Price Shocks and Sovereign Debt Governance: Evidence from the IMF-World Bank Debt Sustainability Framework

Enock W.N. Bulime*

October 2026

Prepared for the Junior IO Scholars Workshop, Fall 2026. Please do not cite without permission.

Abstract

The IMF-World Bank Low-Income Country Debt Sustainability Framework (LIC-DSF) stress-tests commodity-dependent economies by shocking export prices and, where a country imports the same commodity it exports, netting import exposure against the export shock. This netting logic embeds the same structural omissions as the terms-of-trade ratio: it records only relative price movements between the two legs and discards their common movement entirely. Using structural export- and import-price shocks for twenty African economies from 1990 to 2019, I find this distinction matters in two respects. First, the two price components are not symmetric in their debt effects: a ten-percent export-price decline raises gross government debt by 5.7 percentage points of GDP after four years, while an equal import-price rise lowers debt by 3.4 points. A framework that nets these as offsetting reverses the sign of the import-price effect. Second, when export and import prices fall simultaneously, the most common pattern in African commodity data, the netting logic produces a near-zero predicted impact under a terms-of-trade representation, yet gross debt rises by 6.1 percentage points. The export-price stress test in the revised LIC-DSF captures the export leg of this episode and produces a debt increase close to the full common-bust estimate; however, treating import exposure as a negative offset to the export shock—rather than modeling the import leg separately—attenuates the assessed stress precisely when both prices are falling together. Both limitations have direct implications for IO design: the LIC-DSF is the primary instrument through which the IMF and World Bank assess debt sustainability, determine IDA grant allocation, and define the limits on non-concessional borrowing for low-income countries. A stress-test architecture that calibrates commodity exposure through net price movements rather than separate export- and import-price scenarios will underestimate sovereign vulnerability at the moment those institutions are deciding whether and how to intervene.

JEL codes: F34, F53, H63, O19, Q33.
Keywords: debt sustainability; LIC-DSF; terms of trade; common price shocks; international organizations; sovereign debt; Africa.

*Department of Economics, Kansas State University. Email: bulimeenock@ksu.edu. I am grateful to seminar participants for helpful comments. All errors are my own. *Note on abstract submitted to the workshop:* the abstract distributed with the acceptance notification cited a common-bust estimate of 9.1 percentage points from an earlier draft. The current estimate using the final shock identification is 6.1 percentage points; all figures in this paper use the current estimate.

1 Introduction

International organizations govern sovereign debt risk through frameworks that translate economic evidence into operational rules. These frameworks determine who receives concessional finance, on what terms, and how much fiscal adjustment is required in exchange. When the economic representations embedded in a framework do not distinguish between the forces driving commodity-price movements, the framework may produce assessments that systematically understate vulnerability for the class of episodes that matters most. This paper asks when that gap opens, how large it is, and what institutional mechanisms allow it to persist.

The IMF-World Bank Low-Income Country Debt Sustainability Framework (LIC-DSF) is the primary such instrument for low-income and lower-middle-income countries: it determines IDA grant eligibility, sets non-concessional borrowing ceilings, and calibrates the conditionality attached to Fund-supported programs. At its center is a commodity-price stress test that asks how a country's debt path would evolve under an adverse price shock. That stress test is built on the terms-of-trade ratio — a net-price representation that records only the difference between export and import prices. This choice has two implications that are empirically testable. First, it imposes equal and opposite fiscal effects of export- and import-price movements. Second, it registers near-zero fiscal consequence when both prices move together. Both restrictions are rejected for African commodity-dependent economies in Section 5.

Commodity prices co-move across unrelated goods, a pattern documented independently of crisis episodes (Pindyck and Rotemberg, 1990; Byrne, Fazio, and Fiess, 2013). This co-movement is the component that a net-commodity-exposure representation of price risk does not capture as an independent stress scenario; it is also what the LIC-DSF's import-offset provision can attenuate even within an export-price stress test. Between 87 and 97 percent of the 182 economies tracked by the IMF Commodity Terms of Trade database recorded simultaneous export- and import-price declines in 1998, 2009, and 2015 (the share varying across those three years) (Gruss and Kebhaj, 2019). In each of those years, an assessment based solely on the terms-of-trade ratio would have predicted no change in sovereign debt, because the ratio moved only modestly, yet gross debt rose by an average of 6.1 percentage points for the twenty African economies in the sample during those episodes.

Henning (2026, p. 8) proposes eight structural reforms to the LIC-DSF and documents that the best-performing DSA models raise false alarms and miss crises each in roughly one-third of cases. For commodity-dependent economies, this rate of misclassification has a structural source for commodity-dependent economies: the terms-of-trade ratio registers near-zero fiscal impact from common busts, the most common severe price episode in the Africa-20 panel.

I use structural export- and import-price shocks identified by Di Pace, Juvenal, and Petrella (2025) (hereafter DJP) and panel local projections for twenty African economies over 1990–2019 to estimate the separate fiscal consequences of each price component and of their joint movement. A ten-percent export-price decline raises general-government gross debt by 5.7 percentage points of GDP after four years, while an equal import-price rise *lowers* debt by 3.4 points. The LIC-DSF netting logic treats both effects as equal in magnitude and opposite in sign. When both prices fall by ten percent simultaneously (a "common bust"), the terms of trade barely changes and the stress test registers near-zero pressure, yet debt rises by 6.1 percentage points. Common busts are not rare events: they account for 57 percent

of severe external-price episodes in the Africa-20 panel.

Because the LIC-DSF shapes access to concessional finance, debt relief, and program support, underestimating debt stress in common-bust episodes reduces the counter-cyclical response of concessional support: countries enter the most severe price episodes with diminished access to concessional resources, and the institutions providing that support receive a distorted measure of the severity of the shock.

Two prior studies use bilateral price decompositions to study fiscal transmission. Spatafora and Samaké (2012) enter commodity export and import price indices as separate regressors in an Arellano-Bond dynamic panel model for a maximum of 116 LICs and middle-income countries (MICs) (1990–2010), finding that both price components raise deficits and debt; they note that positive co-movement amplifies the fiscal impact, but construct no common-bust scenario and test no symmetry restriction. Levy, Ricci, and Werner (2020) enter export and import prices as separate regressors, each controlling for the other, in panel local projections for a maximum of 194 countries (1970–2017); they find a strong negative export-price effect on debt in EMDC commodity exporters, but do not separately identify structural shocks, do not test the symmetry restriction, and do not construct a common-bust scenario. I extend both studies in three respects: replacing observed commodity price indices with structurally identified bilateral shocks (Di Pace, Juvenal, and Petrella, 2025), formally testing and rejecting the symmetry restriction that any terms-of-trade aggregator imposes, and establishing that the common-bust scenario generates 6.1 percentage points of additional debt at $h = 4$.

