs—baseline macro projections and a battery of stress tests—forecast for five debt burden indicators (PV of PPG external debt to GDP, exports, and fiscal revenues, and PPG external debt service to exports and fiscal revenues) are produced. These in turn are compared against their respective debt thresholds, after which risk signals are determined based on the following aggregation rule:

(i) if none of the forecasted debt burden indicators exceed their corresponding thresholds under the baseline and stress test scenarios, the DSF would signal a low risk of debt distress;

(ii) if all baseline forecasts are below their thresholds but at least one forecast exceeds its threshold under the stress test scenarios, the DSF would signal a moderate risk;

(iii) if at least one baseline debt forecast exceeds its threshold, the DSF would signal a high risk;

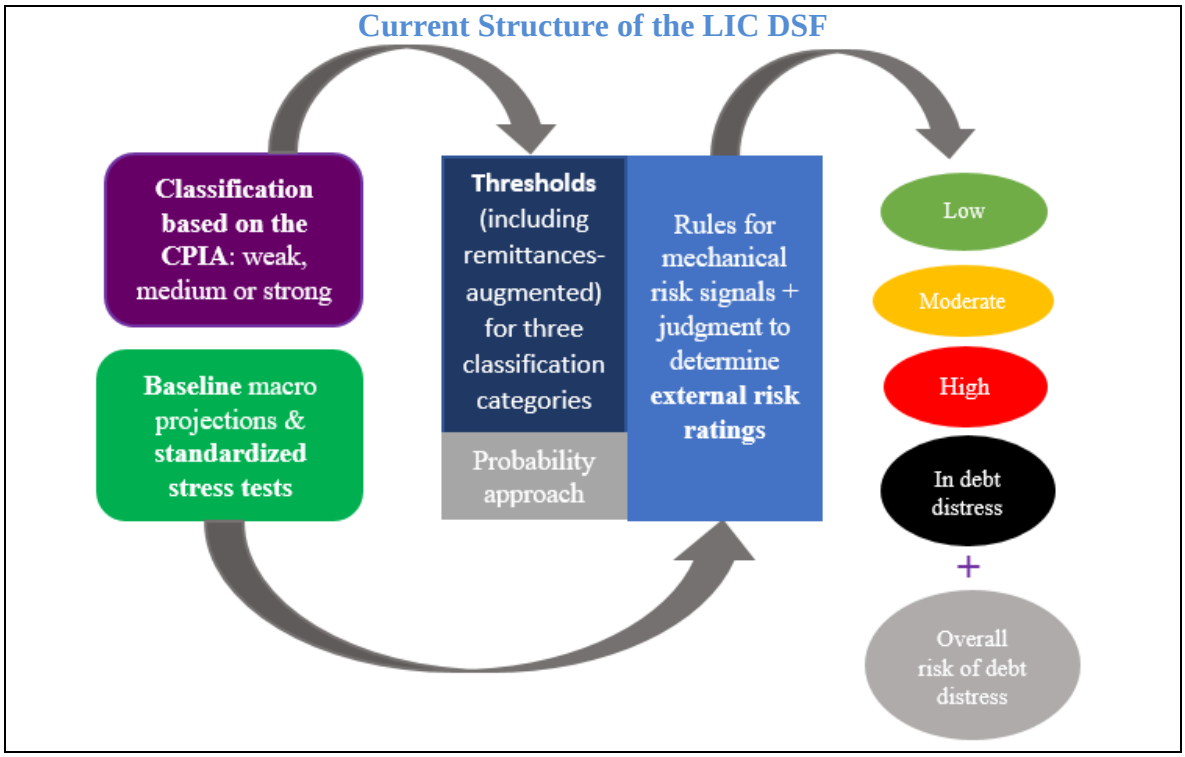
(iv) significant or sustained breach of thresholds, actual or impending debt restructuring negotiations, or the existence of arrears would generally suggest that a country is in debt distress.

4. **Next, the mechanical risk signals are combined with staff judgment.** This helps bringing in country-specific considerations and/or technical elements (e.g., whether breaches are marginal and/or one-off) that cannot be captured by the core model to make a final determination of the risk rating of external debt distress (Low/Moderate/High/In Debt Distress).

5. **In the 2012 review, new features were incorporated to the framework aimed to capture country heterogeneity where relevant.** These included: i) the use of remittances-augmented thresholds, only applied to countries receiving sizable remittances;[1] ii) the introduction

[1]Countries eligible to incorporate remittances into the DSF mechanical analysis have remittances flows greater than 10 percent of GDP and greater than 20 percent of exports of goods and services. Both ratios are measured on a backward-looking, three-year average basis.

of the "probability approach", an alternative technical methodology, that was to be applied at "borderline cases";[2] and iii) qualification of risks stemming from total public debt or private external debt.



[2]A borderline case is defined as one where the largest breach, or near breach, of a threshold under any scenario falls within a 10-percent band around the threshold.

REFERENCES

Abbas, S. M. A., and J. E. Christensen, 2007, "The Role of Domestic Debt Markets in Economic Growth: An Empirical Investigation for Low-Income Countries and Emerging Markets," *IMF Working Paper* No. 07/127, (Washington: International Monetary Fund).

Abbas, S. M. A., N. Belhocine, A. El Ganainy, and M. Horton, 2010, "A Historical Public Debt Database," *IMF Working Paper* No. 10/245, (Washington: International Monetary Fund).

Aisen, A., and D. Hauner, 2008, "Budget Deficits and Interest Rates: A Fresh Perspective," *IMF Working Paper* No. 08/42. (Washington: International Monetary Fund).

Aron, J., R. MacDonald, and J. Muellbauer, 2014, "Exchange Rate Pass-through in Developing and Emerging Markets: A Survey of Conceptual and Policy Issues, and Empirical Findings," *Journal of Development Studies,* 50(1): pp. 101–43.

Aslam, A., S. Beidas-Strom, R. Bems, O. Celasun, S. Kılıç Çelik, and Z. Kóczán, 2016, "Trading on Their Terms? Commodity Exporters in the Aftermath of the Commodity Boom," *IMF Working Paper* No. 16/27. (Washington: International Monetary Fund).

Bahmani-Oskooee, M., and A. Ratha, 2004, "The J-Curve: A Literature Review," *Applied Economics*, 36(13): pp. 1377–98.

Baldacci, E., I. Petrova, N. Belhocine, G. Dobrescu, and S. Mazraani, 2011, "Assessing Fiscal Stress," *IMF Working Paper* No. 11/100. (Washington: International Monetary Fund).

Beers, D. T., and J.-S. Nadeau, 2015, "Database of Sovereign Defaults," *Bank of Canada Technical Report* No. 101.

Berg, A., E. Berkes, C. Pattillo, A. Presbitero, and Y. Yakhshilikov, 2014, "Assessing Bias and Accuracy in the World Bank—IMF's Debt Sustainability Framework for Low-Income Countries," *IMF Working Paper* No. 14/48. (Washington: International Monetary Fund).

Berg, A., E. Buffie, C. Pattillo, R. Portillo, A. Presbitero, and L. Zanna, 2015, "Some Misconceptions about Public Investment Efficiency and Growth," *IMF Working Paper* No. 15/272. (Washington: International Monetary Fund).

Bom, P., and J. Ligthart, 2014, "What Have We Learned from Three Decades of Research on the Productivity of Public Capital?" *Journal of Economic Surveys*, 28(5): pp. 889-916.

Buffie, E., A. Berg, C. Pattillo, R. Portillo, and L. Zanna, 2012, "Public Investment, Growth, and Debt Sustainability: Putting Together the Pieces," *IMF Working Paper* No. 12/144. (Washington: International Monetary Fund).

Calderon, C., E. Moral-Benito, and L. Serven, 2015, "Is Infrastructure Capital Productive? A Dynamic Heterogeneous Approach," *Journal of Applied Econometrics*, 30(2): pp. 177-98.

Catão, L. A. V., and G. M. Milesi-Ferretti, 2014, "External Liabilities and Crises," *Journal of International Economics*, 94: pp. 18–32.

Cespedes, L. F., and A. Velasco, 2013, "Was This Time Different? Fiscal Policy in Commodity Republics," *NBER Working Paper* No. 19748.

Cruces, J. J., and C. Trebesch, 2013, "Sovereign Defaults: The Price of Haircuts," *American Economic Journal: Macroeconomics*, 5(3): pp. 85–117.

Das, U. S., M. G. Papaioannou, and C. Trebesch, 2012, "Sovereign Debt Restructurings 1950- 2010: Literature Survey, Data, and Stylized Facts," *IMF Working Paper* No. 12/203. (Washington: International Monetary Fund).

Gelos, R. G., R. Sahay, and G. Sandleris, 2011, "Sovereign Borrowing by Developing Countries: What Determines Market Access?" *Journal of International Economics*, 83: pp. 243–54.

Giovanni, A., and M. de Melo, 1993, "Government Revenue from Financial Repression," *The American Economic Review*, Vol. 83, No. 4. pp.: 953-63.

International Monetary Fund-The World Bank, 2012, "Revisiting the Debt Sustainability Framework for Low-Income Countries," (January) (Washington).

———, 2013a, "Unification of Discount Rates used in External Debt Analysis for Low Income Countries: Staff Proposals," (October) (Washington).

———, 2013b, "Staff Guidance Note on the Application of the Joint Bank-Fund Debt Sustainability Framework for Low Income Countries," (November) (Washington).

———, 2015, "Public Debt Vulnerabilities in Low-Income Countries: The Evolving Landscape," (November) (Washington).

International Monetary Fund, 2007, "The Changing Dynamics of the Global Business Cycle," Chapter 5, World Economic Outlook (October). (Washington).

———, 2012, "Commodity Price Swings and Commodity Exporters," Chapter 4, World Economic Outlook (April). (Washington).

———, 2014, “Assessing Reserve Adequacy–Specific Proposals,” (December). (Washington).

———, 2015a, “Making Public Investment More Efficient,” (June). (Washington).

———, 2015b, “Exchange Rates and Trade Flows: Disconnected?”, Chapter 3, World Economic Outlook (October). (Washington).

———, 2015c, “Where Are Commodity Exporters Headed? Output Growth in the Aftermath of the Commodity Boom”, Chapter 2, World Economic Outlook (October). (Washington).

———, 2016, “Small States’ Resilience to Natural Disasters and Climate Change: Role for the IMF,” (November) (Washington).

———, 2017, “Estimating the Stock of Public Capital in 170 Countries: January 2017 Update”. http://www.imf.org/external/np/fad/publicinvestment/pdf/csupdate_jan17.pdf.

Kraay, A., and V. Nehru, 2006, "When is Debt Sustainable?" *The World Bank Economic Review*, 20(3): pp. 341–365.

Laeven, L., and F. Valencia, 2013, "Systemic Banking Crises Database," *IMF Economic Review*, 61: pp. 225–70. (Washington: International Monetary Fund).

Ligthart, J., and R. Martin-Suarez, 2011, “The Productivity of Public Capital: A Meta-Analysis,” in Manshanden, W. and W. Jonkhoff, eds. Infrastructure Productivity Evaluation. (New York: Springer).

Mallik, G., and A. Chowdhury, 2001, “Inflation and Economic Growth: Evidence from Four South Asian Countries,” *Asia-Pacific Development Journal*, Vol. 8, Issue 1, pp. 123–35.

Manasse, P., and N. Roubini, 2005, “‘Rules of Thumb’ for Sovereign Debt Crises,” *IMF Working Paper* No. 05/42. (Washington: International Monetary Fund).

Manasse, P., N. Roubini, and A. Schimmelpfennig, 2003, “Predicting Sovereign Debt Crises,” *IMF Working Paper* No. 03/221. (Washington: International Monetary Fund).

Mauro, P. and M. Villafuerte, 2013, “Past Fiscal Adjustments: Lessons from Failures and Successes,” *IMF Economic Review*, 61(2): pp. 379-404. (Washington: International Monetary Fund).

Panizza, U., 2008, “Domestic and External Public Debt in Developing Countries,” *Unctad Discussion Paper* No. 188.

Panizza, U., F. Sturzenegger, and J. Zettelmeyer, 2009, “The Economics and Law of Sovereign Debt and Default,” *Journal of Economic Literature*, 47:3, pp. 651-98.

Pennings, S., 2017, "Long-term Growth Model v4.0: Model Description," available at http://globalpractices.worldbank.org/mfm/Pages/SitePages/MFM Online Tools.aspx

Rand, J., and F. Tarp, 2002, "Business Cycles in Developing Countries: Are They Different?" *World Development*, Vol. 30, No. 12, pp. 2071-88.

Reinhart, C. M., and K. S. Rogoff, 2009, "This Time is Different: Eight Centuries of Financial Folly,*"* Princeton University Press.

———, 2010, "From Financial Crash to Debt Crisis," *American Economic Review*, Vol. 101, No. 5, pp. 1676–706.

Reinhart, C. M., and K. S. Rogoff, 2011, "The Forgotten History of Domestic Debt," *The Economic Journal*, Vol. 121, Issue 552, pp. 319–50.

Reinhart, C. M., and M. B. Sbrancia, 2015, "The Liquidation of Government Debt," *IMF Working Paper* No. 15/7. (Washington: International Monetary Fund).

Spatafora, N., and I. Samake, 2012, "Commodity Price Shocks and Fiscal Outcomes," *IMF Working Paper* No. 12/112. (Washington: International Monetary Fund).

UNCTAD, 2015, "State of Commodity Dependence, 2014," (New York and Geneva: United Nations Conference on Trade and Development).

Yehoue, E., 2005, "International Risk Sharing and Currency Unions: The CFA Zones," *IMF Working Paper* No. 05/95. (Washington: International Monetary Fund).

ANNEXES

Annex I. Underlying Analysis on Stress Tests

1. **This annex summarizes the analysis used to examine the magnitude/duration of standardized and tailored shocks as well as interactions among key macro variables.** It also describes the set of countries that would be required to undertake tailored stress tests.

A. Performance of Stress Tests in Recent LIC DSAs

2. **The DSF standardized stress tests apply temporary shocks to a set of macroeconomic variables.** The magnitude of these shocks is based on the historical behavior of these variables in each country. Specifically, most stress test shocks set the relevant variable (GDP growth, export growth, primary balance, and current transfers and FDI flows) at *one standard deviation* below its historical average for two consecutive years. Both the averages and the standard deviations are calculated using each country's historical annual data over the most recent ten-year period. The exchange rate shock is the exception, which is defined as a one-off 30 percent nominal depreciation (on top of the baseline projections) of the domestic currency for all countries regardless of their historical exchange rate movements.

3. **Setting adverse shocks at one standard deviation remains appropriate.** Shocks of this magnitude have a chance of about 16 percent to occur each year, or around once every six years on average.[1] This time horizon coincides with the length of business cycles in major economies. The average length of business cycles in the US, for example, is about 5¾ years. Therefore, the magnitude of a one standard deviation shock would be representative of those observed under a typical recession.[2]

4. **Staff examined whether the standardized stress tests in past DSAs are indeed delivering realistic shocks (i.e., of about one standard deviation) as envisaged.** Specifically, using all LIC DSAs produced since 2008, staff tested whether the shocks generated by stress tests during the second and third year of projections (or first year of projection in the case of the FX shock) indeed represented shocks at around the 16th percentile of the realized distributions. Results show that for most shocks—including the GDP growth shock, export growth shock, and the non-debt creating flows shock—stress tests worked as envisaged; the post-shock values in the

[1]For variables that follow normal distributions, a one standard deviation adverse shock from the mean will yield a post-shock value at about the 16th percentile. Given that the distributions of economic variables tend to have fatter tails than normal distributions, the likelihood of shocks of magnitudes at or above one standard deviation could be slightly higher than 16 percent.

[2]Some studies have suggested that business cycles may be shorter in developing countries than in major economies (Rand and Tarp, 2002; IMF, 2007). For a country with shorter business cycles, the one standard deviation shocks would represent more severe shocks than those observed in a typical business cycle.

stress test scenarios turned out to be at around the 16-20th percentile of the actual outturns (Table AI.1).

Table AI.1. Standardized Stress Tests: Post-shock Values and Outturns

	Actual < Post-Shock Value	Actual > Post-Shock Value
	(in percent)	
GDP Growth Shock	18.2	81.8
Exports Growth Shock	19.9	80.1
Primary Balance Shock	**31.4**	**68.6**
Non-Debt Flows Shock	15.3	84.7
Exchange Rate Shock (annualized)	16.6	83.4

Sources: LIC DSAs during 2008–15, and Fund staff calculations. For the exchange rate shock, WEO data covering a longer period (1990–2015) was used.

5. **However, the primary balance shock appears to be too small compared to other shocks.** There is a much higher chance—at about 31 percent—that the observed primary deficit was higher than the post-shock value used in the stress test scenario. This translates into only about ½ standard deviation below the sample mean, assuming a normal distribution. This suggests that the primary balance shock have been milder than envisaged in the past DSAs.

6. **The magnitude of the exchange rate shock is broadly in line with that of other shocks.** Staff tested whether an exchange rate depreciation at an annualized rate of 15 percent is in line with other shocks, at around the 16th percentile of the realized distributions.[3] Using data of LICs over 1990–2015,[4] results show that a 15 percent nominal depreciation lies at about the 17th percentile of the actual distribution of exchange rate changes, which is in line with other shocks.

7. **Bias in historical averages appears to be the main driver for the underestimated primary balance shock.** As post-shock values in DSA stress tests are set at one standard deviation below historical averages, a key assumption is that the 10-year historical averages and standard deviations are good predictors for the performance of the relevant indicators in the projection years. This assumption did not hold for the primary balance: outturns of primary balances turned out to be lower than historical averages in about 2/3 of all the observations (Table AI.2). On average, the actual primary balance to GDP ratio was 1.2 percentage points lower than historical averages.

[3]The exchange rate shock is 30 percent depreciation in one year, while other shocks normally last for two years. To compare the magnitude of the exchange rate shock with other shocks, the assumed depreciation needs to be annualized over two years, or about 15 percent annually.

[4]A longer period is justified because exchange rate shocks appear to happen less frequently.

Table AI.2. Standardized Stress Tests: Comparing Historical Averages and Baseline Projections to Outturns				
	Actual < Historical Average	Actual > Historical Average	Actual < Baseline Projection	Actual > Baseline Projection
	(in percent)			
GDP Growth	46.8	53.2	52.8	47.2
Exports Growth	57.8	42.2	48.9	51.1
Primary Balance	**67.6**	**32.4**	56.2	43.8
Non-Debt Flows	40.3	59.7	47.5	52.5

Sources: LIC DSAs during 2008-15, and Fund staff calculations.

8. **To correct for the possible bias from historical averages, an alternative approach would be to derive post-shock values from baseline projections.** The baseline projections turned out to have better predictive power for future performance than historical averages, having smaller median forecast errors as well as smaller standard deviations in forecast errors (Table AI.3). However, basing the shock scenario solely on baseline projections may have shortcomings. In particular, it would not work as envisaged when baseline projections are too optimistic. An effective approach would be to set post-shock values at one standard deviation below the historical average, or the projected value under the baseline, whichever is lower, so to minimize possible biases from historical averages, while at the same time limiting the potential bias from overly optimistic baseline projections. Indeed, calculations using the same sample show that once this approach is applied, the post-shock value of primary balance would be at about the 16th percentile of the distribution of actual outturn, in line with other shocks.

Table AI.3. Predicative Power of Historical Averages and Baseline Projections				
	Historical Average		Baseline Projection	
	Median Forecast Error	Std. Dev. of Forecast Error	Median Forecast Error	Std. Dev. of Forecast Error
GDP Growth	0.3	5.5	-0.1	4.5
Exports Growth	-2.5	21.6	0.3	19.3
Primary Balance	1.1	4.6	0.4	4.5
Non-Debt Flows	0.4	7.0	0.1	7.9

Sources: LIC DSAs during 2008-15, and Fund staff calculations.

9. **An event analysis was also conducted to evaluate the appropriateness of the assumed duration of shocks (Table AI.4).** Using the WEO database, the study compares the shocks used in the current DSF with the distribution of outturns over the 1995-2015 period to assess whether the duration of stress test shocks was adequate. Specifically, the study first identified "events" where an outturn has fallen below a shock as used in the existing DSF (i.e., 10-year historical average minus one standard deviation). Results indicate that a shock duration of about 2 years remains broadly appropriate.

Table AI.4. Appropriateness of Shock	
	Persistence (years)
Ex-B1. Real GDP growth	1.3
Ex-B2. Exports	1.3
Ex-B4. Other flows	1.8
Ex-B6. Depreciation (>30%)	1.9
Pub-B2. Primary balance	1.8

Source: Fund staff calculations.

B. Macro-linkages in Stress Tests

10. **Interactions of key macro variables under the standardized stress tests were identified using event studies and complemented with the literature**. The event study analyzed the behavior of key variables – including FX nominal depreciation, real GDP growth, GDP deflator, and export growth – under the relevant economic shocks. To remove the impact of outliers and appropriately capture the characteristics of the interactions, events are defined as cases where an outturn falls near the relevant cutoff for shocks (i.e. its historical average minus one standard deviation). For example, events for the growth shock are defined as those where actual GDP growth falls between a historical average of minus 1.25 and 0.75 standard deviations. Based on this sample of events, elasticities were calculated. These calculations suggest that inflation decreases with an elasticity to real growth of 0.6 and real GDP growth falls with an elasticity to exports of 0.8. Exchange rate pass-through to the GDP deflator is estimated at 0.3 in the year of the shock (Table AI.5). Other interactions are taken from the literature (e.g. Aisen and Hauner (2008) on the domestic borrowing costs-primary balance interaction).

Table AI.5. Key Macro Interactions			
Type of shocks	**RGDP**	**GDP deflator**	**Exports**
Real GDP growth	-	0.6	-
Exports	0.8	-	-
Depreciation (+: depreciation)	-	0.3	0.2

Source: Fund staff calculations.

11. **Macro interactions would focus on first-round effects.** Rather than introducing all interactions under a given shock, which would render all scenarios akin to the combined shock, staff prefers to focus on key first-round effects to facilitate the interpretation of the responses to

the main shock. Proposed macro-interactions and re-calibrated standardized stress tests are reported in Table AI.6. a–b.

C. Tailored Stress Tests

12. **The design of tailored stress tests rests on a variety of specialized data sources.** To gauge appropriate triggers and scenario parameters, staff employed a variety of external databases to inform the design of these scenario stress tests. This section elaborates on the design of individual tailored stress tests.

a) Natural disasters[5]

- Trigger:
    - *Group 1*: Small-state LICs identified as vulnerable to natural disasters in IMF (2016). These are 14 countries: Comoros, Dominica, Grenada, Kiribati, Maldives, Micronesia, Samoa, São Tomé and Príncipe, Solomon Islands, St. Lucia, St. Vincent and the Grenadines, Tonga, Tuvalu, and Vanuatu.
    - *Group 2*: LICs that meet a *frequency* (around 2 disasters every 3 years) and *economic losses* (above 5 percent of GDP per year) criteria, based on the EM-DAT database during 1950-2015. These criteria lie between the 75th and 90th percentile of the respective distributions.[6] It identifies 9 countries: Bangladesh, Haiti, Honduras, Nepal, Madagascar, Mozambique, Myanmar, Nicaragua, and Tajikistan.
- Default scenario:
    - Size: A one-off shock to public debt of 10 percent of GDP in the first year of projection. This corresponds to the median change in the public debt to GDP ratio one year after the natural disaster from its pre-shock level, across all episodes with measured economic losses of at least 5 percent of GDP. This captures a range of fiscal strains, including increased spending inefficiencies and reconstructions needs, as well as, the materialization of contingent liabilities.
    - Interactions: Real GDP growth and exports are lowered by 1.5 and 3.5 percentage points, respectively, in the year of the shock, based on staff event analysis comparing the median growth during the year when the natural disaster took place and the median growth over the preceding 10 years, across the same sample used to identify the size of the shock (Table AI.7).

[5]Pandemics were excluded in the analysis for determining the default scenario settings given their rare and infrequent nature. Nevertheless, the tool could be customized to assess their impact where relevant.

[6]These distributions are highly skewed. For example, the median, 75th and 90th percentiles of the distribution of economic losses across all LICs is equal to 0.3, 5 and 10 percent of GDP, respectively.

Table AI.6. Proposed Changes to External and Public DSA

(a) Alternative Scenarios

Current External DSA	Current Public DSA	Staff's proposal for both External and Public DSAs
	Alternative scenarios (permanent shocks)	
A1. Historical Real GDP growth, GDP deflator, non-interest current account, and net FDI flows set to their historical averages 1/	**A1. Historical** Primary balance-to-GDP ratio and real GDP growth set to their historical averages	**A1. Historical** Real GDP growth, primary balance-to-GDP ratio, GDP deflator, non-interest current account, and net FDI flows set to their historical averages
A2. External financing External borrowing assumed to be less concessional (by 200 basis points)		Revised as a tailored stress test
	A2. Primary balance Primary balance-to-GDP ratio set to its value in the first year of the projection period	Drop, to streamline as already partially reflected in B2. Permanent shocks can be modeled using the customize scenario where relevant
	A3. Lower real GDP growth Real GDP growth lowered by a fraction of its standard deviation	Drop, to streamline as already partially reflected in B1. Permanent shocks can be modeled using the customize scenario where relevant

Table AI.6. Proposed Changes to External and Public DSA (concluded)

(b) Bound Tests

External and Public DSAs: Standardized Stress Tests

Current External DSA	Current Public DSA	Staff's proposal for both External and Public DSAs	Staff's proposal for interactions among variables
		Bound tests (temporary shocks)	
B1. Real GDP growth Real GDP growth set to its historical average minus one standard deviation	**B1. Real GDP growth** Real GDP growth set to its historical average minus one standard deviation	**B1. Real GDP growth** Real GDP growth set to its historical average minus one standard deviation, or the baseline projection minus one standard deviation, whichever is lower for the second and third years of the projection period	- Inflation to decrease with an elasticity to real growth of 0.6 2/ - Primary balance deteriorates as the revenue to-GDP ratio remains the same as in the baseline, but the ratio of non-interest expenditures to GDP increases as the level of spending is kept the same as in the baseline
	B2. Primary balance Primary balance-to-GDP ratio set to its historical average minus one standard deviation	**B2. Primary balance** Primary balance-to-GDP ratio set to its historical average minus one standard deviation, or the baseline projection minus one standard deviation, whichever is lower in the second and third years of the projection period	- Domestic borrowing cost to increase by 25 basis points per 1 percent of GDP worsening of the primary balance for the case of LICs with domestic market financing 3/; and - For market-access countries, external commercial borrowing cost to increase by 100 basis points per 1 percent of GDP worsening of the primary balance, or 400 basis points, whichever is lower 2/
B2. Exports Nominal export growth (in USD) set to its historical average minus one standard deviation		**B3. Exports** Nominal export growth (in USD) set to its historical average minus one standard deviation, or the baseline projection minus one standard deviation, whichever is lower in the second and third years of the projection	- Real GDP growth rate is lowered with an elasticity to exports of 0.8 2/
B3. Deflator Domestic GDP deflator (in USD) set to its historical average minus one standard deviation		Drop, as this bound test was rarely identified as the most extreme shock in past DSAs and overlaps with the depreciation shock	
B4. Other flows Current transfers-to-GDP and FDI-to-GDP ratios set to their historical average minus one standard deviation		**B4. Other flows** Current transfers-to-GDP and FDI-to-GDP ratios set to their historical average minus one standard deviation, or baseline projection minus one standard deviation, whichever is lower in the second and third years of the projection period	
B5. Combination of B1 through B4 Each variable set to its historical average minus half a standard deviation	**B3. Combination of B1 and B2** Real GDP growth and primary balance-to-GDP ratio set to their historical average minus half a standard deviation	**B6. Combination of B1 through B5** Apply all individual shocks (B1 through B5) at half of the magnitude in the second and third years of the projection period	Apply interactions in each individual shock scenarios
B6. Depreciation One-time 30 percent nominal depreciation of the domestic currency in the first year of the projection period	**B4. Depreciation** One-time 30 percent nominal depreciation of the domestic currency in the first year of the projection period	**B5. Depreciation** One-time 30 percent nominal depreciation of the domestic currency in the first year of the projection period, or the size needed to close the estimated real exchange rate overvaluation gap, whichever is larger	- Real net exports as a percent of GDP increases with an elasticity to real depreciation of 0.15, starting in the year following the shock 4/ - Pass-through to inflation with an elasticity of 0.3 in the year of the shock 2/
	B5. Other debt-creating flows One-time increase in other debt-creating flows amounting to 10 percent of GDP in the second year of the projection period	**B7. Contingent liability shock** One-time increase in other debt-creating flows amounting to 5 percent of GDP in the second year of the projection period 5/	

1/ Throughout this table, historical averages refer to averages over the last 10 years.
2/ Based on staff event analysis.
3/ Aisen and Hauner (2008) offers evidence for EMs.
4/ IMF (2015e) finds that a 10 percent real depreciation leads to 1.5 percentage point improvement in real net exports to GDP starting in the year following the shock in a sample of 60 developed and developing countries.
5/ Based on the average increase in debt-to-GDP ratios observed for 44 banking crisis episodes for LICs since the 1980s, as in Laeven and Valencia (2012).

Table AI.7. Event Analysis of Natural	
Sample:	Economic losses > 5 percent of GDP
Real GDP growth (median)	
Event year	2.2
10-year average	3.7
Sample size	47
Export growth (median)	
Event year	7.4
10-year average	10.8
Sample size	45
Source: Fund staff calculations.	

b) **Contingent liability supplement:** The re-calibrated standardized contingent liability shock would be supplemented with a tailored shock triggered when there are unaccounted dimensions of the non-financial public sector in the debt concept reported in the DSA (e.g., other parts of general government, public private partnership (PPP) exposures, or unaccounted SOE exposures) or when financial sector risks to the sovereign balance sheet are deemed larger than the size modeled under the standardized contingent liability shock.

- The general government component and financial sector risk supplement would be defined by DSF users based on country-specific circumstances (default shocks would be set to zero).

- Trigger for PPP shock:

  - PPP capital stock above 3 percent of GDP (equivalent to the median across LICs based on the WB's Private Participation in Infrastructure database over the period 1990–2015). This qualifies 26 countries: Benin, Bhutan, Burundi, Cape Verde, Cambodia, Rep. of Congo, Cote d'Ivoire, Djibouti, Dominica, Ghana, Grenada, Honduras, Lao PDR, Liberia, Mali, Mauritania, Mozambique, Nepal, Nicaragua, São Tomé and Príncipe, Senegal, St. Vincent and the Grenadines, Tajikistan, Togo, Uganda, and Zambia.

- Default scenario for PPP shock:

  - A permanent increase in external PPG debt equivalent to 35 percent of the country's PPP capital stock in the second year of projection (proxying for the present value of direct and potential future fiscal costs from PPP distress and/or cancellations). This estimate models a shock to ¼ of the PPP portfolio (assuming that the rest of the portfolio is not highly correlated) plus 10 percent of penalties. Recent experiences (e.g., Portugal) have shown that even in advanced economies where PPP portfolios have many contracts, concentration and correlation among them is significant. In the case of EMs and LICs where PPP portfolios have fewer contracts, this issue is more

relevant, meaning that one contract going bad could mean 70/80 percent of the portfolio being in distress.

- Trigger for SOE shock:

    - Based on a Fund staff survey conducted in 2016, LICs subject to this shock are those whose responses indicated a partial coverage or unavailable data of SOE external guaranteed debt. The new DSF template would continue monitoring SOE liabilities to inform this trigger going forward. This qualifies 37 countries: Afghanistan, Bangladesh, Benin, Cape Verde, Cambodia, Cameroon, Rep. of Congo, Djibouti, Dominica, Eritrea, Gambia, Haiti, Kyrgyz Republic, Lao PDR, Lesotho, Liberia, Madagascar, Malawi, Maldives, Mali, Mauritania, Moldova, Nepal, Papua New Guinea, Samoa, São Tomé and Príncipe, Senegal, Sierra Leone, Solomon Islands, Somalia, South Sudan, Sudan, Uganda, Uzbekistan, Yemen, and Zimbabwe.