I make four contributions. First, I document the specific assessment gap that a net-price commodity architecture creates, using the LIC-DSF as the primary institutional case. Second, I translate the econometric estimates into the framework's operational language: the stress-test scenarios, the debt-path projections, and the threshold comparisons that drive DSA ratings. Third, I establish that export-price-centric architecture is not unique to the LIC-DSF: a review of seventeen sovereign debt assessment frameworks finds that none employs a bilateral export-import price decomposition. Fourth, I use the six-percentage-point common-bust estimate to quantify two IO mechanisms — normalization of deviance and constitutive authority — that Section 2 develops through three observable implications.

The LIC-DSF provides an unusually clean empirical test of when IOs revise their diagnostic frameworks and when they do not: the assessment gap is measurable, the staff have demonstrated analytical capacity, and three review cycles provide a structured test of institutional persistence. The argument proceeds in eight stages. Section 2 develops the IO theory and observable implications. Section 3 describes the LIC-DSF's architecture, decision rules, and the role of commodity stress tests. Section 4 develops the debt-accounting framework that makes the ratio's restrictions precise. Section 5 presents the identification strategy. Section 6 reports the empirical results, robustness checks, and transmission mechanisms. Section 7 maps the findings to specific DSA template cells and illustrates the gap between what the framework projects and what the evidence implies. Section 8 discusses implications for IO design. Section 9 concludes.

2 Why Technically Capable IOs Are Slow to Revise

2.1 Normalization of Deviance and Constitutive Authority

Barnett and Finnemore (1999, 2004) identify two mechanisms through which technically capable institutions systematically, rather than randomly, fail to revise diagnostic frameworks that produce wrong outputs.

The first mechanism is *normalization of deviance*: repeated application of a procedure that produces acceptable-looking outputs embeds that procedure as legitimate, insulating it from scrutiny of whether the outputs are accurate. Each review cycle in which IMF and World Bank staff apply the terms-of-trade stress test and generate ratings that clear normal scrutiny confirms the procedure's legitimacy, not its accuracy. The 2026 review is the third such cycle since 2005.

The second mechanism is *constitutive authority*: the classifications an IO produces do not merely describe the world; they constitute it. A "Low Risk" DSA rating does not report a finding; it creates an institutional fact that reshapes how creditors extend exposure, how donors allocate budgets, and how borrowing governments set fiscal targets. When the stress test produces "Low Risk" during a common bust—because the import-offset provision attenuates the assessed export-price stress or because the ToT ratio registers near-zero movement—the rating certifies the absence of distress, with effects that compound the fiscal deterioration the rating understates. The six-percentage-point debt estimate in Section 6 is the quantitative signature of this mechanism.

2.2 Epistemic Communities and the Prebisch-Singer Prior

The design gap persists across three review cycles not because staff lack the capacity to address it, but because the analytical tradition staff were trained in excludes import-price co-movement from the framework's analytical scope. Haas (1992) defines an epistemic community as a network of professionals sharing a causal model that they carry into institutional positions, where it shapes problem definition rather than functioning as one hypothesis among others. The export-price-centric view of commodity risk is that causal model.

It originates in Prebisch (1950) and Singer (1950), who documented that primary-commodity export prices decline secularly relative to manufactured goods and framed the commodity problem entirely from the export side. The possibility that both price components might fall together with independent fiscal consequences is absent from the Prebisch-Singer framework. Chwieroth (2010) documents how this professional consensus diffuses across institutional positions: staff trained in this tradition carry its assumptions into the institution, where they operate as technical defaults rather than debated choices. The diagnostic question encoded in the framework, "how much will resource-export revenues fall?", reflects the question Prebisch-Singer asks, not the question relevant to DSA: "how much will sovereign debt rise?"

2.3 Three Observable Implications

OI-1 (Normalization of deviance). If three review cycles operate as normalization, each should document the stress-test's outputs as satisfactory while leaving the common-bust

element unchanged. The test is not whether the framework was revised in any dimension (it was, in 2017–18 and 2026) but whether the common-bust element was retained. Three unchanged cycles despite quantifiable evidence of failure are the observable signature of normalization, not ordinary learning.

OI-2 (Constitutive authority). If DSA ratings exercise constitutive authority, then "Low Risk" and "Moderate Risk" classifications during common-bust episodes should reshape creditor, donor, and co-financier behavior. Financing flows should track the DSA rating rather than the underlying fiscal trajectory. An assessment gap of this kind should produce compounding effects: reduced grant access raises the stock of concessional debt, which raises debt-service burdens, which strains the fiscal position the next DSA assesses.

OI-3 (Epistemic community). If the export-price-centric prior is a field-wide analytical commitment rather than an institution-specific model choice, the common-bust structural omission should appear across independent organizations: multilateral development banks, the IMF's own external assessment tools, and private sovereign rating agencies. A principal-agent failure would be correctable by changing the principal; a resource constraint would be absent in better-resourced agencies; a simple error would be caught and corrected by independent organizations working separately. Seventeen independently produced frameworks all embedding the same ToT ratio is the observable implication of shared epistemic origin.

2.4 Alternative Explanations

Principal-agent slack. Agency-slack accounts (Copelovitch, 2010; Stone, 2011) predict that design gaps reflect political preferences of major shareholders exercised through agenda control and informal influence. This does not explain OI-3: private rating agencies (Fitch, Moody's, S&P) have no IMF principals. If all three embed the same structural omission as the LIC-DSF, the common cause cannot be principal preferences at one IO. The epistemic-community account predicts the cross-institutional pattern; the agency-slack account, which requires a common principal, cannot.

Measurement constraints. A measurement-constraint account argues that import price indices for developing countries were sparse in the early 2000s, making bilateral decomposition impractical. That constraint, genuine in the early 2000s, no longer obtains. The IMF's own External Balance Assessment constructs bilateral commodity import price indices across 42 categories for a wide sample of economies. The DJP database provides bilateral structural shocks for 38 developing economies. The 2026 reform used local projections to validate revenue and GDP responses to commodity shocks (Annex V), the same estimator used for debt responses, but did not extend to estimating debt responses. The constraint is not measurement; it is what the epistemic framework defines as the relevant dependent variable.