- Default scenario for SOE shock:

    - One-off permanent increase of 2 percent of GDP in public debt, starting in the second year of projection, based on the median SOE PPG external liabilities.

c) Commodity Price Shock

- Trigger:

    - LICs whose commodity exports represent at least 80 percent of merchandise exports, based on data from UNCTAD (36 countries): Afghanistan, Benin, Burkina Faso, Cameroon, Central African Republic, Chad, Dem. Rep. of Congo, Rep. of Congo, Cote d'Ivoire, Eritrea, Ethiopia, Ghana, Guinea, Guinea-Bissau, Guyana, Kiribati, Lao PDR, Malawi, Maldives, Mali, Mauritania, Micronesia, Mozambique, Myanmar, Papua New Guinea, Rwanda, Solomon Islands, South Sudan, Sudan, Sierra Leone, Somalia, Tanzania, Timor-Leste, Yemen, Zambia, and Zimbabwe.

- Default scenario:

    - Size and duration: Export-to-GDP ratio is shocked by a commodity price gap in the first year of projection, which closes over 6 years.[47] The price gap is defined as the difference between the baseline commodity price in the first year of projection and a specific moment of the forecast price distribution, i.e., the lower end of the 68 percent confidence interval (equivalent to a minus one SD) from RES commodity price forecast distributions for fuel (WTI crude oil price projections, Figure AI.1) and non-

[6]Event analysis conducted by staff and evidence from IMF (2015c) show that commodity cycles tend to be persistent.

fuel (wheat and copper). The price gap for fuel and non-fuel exports is multiplied by their respective shares in total commodity exports.

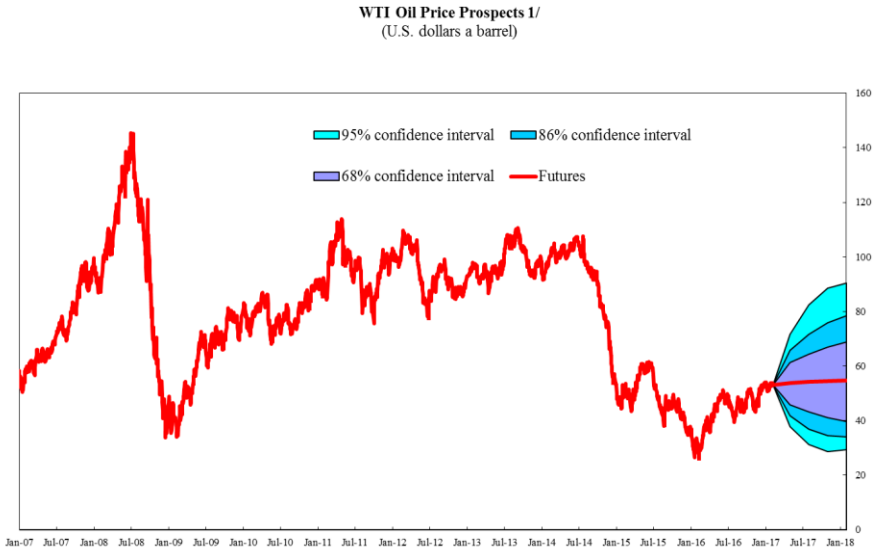
- Interactions:

  - Real GDP growth is reduced by 0.5 percentage points, and fiscal revenues-to-GDP are reduced by 0.75 percentage points in each of the first three years of projections for each 10-percentage point contraction of commodity prices. This gap converges to the baseline in 6 years. The size and duration of these responses were informed by the analysis of episodes of commodity price busts in a sample of 34 commodity-intensive LICs during 1990-2015.[48] The elasticities are within the range of estimates found in the literature (e.g., IMF (2012, 2015c), Spatafora and Samake (2012), Céspedes and Velasco (2013)), and in line with those used in the IMF's Vulnerability Exercise for LICs. Country teams may customize these default scenario parameters in close discussion with country authorities regarding the commodities for which the shock would be applied, the size of the price shock, and responses of real GDP growth and fiscal revenues to the price shock.

  - GDP deflator is reduced by the impact of the commodity price gap in the first year of the shock, converging to the baseline in 6 years.

[7]The exercise proceeded in four steps: (i) construct a country-specific commodity price index, defined as the weighted average of food, raw materials, metals and fuel, where the weights are given by the shares of the commodity groups in the country's exports; (ii) extract the cyclical component of the country-specific index using the HP filter (a longer sample 1960-2030 was used in this step to minimize tail problems); (iii) for each country, identify episodes of commodity price bust using the cyclical component, where a bust is defined as a fall below trend in the country-specific commodity price for at least two consecutive years and exceeding one (country-specific) standard deviation over the period 1990–2015; and (iv) study the responses of GDP growth and government revenues around the commodity price bust episodes.

Figure AI.1. Confidence Intervals of Commodity Price



Sources: Bloomberg L.P.; and IMF staff calculations.
1/ Derived from prices of futures options October 18, 2013.

d) Market-financing shock

- Trigger:
    - LICs that either: (i) are identified as "frontier" according to the prevailing IMF definition; (ii) have outstanding Eurobonds; and (iii) meet the market access criterion for graduation from the PRGT but have not graduated due to short-term vulnerabilities. These criteria identify 17 countries (Table AI.8).

Table AI.8. Countries Meeting Criteria for Market-financing Shock

Countries	Frontier LICs 1/	Past Bond issuers	PRGT Market Access Criteria 2/
Bangladesh	✓		
Cabo Verde			✓
Cameroon		✓	
Rep. of Congo		✓	✓
Cote d'Ivoire	✓	✓	✓
Ethiopia		✓	
Ghana	✓	✓	✓
Honduras		✓	
Kenya	✓	✓	
Maldives			✓
Mozambique	✓	✓	
Papua New Guinea	✓		
Rwanda		✓	
Senegal	✓	✓	
Tanzania	✓	✓	
Uganda	✓		
Zambia	✓	✓	

1/ As defined in the IMF (2014) LIDC report.

2/ These are countries that meet the market access criteria as in IMF (2015), but did not graduate from PRGT given their short-term vulnerabilities.

- Default scenario:
    - Size and duration: Based on a staff event analysis of the median responses around external debt distress episodes during 1995-2015, the scenario consists of a 400 bps increase (sustained for 3 years) in the cost of new external commercial borrowing, FX depreciation equivalent to 15 percent, and shortening of maturities of new commercial external borrowing (to 5-year maturity, with grace periods adjusted according to the type of instrument).

Annex II. Select Case Studies

1. **The DSF has been successful in signaling impending debt difficulties in several country cases.** Examples include Sri Lanka (downgraded to high risk of debt distress in 2008, a year ahead of a request for a Fund program with exceptional access to help restore debt sustainability); Chad (rated at high risk of debt distress since 2012, undertaking a debt restructuring with a major commercial creditor in 2015 and currently running sizable external arrears); and Mongolia (downgraded to high risk of debt distress in 2015, with subsequent debt service difficulties and a 2017 request for official financial support and a debt exchange).

2. **Other country cases, however, point to areas in which the framework needs strengthening**:

- *Securing stronger baseline debt projections.* Optimism in macro projections, particularly on growth and fiscal adjustment, has impeded the DSF from providing early warnings of debt distress in some cases. In Ghana, the downgrade to high risk of external debt distress in 2015 was partly driven by lower-than-projected growth and large fiscal slippages vis-à-vis projections. Similarly, large fiscal deviations were behind the downgrades in St. Lucia (2011), Dominica (2014), and Maldives (2016). Public investment scaling-up also seems to have introduced macroeconomic forecast risk (e.g. Djibouti and Ghana, 2015) pointing to a need for greater scrutiny of assumptions in this area.

- *Stress tests.* The existing standardized stress tests have not captured the relevant risks for some countries. For example, while market-financing risks in Sri Lanka and Ghana were flagged in the DSA narrative (2006-07 and 2012-14, respectively), the impact of market-related stress was not simulated in the DSAs. Similarly, as there are no dedicated tools to analyze debt sustainability risks from natural disasters, commodity price shocks, or contingent liabilities, such risks were only infrequently analyzed in countries vulnerable to such shocks (e.g., natural disasters in Samoa and commodity price shock in Malawi).

- *Determination of risk ratings.* First, in some cases, identified risks have not appropriately informed risk ratings. Maldives (2015) was deemed to be at moderate risk of debt distress, when there was substantial uncertainty regarding the accuracy of data and rising domestic debt vulnerabilities; these concerns soon were reflected in an external debt problem. For Ghana, high public gross financing needs (GFN) projected since 2010 did not inform the risk rating, which was maintained at moderate until 2014. Second, steady erosion of debt positions over time (while remaining within the moderate risk category) may not be given adequate attention in describing debt vulnerabilities (as in Central African Republic 2010-12, Samoa 2010-12, and Ghana 2010-13).

- *Factors not captured in the present model.* Cases where the DSF predicted crises that did not occur ("false alarms") or are yet to occur seem to be linked to a combination of factors, including: (i) limitations of the framework in accounting for country-specific factors (e.g., Sao Tome and Principe or Burundi, countries with relatively comfortable reserve coverage); and

(ii) countries with specific features (e.g., post-civil conflict) that cannot be automatically brought into a formal framework (e.g., Afghanistan).

Table AII.1. Select Case Studies

(a) Country Examples of Missed Calls

Country	Crisis Trigger	Hints in the DSA	Problem	How DSA review would address them?
Ghana (2012-2014)	Accumulation of large fiscal and external imbalances led to a slowdown in growth. Efforts to achieve fiscal consolidation in mid-2013 were undermined by policy slippages, external shocks and rising interest cost. By mid- 2014, the net international reserves position had further weakened and the exchange rate depreciated sharply, fueling inflationary pressures. By then public debt continued to rise at an unsustainable pace.	DSAs prior to Ghana's downgrade to high risk emphasized several of the risks that ultimately triggered the dowgrade: i) the risks from growing GFN; ii) vulnerabilities from shift in financing mix away from concessional sources; iii) risks resulting from greater participation of foreign investors in the domestic debt market and the issuance of Eurobonds.	i) Baseline oversight: significant forecast errors driven by optimism in the forecast of growth, primary balance, exports and exchange rate; ii) Unanticipated shocks: standard stress tests unable to anticipate the large exchange rate depreciation or the primary balance shock (both magnitude and spillovers to market financing cost); iii) Lack of market risk analysis: debt service indicators did not fully capture rollover risks; stress tests did not flag risks from rising borrowing costs in advance.	Realism tool to flag inconsistency in key macro projections, including on ambitious growth projections and fiscal consolidation plans. Size of the standardized primary balance shock will be strenghtened, following staff's event anlaysis. Revised debt service theresholds are now tighter to reflect greater liquidity risks in LICs' debt portfolio.
Malawi (2010-2012)	Over 2010-2012, chronic BOP difficulties due to shortages of foreign exchange and critical imports, coupled with increased government expenditure at a time when external grants fell sharply and revenue performance deteriorated, led to a significant increase in public debt. The Kwacha was devalued and left to float, among a mix of other policy measures under a new ECF in 2012.	The 2010 DSA cautioned of the vulnerability to a shock on tobacco prices, the country's main export commodity. A country-specifc shock simulating a permanent one-third fall in tobacco prices projected a breach in the debt to exports ratio.	The 2010 DSA did not envisage the decline in external grants, which together with a drop in tobacco prices led to a siginificant deterioration in public debt. While the DSA flagged risks from increasing domestic debt burden, these were subsumed by optimistic fiscal assumptions and ultimately did not inform the risk assesment.	Introduction of a commodities price stress test to formally model the debt and growth outlook for commodity exporters under adverse prices. Realism tool to flag inconsistency in key macro projections, including on ambitious growth projections and fiscal consolidation plans. Incorporation of signal from total public debt benchmark in the external risk rating.
Mozambique (2009-2014)	Mozambique was under the PSI when it requested a 12-mo. External Shock Facility to forestall a projected large decline in reserves due to reduced FDI and other capital flows following the global financial crisis. During this time, the government significantly scaled up investments financed by nonconcessional loans. By 2015, the economy was suffereing from fiscal slippages and a series of external shocks resulting in a large depreciation. Debt service to export was projected to almost double in three years. A Standby Credit Facility was requested to supplement the PSI.	Although the risk rating remained low from 2009 to 2012, the debt projection was continually revised upward due to the heavy public investment schedule. Breaches of the PV of debt/GDP threshold for 8 years in the most extreme scenario were deemed temporary. In 2013, downgrade to moderate risk due to (i) lower discount rate, (ii) increase in NCB to finance public investment, and (iii) deterioration in the current account due to natural gas exploration. In 2015, the risk rating remained moderate, but risks from nonconcessional borrowing were highlighted.	During this time, even though the medium term debt service projection was continually revised up, none of the debt service indicators breached their thresholds. In addition, while breaches to the PV debt/GDP ratios appeared in every DSA, the use of judgement for rataining the low risk rating may have unduly delayed the downgrade to moderate risk, therby providign a flase sense of comfirt regardin the true extent of debt vulnerabilities.	Revised debt service theresholds are now tighter to reflect evolving financing landscape. A new market risk module would suggest early warning benchmarks on gross financing needs to better capture rollover risks. Redesign of standard stress test framework, in particular having the primary balance shock factored in the assessment of the external risk of debt distress.

Table AII.1. Select Case Studies (continued)

(b) Country Examples of False Alarms

Country	Trigger of false alarm	Problem	How DSA review would address them?
Burkina Faso (2008-2011)	Burkina Faso's CPIA worsened to medium in 2008, tightening debt thresholds, and leading to the high risk rating. In 2012, revised projections for gold export volumes moved CPIA back to moderate.	The worsening of the CPIA immediately tightened thresholds, leading to a worseing of the risk rating without any evident deterioration in the macroeconomic outlook.	Introduction of a composite indicator, to reflect a range of country specific characteristics such as growth, and international reserves, reducing reliance on the CPIA score.
Sao Tome and Principe (2008-2012)	Despite reaching HIPC completion in 2007, risk rating worsened to high in 2008 as delays in oil production and exports deteriorated the macro outlook. The rating remained high as delays in oil production where compounded by the narrow export base and prospects of less concessional financing.	Breaches under the baseline were primarily in stock indicators (most notably debt-to-exports), while debt service inidcators remained comfortably below their respective thresholds. Every DSA vintage projected a rapid and significant drop in the stock ratios once oil revenues were assumed to come into stream.	The composite indicator will help factor in country specific circumstances (such as reserve level) to better reflect a country's capacity to repay. Recalibrated thresholds will allow for higher stock indicators, adressing some of the concerns for countries with narrow export bases.
Burundi (2009-2012)	Burundi's narrow export base is its most significant vulnerability. Breaches in the debt-to-export ratio in the baseline has put the country at high risk of debt distress since 2009, even after receiving substantial debt relief. All other baseline indicators have been below their thresholds.	Countries with a narrow export base often see protracted breaches in the PV of debt-to-export threshold even if the countries are not headed into debt distress.	Recalibrated thresholds will allow for higher stock indicators (while debt service thresholds tighter) to better balance the signals coming from stock and flow indicators.

Table AII.1. Select Case Studies (concluded)

(c) Other Country Examples: The Case of Incorporating Megaprojects and Anticipating Commodity Performance

Country	Developments over the period	Hints in the DSA	Problem	How DSA review would address them?
Mongolia (2009-2015)	Macro projections were consistently optimistic compared to actual fiscal outturn and mining sector developments (timing of mining projects and export revenues), resulting in significant forecast errors in most debt stock indicators. Baseline assumptions were continually revised leading to risk rating deterioration-- to moderate in 2014, high in 2015. By the 2015 DSA, baseline assumptions differed greatly from the previous DSA, reflecting more realistic fiscal and mining revenue assumptions, and the borrowing required to meet the BOP gap. Unanticipated large debt issuances by the Development Bank of Mongolia contributed to the downgrades.	In 2014, the team prepared a DSA with a weak and strong policy scenarios. Though significant risks to external debt sustainability were illustrated by the weak policy scenario in the DSA, the more optimistic scenario was taken as the baseline. In addition, while alternative scenarios in the public DSA illustrated a steadily rising trend of public debt ratios, underscoring the importance of fiscal assumptions for the projected debt path, this was not used to inform final risk ratings.	i) Baseline oversight: Projections in the DSAs did not envision delays in mining projects, or fiscal slippages vis-a-vis the newly legislated fiscal framework; ii) Unanticipated shocks: off-budget borrowing accounted for much of the the increase in total public debt. Limited debt coverage limits the team's ability to account for these risks; iii) Inability to factor in fiscal risks in external risk rating: even though the public DSA signaled significant risks, lack of primary balance shock analysis in the external DSA limited the link between this risk and the external risk ratings.	Realism tool to flag inconsistency in key macro projections, including ambitious growth projections and fiscal consolidation plans. Redesign of standard stress test framework, in particular with the primary balance shock factored in the assessment of the external risk of debt distress. Incorporation of signal from total public debt benchmark in the external risk rating. Introduction of a commodity price stress test to formally model the debt and growth outlook for commodity exporters under adverse prices.