Genuine uncertainty. A genuine-uncertainty account argues that before this paper, the bilateral fiscal transmission was unknown, and the framework's designers made a reasonable inference. This account is inconsistent with the available historical evidence: Pindyck and Rotemberg (1990) established in 1990 that commodity prices moved together in excess of what macroeconomic fundamentals predicted. Byrne, Fazio, and Fiess (2013) confirmed the common factor with nonstationary panel methods on a broad sample. The co-movement of commodity prices has been documented since 1990. The staff who designed the 2005

framework and each subsequent revision worked within an institution that published the World Economic Outlook's commodity price tables. The availability of this co-movement evidence, combined with the framework's failure to incorporate it, is what the epistemic-community account must explain.

3 The LIC-DSF: Architecture, Stakes, and Design Logic

The IMF and World Bank established the Low-Income Country Debt Sustainability Framework jointly in 2005 and substantially revised it in 2017–2018 (IMF and World Bank, 2018). It assesses sovereign debt vulnerability for countries that access the International Development Association (IDA) window of the World Bank and the Poverty Reduction and Growth Trust (PRGT) of the IMF. As of 2026, the framework covers 70 countries across sub-Saharan Africa, South Asia, and other low-income regions.

The framework produces a *debt sustainability analysis* (DSA) for each country. A DSA has three outputs: (i) a baseline debt-path projection; (ii) a set of stress tests that project debt under adverse scenarios; and (iii) a composite risk rating, Low, Moderate, High, or In Distress, based on how debt-path projections compare against country-specific thresholds derived from a Debt Carrying Capacity (DCC) assessment.

3.1 What the Rating Controls

IDA grant eligibility. IDA allocates resources through grants and concessional loans and determines the grant share entirely by the risk rating. Under IDA20+ rules (in effect since July 2022), countries rated High Risk or In Distress receive a 100 percent grant share of their IDA allocation. Countries rated Low or Moderate Risk receive zero grants (100 percent concessional credits), except IDA-only small states at Moderate Risk, which retain a 50 percent grant share. Under IDA19, which covered most of the 2005–2021 period, Moderate Risk countries received 50 percent grants and High Risk received 100 percent. A country assessed as Moderate Risk in a common-bust year, when a two-price stress scenario would project High Risk, receives zero grants under IDA20+ rules instead of 100 percent. As of March 2026, 23 of 67 PRGT-eligible countries are rated High Risk and 9 are In Distress, so the correct classification of 32 countries determines full grant access (International Monetary Fund, 2026).

Non-concessional borrowing ceilings. Both the IMF's Debt Limits Policy (DLP) and the World Bank's Non-Concessional Borrowing Policy (NCBP) rely on the Debt Sustainability Framework (DSF) to define the room for non-concessional commercial borrowing (IMF and World Bank, 2018). Countries at higher risk ratings face tighter or zero NCB headroom, restricting commercial borrowing when their fiscal position appears vulnerable. A country whose DSA understates its vulnerability operates under a looser NCB ceiling than its actual position warrants: commercial borrowing is permitted at the moment the balance sheet is most fragile. Conversely, a country assessed at High Risk when a two-price scenario would project Moderate Risk faces NCB restrictions that may foreclose investment-grade commercial opportunities even when its fundamentals are sound.

G20 Common Framework. The Common Framework for Debt Treatment Beyond the Debt Service Suspension Initiative (DSSI), endorsed by the G20 and Paris Club in 2020, uses LIC-DSF assessments as the basis for restructuring negotiations. Countries that trigger treatment under the Common Framework (Zambia, Ethiopia, Ghana, Chad) do so in part because their DSA ratings indicate debt in distress or at high risk. Ghana's 2015 re-rating to High Risk, during the 2014–16 commodity bust, preceded its 2022 default and Common Framework entry by seven years.

Zambia's memorandum of understanding under the G20 Common Framework, reached in June 2023, references debt-carrying capacity thresholds in defining the terms of relief: the agreed restructuring parameters are conditioned on Zambia's DCC composite index remaining above specified levels through the projection horizon. The Composite Indicator of Debt-Carrying Capacity (CI-DCC) is defined as $\text{CI-DCC} = 0.377 \cdot \text{CPIA} + 4.139 \cdot \text{Real Growth} + 0.241 \cdot \text{Reserve Coverage} + 3.5 \cdot \text{Remittances} - 0.077 \cdot (\text{Interest Rate} \times \text{External Debt Share}) - 2.634 \cdot \text{Output Volatility}$. No commodity import price variable appears in the formula. A common bust that compresses export revenues, raises import costs, and slows growth enters the DCC calculation only through the growth and volatility terms; the bilateral price channel is absent. A country whose debt trajectory was shaped by a common-bust episode can therefore move from High Risk to In Distress while the DCC composite, which governs the restructuring target, incorporates no direct commodity price variable.

3.2 The Commodity Stress Test

The 2018 Guidance Note specifies that the tailored commodity-price stress test applies to "LICs where commodities constitute at least 50 percent of total exports (goods and services) over the previous three-year period, per the data provided in the DSF template" (IMF and World Bank, 2018, Table 10). Of 67 LIC-DSF countries for which commodity export data are available, 58–87 percent have primary commodity export shares exceeding this threshold (the share varying with reference year and data vintage) (UNCTAD, 2026). When triggered, the default shock captures a one-standard-deviation decline in export prices, with the price gap converging to the baseline over six years. The interactions are calibrated as follows: 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 three years starting from year 2, per 10-percentage-point contraction in commodity prices. Where the same commodity is both a major export and a major import, the framework permits, but does not require, an import mitigation adjustment (the G_{49} and G_{50} cells in the DSA workbook's "Input 6" sheet) to reduce the net export share entering the composite shock.

The test incorporates one price, the export price, and assigns the import price either no weight or a mitigating weight that attenuates the export shock. In a common bust, export prices fall, so the C3 stress test is triggered and captures the export leg of the fiscal response. The structural estimates in Section 6 imply this export leg accounts for approximately 5.7 of the 6.1 percentage points of debt accumulation at $h = 4$. The gap relative to the ToT ratio—which assigns zero fiscal consequence when both prices fall proportionally—therefore arises from two distinct sources. First, the ToT ratio suppresses the common component by construction; the C3 test avoids that suppression. Second, the netting provision, where applied, reintroduces a net-price logic: by treating import exposure as a mitigant, it reduces the modeled export shock in precisely the scenarios where both prices are declining together.

A framework that permits but does not require that adjustment cannot guarantee that common-bust stress is fully assessed.

The calibration horizon also does not match commodity price cycle dynamics. Cashin, McDermott, and Scott (2002) analyze boom and slump episodes across a large set of world commodity prices and find that, for most commodities, price slumps last longer than price booms and that the price decline during a slump is larger in amplitude than the subsequent rebound. The probability that a slump ends is independent of its duration; slumps do not exhibit mean-reversion within the six-year convergence window the LIC-DSF stress test assumes. The LIC-DSF stress test applies a one-period shock converging to baseline over six years, a design that assumes the price episode has converged to the baseline within the six-year convergence window.

3.3 The September 2026 Review

The 2026 reform's Annex V recalibration of the commodity shock constructed country-specific commodity price indexes as "weighted averages of raw materials, energy, food and metals prices, with weights given by each country's share in individual country exports" (IMF and World Bank, 2026, Annex V, para. 8). The 2026 review then estimated local projection regressions on these export-weighted price indexes to validate the GDP growth and revenue elasticities embedded in the stress test. No import-weighted price index was constructed, and no import-price local projection was estimated.

The IMF and World Bank completed a review of the LIC-DSF in September 2026 (IMF and World Bank, 2026). After the IMF and World Bank consulted 50 external organizations and engaged both Executive Boards, the 2026 Policy Paper stated: "the commodity price shock and market financing shock have been reviewed and would be retained" (para. 58). The review recalibrated the retained shock as a shock to export *growth* rather than export price level and added a Gross Financing Needs (GFN) ratio as a new liquidity threshold alongside the existing debt-to-GDP and PV ratios. It did not add a joint export-import price stress scenario, and it did not modify the netting logic for common price movements. The reform's own summary table (Box 1) lists the new stress tests: an updated natural disaster test, a new banking crisis stress test, and a new domestic financing stress test (IMF and World Bank, 2026, Box 1).

A second architectural decision was preserved: the CI-DCC, which determines a country's DCC group and thus which thresholds apply to its debt ratios, was recalibrated with six macroeconomic variables but no import price exposure term. Oil and food importers with high import-price exposure carry no higher CI-DCC penalty than economies without that exposure. The reform states that output volatility "captures differences in LICs economic structure and exposure to shocks, as more diversified economies ... are generally more resilient to external shocks than countries with economies dominated by a small number of industries, such as those of commodity exporters" (IMF and World Bank, 2026, Annex I, para. 10). Output volatility proxies for commodity exposure, but only the export side: an economy with low export concentration but high food or fuel import dependence carries the same CI-DCC score as one without it.

A parallel asymmetry appears in Annex V's recalibration of the stress-test interaction coefficients. The reform finds that currency depreciation has a near-zero effect on export volumes in LICs, attributing this to "high import content in export production, rising im-

ported input costs following depreciation, and commodity price-taking behavior" (IMF and World Bank, 2026, Annex V, para. 6). The 2026 review used that finding to recalibrate the depreciation-to-export-growth interaction coefficient (Table V.2 of the Annexes), but did not extend the commodity stress test to incorporate import price shocks.

The reform's own back-testing additionally identified "some optimism bias in revenue and exports forecasts over longer horizons" (the scale of which the reform report does not quantify), indicating that the baseline forecast challenge (Gaudin, Harnoys Vannier, and Kessler, 2024) and the stress-test architecture gap are distinct and compounding. Back-testing covering LIC-DSAs issued between 2018 and December 2025 found that one in five High Risk countries had no threshold breaches over the ten-year forecast horizon; they were classified solely on the basis of long-term climate risk judgments with no mechanical model trigger (IMF and World Bank, 2026, Figure 6). The coexistence of false positives from long-term climate judgment and false negatives arising from the common-bust omission indicates that classification errors are bidirectional.

4 A Debt-Accounting Framework

4.1 The Budget Identity and the Two Price Legs

To make the ratio's restrictions precise, I follow Di Pace, Juvenal, and Petrella (2025) in treating export and import prices separately. Let $P_{x,it}$ and $P_{m,it}$ be the export and import price indices of country i in year t, with log changes $x_{it} \equiv \Delta \ln P_{x,it}$ and $m_{it} \equiv \Delta \ln P_{m,it}$. The terms of trade is $\mathrm{ToT}_{it} \equiv P_{x,it}/P_{m,it}$, so $\Delta \ln \mathrm{ToT}_{it} = x_{it} - m_{it}$.

Gross government debt D_{it} evolves as

$$D_{it} = (1 + r^d_{it})\, D_{i,t-1} - PB_{it} + SFA_{it}, \tag{1}$$

where r^d_{it} is the effective interest rate, PB_{it} the primary balance, and SFA_{it} the stock-flow adjustment collecting valuation changes and off-budget transactions. The primary balance decomposes as

$$PB_{it} = R^x_{it}(P_{x,it}) + R^o_{it} - S^m_{it}(P_{m,it}) - S^o_{it}, \tag{2}$$

where R^x_{it} is commodity-linked revenue (royalties, profit taxes, state-owned enterprise transfers); S^m_{it} is import-priced expenditure (procurement, subsidies); and R^o_{it}, S^o_{it} collect other revenues and spending. The two price-sensitive components need not satisfy any equality.

4.2 The Terms-of-Trade Restriction and Its Limits

A first-order expansion gives

$$\frac{\partial \Delta^{(h)} y_{it}}{\partial \Delta \ln P_{j,it}} = \underbrace{\frac{B^{(h)}_{j,i}}{\bar{Y}_{i,t+h}}}_{\text{primary balance}} + \underbrace{\left[\frac{V^{(h)}_{j,i}}{\bar{Y}_{i,t+h}} - \frac{\bar{D}_{i,t+h}}{\bar{Y}^2_{i,t+h}} N^{(h)}_{j,i}\right]}_{\text{remainder}}, \quad j \in \{x, m\}, \tag{3}$$

where $B^{(h)}_{j,i}$ is the primary-balance contribution, $V^{(h)}_{j,i}$ collects other numerator effects, and $N^{(h)}_{j,i}$ is the nominal GDP response. A price change can worsen the primary balance while com-

pressing the debt ratio if the nominal-GDP elasticity exceeds the primary-balance elasticity in absolute value (the import-side finding below).