Annex III. Update of the Core Debt Distress Model and Re-estimation of Debt Thresholds and Benchmarks

1. **This annex provides details on the reforms to the DSF aimed at improving its predictive performance.** In particular, it covers the following:

- Revised methodology for identifying episodes of external debt distress.
- Enhanced specification of the core debt distress model.
- Improving the framework for determining external risk ratings.
- New approach to estimate total public debt benchmarks.

A. Identification of External Debt Distress Episodes

2. **A revised methodology to identify external debt distress episodes aims to better capture the diversity of debt crises across LICs and debt difficulties associated with liquidity pressures** (Figure AIII.1). Specifically, the revised methodology: (i) allows for a more flexible assumption on the duration of debt distress episodes; (ii) introduces a new set of distress signals; (iii) provides a better treatment of distress signals based on IMF disbursements; and (iv) proposes a better treatment of debt restructuring signals, including those related to the HIPC Initiative.

Figure AIII.1. Existing and Revised Methodology for Identifying External Debt Distress

Signals of debt difficulties	Existing definition	New definition
Arrears on PPG external debt	If larger than 5 percent of PPG external debt should last at least 3 consecutive years to qualify as an episode	If larger than 5 percent of PPG external debt for at least 3 consecutive years
Restructuring of Paris Club debt	Assumed to last for 3 consecutive years for both HIPC and non-HIPC debt treatments	Assumed to start the year before the debt deal (negotiation phase) No assumption on duration for non-HIPC debt treatments HIPC treatments are bunched in a single episode
Large IMF programs	Cumulative disbursements > 50 percent of quota Financing under SBAs and EFFs	Disbursements in the first 6 months > 30 percent of quota SBAs and EFFs and LIC facilities (ECFs, SCFs, PRGFs)
Restructuring of commercial debt		Assumed to start the year before the debt deal (negotiation phase) No assumption on duration
Defaults on PPG external debt		No assumption on starting point or duration

3. **Duration of debt distress episodes.** To better reflect the changing financing landscape facing LICs, the revised methodology would no longer require observing distress signals for at least three consecutive years to identify debt distress episodes. Episodes that last only one or two years would be included, except when the episode is solely driven by the occurrence of arrears, for which at least three consecutive years would continue to be required (this helps rule out episodes associated with one-off and temporary occurrence of external arrears, including for technical reasons, that do not necessarily signal debt distress). Thus, *an episode of debt distress is*

defined as a period in which at least one distress signal is observed, whereas a non-distress episode continues to be defined as a non-overlapping three-year period in which none of the distress signals is observed.

4. **Introducing a richer set of distress signals.** Taking advantage of recent debt restructurings databases, the new methodology expands the set of distress signals to include restructurings of public external debt held by private foreign creditors (Cruces and Trebesch, 2013) as well as instances of outright defaults (Standard & Poor's and Catão and Milesi-Ferretti, 2014). Stand-alone restructurings of commercial debt would help identify crisis episodes in countries with increasing reliance on market financing, whereas outright defaults would help better predict the onset of a crisis when they preceded the occurrence of other distress signals such as the emergence of external arrears.

5. **Better treatment of signals based on IMF disbursements.** Rather than considering large IMF disbursements on a cumulative basis as in the current methodology, which may issue false alarms or fail to properly capture the onset of debt difficulties, the proposed criterion focuses on large upfront financing disbursements as they are generally associated with situations where countries are facing large financing needs and have no alternative financing sources. This is measured as IMF disbursements during the first six months of a Fund-supported program that are larger than 30 percent of quota, including both GRA and PRGT disbursements. This cutoff is twice as large as the current "access norm" to Fund resources (i.e., the typical size of financing), above the median "access norm" since late 1980s, and corresponds to the 75th percentile of disbursements to LICs during the first six months of Fund-supported programs under the GRA and PRGT since 1970.[1]

6. **Better treatment of debt restructurings.** To capture typical negotiation delays before the settlement of a restructuring deal and the fact that the decision to restructure often comes too late, it is assumed that debt difficulties leading to debt restructurings start in the year preceding the restructuring deal. In addition, it is also proposed to treat multi-round restructurings between decision and completion points under the HIPC Initiative as a single episode of debt distress, as in general all the different rounds are associated with the same debt problem.

7. **Together, the proposed changes identify a larger and more diverse set of debt distress episodes in LICs.** The revised methodology identifies a total of 98 debt distress episodes in LICs (compared to only 76 under the existing methodology) between 1970 and 2015, of which about a third was triggered by IMF signals, a third by the occurrence of external arrears, and a third by debt restructurings, defaults or a combination of signals (Table AIII.1). The distribution of these episodes over time is in line with what the literature has found for emerging economies, i.e., some clustering in the 1980s, a smaller incidence during the commodity boom of the 2000s, and a spike

[1]To take into account the 2009 reform of the Fund's financial facilities for LICs, as well as blend financial assistance, disbursements of Poverty Reduction and Growth Trust (PRGT) resources under the Extended Credit Facility (ECF) and the Stand-by Credit Facility (SCF) are also included. For consistency, their predecessors are also considered but rarely binding.

at the onset of the Global Financial Crisis (10 episodes of external debt distress are identified since 2008 under the proposed new methodology, while only one would be captured under the existing methodology). Relaxing the assumption on duration results in an average duration of debt distress episodes of just over a decade (compared to 18 years under the existing methodology).

Table AIII.1. External Debt Distress Episodes in LICs, 1970–2015

Signal	Old definition	New definition
Number of episodes	**76**	**98**
Of which triggered[1/] by:		
IMF disbursements	7	35
Arrears	51	32
Defaults		1
Restructurings	15	22
Some combination of the above 2/	3	8
Median duration in years 3/	**18**	**12**

1/ A signal is considered a "trigger" when it is the only one observed in the first year of an episode.
2/ Multiple signals are observed in the first year of an episode.
3/ For the purposes of this calculation, crises that are less than three years apart from the initial episode are treated as a single episode.

II. Enhanced Specification of the Core Debt Distress Model

8. **An expanded specification of the core model is warranted to better predict external debt distress using the revised identification methodology.** The existing model predicts debt distress less well under the revised identification methodology. Thus, the specification of the probit model is expanded to include two important proxies of capacity to repay—international reserves (scaled by imports) and remittances (scaled by nominal GDP)—as well as a proxy for global shocks—world growth (see Box AIII.2 for description of the data).[2] The addition of these new explanatory variables strengthens the model's capacity to predict external debt distress. These additional controls have been used extensively in the literature on sovereign debt crises and sovereign borrowing, including Manasse and others (2003), Manasse and Roubini (2005), Panizza and others (2009), Gelos and others (2011), and Catão and Milesi-Ferretti (2014).

9. **The expanded specification of the core model provides a better description of the data.** The new model is estimated in a dataset that pools multiple episodes for each country. The estimation used a sample covering 80 LICs during 1970–2014 (see detailed data description in Box AII.2). The estimated probit models have reasonable statistical and economic properties (Table AIII.2). The likelihood of debt distress is positively correlated with the level of indebtedness, and negatively correlated with the quality of institutions and policies (measured by the CPIA), and with other country-specific factors (country growth, reserves, remittances). Favorable external conditions (world growth) exert an important impact on the probability of debt

[2] Reserves are included with a linear and a quadratic term to pick up plausible nonlinearities in the relationship between reserves and debt distress. Also, see Box AIII.3 for a discussion of issues on reserve measurement in currency unions.

distress. Note that the inclusion of world growth in the regressions tends to reduce the statistical significance of domestic growth, reflecting the synchronization of business cycles across countries. Also, note that the reserves' non-linear term has the expected positive sign, implying that above certain coverage, additional accumulation of reserves contributes less to reducing the probability of debt distress.[3][4]

Table AIII.2. Probit Regressions, Summary of Baseline

Controls	PV of debt as a percent of		Debt service as a percent of	
	GDP	Exports	Revenue	Exports
Debt burden indicator	1.541***	0.359***	3.745***	3.541***
CPIA	-0.400***	-0.381***	-0.362***	-0.395***
Domestic growth	-3.081*	-2.853*	-3.001	-1.942
Reserves	-4.223***	-4.591***	-3.696**	-3.699***
Reserves^2	3.953*	4.582**	3.743	3.683*
Remittances	-2.235**	-2.282***	-1.635*	-1.934**
World growth	-12.40***	-14.09***	-13.84***	-13.75***
Constant	1.310***	1.331***	0.979*	1.148**
Observations	409	403	343	380
Pseudo R-squared	0.169	0.174	0.190	0.184
Log-likelihood	-150.1	-145.7	-116.8	-135.2
BIC	348.4	339.5	280.2	317.9

***, **, * denote significance at 1%, 5%, and 10%, respectively, based on robust standard errors.

10. **Overall, the new specification of the core model outperforms the existing one.** Table AIII.3. a–b compares the new model specification (last column in each table) with the model estimated in the 2012 review (second column), the old model re-estimated with the updated data (third column), and the same model re-estimated with the updated data but excluding the sample of MICs and adding the new set of debt distress episodes. Despite the smaller number of observations in the new model (a result of removing the MICs sample), it tends to fare better in terms of goodness of fit (i.e., R-squared, log-likelihood, and BIC).

[3] The estimated non-linear term for reserves does not allow for an unbounded contribution to the composite indicator from reserve accumulation (implying a positive but decreasing marginal contribution up to about 6 months of imports, covering the common range of reserve adequacy assessments for LICs, i.e. 3-5 months of imports).

[4] Note that the table does not present results for the PV of debt-to-revenue as staff's analysis confirms the indicator does not contribute to improve the framework's predictive performance (its inclusion or exclusion leaves broadly unchanged type I and II errors).

Table AIII.3. Probit Regressions

(a) Debt Stock Indicators

	PV of debt-to-GDP					PV of debt-to-exports			
Sample period	1970-07	1970-14	1970-14	1970-14	Sample period	1970-07	1970-14	1970-14	1970-14
Sample of countries	LICs+MICs	LICs+MICs	LICs	LICs	Sample of countries	LICs+MICs	LICs+MICs	LICs	LICs
Definition of debt distress	Old	Old	New	New	Definition of debt distress	Old	Old	New	New
Model specification	Old	Old	Old	New	Model specification	Old	Old	Old	New
Controls					Controls				
Debt burden indicator	1.912***	1.763***	1.436***	1.541***	Debt burden indicator	0.319***	0.335***	0.294***	0.359***
Debt burden indicator x MIC	0.232	0.396			Debt burden indicator x MIC	0.0511	0.259***		
CPIA	-0.603***	-0.572***	-0.532***	-0.400***	CPIA	-0.421***	-0.593***	-0.475***	-0.381***
Domestic growth	-6.136***	-4.414***	-4.547***	-3.081*	Domestic growth	-7.447***	-3.886***	-4.256***	-2.853*
Reserves				-4.223***	Reserves				-4.591***
Reserves^2				3.953*	Reserves^2				4.582**
Remittances				-2.235**	Remittances				-2.282***
World growth				-12.40***	World growth				-14.09***
Constant	0.562*	0.435	0.579	1.310***	Constant	0.0433	0.448	0.403	1.331***
Observations	740	893	455	409	Observations	557	894	445	403
Pseudo R-squared	0.188	0.156	0.108	0.169	Pseudo R-squared	0.175	0.178	0.0893	0.174
Log-likelihood	-238	-276.4	-183.1	-150.1	Log-likelihood	-175.4	-269.3	-177.9	-145.7
BIC	509	586.8	390.7	348.4	BIC	382.4	572.5	380.1	339.5

***, **, * denote significance at 1%, 5%, and 10%, respectively, based on robust standard errors.

***, **, * denote significance at 1%, 5%, and 10%, respectively, based on robust standard errors.

(b) Debt Service Indicators

	Debt service-to-revenue					Debt service-to-exports			
Sample period	1970-07	1970-14	1970-14	1970-14	Sample period	1970-07	1970-14	1970-14	1970-14
Sample of countries	LICs+MICs	LICs+MICs	LICs	LICs	Sample of countries	LICs+MICs	LICs+MICs	LICs	LICs
Definition of debt distress	Old	Old	New	New	Definition of debt distress	Old	Old	New	New
Model specification	Old	Old	Old	New	Model specification	Old	Old	Old	New
Controls					Controls				
Debt burden indicator	4.877***	3.174***	4.739***	3.745***	Debt burden indicator	3.398***	3.768***	3.704***	3.541***
Debt burden indicator x MIC	-1.730**	-0.900			Debt burden indicator x MIC	-0.105	-0.479		
CPIA	-0.365***	-0.409***	-0.416***	-0.362***	CPIA	-0.448***	-0.461***	-0.430***	-0.395***
Domestic growth	-7.386***	-4.500***	-2.724	-3.001	Domestic growth	-5.550***	-3.455***	-2.553	-1.942
Reserves				-3.696**	Reserves				-3.699***
Reserves^2				3.743	Reserves^2				3.683*
Remittances				-1.635*	Remittances				-1.934**
World growth				-13.84***	World growth				-13.75***
Constant	-0.434	-0.0998	-0.168	0.979*	Constant	-0.00442	-0.0684	0.0593	1.148**
Observations	543	688	361	343	Observations	659	762	408	380
Pseudo R-squared	0.224	0.146	0.151	0.190	Pseudo R-squared	0.186	0.182	0.111	0.184
Log-likelihood	-154.4	-209.5	-127.9	-116.8	Log-likelihood	-203.2	-221.4	-154.5	-135.2
BIC	340.3	451.7	279.3	280.2	BIC	438.8	475.9	333.1	317.9

***, **, * denote significance at 1%, 5%, and 10%, respectively, based on robust standard errors.

***, **, * denote significance at 1%, 5%, and 10%, respectively, based on robust standard errors.

11. **The new model survives relevant robustness checks**. Staff subjected the new model to several robustness exercises. Table AIII.4 shows some of the results.

- Subsample 1980-2014. Staff re-estimated the model excluding the 1970s where data had lower quality and LICs were less integrated into the world economy. All coefficients retain their expected sign but CPIA and remittances lose statistical significance in one model each.

- *Adding real GDP per capita*. Higher income levels are typically associated with better policies and institutions and higher capacity to repay. Hence, one would expect that the CPIA would lose statistical significance by including real GDP per capita in the regressions. However, this variable comes with the wrong sign and is not generally significant. The coefficients on the CPIA and on the other variables remain statistically significant and are somewhat larger than in the baseline.

- *Adding trade openness*. Staff considered trade openness, as it has been found by the literature to be a determinant of the probability of default in emerging economies (e.g., Panizza and others, 2009). In the specifications without remittances, openness was statistically significant, in line with the literature. However, it tends to lose statistical significance when considered jointly with remittances, in part reflecting their high correlation, and the baseline model with only remittances has a better predictive performance.[5] Furthermore, it is less clear that openness is a good measure of capacity to repay in small open developing states with undiversified production bases. In these countries, openness typically reflects large imports, which are inelastic (e.g., fuel, food, and raw materials) and thus more difficult to adjust in times of distress. Staff also tested whether including an alternative definition of openness (defined as exports + imports net of FDI, scaled by GDP) would weaken the statistical significance of remittances, given the strong correlation between these two variables. While remittances are no longer significant in two models (PV of debt-to-GDP and debt service-to-revenue), openness is not significant in any of the models.

- *Replacing remittances with a broader measure of FX income*. Replacing remittances with a broader measure of foreign exchange, defined as exports plus remittances, yields results that are comparable to the baseline. Despite favorable goodness of fit, the added variable is not significant in one model (debt service-to-exports).

- *Replacing world growth with a proxy for country risk premium*. Staff also tested whether global rates are more relevant than world growth to predict debt distress in LICs. Staff considered Moody's corporate Baa spread over the U.S. 10-year treasury as a proxy for country risk premium. Cruces and Trebesch (2013) used similar measures to proxy for borrowing conditions facing developing countries. The estimated coefficient has the expected sign but, interestingly, it is statistically significant in the specifications that include debt service indicators but not in those including debt stock variables. However, when this variable is included together with world growth, it is not statistically significant, while world growth remains economically and statistically relevant across the four probit regressions.

- *Adding an indicator of conflicts*. To test whether conflict helps to predict debt distress in LICs, staff added to the baseline a dummy variable measuring armed conflicts. This variable has the expected sign in three of the four models but is not statistically significant.

[5]Better predictive performance refers to a lower in-sample weighted sum of type I and II errors.

Table AIII.4. Probit Regressions, Robustness Checks

Control variables	New baseline	Subsample 1980-14	Real GDP per capita	Openness (X+M-FDI)	FX income (exp.+rem.)	Risk premium	Armed conflicts	New baseline	Subsample 1980-14	Real GDP per capita	Openness (X+M-FDI)	FX income (exp.+rem.)	Risk premium	Armed conflicts
	PV of debt-to-GDP							**PV of debt-to-exports**						
Debt burden indicator	1.541***	1.498**	1.406**	1.884***	1.761***	1.453***	1.529***	0.359***	0.315**	0.437***	0.319**	0.297**	0.338***	0.364***
Institutions	-0.400***	-0.227	-0.425***	-0.407***	-0.377***	-0.397***	-0.400***	-0.381***	-0.222*	-0.471***	-0.415***	-0.377***	-0.382***	-0.379***
Domestic growth	-3.081*	-3.353*	-3.363**	-2.437	-3.041*	-3.108*	-3.041*	-2.853*	-3.151	-2.835	-2.936	-2.996*	-2.755	-2.856*
Reserves	-4.223***	-4.150***	-4.148***	-4.086***	-4.549***	-4.533***	-4.227***	-4.591***	-4.716***	-4.473***	-4.527***	-5.214***	-5.028***	-4.616***
Reserves^2	3.953*	4.457**	3.880*	4.150*	4.263**	4.130**	3.956*	4.582**	5.229***	4.393**	4.973**	5.325***	4.876**	4.623**
Remittances	-2.235**	-1.608*	-2.369**	-1.684		-2.091**	-2.227**	-2.282***	-1.678*	-2.663**	-2.284**		-2.164**	-2.287***
World growth	-12.40***	-18.30***	-12.41***	-14.50***	-14.73***		-12.29***	-14.09***	-18.47***	-14.54***	-14.01***	-14.95***		-14.24***
Log GDP per capita			0.0842							0.278**				
Openness				-0.450							0.00913			
FX income					-1.463***							-0.843**		
Risk premium						13.86							19.41	
Dummy for conflicts							0.119							-0.230
Constant	1.310***	0.782	0.796	1.515***	1.673***	0.605	1.301***	1.331***	0.888*	-0.574	1.434***	1.656***	0.471	1.336***
Observations	409	324	383	357	403	409	409	403	318	377	351	395	403	403
Pseudo R-squared	0.169	0.151	0.173	0.177	0.191	0.156	0.170	0.174	0.153	0.189	0.165	0.174	0.160	0.174
Log-likelihood	-150.1	-113.8	-141.6	-127.6	-144	-152.5	-150.1	-145.7	-111.3	-135.4	-127.3	-141.7	-148.2	-145.6
BIC	348.4	273.9	336.8	308.2	336.1	353.2	354.3	339.5	268.7	324.2	307.3	331.3	344.4	345.3
	Debt service-to-revenue							**Debt service-to-exports**						
Debt burden indicator	3.745***	4.238***	3.753***	4.322***	4.201***	3.858***	3.757***	3.541***	3.645***	3.972***	4.164***	3.211***	3.674***	3.535***
Institutions	-0.362***	-0.357**	-0.418***	-0.328**	-0.321***	-0.363***	-0.359***	-0.395***	-0.320**	-0.502***	-0.490***	-0.377***	-0.403***	-0.396***
Domestic growth	-3.001	-2.593	-3.978**	-1.349	-3.099	-3.441*	-2.925	-1.942	-1.910	-1.883	-1.062	-2.354	-2.184	-1.928
Reserves	-3.696**	-3.443*	-3.671**	-3.873**	-3.886**	-3.957**	-3.684**	-3.699***	-3.439**	-3.430**	-3.608**	-4.260***	-4.055***	-3.694***
Reserves^2	3.743	3.914	3.741*	4.512*	3.926*	3.845*	3.717	3.683*	3.755	3.395*	4.065*	4.315*	3.922*	3.678*
Remittances	-1.635*	-1.342	-1.787*	-1.240		-1.753*	-1.629*	-1.934**	-1.657*	-2.127**	-2.056**		-1.990**	-1.932**
World growth	-13.84***	-17.05***	-15.27***	-17.33***	-16.31***		-13.69***	-13.75***	-16.57***	-14.33***	-14.21***	-14.44***		-13.71***
Log GDP per capita			0.0922							0.263*				
Openness				-0.119							0.167			
FX income					-1.199***							-0.462		
Risk premium						26.47**							25.61**	
Dummy for conflicts							0.204							0.0729
Constant	0.979*	0.867	0.550	0.868	1.201**	-0.0733	0.949*	1.148**	0.872	-0.575	1.256**	1.289**	0.150	1.145**
Observations	343	288	328	308	338	343	343	380	304	356	337	373	380	380
Pseudo R-squared	0.190	0.187	0.208	0.207	0.217	0.184	0.191	0.184	0.168	0.200	0.194	0.181	0.176	0.185
Log-likelihood	-116.8	-94.35	-110.9	-101.6	-110.9	-117.7	-116.7	-135.2	-104.6	-125.3	-119.6	-132	-136.6	-135.2
BIC	280.2	234	274	254.7	268.5	282.2	285.9	317.9	255	303.6	291.5	311.4	320.8	323.8

***, **, * denote significance at 1%, 5%, and 10%, respectively, based on robust standard errors.

III. Improving the Framework for Determining External Risk Ratings

Approach to Deriving Debt Thresholds

12. **Debt thresholds should be derived in a manner that is consistent with the DSF's aggregation rule to determine risk signals.** In the 2012 review, thresholds were derived individually, without regard to the information content in other debt burden indicators for predicting external debt distress, thus likely introducing conservative bias into the framework (see Box AIII.1). To overcome this issue, staff proposes an algorithm to derive debt thresholds jointly for all debt burden indicators, taking as given a transparent policy choice for the weights of missed crises and false alarms. This is fully consistent with the way the framework is used in practice, i.e., any single breach of a debt threshold under the baseline scenario signals high risk of external debt distress, whereas any single breach under stress tests signals moderate risk. Otherwise, countries with no breaches under the baseline and stress tests are signaled as low risk.

Box AIII.1. Existing Approach for Deriving Debt Thresholds

The existing approach derives debt thresholds individually (i.e., without regard to the information of other debt burden indicators to predict external debt distress). The process can be broken down into the following three steps:

Step 1: Estimate a probit model for each debt burden indicator using a sample of LICs and MICs during the period 1970-2007. The 2012 model included an interaction term to allow the coefficient on the debt burden indicator to vary between LICs and MICs:

$$P(y = 1|controls) = \Phi(\beta_0 + \beta_1 debt\ burden + \beta_2 MIC * debt\ burden + \beta_3 growth + \beta_4 CPIA)$$

where $P(.)$ is the probability of external debt distress; y is a binary variable that takes the value of 1 if the country experiences external debt distress and zero otherwise; Φ is the standard normal cumulative distribution function; β's are the coefficients to be estimated; "debt burden" is the corresponding debt burden indicator; "MIC" is a dummy variable that takes the value of 1 for MICs and zero otherwise; "growth" is the real GDP growth; and "CPIA" is the World Bank's measure of the quality of policies and institutions.

Step 2: Estimate an optimal cutoff probability of external debt distress *for each probit model individually*. Each optimal cutoff probability was calculated in four steps:

(i) calculate the fitted probability of debt distress, $\hat{P}$, using the estimated model and data for each episode;

(ii) classify episodes as distress or non-distress by comparing $\hat{P}$ with candidate cutoff probabilities, $\bar{P}$, such that a distress (non-distress) episode is predicted correctly whenever $\hat{P} \geq \bar{P}$ ($\hat{P} < \bar{P}$);

(iii) compare the classification generated in (ii) with actual outcomes to determine type I (failure to predict distress that occurred) and type II (incorrectly predicted distress when it did not occur) errors; and

(iv) select the optimal cutoff probability, $\bar{P}^*$, that minimizes a loss function equal to the weighted sum of type I and type II errors, for appropriately chosen weights.

In 2012, staff restricted the cutoff probabilities to certain ranges, $\bar{P}_{min} < \bar{P} < \bar{P}_{max}$, typically in the neighborhood of the unconditional probability of debt distress. Staff also considered weights on type I (type II) error varying from 0.5 (0.5) to 0.75 (0.25), that is, the relative weight of type I ranged from 1 to 3. Accordingly, each final optimal cutoff probability was not a point estimate but the average of the cutoff probabilities that minimized type I and type II errors over the different weights.

Step 3: Finally, given the estimated optimal cutoff probabilities, $\bar{P}^*$, debt thresholds were derived by inverting each estimated probit model, after fixing the CPIA to one of its three cutoffs (weak = 3.25, medium = 3.50 and strong = 3.75), and setting real GDP growth to the historical average for LICs:

$$Threshold^c = \left(\Phi^{-1}(\bar{P}^*) - \left(\hat{\beta}_0 + \hat{\beta}_3 \overline{growth} + \hat{\beta}_4 CPIA^c\right)\right) / \hat{\beta}_1$$

where "c" denotes each of the CPIA cutoffs, hats denote estimated values, Φ^{-1} is the inverse of the CDF, and $\overline{growth}$ is average growth for LICs.

13. **The algorithm for deriving debt thresholds uses an aggregation rule to combine information from the probit regressions, and selects the debt thresholds to minimize a prediction loss function that penalizes Type I and Type II errors.** The process can be broken down into the following five steps:

Step 1: Estimate a probit model for each debt burden indicator:

$$P(Debt\ Distress) = \Phi\left(\alpha_j + \gamma_j d_j + \sum_{k=1}^{6} \beta_{j,k} X_k\right)$$

where d_j represents the four different debt burden indicators (PV of debt to GDP, PV of debt to exports, debt service to revenues, and debt service to exports) and X_k represents the non-debt explanatory variables included in the probit regressions (CPIA, country growth, reserves, squared reserves, remittances, and world growth).

Step 2: Construct a composite indicator (CI) reflecting the contributions of all the non-debt explanatory variables to the risk of debt distress. The CI is defined as:

$$CI = \sum_{k=1}^{6} \overline{\beta_k}\, \bar{X}_k$$

where $\overline{\beta_k}$ is the average slope coefficient across the four probit regressions for each explanatory variable and $\bar{X}_k$ is a 10-year average of each of the explanatory variables prior to the episode. Benchmark values for the strong (medium) (weak) categories are set at the 75th (50th) (25th) percentiles of the distribution of the CI over the most recent 2005-14 period.

Step 3: Given a cutoff probability for each debt burden indicator, p_j^*, "invert" each of the estimated probit regressions to obtain debt thresholds $\bar{d}_j(p_j^*)$ as a function of the CI:

$$\bar{d}_j(p_j^*) = (\Phi^{-1}(p_j^*) - \alpha_j - CI)/\gamma_j$$

Evaluating these thresholds at the 75th (50th) (25th) percentiles of the distribution of the CI over the most recent 2005–14 period results in debt thresholds for the strong/medium/weak categories for each of the debt burden indicators.

Step 4: Define a prediction rule for predicting debt distress that is consistent with the practice of the DSF. Specifically, the DSF signals a high risk of debt distress ($\hat{y} = 1$) if any of the thresholds is breached, i.e.:

$$\begin{cases} 1, & if\ d_j > \bar{d}_j(p_j^*), for\ ANY\ debt\ indicator\ j \\ 0, & if\ d_j < \bar{d}_j(p_j^*), for\ ALL\ debt\ indicators\ j \end{cases}$$

In determining whether the threshold is breached for a particular episode, the observed debt level is compared with the debt thresholds for the strong/medium/weak categories, where the latter

determination is based on comparing the country-specific CI with the benchmarks described in the previous step: a country is compared with the debt threshold for the weak category if its CI is below the 25th percentile of the 2005-14 average CI, with the threshold for the strong category if its CI is above the 75th percentile, and with the threshold for the medium category otherwise.

Step 5: Select the cutoff probabilities p_j^* for the four debt burden indicators, and the corresponding predictions of debt distress, to minimize a prediction loss function that penalizes Type I error ("missed calls") and Type II error ("false alarms") with transparent weights:

$$Loss\ Function = \omega\ Type\ I\ error + (1 - \omega)\ Type\ II\ error$$

The weight on Type I error is set at $\omega = 0.67$.

This procedure results in the proposed debt thresholds shown in Table AIII.5.

14. **The re-estimated external debt thresholds imply some additional room for countries to borrow, provided they manage their debt service well.** The new framework would maintain or increase debt stock thresholds for all countries. Countries which retain a weak or medium classification under the new classification scheme would see lower debt service thresholds (Table AIII.5). Since the present framework had set debt service thresholds at what were typically non-binding levels (see Section II.A in the Board paper), it is not surprising that the improved methodology produces lower debt service thresholds.

Table AIII.5. Existing and Re-estimated External Debt Thresholds 1/

Country	PV of debt-to-GDP		PV of debt-to-exports		Debt service-to-revenue		Debt service-to-exports	
Classification	Old	New	Old	New	Old	New	Old	New
Weak	30	30	100	140	18	14	15	10
Medium	40	40	150	180	20	18	20	15
Strong	50	55	200	240	22	23	25	21

Source: Fund staff calculations.
1/ Debt stock thresholds are rounded; debt service thresholds are point estimates.

15. **The new framework's predictive performance improves over the status quo.**

- In comparison with the existing framework, the proposed changes under this review clearly improve its predictive performance (see second and last columns of Table AIII.6). The rate of false alarms is significantly reduced by 10 percentage points while DSF's capacity to anticipate debt distress is improved.

- The new baseline also improves, in terms of the balance across errors and overall predictive performance (i.e., lower loss function), upon the approach in which the model specification remains unchanged but it is re-estimated using the updated sample (1970–2014) and the new methodology to identify debt distress episodes (third column of Table AIII.6). Such approach would imply a significant tightening of debt thresholds, leading to a much larger

rate of false alarms (about 70 percent) and an excessively low rate of missed crises (about 6 percent).

Table AIII.6. Predictive Power of Existing and Revised Framework

Statistics	Based on the DSF's aggregation rule			
	Old 1/	Old updated 2/	Old modified 3/	New 4/
Loss function 5/	0.28	0.25	0.27	0.24
Type I error	0.18	0.08	0.06	0.18
Type II error	0.48	0.60	0.68	0.38

1/ Existing framework (2012 LIC DSF review).
2/ Existing framework after re-estimating debt thresholds using existing methodology, new sample (1970-14, LICs+MICs) and old definition of debt distress.
3/ Existing framework after re-estimating debt thresholds using existing methodology, new sample (1970-14, only LICs) and new definition of debt distress.
4/ Based on all proposed changes to the existing framework in this review.
5/ Weighted sum of errors (weight on type I error equalt to 67 percent).

Enhanced Country Classification

16. **Countries would continue to be classified as weak, medium or strong based on an expanded measure of repayment capacity.** Country classification would be based on a significantly expanded measure of capacity to repay—their country-specific CI—, thus moving away from relying exclusively on the CPIA.