Consider the panel local projection

$$\Delta^{(h)} y_{it} = \alpha_i^{(h)} + \beta_x^{(h)}\, x_{it} + \beta_m^{(h)}\, m_{it} + \mathbf{w}_{it}'\, \boldsymbol{\gamma}^{(h)} + \varepsilon_{it}^{(h)}, \tag{4}$$

where $\alpha_i^{(h)}$ is a country fixed effect and $\mathbf{w}_{it}$ a vector of lagged controls. The restriction that any terms-of-trade mapping imposes is

$$H_0: \quad \beta_x^{(h)} + \beta_m^{(h)} = 0 \quad \text{for each } h. \tag{5}$$

Rotating the prices into a relative direction $v_{it}^r = x_{it} - m_{it}$ and a common direction $v_{it}^c = x_{it} + m_{it}$ reveals two distinct failure modes. A terms-of-trade regression that includes only v_{it}^r assigns its coefficient to movements in either direction, reversing the sign of the import response when H_0 fails, and assigns zero fitted effect to v_{it}^c. When both prices fall by the same proportion, $v_{it}^r = 0$ and the regression predicts no change in debt, regardless of how much debt actually rises.

Proposition 1 (Two limitations of the ToT ratio)**.** *When H_0 fails: (i) the terms-of-trade regression assigns the wrong sign to import-price movements if $\beta_m^{(h)}$ has the opposite sign from $-\beta_x^{(h)}$; and (ii) a simultaneous equal-proportion decline in both prices, a common bust, is predicted to leave debt unchanged, even if the sum $\beta_x^{(h)} + \beta_m^{(h)} \neq 0$ implies a substantial debt increase.*

The IMF's own research staff has identified commodity price shocks as among the leading determinants of LIC debt distress. Bal Gündüz (2017) estimates a model of debt distress determinants for low-income countries and ranks adverse shocks to real non-oil commodity prices among the three most prominent risk factors, alongside weak property-rights institutions and high initial debt burdens. She further finds that persistent common commodity shocks can generate clustering of debt distress episodes across countries, a pattern consistent with the common-bust mechanism. This finding is inconsistent with a stress test built on the terms-of-trade ratio, which produces a zero predicted impact when export and import prices fall in parallel.

Schmitt-Grohé and Uribe (2018) estimate structural VARs for 38 low-income and emerging-market economies and find that terms-of-trade shocks explain, on average, 10 percent of movements in aggregate output, compared with the 30 percent attributed to such shocks by calibrated business-cycle models, a gap the authors describe as a disconnect between theoretical and empirical models. The scalar ratio, P_x/P_m, is the observable in both exercises: it collapses two potentially co-moving price series into a single index, so any shock that moves export and import prices in the same direction registers as a smaller terms-of-trade movement than the underlying price-level changes warrant. To the extent that common busts are a structural feature of commodity markets, as established by Pindyck and Rotemberg (1990) and Byrne, Fazio, and Fiess (2013), the ratio will mechanically understate co-movement as a source of fiscal fluctuations, and frameworks built on it embody the same measurement error.

5 Empirical Strategy

5.1 Specification

I estimate equation (4) by panel local projection (Jordà, 2005; Jordà and Taylor, 2025) for each horizon $h = 0, \ldots, 8$ using the Africa-20 panel over 1990–2019. The control vector $\mathbf{w}_{it}$ contains two lags of the debt ratio, both price changes, domestic real GDP growth, CPI inflation, the log change in the real effective exchange rate, and world real GDP growth. Both prices enter jointly; omitting one assigns the global factor to the remaining price and biases the symmetry test.

The long-difference specification $\Delta^{(h)} y_{it} = y_{i,t+h} - y_{i,t-1}$ follows Jordà and Taylor (2025). I retain country fixed effects rather than difference them out, because each country's commodity structure is correlated with its price exposure. Standard errors cluster by country, permitting arbitrary within-country serial correlation (Cameron, Gelbach, and Miller, 2008). Driscoll-Kraay intervals complement the clustered estimates for cross-sectional dependence (Driscoll and Kraay, 1998). Montiel Olea and Plagborg-Møller (2021) establish that lag-augmented local projections with standard critical values are asymptotically valid uniformly across both stationary and integrated processes and across a wide range of horizons; I implement this by including two lags of each control variable, which also guards against the horizon-increasing bias identified by Teulings and Zubanov (2014).

5.2 Structural Identification

The observed price changes share a dominant global factor, accounting for 68 percent of the joint variance in raw price innovations. A global demand shock reduces both prices simultaneously, generating a correlation between the joint price decline and global conditions that also affect African fiscal accounts through aid flows, remittances, and borrowing costs. I address this by replacing the observed price changes with structural export- and import-price shocks from Di Pace, Juvenal, and Petrella (2025) (hereafter DJP). DJP identify these shocks by imposing sign restrictions, relative-elasticity conditions, and narrative sign patterns on bilateral vector autoregressions for each of 38 developing economies. Their orthogonalization separates export-specific, import-specific, global-demand, and global-supply disturbances.

The structural specification replaces the price changes in equation (4) with the DJP posterior-median shocks $u_{x,it}$ and $u_{m,it}$, each scaled to unit standard deviation. The common direction still accounts for 56 percent of the joint variance of the shocks, down from 68 percent in raw price innovations, because the identification requires each shock to dominate its own price but does not impose zero cross-price correlation. Cross-country synchronization reinforces this: the mean pairwise correlation of the common component $u_x + u_m$ across Africa-20 economies is 0.673, against 0.057 for the relative component $u_x - u_m$ (Appendix Table 17). The common shock is both the dominant direction in variance and the most synchronized across countries, the direction the terms-of-trade ratio discards. Causal results are conditional on DJP's maintained identification assumptions.

The common-bust scenario requires special construction because each structural shock also moves the other price. Let $\mathbf{A}$ be the 2×2 impact matrix whose element a_{jl} is the response of $\Delta \ln P_{j,it}$ to a unit shock $u_{l,it}$, estimated at $\hat{g}_{x,0} = 0.096$ and $\hat{g}_{m,0} = 0.036$ on the diagonal. A single-leg scenario that moves export prices by ten percent is constructed by solving for

the shock magnitude s such that $a_{xx} \cdot s = 0.10$; the debt response is then $\hat{\beta}_{u_x}^{(h)} \cdot s$. A common bust cannot be constructed by summing the two single-leg scenarios because each shock also moves the other price. The scenario instead solves $\mathbf{As} = \mathbf{p}$ for the shock combination $\mathbf{s}$ that drives both prices down by 0.10 log points simultaneously, and its debt response is $\mathbf{b}^{(h)\prime}\mathbf{A}^{-1}\mathbf{p}$ with $\mathbf{b}^{(h)} \equiv (\hat{\beta}_{u_x}^{(h)}, \hat{\beta}_{u_m}^{(h)})'$.

The twenty African economies are the African subset of DJP's 38-economy panel; selection is based on bilateral price-data coverage (minimum 15 years of both export- and import-price observations), not debt-distress status or LIC-DSF eligibility. Twelve appear on the World Bank LIC-DSF country list. The panel ends in 2019, capturing debt trajectories through the last pre-pandemic cycle but excluding the post-2019 Common Framework restructuring cases (Zambia, Ethiopia, Ghana's 2022 default).

Export and import price indices are DJP's bilateral country-specific series, each weighting 62 disaggregated commodity and manufacturing price series by the country's time-varying trade structure. General-government gross debt, revenue, expenditure, the primary balance, and interest payments come from the April 2025 IMF World Economic Outlook (International Monetary Fund, 2025). Gross debt is the debt stock assessed under the LIC-DSF framework. Nominal GDP, the GDP deflator, and the real effective exchange rate come from the Global Macro Database (Müller, Xu, Lehbib, and Chen, 2025). Effective capital- and labor-income tax rates come from Bachas, Fisher-Post, Jensen, and Zucman (2026). Sovereign restructuring events come from the Horn, Reinhart, and Trebesch dataset and the Asonuma and Trebesch dataset.

6 Empirical Results

6.1 Asymmetric Debt Responses and the Symmetry Rejection

Panel local projections on the structural price shocks yield asymmetric estimates for the two price components: the export coefficient is 4.4 times as large as the import coefficient at $h = 4$. A one-standard-deviation positive export-price shock reduces gross government debt at every horizon through $h = 4$, by 5.49 percentage points of GDP at $h = 4$ ($SE = 1.19$; Table 3). A one-standard-deviation positive import-price shock also reduces debt, by 1.24 points at $h = 4$ ($SE = 0.70$), with the effect individually significant at the 10 percent level at $h = 1$, $h = 3$, and $h = 4$ and at the 5 percent level at $h = 2$.

Rescaling to a ten-percent own-price movement using the impact matrix formula, a ten-percent export-price decline raises debt by 5.7 percentage points after four years, while a ten-percent import-price rise lowers debt by 3.4 points (Table 4). A terms-of-trade regression assigns +3.9 percentage points to each: the the opposite sign on the import leg and an inaccurate magnitude on both.

The restriction $H_0 : \beta_x^{(h)} + \beta_m^{(h)} = 0$ is rejected at the 1 percent level from $h = 0$ through $h = 4$ in both the Africa-20 sample and in DJP's broader 38-economy panel under the reduced-form specification. The rejection reflects the asymmetry in magnitudes: the export-price coefficient is negative and statistically significant throughout, while the import-price coefficient, though negative, is never large enough in absolute value to satisfy $\hat{\beta}_m = -\hat{\beta}_x$ at any horizon where the export coefficient is significant. The pattern is robust to dropping outcome years after 2019, split-panel jackknife estimation, and every lag-order specification

from one to four lags.

The netting logic embedded in the framework implies that when a commodity is both exported and imported, the import channel buffers the fiscal damage from the export shock. The symmetry rejection implies that no such buffer operates at the magnitudes the framework's netting restriction requires.

6.2 The Common-Bust Gap

When both export and import prices fall by ten percent simultaneously, the structural estimates imply a gross debt increase of 6.1 percentage points of GDP at $h = 4$ (95-percent interval $[2.7, 9.6]$, wild cluster bootstrap). The terms-of-trade regression assigns zero: the ratio is unchanged, so its prediction is flat. The interval excludes zero from $h = 1$ onward.

Common busts are not tail events: in the Africa-20 panel, 30.3 percent of country-years have both the structural export and import shocks negative. Defining a severe price episode as one in which the terms of trade or the sum of the two price changes declines by more than ten percent, 57 percent of such episodes are common busts that a terms-of-trade screen would not detect. In the nine country-years in which both prices fell by more than ten percent, the median four-year debt increase was 13.1 percentage points of GDP with no controls, 2.1 times the 6.1-point common-bust scenario estimate, consistent with the co-occurrence of adverse fiscal conditions during those extreme episodes.

In 2001 and 1998, 18 and 16 of the 20 sample economies were simultaneously in a common bust; the panel-mean terms-of-trade change was positive in both years (+0.177 in 2001 and +0.125 in 1998; Appendix Table 16), because both price components contracted proportionally, leaving the ratio near its pre-shock level. The 2015 and 2019 episodes follow the same pattern: 15 of 20 economies entered a common bust in 2015, with the terms-of-trade change negative but 5.1 times smaller in absolute value than the common component, and all 20 economies entered a common bust in 2019, with the terms-of-trade change 30.7 times smaller in absolute value than the common component. Any LIC-DSF commodity stress test run in the year before these episodes would have shown no commodity pressure from the common direction.

The debt accumulation links to actual sovereign distress. Using the Horn, Reinhart, Trebesch and Asonuma, Trebesch panels covering bank, bond, preemptive, and post-default restructurings, I identify restructuring entry for sixteen of the twenty sample economies. A positive structural export-price shock is associated with a higher probability of entering a debt restructuring within one to three years ($\hat{\beta}_{u_x} = 0.063$, $p = 0.024$ at $h = 1$, linear probability model; Appendix Table 6), consistent with a boom-bust dynamic in which elevated export prices precede a borrowing surge that becomes unsustainable. Import shocks are insignificant at all horizons.

6.3 Fiscal Transmission Mechanisms

A fiscal-accounting decomposition separates the debt response into a primary-balance contribution and a remainder capturing nominal-GDP effects, valuation changes, and off-budget transactions.

For common busts, debt rises through nominal GDP compression, not fiscal deterioration. The primary balance contribution at $h = 4$ is -1.2 pp; debt rises in ratio terms because

the denominator contracts (Appendix Table 11). The budget share is negative at every horizon in both the full sample and the sample that excludes the two major crisis episodes (2008–09 and 2014–16), so the debt response is not attributable to the crisis episodes. Excluding those episodes, the common-bust debt effect is larger (10.4 pp), not smaller. Debt accumulation operates predominantly through a remainder channel (nominal GDP compression, debt valuation effects, and financing pressure): the primary balance contributes a small partially offsetting reduction (-1.2 pp at $h = 4$) while the remainder accounts for the full net debt increase. The LIC-DSF commodity stress test evaluates fiscal flows, export-price gaps, and revenue projections; it does not model these channels.