17. **The CI would be calculated based on historical data and projections.** Given that historical data is not always a good predictor of future performance, staff propose to enhance the information content underlying country classification by calculating the country-specific CI based on the latest five years of historical data and the first 5 years of projections. Mixing historical data and forecasts allows the framework to capture ongoing changes in countries' fundamentals in a forward-looking fashion while keeping country classification sufficiently stable.[6]

18. **Despite the added new features, the proposed approach to classify countries is simple and can be summarized as follows.** Staff defines the CI as the weighted sum of the non-debt determinants of debt distress (CPIA, country growth, reserves, squared reserves, remittances, world growth), where the weights are given by the average estimated coefficients across the probit models. More specifically, country classification is done in three steps as follows:

- Calculate the CI for each country *i*:

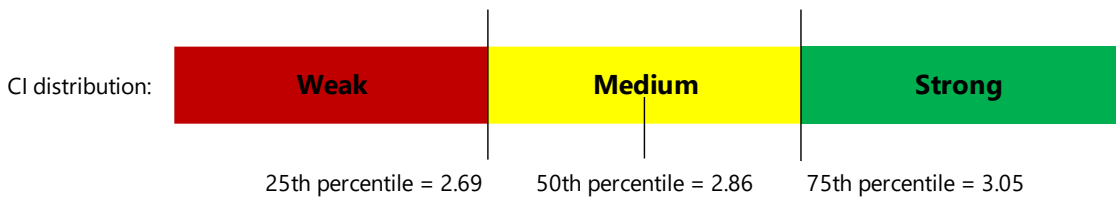
$$CI_i = \sum_{k=1}^{6} \overline{\beta_k}\, \overline{X}_{i,k}$$

[6]Forecasts of the additional variables in the CI are routinely produced in the WEO database and individual DSAs. Recognizing its slow-moving nature, the forecast for the CPIA rating would consist of its most recent value.

where k denotes the control variables, $\overline{\beta_k}$ denotes the average coefficient for the k control across the four probit models, and $\bar{X}_{i,k}$ denotes the 10-year average of each control, covering the latest 5 historical and first 5 projections years.

- Compare the country CI with the CI cutoffs, estimated as the 25th (CI_{25}) and 75th (CI_{75}) percentiles of the its distribution in the last ten years of the regression sample (2005-2014) (see Figure AIII.2, which also shows the range of contributions to the composite indicator across LICs). Countries are classified by their debt-carrying capacity as follows:
  - *Weak* if $CI_i < CI_{25}$,
  - *Medium* if $CI_{25} \leq CI_i \leq CI_{75}$,
  - *Strong* if $CI_i > CI_{75}$

Figure AIII.2. Illustration of the New Approach to Classify Countries



Contribution to the CI across LICs	Min	Max
World growth	0.47	0.47
CPIA	0.76	1.54
Reserves 1/	0.03	1.03
Remittances	0.00	0.31
Country growth	-0.01	0.24

1/ Including reserves' linear and non-linear terms.

IV. New Approach for the Estimation of Total Public Debt Benchmarks

19. **Considering data constraints, staff relied on a new set of signals to better identify domestic debt distress episodes and estimate total public debt benchmarks.** In the absence of comprehensive and systematic data on domestic arrears, staff used not only outright defaults (which are rare events in the sample) but also relied on alternative measures of *de facto domestic* defaults. *De facto* domestic defaults are broadly defined as episodes of inflation and/or consistently low real interest rates (indicative of non-structural domestic financial repression) leading to an erosion of the real value of domestic debt over time. These types of debasement of domestic debt or unorthodox forms of raising fiscal revenues have been well documented in the literature and have been used not only in developing countries but also in advanced economies in the 19th and early 20th centuries (e.g., Giovannini and de Melo (1993), Reinhart and Rogoff (2011), Reinhart and Sbrancia (2015)).

20. **Staff opted for several proxies for *de facto* domestic default.** Episodes of unexpected inflation or inflation surprises would be an ideal proxy for *de facto* defaults. However, this would require comparing actual with expected inflation, with the latter not available for most LICs. Hence, staff followed the literature and used spikes in inflation (i.e., inflation above a certain cutoff) to proxy for inflation surprises that coincide with large government financing by the central bank but not with significant changes in the typical drivers of inflation such as spikes in food and oil prices (proxying for negative supply shocks) and large positive output gaps (proxying for positive demand shocks). Besides inflation surprises, governments often force below-market interest rates on domestic lenders, contributing to the erosion of the real value of domestic debt service. Therefore, our proposed proxies for *de facto* domestic default cover not only episodes of high inflation but also those associated with persistent (but non-structural) negative real interest rates due to high inflation and/or financial repression. Staff selected a total of 18 different definitions of domestic debt distress based on these two proxies (Table AIII.7), of which:

- 16 definitions are based on negative real interest rates (RIR), of which four are unconditional, 10 control for the size of domestic debt, and two control for the size of domestic-debt-to-revenue ratio; and

- two definitions are based on episodes of inflation that coincide with large government financing by the central bank but not with significant changes in the typical drivers of inflation such as spikes in food and oil prices (proxy for negative supply shocks) and large positive output gap (proxy for positive demand shocks).

Table AIII.7. Domestic Debt Distress Episodes in LICs, 1970–2015

Signal	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Total episodes	**116**	**81**	**128**	**91**	**80**	**53**	**88**	**66**	**97**	**67**	**125**	**89**	**111**	**83**	**66**	**43**	**45**	**35**
Of which triggered by:																		
Outright default	18	18	19	20	19	19	19	20	20	21	19	20	19	20	19	21	22	23
Financial repression	97	62	109	71	61	34	69	46	77	46	106	69	92	63	46	22	23	12
Both	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0

21. **Different from the 2012 review, instead of estimating a separate probit model for total public debt, staff followed the noise-to-signal approach to derive benchmarks for the PV of total public debt.** The estimation of a separate probit model for total public debt would lead to a composite indicator potentially different from the one derived in the external debt block, even after assuming the same set of controls. The composite indicators for the public and external debt blocks would certainly be different if domestic determinants of the likelihood of domestic debt distress had been incorporated in the probit regressions (e.g., fiscal balance, size of the domestic banking sector, etc.). Different composite indicators would in turn lead to different and possibly conflicting country classifications, thus potentially increasing the complexity and reducing the transparency of the framework. To get around these technical considerations, staff used the noise-to-signal (NTS) approach to derive benchmarks for the PV of total public debt.

22. **The proposed strategy to find the best benchmarks for the PV of total public debt under the NTS approach ensures full consistency with the external block.** The approach consisted of five steps:

- First, an indicator of overall public debt distress is defined by combining the external and domestic debt distress episodes.
- Second, countries were assigned the same classification as in the external block.
- Third, an expanded aggregation rule was imposed to signal the overall risk of debt distress, based on the joint information from five debt burden indicators for predicting total public debt distress (i.e., the four from the external block, whose thresholds are taken as given as previously estimated, plus the PV of total public debt-to-GDP). This augmented aggregation rule avoids potential inconsistencies arising from treating the two blocks separately, such as high risk of external debt distress and at the same time a low risk of total public debt distress.
- Fourth, staff defined Type I and Type II errors in the same was as was done for the external debt block described above.
- Finally, the three debt benchmarks for total public debt for the weak/medium/strong categories were jointly chosen to minimize the loss function.

Therefore, although staff uses a different method than in 2012 to derive the benchmarks for total public debt, the aggregation rule is symmetric to that used in the external block, and consistent with the way the framework is used in practice, i.e., the framework signals high risk of debt distress (no high risk) if any (none) of the five debt burden indicators breaches (does not breach) its corresponding threshold/benchmark.

23. **The new benchmarks are broadly in line with the existing benchmarks.** As can be seen in Table AIII.8, the new benchmarks are slightly lower across the classification categories, with the one for the strong category being equal to the high-risk benchmark in the MAC DSA (70 percent of GDP). It is difficult to precisely decompose these changes as they reflect a multitude of factors, including differences in the definition of total public debt distress, in the algorithm to derive the benchmarks, in the weighting scheme, and in the aggregation rule. However, as was the case with the external block, the new benchmarks also imply a lower rate of false alarms compared to the one achieved by the existing framework at the time of the 2012 review.

Table AIII.8. Existing and Proposed Benchmarks for the PV of Total Public Debt

Variable	Benchmarks			Type I error	Type II error	Loss function
	weak	medium	strong			
2017 review 1/	35	55	70	0.24	0.44	0.31
2012 review 2/	38	56	74	0.19	0.52	0.31

1/ Based on NTS approach, using 0.67 weight on type I error. Benchmarks are rounded up to the nearest 5 percent.

2/ Based on the probit model and weight on type I error ranging from 0.5 to 0.75.

Box AIII.2. Data Description

Staff relied on several data sources to construct the debt burden indicators and hence maximize the number of observations and better identify the econometric model. As in 2012, data on external debt stock is taken from the World Bank's Debt Reporting System, in which the PV of external debt is calculated by discounting the stream of debt service using time-varying discount rates (commercial interest reference rates—CIRRs). Data on external debt service comes from the World Bank's World Development Indicators (WDI) and International Debt Statistics (IDS). Since this data is on a paid basis, staff adjusted it by the accumulation of arrears to estimate the debt service on a due basis. Staff combined several data sources to construct the data on government revenues, including the IMF's World Economic Outlook database (WEO), Government Finance Statistics (GFS), WDI, and data compiled by Fund staff. Similarly, staff relied on the following multiple sources for the raw data on domestic and total public debt: WEO, several DSA vintages, Abbas and Christensen (2009), Abbas and others. (2010), Panizza (2008), Reinhart and Rogoff (2010).[1] Since the coverage of the government perimeter varies across countries, staff used the widest coverage available (typically general government and central government).

Similarly, staff combined information from inside and outside the Fund to construct the debt distress episodes. Data on Fund disbursements comes from the IMF's International Finance Statistics (IFS) and the Fund's Finance Department. Staff used data on external arrears from the World Bank's Global Development Finance (GDF) and Bank of Canada's Database on Sovereign Defaults, compiled by Beers and Nadeau (2015).[2] Data on Paris Club restructurings comes from Das and others. (2011), Fund staff, and the Paris Club website,[3] whereas the data on commercial debt restructurings is from Cruces and Trebesch (2013),[4] complemented by Fund staff. Staff used information on external defaults from Standard and Poor's and from Catão and Milesi-Ferretti (2014), whereas information on domestic defaults comes from Reinhart and Rogoff (2009, 2011), complemented by Fund staff.

Staff also combined the usual data sources for the macroeconomic time series. The series on GDP, real GDP growth, exports, imports, international reserves, openness, inflation, real exchange rate, central bank financing, remittances, real GDP per capita, world growth, and the series used in the robustness tests were constructed by combining information from one or more of the following sources: WEO, WDI, IFS, United Nations, Penn World Tables,[5] Maddison Statistics,[6] and World Bank's Migration and Remittances Data.[7] In the specific case of remittances, staff approximated missing observations by using information on nominal GDP growth of the major source country. Bilateral remittances data used to identify the major source country was obtained from the World Bank. Data on natural disasters and conflicts comes from The International Disaster Database (EM-DAT) and UCDP/PRIO Armed Conflict Dataset.[8] The World Bank's CPIA (Country Policy and Institutional Assessment index) is a key measure of the quality of institutions and policies used in the regressions.[9] Staff also considered the political risk rating from the International Country Risk Guide (ICRG) for the robustness tests.

The final sample comprises 80 LICs and covers 45 years (1970–2014). The sample potentially comprises 3600 country-years but data availability and the identification of distress and non-distress episodes constrains the actual sample size used in the regressions. LICs are defined as IDA-only countries as in the 2012 review. However, unlike the 2012 review, countries that graduate (reverse graduate) from IDA status are treated as LICs if they spent at least half of the sample years under that status. Otherwise they are treated as middle-income countries and excluded from the sample.

Box AIII2. Data Description (concluded)
LICs Sample

Country	Iso Code	IFS Code	% of time as LIC	Country	Iso Code	IFS Code	% of time as LIC
Afghanistan	AFG	512	100	FYR Macedonia	MKD	962	73
Albania	ALB	914	87	Madagascar	MDG	674	100
Angola	AGO	614	100	Malawi	MWI	676	100
Armenia	ARM	911	100	Maldives	MDV	556	100
Bangladesh	BGD	513	100	Mali	MLI	678	100
Benin	BEN	638	100	Mauritania	MRT	682	100
Bhutan	BTN	514	100	Moldova	MDA	921	100
Bolivia	BOL	218	100	Mongolia	MNG	948	100
Burkina Faso	BFA	748	100	Montenegro	MNE	943	87
Burundi	BDI	618	100	Mozambique	MOZ	688	100
Cabo Verde	CPV	624	100	Myanmar	MMR	518	100
Cambodia	KHM	522	100	Nepal	NPL	558	100
Cameroon	CMR	622	73	Nicaragua	NIC	278	80
Central African Republic	CAF	626	100	Niger	NER	692	100
Chad	TCD	628	100	Nigeria	NGA	694	58
Comoros	COM	632	100	Papua New Guinea	PNG	853	58
Congo, Dem. Rep.	ZAR	636	100	Rwanda	RWA	714	100
Congo, Rep.	COG	634	76	Samoa	WSM	862	100
Côte d'Ivoire	CIV	662	60	São Tomé and Príncipe	STP	716	100
Djibouti	DJI	611	100	Senegal	SEN	722	100
Dominica	DMA	321	100	Serbia	SRB	942	87
Equatorial Guinea	GNQ	642	53	Sierra Leone	SLE	724	100
Eritrea	ERI	643	100	Solomon Islands	SLB	813	100
Ethiopia	ETH	644	100	Somalia	SOM	726	100
Gambia	GMB	648	100	Sri Lanka	LKA	524	100
Georgia	GEO	915	100	St. Kitts and Nevis	KNA	361	56
Ghana	GHA	652	100	St. Lucia	LCA	362	100
Grenada	GRD	328	100	St. Vincent and the Grenadines	VCT	364	100
Guinea	GIN	656	100	Sudan	SDN	732	100
Guinea-Bissau	GNB	654	100	Tajikistan	TJK	923	100
Guyana	GUY	336	100	Tanzania	TZA	738	100
Haiti	HTI	263	100	Timor-Leste	TMP	537	100
Honduras	HND	268	78	Togo	TGO	742	100
Kenya	KEN	664	100	Tonga	TON	866	100
Kiribati	KIR	826	100	Uganda	UGA	746	100
Kosovo	KSV	967	100	Vanuatu	VUT	846	100
Kyrgyz Republic	KGZ	917	100	Vietnam	VNM	582	100
Lao PDR	LAO	544	100	Yemen	YEM	474	100
Lesotho	LSO	666	100	Zambia	ZMB	754	100
Liberia	LBR	668	100	Zimbabwe	ZWE	698	82

[1]The latter is available at http://www.carmenreinhart.com/data/browse-by-topic/topics/9/.
[2]http://www.bankofcanada.ca/2014/02/technical-report-101/.
[3]http://www.clubdeparis.org/en/.
[4]https://sites.google.com/site/christophtrebesch/data.
[5]http://www.rug.nl/research/ggdc/data/pwt/.
[6]http://www.ggdc.net/maddison/maddison-project/data.htm.
[7]http://www.worldbank.org/en/topic/migrationremittancesdiasporaissues/brief/migration-remittances-data
[8]http://www.emdat.be/database and https://www.prio.org/Data/Armed-Conflict/UCDP-PRIO/, respectively.
[9]To maximize the number of CPIA observations, missing CPIA data in the early 1970s are taken from staff's 2012 estimations. These values were fitted using a simple regression model in which the annual CPIA score is explained by three covariates: country real GDP growth rate relative to the sample average; annual inflation rate relative to the sample average; and the lead value of the annual CPIA score.

Box AIII.3. Reserve Measurement in Currency Unions

Fifteen LICs are members of three currency unions: *Eastern Caribbean Currency Union* (ECCU): Dominica, Grenada, and St. Vincent and the Grenadines; *Central African Economic and Monetary Community* (CEMAC): Cameroon, Central African Republic, Chad, and Republic of Congo; and *West African Economic and Monetary Union* (WAEMU): Benin, Burkina Faso, Cote d'Ivoire, Guinea-Bissau, Mali, Niger, Senegal, and Togo. The ECCU's common currency is the Eastern Caribbean dollar, which is pegged to the U.S. dollar, whereas the currency of the two African monetary zones ("Franc Zone") is the CFA franc, pegged to the euro.