The revenue channel dominates export-price shock transmission. A one-standard-deviation positive export shock raises government revenue at $h = 1$ and $h = 2$, with corporate income tax receipts and effective capital-income tax rates rising simultaneously, consistent with transmission through commodity-linked capital income rather than labor income. The primary-balance contribution accounts for 43 percent of the total debt response at $h = 4$ (between 38 and 62 percent across horizons) and is the largest single contributor; adverse export-price shocks also raise domestic issuance and increase debt through channels outside the primary balance.

Import-price shock estimates are structurally atypical: a positive import shock worsens the primary balance while reducing the debt ratio. The primary-balance contribution raises debt by 0.97 points at $h = 4$, while the remainder reduces it by 2.21 points. The evidence from a restricted sample excluding high-inflation episodes points to the GDP deflator: a higher import price index raises nominal GDP, compressing the debt-to-GDP ratio in the denominator even as the numerator worsens through procurement and subsidy costs. Interactions with predetermined foreign-currency debt shares are insignificant at all horizons, ruling out a foreign-currency revaluation channel. This denominator channel is not captured by the LIC-DSF and does not operate through any fiscal rule: no fiscal rule or intervention mechanism in the data offsets it in response to import prices.

For common busts, the four-year debt increase at $h = 4$ divides nearly equally between numerator and denominator contributions (3.05 percentage points through numerator effects and 3.07 through denominator effects in the restricted sample). The numerator contribution operates through domestic issuance, the channel the revised framework's new gross-financing-needs (GFN) threshold is designed to capture.

6.4 Identification Robustness

Posterior draws. Re-estimating the Africa-20 bilateral VARs with DJP's narrative sign restrictions and projecting debt on 50,000 posterior draws, the export coefficient is negative in 99.6 to 100 percent of draws and symmetry is rejected at the 10 percent level in 97.8 to 99.5 percent of draws.

Shift-share instruments. Three shift-share designs interact predetermined trade shares with world commodity prices, sharing no input with DJP's structural rotation. All three have strong first stages (Cragg-Donald $F > 10$ in each design); the Anderson-Rubin test rejects joint irrelevance of the two prices ($p = 0.0001$); the first two reject symmetry at every horizon; the third implies a common-bust response of 8.2 points ($SE = 2.7$), 2.1 percentage points above the structural estimate of 6.1 points.

Out-of-sample global test. A BACI/Pink Sheet shift-share instrument extended to all 51 commodity-dependent developing economies with at least 15 years of debt coverage rejects $H_0 : \beta_x + \beta_m = 0$ at $p < 0.001$ at every horizon $h = 0$–4 (Appendix Table 12). Panel B excludes all 13 African economies in the main sample, leaving 38 non-Africa-overlap countries spanning all four World Bank income groups and major non-African exporters (Russia, Chile, Indonesia, Kazakhstan). The rejection holds in the out-of-sample panel at $p < 0.001$ at every horizon. The symmetry failure that motivates the LIC-DSF critique is not specific to African economies, to low-income countries, or to DJP's structural identification assumptions.

Reduced-form global bilateral LP. A reduced-form bilateral panel local projection across 153 countries (1990–2018), estimated with country fixed effects and controls for inflation, real GDP growth, and the real effective exchange rate, finds a combined coefficient of -34.0 percentage points on the sum of export and import price changes at $h = 4$ ($p = 0.042$), ranging from -10.1 to -34.0 across horizons ($h = 0$ to $h = 4$). Because both prices fall in a common bust, this negative combined coefficient implies debt rises — consistent in direction and order of magnitude with the Africa-20 structural estimate. The fiscal channel is not specific to the African sample.

WTI instrument. For the six oil-exporting economies, a West Texas Intermediate price instrument, requiring no DJP structure, raises revenue, improves the primary balance, and lowers debt at every horizon, consistent in sign and comparable in order of magnitude with the Africa-20 structural estimates.

Inference under restricted wild cluster bootstrap produces p-values consistent with the asymptotic ones. Excluding any single country leaves the $h = 2$ reduced-form import coefficient between -79.0 and -59.5 (percentage points of GDP per standard-deviation shock). Adding the World Uncertainty Index (Ahir, Bloom, and Furceri, 2022) as a control attenuates the import-price coefficient (from -66.7 to -48.8 percentage points of GDP per standard-deviation shock, at $h = 2$); the symmetry rejection is maintained. The only specification that overturns the rejection adds proxies for U.S. monetary conditions that reduce the usable sample by approximately half; on that smaller sample the baseline still rejects ($p = 0.015$).

The symmetry rejection is heterogeneous across commodity groups. Splitting the Africa-20 panel into oil exporters (Algeria, Gabon, Equatorial Guinea, Nigeria, Sudan, Chad) and agricultural and mineral exporters (all others), the symmetry rejection is robust for agricultural and mineral exporters ($p \leq 0.008$ at $h = 2$ and $h = 4$) but not significant for oil exporters. The LIC-DSF tailored stress test is designed primarily for commodity-dependent economies, a category dominated in sub-Saharan Africa by agricultural and mineral exporters (coffee, cocoa, cotton, copper, gold, tobacco, vanilla) rather than petroleum. The finding implies that the stress-test architecture is least reliable for the agricultural and mineral exporters (coffee, cocoa, copper, gold) that the tailored stress test was designed to cover.

7 Mapping the Common-Bust Gap to DSA Practice

7.1 Stress-Scenario Comparisons at the Threshold

The LIC-DSF commodity stress test produces a projected debt-to-GDP path under an adverse price scenario and compares it against country-specific thresholds. The framework implicitly assumes that a one-standard-deviation export-price shock is the binding constraint on sovereign debt.

The structural estimates are inconsistent with this assumption. I first decompose price variation into the common direction ($u_x + u_m$, which the ratio discards) and the relative direction ($u_x - u_m$, which it retains). Structural debt local projections on each direction indicate that the common direction accounts for 61–85 percent of the combined semi-partial R^2 attributable to price shocks in debt regressions, at horizons $h = 0$, 2, and 4 respectively (Appendix Table 10). By using the ToT ratio, a terms-of-trade representation does not capture the dominant direction of price-shock variation for sovereign debt. Table 1 translates this into the DSA's natural unit:

Table 1: Debt Response Scenarios at $h = 4$: Evidence versus Framework

Scenario	ToT change	Debt Δ (pp GDP)	Framework treatment
Export bust (-10% P_x)	-10%	$+5.7$	Stressed (default)
Import cost ($+10\%$ P_m)	-10%	-3.4	Buffers export stress
Common bust (-10% both)	0%	$+6.1$	Not stressed

Notes: Debt response estimated from structural DJP price shocks, Africa-20 panel, 1990–2019. Common-bust interval: $[2.7, 9.6]$, wild cluster bootstrap. Import-cost interval: $[-7.2, 0.4]$.