Conceptually, the union-wide reserve pooling or country-level reserves may be the appropriate measure of insurance against shocks depending on union and country-specific circumstances. Countries may gain greater balance of payments protection from reserve pooling. When a common central bank holding an adequate level of reserves can allocate adequate liquidity within the union, there would generally be less need for members to cover their reserve needs through their own reserve holdings. However, there may be factors that limit the scope and gains from reserve pooling, such as issues with the financial architecture of the union (e.g., absence of a banking union or the possibility that liquidity may not be allocated efficiently within the union to stem financial pressures) and the synchronization of shocks among members (i.e. lack of sufficient economic diversification leading to the synchronization of business cycles and correlated shocks).[1 2] In addition, members may free ride by benefitting from the advantages of the union while not contributing to its costs, up to a point in which they lose effectively access to the regional pool. It is in this case where the economically-relevant measure of insurance is the member's own reserve holdings.

For the purposes of estimating the probit model, the use of imputed reserves appears to be a robust approach. Given data constraints on using the union's pooled reserves to union import ratio (specifically with adjustments for intra-union imports), staff opted to use the members' imputed reserves for estimating the probit baseline model. Staff also considered the implications of using a proxy of regional reserve coverage (i.e., sum of members' imputed reserves-to-sum of member's imports ratio) and found it to have a secondary impact on the probit model (next Table) and no meaningful effect on the weights for the calculation of the composite indicator, suggesting that the imputed-reserves estimation approach is robust for the purposes of setting up the framework.

For the purposes of using the framework for currency union members, the use of the union-wide reserve coverage for classifying countries' debt-carrying capacity would generally be appropriate, with some exceptions. As noted, the model appears robust to such an approach. However, testing suggests that some currency union members that effectively have lost access to the pool (e.g. due to large advances) could be upgraded. In these circumstances, stronger performance of the framework would require that a member's classification be determined by the level of imputed reserves. The Staff Guidance Note would provide a more detailed assessment as to a trigger for such a switch.

Box AIII.3. Reserve Measurement in Currency Unions (concluded)

Alternative Model Specification

(based on a proxy of regional reserve coverage for currency unions)

Controls	PV of debt as a percent of		Debt service as a percent of	
	GDP	Exports	Revenue	Exports
Debt burden indicator	1.614***	0.361***	3.717***	3.591***
CPIA	-0.396***	-0.380***	-0.355***	-0.388***
Domestic growth	-3.120*	-2.926*	-3.051	-1.989
Reserves_CU	-4.242***	-4.458***	-3.983**	-3.876***
Reserves_CU^2	4.069**	4.520**	4.183*	4.019*
Remittances	-2.288**	-2.355***	-1.648*	-1.943**
World growth	-12.42***	-14.07***	-13.48***	-13.76***
Constant	1.276***	1.303***	0.979*	1.129**
Observations	410	404	343	381
Pseudo R-squared	0.166	0.166	0.191	0.184
Log-likelihood	-151.0	-147.2	-116.6	-135.4
BIC	350.1	342.4	279.9	318.4

***, **, * denote significance at 1%, 5%, and 10%, respectively, based on robust standard errors.

[1]See IMF (2014).

[2]Yehoue (2015) found evidence of small gains from reserve pooling in the CFA zone, providing evidence of little risk sharing or shock smoothing.

Annex IV. Realism Tools for Assessing Baseline Growth Projections

A. Public Investment Scaling-up

1. This section describes the assumptions, methodology, and data requirements for the proposed tool for assessing the realism of baseline growth projections in the context of public investment scaling-up.

2. **The goal of this tool is modest: to provide a simple Excel-based diagnostic that is easy to understand, implement, and apply in the DSF template.** It is not intended to be a substitute for more sophisticated analysis of the growth effects of public investment, which can be done outside the confines of the DSA, as a means of informing Fund-supported programs, World Bank growth diagnostics, and the policy dialogue more generally. Available tools supported by Bank and Fund staff for this purpose include the IMF's Debt-Investment-Growth model (see Buffie and others, 2012), and the World Bank's Long-Term Growth model (see Pennings, 2017).

3. **This tool takes as given the projected path of public investment and growth, and decomposes the projected growth rates into a simulated (i) contribution of the increase in the government capital stock due to public investment, and (ii) contribution from other sources.** The value of this simple decomposition is that it makes explicit the implicit assumptions made by the country team preparing the DSA about the contribution to growth of public investment scaling-up, and the contribution of other sources of growth. This in turn can be used to assess the realism of baseline growth projections by comparing these two sources of growth with historical data or other relevant reference points for the country.

4. **This decomposition is performed using a standard growth-accounting methodology, and has very limited data requirements.** In addition to the projected paths of public investment and growth, the only other information required is an estimate of the stock of government capital, an estimate of the output elasticity of government capital, and an estimate of public investment efficiency (if estimated to be different than historical efficiency). Estimates of the stock of government capital are now available for 170 countries in a recent FAD database (see IMF, 2017). The output elasticity is based on existing literature and public investment efficiency can be calibrated based on recent FAD estimates (see IMF, 2015a).

5. **The growth accounting methodology is based on a standard functional form assumption for the aggregate production function.** Specifically, the aggregate production function is assumed to be isoelastic in government capital:

(1) $$Y_t = G_t^{\beta} F(A_t, K_t, H_t)$$

where Y_t is real GDP; G_t is government capital; β is the output elasticity of government capital, and $F(A_t, K_t, H_t)$ is a function of productivity (A_t), private capital (K_t), and human capital (H_t) that does not need to be fully specified. Given this functional form, there is a simple decomposition of output growth:

(2) $$\frac{Y_t - Y_{t-1}}{Y_{t-1}} = \beta \frac{G_t - G_{t-1}}{G_{t-1}} + \varepsilon_t$$

where the first term $\beta \frac{G_t - G_{t-1}}{G_{t-1}}$ represents the contribution of changes in government capital to growth, and the second term ε_t represents the contribution of all other factors and productivity to growth. This decomposition captures only the direct effects of changes in government capital on growth, and does not consider possible endogenous responses of productivity or private factors of production to increases in government capital. Country teams interested in pursuing these richer channels can do so with the help of more sophisticated models such as those referenced above. Based on existing literature, this tool sets the output elasticity of government capital at $\beta = 0.15$.[1]

6. **The growth rate of the government capital stock can be inferred from the path of public investment, together with assumptions on an initial stock of government capital, depreciation rates, and public investment efficiency.** The accumulation equation for government capital is:

(3) $$G_{t+1} = (1 - \delta)G_t + \phi_F i_{Gt}$$

where δ is the depreciation rate (set at 0.05), ϕ_F is the public investment efficiency (that can be calibrated using FAD's estimates), and i_{Gt} is the public investment path. This equation is used to project the path of the government capital stock, from which growth rates of the government capital stock can be calculated.

7. **The analysis can accommodate assumptions about the "efficiency" of past and future government investment.** In this context, "efficiency" refers to the fraction of a dollar of government investment that turns into government capital. Incorporating efficiency involves (i) replacing the projected future investment rate with $\phi_F i_{Gt}$, and (ii) scaling down the historical capital stock series by a factor ϕ_H. The efficiency parameters $0 < \phi_F \leq 1$ and $0 < \phi_H \leq 1$ reflect judgment as to the efficiency of future and past investment. The adjustment for historical efficiency is important: if past efficiency of public investment was low, then the accumulated investment flows on which the historical capital stock series are based are an over-estimate of the actual stock of government capital, and the marginal productivity of an additional dollar of government investment is underestimated (see Berg and others (2015)).[2] While DSF users can separately specify the historical and future efficiency parameters, the analysis in the realism tool depends only on the ratio of the two, ϕ_F/ϕ_H. The benchmark assumption is that $\phi_F = \phi_H = 1$

[1]Bom and Ligthart (2014) perform a meta-analysis of 68 studies on the productivity of government capital and find an average output elasticity of government capital among mostly OECD countries of 0.11. Ligthart and Martin-Suarez (2011) perform a similar analysis, finding an average output elasticity around 0.15. More recent econometric analysis of the relationship between physical infrastructure stocks and output in low-income developing countries in Calderon, Moral-Benito, and Serven (2015) suggests output elasticities ranging from 0.06 to 0.18.

[2]Recent work has shown, for the growth analysis of public investment scaling-up, that it is critical to distinguish between levels and rates of change in efficiency (see Berg et. al., 2015). A change in efficiency would increase output even absent an increase in public investment spending and would increase the growth impact of any increase in public investment spending.

and departures from the benchmark, particularly those that imply an improvement in efficiency, would require careful justification by country teams.

B. Fiscal Adjustment Impact on Growth

8. This section describes the proposed tool to illuminate the impact of planned fiscal adjustment on growth projections under a range of plausible fiscal multipliers, allowing a comparison with the baseline projected growth path.

9. **Overly optimistic growth projections could undermine adjustment plans.** Negative growth surprises have been identified as the main factor derailing fiscal consolidation (Mauro and Villafuerte, 2013). While a useful check for the consistency of growth and primary balance projections is to uncover the growth path consistent with a neutral fiscal stance and assess its realism, such an approach is challenging in the case of LICs. This is because LICs' borrowing capacity is very weak in the absence of fiscal adjustment, casting doubts about the likelihood of such a counterfactual.[3]

10. **This tool displays the range of possible projected growth paths consistent with different fiscal multipliers.** Given the country team's assumed fiscal multiplier, the tool calculates the underlying projected growth rates (i.e. the counterfactual that would materialize absent fiscal adjustment). For any projected growth and fiscal adjustment paths under the baseline scenario, the underlying projected growth path can be computed as follows:

Underlying projected growth =

Projected growth (including fiscal adjustment) + change in growth due to fiscal adjustment

11. **The change in growth due to fiscal adjustment depends on the size and persistence of the country team's assumed fiscal multiplier.** The tool assumes an AR(1) process to model the evolution of the fiscal multiplier:

(4) $$m_t = m * p^{t-1}$$

where m is the size of the impact multiplier, defined as the percent change of real GDP from a one percentage point adjustment in the structural primary balance (as percent of GDP) during the first year of full impact. Given that the effect on growth of fiscal adjustment can take time to be fully realized, the first year of full impact is set as the year after the implementation. The impact during the implementation year itself is assumed to be half of the impact multiplier. The persistence of the multiplier is captured by p, the autocorrelation coefficient (default value is set at 0.6).

[3]This is not the case of countries with market access, in which such a counterfactual consistent with a neutral fiscal stance has reasonable odds, and therefore can be used to assess the realism of growth projections as is done in the MAC DSA.

12. **Once the underlying projected growth is uncovered, the tool calculates framework-consistent projected growth paths assuming different values of the fiscal multiplier.** This allows a comparison between the projected growth path under the baseline scenario and alternative projected paths under plausible values for the fiscal multiplier.

Annex V. Reflecting Market-Related Risks

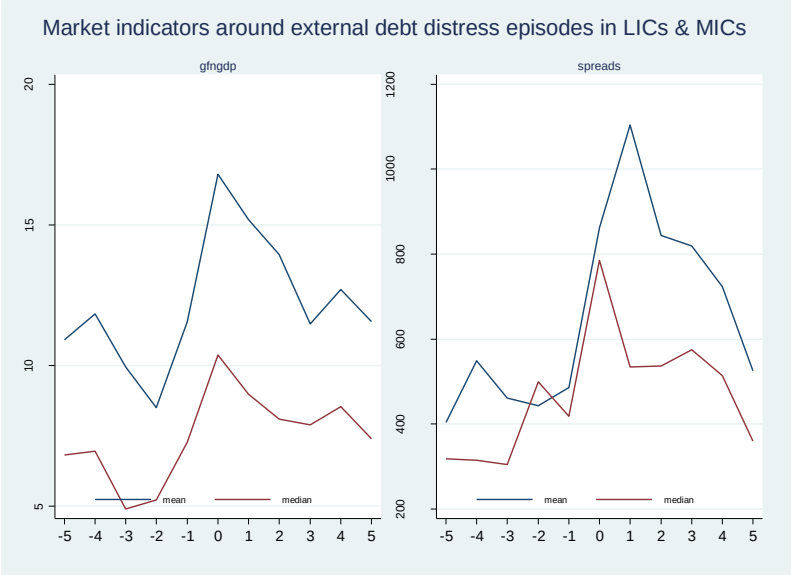
1. Changes in the financing landscape over the last decade have led to rising shares of non-concessional and market borrowings in LICs. These changes have been particularly pronounced for frontier LICs, which have been able to tap international capital markets while attracting increased non-resident participation in domestic debt markets. LICs with market access are thus exposed to shorter maturities, higher and more volatile financing costs, more frequent spikes in financing needs, and greater scrutiny from investors. Liquidity risk and market perception, therefore, are important dimensions of debt sustainability risks for this group of countries.

2. The module to reflect market-related risks would apply to LICs deemed to have market access. These are the same set of countries identified for the tailored market-financing stress test. For these countries, this annex shows that including information on public gross financing needs (GFN) and EMBI spreads, significantly increases the predictive power of the framework.

Market risk indicators

3. Two key indicators—public gross financing needs (GFN) and the EMBI spreads—will be used to flag exposure to market-financing risks. Stress in these indicators can reflect policy slippages or increased uncertainty which affect countries' borrowing costs. Empirically, both indicators tend to move together in the run-up to external debt distress, peaking in the first and second years of the debt distress episodes. Median GFN peaks at around 10 percent of GDP (average peaks at 15 percent of GDP). The EMBI spreads exceed 400bps in the run-up to the crisis and peak at round 800 bps (or over 1000bps if the mean is considered) (Figure AV.1).

Figure AV.1. GFN and EMBI Spreads Around External Debt Distress Episodes



Data and methodology

4. We used annual data on GFN/GDP and EMBI spreads since the mid-1990s for frontier LICs and EMs for which data on these two indicators were available. Extending the data to cover EMs helped to increase the number of observations for EMBI spreads, for which only a few frontier LICs have data. Grouping LICs and EMs for the purposes of market risk assessment is also logical as frontier LICs are benchmarked against EMs. The most recent PRGT graduates—Nigeria, Vietnam, and Mongolia—were all frontier LICs. Also, the median dynamics of GFN and EMBI spreads have been somewhat similar across EMs and frontier LICs, including around the global financial crisis (Figure AV.2).

Figure AV.2. GFN and EMBI spreads in LICs and EMs

5. Joint early warning benchmarks for the GFN and EMBI spreads were derived using the bivariate signal approach. In this exercise, the indicators correctly predict debt distress when at least one exceeds a given benchmark in the period before debt distress occurs; while they correctly predict normal time when both indicators are below the benchmark and no distress occurs.[1] A "false alarm" is when one of the indicators exceeds its benchmark but no distress occurs; and a "missed crisis" is when neither indicator exceeds its benchmark but distress occurs. The benchmarks for GFN and EMBI spreads are jointly chosen to minimize the loss function, which is the weighted sum of missed crises ("Type I errors") and false alarms ("Type II errors"), with the weight of 67 percent on Type I errors, the same weight used to derive the external debt thresholds. Other potential indicators such as debt service/reserves, change in domestic debt, and external gross financing needs were also examined, but they delivered poor in-sample predictive performance.

6. The GFN-EMBI spread pair of (14, 570) yields the best combination of Type I and Type II errors. Under these two benchmarks, the type I error and type II error rates are 18 percent

[1]The definition of debt distress is the one used for deriving the four external debt burden thresholds in the LIC DSF.

and 48 percent, respectively, in the ballpark of the statistics obtained for the external debt thresholds. The selected pair is also comparable to the high-risk benchmarks currently used in the MAC-DSA.

7. The GFN and EMBI spread indicators add information not captured by the other external debt burden indicators for the sample of frontier LICs and EMs. Including the GFN and EMBI spread in the multivariate signal approach exercise to predict debt distress (taking the external debt thresholds as previously estimated) significantly improves the loss function and the rate of Type I error. Table AV.1 illustrates the results for the sample of frontier LICs and EMs used to derive the GFN-EMBI spread benchmarks. The results confirm that for countries that have market access, the GFN and EMBI contain additional information for flagging potential debt distress.

Table AV.1. Overall predictive performance of model with and without GFN-EMBI

Indicator	PV Debt/GDP	PV Debt/ Exports	Debt Service/ Revenue	Debt Service/ Exports	GFN/GDP	EMBI Spread	DSF's Aggregation Rule
Sample used in signal approach for GFN-EMBI spreads; 1995-2015; Frontier LICs and MICs							
Loss function 1/	0.68	0.60	0.50	0.58			0.48
Type I error	0.93	0.87	0.60	0.73			0.53
Type II error	0.17	0.07	0.29	0.26			0.37
Loss function 1/	0.68	0.60	0.50	0.58	0.56	0.35	0.28
Type I error	0.93	0.87	0.60	0.73	0.80	0.47	0.20
Type II error	0.17	0.07	0.29	0.26	0.07	0.11	0.44

1/ Weighted sum of the type I error and type II error rates, with type I error weight of 67 percent.

Annex VI. Enhanced Information on the Moderate Risk Category

2. **Countries rated as facing a moderate risk of debt distress are numerous and display a great diversity of debt vulnerabilities.** As a result, a tool to characterize the extent of debt vulnerabilities in countries within the moderate risk category could enhance the information value of the DSF. The proposed methodology would improve the understanding of the moderate risk category without operational implications in the application of Fund/Bank debt policies or additional data needs.

2. **The robustness of the debt position of a country at moderate risk of debt distress depends on the "space to absorb shocks" without being downgraded.** Countries at moderate risk are those whose baseline debt burden indicators are below their respective thresholds while stress test scenarios push the indicators above the thresholds. Countries would be downgraded from moderate to high risk when shocks lead to thresholds breaches under the baseline scenario.

3. **The distance between the baseline debt burden indicators and their thresholds is a measure of the extent of "space to absorb shocks" for a country to absorb shocks without being downgraded.** Staff proposes to use this "space to absorb shocks" measure as the basis for differentiating the extent of debt vulnerabilities among countries at moderate risk of debt distress, by assessing the adequacy of such "space" against potential shocks.

4. **Using DSAs produced since the LIC DSF inception, staff calculated the distribution of observed shocks that led to a rating downgrade to high risk of debt distress.** Such shocks are calculated as the observed change in debt burden indicators (peak in debt after and before the shock leading to a downgrade) in percentage of the respective threshold.

$$Shock_{i,t} = \frac{d_{i,t} - d_{i,t-1}}{threshold_i},$$

$$where\ d_i : \frac{PV\ of\ debt}{GDP}, \frac{PV\ of\ debt}{Exports}, \frac{Debt\ service}{GDP}, \frac{Debt\ service}{Exports}$$

Potential shocks are calibrated from their observed distributions (Figure AVI.2). For debt stock indicators (PV of debt to GDP and PV of debt to export), the median shock is around 20 percent while shocks in the upper quartile are those larger than 40 percent. For debt service indicators (debt service to GDP and debt service to export), the median shock is around 12 percent and shocks in the upper quartile are those larger than 35 percent.

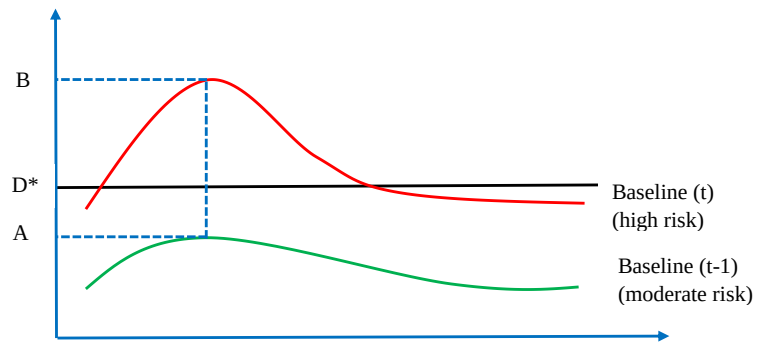
5. **The proposed tool compares the country-specific measure of "space" to absorb shocks with benchmarks derived from the observed distribution of shocks (Figure 2).** In this regard, countries facing a moderate risk of debt distress would be characterized as having:

- *"Limited space to absorb shocks"* where at least one baseline debt burden indicator is close enough to their respective threshold so that the median observed shock would downgrade it to high risk (i.e., the country-specific distance to threshold is less 20 percent of the threshold for debt stock indicators, and 12 percent for debt service indicators).

- *"Substantial space to absorb shocks"* where all baseline debt burden indicators are well below their respective thresholds, such that only shocks in the upper quartile of the observed distribution of shocks would downgrade the country to high risk of debt distress (i.e., the country-specific distance to thresholds across debt indicators is higher than 40 percent of the threshold for debt stock indicators and 35 percent for debt service indicators).

Figure AVI.1. Characterization of the Moderate Risk Category

(a) Illustrative Rating Downgrade from Moderate to High Risk



(b) Distribution of Observed Shocks Driving Downgrade from Moderate to High Risk

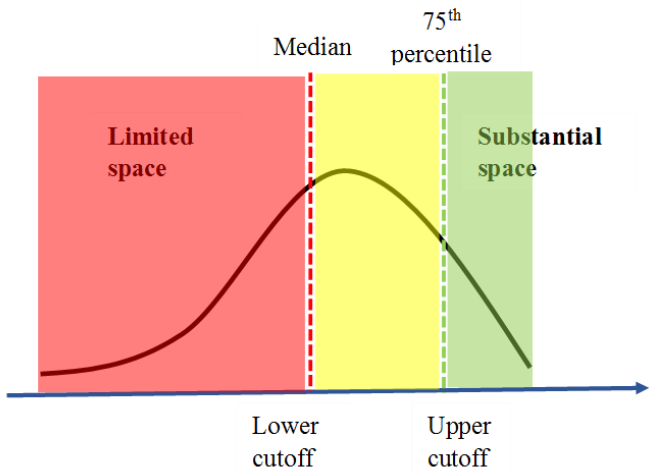
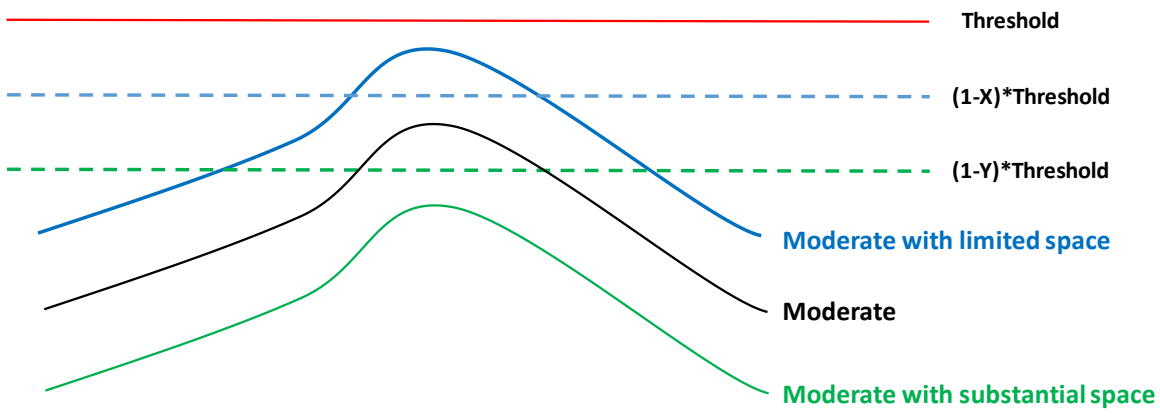


Figure AVI.2.
Illustration of Proposed Approach



Note: For the PV debt/GDP and PV debt/exports thresholds, X is 20 percent and Y is 40 percent. For debt service/exports and debt service/revenue thresholds, X is 12 percent and Y is 35 percent.

6. **The proposed methodology would increase the information value of the DSF without creating additional data needs and operational implications to Fund/Bank debt policies.** The Fund's Debt Limits Policy and the World Bank's Non-Concessional Borrowing Policy will be unaffected by the characterization of "space to absorb shocks" for countries in the moderate risk category. The rules under the two policies relevant for countries at moderate risk of debt distress would continue to apply, regardless of whether countries are characterized as having limited or substantial space. It is important to emphasize that this tool is a counterfactual exercise, and does not constitute a view on the likelihood of shocks occurring that could lead to a rating downgrade.

References

Abbas, S. M. A., and J. E. Christensen, 2007, "The Role of Domestic Debt Markets in Economic Growth: An Empirical Investigation for Low-Income Countries and Emerging Markets," IMF Working Paper No. 07/127, (Washington: International Monetary Fund).

Abbas, S. M. A., N. Belhocine, A. El Ganainy, and M. Horton, 2010, "A Historical Public Debt Database," IMF Working Paper No. 10/245, (Washington: International Monetary Fund).

Aisen, A., and D. Hauner, 2008, "Budget Deficits and Interest Rates: A Fresh Perspective," IMF Working Paper No. 08/42. (Washington: International Monetary Fund).

Aron, J., R. MacDonald, and J. Muellbauer, 2014, "Exchange Rate Pass-through in Developing and Emerging Markets: A Survey of Conceptual and Policy Issues, and Empirical Findings," *Journal of Development Studies,* 50(1): pp. 101–43.

Aslam, A., S. Beidas-Strom, R. Bems, O. Celasun, S. Kılıç Çelik, and Z. Kóczán, 2016, "Trading on Their Terms? Commodity Exporters in the Aftermath of the Commodity Boom," IMF Working Paper No. 16/27, (Washington: International Monetary Fund).

Bahmani-Oskooee, M., and A. Ratha, 2004, "The J-Curve: A Literature Review," *Applied Economics,* 36(13): pp. 1377–98.

Baldacci, E., I. Petrova, N. Belhocine, G. Dobrescu, and S. Mazraani, 2011, "Assessing Fiscal Stress," IMF Working Paper No. 11/100. (Washington: International Monetary Fund).

Beers, D. T., and J.-S. Nadeau, 2015, "Database of Sovereign Defaults," *Bank of Canada Technical Report* No. 101.

Berg, A., E. Berkes, C. Pattillo, A. Presbitero, and Y. Yakhshilikov, 2014, "Assessing Bias and Accuracy in the World Bank—IMF's Debt Sustainability Framework for Low-Income Countries," IMF Working Paper No. 14/48. (Washington: International Monetary Fund).

Berg, A., E. Buffie, C. Pattillo, R. Portillo, A. Presbitero, and L. Zanna, 2015, "Some Misconceptions about Public Investment Efficiency and Growth," IMF Working Paper No. 15/272. (Washington: International Monetary Fund).

Bom, P., and J. Ligthart, 2014, "What Have We Learned from Three Decades of Research on the Productivity of Public Capital?" *Journal of Economic Surveys,* 28(5): pp. 889–916.

Buffie, E., A. Berg, C. Pattillo, R. Portillo, and L. Zanna, 2012, "Public Investment, Growth, and Debt Sustainability: Putting Together the Pieces," IMF Working Paper No. 12/144. (Washington: International Monetary Fund).

Calderon, C., E. Moral-Benito, and L. Serven, 2015, "Is Infrastructure Capital Productive? A Dynamic Heterogeneous Approach," *Journal of Applied Econometrics,* 30(2): pp. 177–98.

Catão, L. A. V., and G. M. Milesi-Ferretti, 2014, "External Liabilities and Crises," *Journal of International Economics,* 94: pp. 18–32.

Cespedes, L. F., and A. Velasco, 2013, "Was This Time Different? Fiscal Policy in Commodity Republics," NBER Working Paper No. 19748.

Cruces, J. J., and C. Trebesch, 2013, "Sovereign Defaults: The Price of Haircuts," *American Economic Journal: Macroeconomics,* 5(3): pp. 85–117.

Das, U. S., M. G. Papaioannou, and C. Trebesch, 2012, "Sovereign Debt Restructurings 1950–2010: Literature Survey, Data, and Stylized Facts," IMF Working Paper No. 12/203. (Washington: International Monetary Fund).

Gelos, R. G., R. Sahay, and G. Sandleris, 2011, "Sovereign Borrowing by Developing Countries: What Determines Market Access?" *Journal of International Economics,* 83: pp. 243–54.

Giovanni, A., and M. de Melo, 1993, "Government Revenue from Financial Repression," *The American Economic Review,* Vol. 83, No. 4, pp. 953–63.

International Monetary Fund-The World Bank, 2012, "Revisiting the Debt Sustainability Framework for Low-Income Countries," (January) (Washington).

———, 2013a, "Unification of Discount Rates used in External Debt Analysis for Low Income Countries: Staff Proposals," (October) (Washington).

———, 2013b, "Staff Guidance Note on the Application of the Joint Bank-Fund Debt Sustainability Framework for Low Income Countries," (November) (Washington).

———, 2015, "Public Debt Vulnerabilities in Low-Income Countries: The Evolving Landscape," (November) (Washington).

International Monetary Fund, 2007, "The Changing Dynamics of the Global Business Cycle," Chapter 5, World Economic Outlook (October). (Washington).

———, 2012, "Commodity Price Swings and Commodity Exporters," Chapter 4, World Economic Outlook (April). (Washington).

———, 2014, "Assessing Reserve Adequacy—Specific Proposals," (December). (Washington).

———, 2015a, "Making Public Investment More Efficient," (June). (Washington).

———, 2015b, “Exchange Rates and Trade Flows: Disconnected?” Chapter 3, World Economic Outlook (October). (Washington).

———, 2015c, “Where Are Commodity Exporters Headed? Output Growth in the Aftermath of the Commodity Boom”, Chapter 2, World Economic Outlook (October). (Washington).

———, 2016, “Small States’ Resilience to Natural Disasters and Climate Change: Role for the IMF,” (November) (Washington).

———, 2017, “Estimating the Stock of Public Capital in 170 Countries: January 2017 Update”. http://www.imf.org/external/np/fad/publicinvestment/pdf/csupdate_jan17.pdf.

Laeven, L., and F. Valencia, 2013, "Systemic Banking Crises Database," *IMF Economic Review*, 61: pp. 225–70. (Washington: International Monetary Fund).

Ligthart, J., and R. Martin-Suarez, 2011, “The Productivity of Public Capital: A Meta-Analysis,” in Manshanden, W. and W. Jonkhoff, eds. Infrastructure Productivity Evaluation. (New York: Springer).

Mallik, G., and A. Chowdhury, 2001, “Inflation and Economic Growth: Evidence from Four South Asian Countries,” *Asia-Pacific Development Journal*, Vol. 8, Issue 1, pp. 123–35.

Manasse, P., and N. Roubini, 2005, “‘Rules of Thumb’ for Sovereign Debt Crises,” IMF Working Paper No. 05/42. (Washington: International Monetary Fund).

Manasse, P., N. Roubini, and A. Schimmelpfennig, 2003, “Predicting Sovereign Debt Crises,” IMF Working Paper No. 03/221. (Washington: International Monetary Fund).

Mauro, P. and M. Villafuerte, 2013, “Past Fiscal Adjustments: Lessons from Failures and Successes,” *IMF Economic Review*, 61(2): pp. 379–404. (Washington: International Monetary Fund).

Panizza, U., 2008, “Domestic and External Public Debt in Developing Countries,” *Unctad Discussion Paper* No. 188.

Panizza, U., F. Sturzenegger, and J. Zettelmeyer, 2009, “The Economics and Law of Sovereign Debt and Default,” *Journal of Economic Literature*, 47:3, pp. 651–98.

Pennings, S., 2017, “Long-term Growth Model v4.0: Model Description,” available at http://globalpractices.worldbank.org/mfm/Pages/SitePages/MFM Online Tools.aspx

Rand, J., and F. Tarp, 2002, “Business Cycles in Developing Countries: Are They Different?” World Development, Vol. 30, No. 12, pp. 2071–88.

Reinhart, C. M., and K. S. Rogoff, 2009, “This Time is Different: Eight Centuries of Financial Folly,” Princeton University Press.

———, 2010, “From Financial Crash to Debt Crisis,” *American Economic Review*, Vol. 101, No. 5, pp. 1676–706.

Reinhart, C. M., and K. S. Rogoff, 2011, “The Forgotten History of Domestic Debt,” *The Economic Journal*, Vol. 121, Issue 552, pp. 319–50.

Reinhart, C. M., and M. B. Sbrancia, 2015, “The Liquidation of Government Debt,” IMF Working Paper No. 15/7. (Washington: International Monetary Fund).

Spatafora, N., and I. Samake, 2012, “Commodity Price Shocks and Fiscal Outcomes,” IMF Working Paper No. 12/112. (Washington: International Monetary Fund).

UNCTAD, 2015, “State of Commodity Dependence, 2014,”(New York and Geneva: United Nations Conference on Trade and Development).

Yehoue, E., 2005, “International Risk Sharing and Currency Unions: The CFA Zones,” IMF Working Paper No. 05/95. (Washington: International Monetary Fund).