The framework stresses the export bust, the first row, and treats the import-cost scenario as a mitigant that reduces the net shock. The import-cost scenario, however, *lowers* debt (-3.4 points), so netting it against the export shock understates stress: a negative-debt scenario is treated as a positive-debt offset.

The common-bust scenario in the third row carries a debt increase ($+6.1$ points) larger than the default export-bust stress, yet registers a terms-of-trade change of zero. A DSA that relies on the ToT ratio would assign no commodity stress to this scenario. The export-price stress test detects the export leg and produces a debt increase close to the full common-bust estimate ($+5.7$ points); however, where the optional import-offset provision is applied, that assessed stress is reduced by netting import exposure against the export shock, understating the combined fiscal pressure precisely when both prices are falling.

7.2 The C3 Template in Practice: Coverage and Timing

The FinDevLab LIC-DSA Database, covering 606 DSAs across 68 countries from 2013 to 2024, permits a direct comparison between the DSF's own commodity stress calibration and the empirical increment (Gaudin, Harnoys Vannier, and Kessler, 2024). Of the 142 post-2017 DSAs containing a tailored commodity price stress scenario (the C3 scenario in the revised template), the 62 Moderate-rated assessments for which Debt-Carrying Capacity tier data

are available imply an average commodity stress peak of 26.5 percent of external debt to GDP, against an average DCC-determined threshold of 36.4 percent, with an average increment above the baseline projection of +6.7 percentage points. This increment of +6.7 percentage points is within 0.6 points of the 6.1 percentage point common-bust estimate from Table 1: the DSF's own stress scenario projects a debt deterioration consistent with the empirical estimate.

What the threshold structure does not register, however, is that 91 percent of Low- and Moderate-rated countries in the sample, 65 DSAs for which DCC data permit classification, have commodity stress peaks below the applicable DCC-determined threshold (30 percent for Weak, 40 percent for Medium, 55 percent for Strong capacity). The framework therefore classifies 91 percent of these economies as safe even under a scenario that projects a deterioration numerically similar to the common-bust estimate (within 0.6 percentage points). The diagnostic gap is not the magnitude of the stress increment but the direction of the scenario: the C3 template simulates an export-price fall or an import-price rise, not both simultaneously. The common-bust channel adds the same increment a second time, under a scenario the framework does not evaluate.

The retained Annex V commodity stress module calibrates debt impact via a revenue-elasticity path: each 10 percentage-point contraction in the commodity export price reduces fiscal revenues by 0.75 percentage points of GDP, applied over the five-year projection horizon, implying a cumulative debt accumulation of 3.75 pp (5×0.75 pp) at the evaluation horizon. The composition-aware common-bust path implies 8.2–8.5 pp, depending on identification strategy. The 4.8-point difference represents the import-price channel that the export-weighted aggregator does not capture.

A further architectural problem is timing. The import-price effect on sovereign debt peaks at $h = 2$ (4.1 pp) before partially reversing by $h = 4$ (3.4 pp). Annex V evaluates debt sustainability over a five-year projection horizon, a window long enough that the peak import-price effect has materialized and begun to fade before the framework evaluates whether a stress scenario would push the debt path above a threshold.

7.3 Three Global Common Busts: 1998, 2009, 2015

Three global commodity episodes illustrate the gap between projected and realized debt paths. In 1998, the Asian financial crisis and Russian default produced a synchronized collapse in commodity prices: 16 of 20 Africa-20 economies entered a common bust, yet the panel-mean terms-of-trade change was positive (+0.177; Appendix Table 16), because both price components fell proportionally. In 2009, the global financial crisis produced a similar pattern: 14 of 20 economies experienced declines in both price indices exceeding five percent.

In each of these years, an LIC-DSF commodity stress test would have registered near-zero additional pressure. A common-bust stress scenario would have projected an additional 6.1 percentage points of debt on average across sample economies that experienced the episode, in the years when commodity exposure was most synchronized.

The Annex V export-bust stress test was predominantly identifying common-bust episodes without classifying them as such. Applying an HP filter to isolate years of export-price decline exceeding the HP-filter threshold in the Africa-20 panel, 76 percent of those bust episodes also had import prices falling simultaneously, meaning the framework's stress measure captured only the single-leg channel, attributing the full fiscal damage to the export-revenue

component. The framework was identifying the correct countries in the relevant years but attributing the pressure to the wrong mechanism, and therefore projecting only the export-revenue channel (≈ 3.75 pp at $h = 4$) instead of the full two-leg response (8.2–8.5 pp under both DJP structural and reduced-form shift-share identification).

The assessment gap operates in both directions. In episodes where the terms-of-trade ratio registers a large negative value driven by rising import prices with flat export prices, the framework projects debt accumulation, but debt actually falls by 3.4 pp because the GDP-denominator channel offsets the budget deterioration. The framework simultaneously overstates debt accumulation in import-cost-driven episodes and understates it in common-bust episodes; the configuration that generates a false negative is also the one associated with the largest realized debt accumulation.

7.4 The 2014–16 Bust: Realized Versus Projected Debt Paths

The 2014–16 bust offers the only episode in the Africa-20 sample for which a composition-aware projection can be compared against a realized four-year debt path. Fourteen country-years in the Africa-20 sample entered that episode with both prices falling by more than five percent. The asymmetric composition map, which uses the sign-split structural coefficients reflecting that bust responses exceed boom responses, projected a four-year average debt increase of 23.2 percentage points. Realized debt rose by 23.5 points (Appendix Table 9). The linear composition map, which pools boom and bust responses symmetrically, projected 14.4 points and understated realized debt by 9 points. The ToT-based projection would register near-zero for the common component and could not account for the realized trajectory at all. The 2014–16 episode produced more LIC-DSF rating deteriorations to High Risk or In Distress than any other two-year window in the framework's operating history from 2005 to 2019 (International Monetary Fund, 2026) (Ghana reached High Risk in 2015; Zambia in 2017). The aggregate rating distribution confirms the scale of delayed recognition: the count of Low-risk countries in the LIC-DSF listing fell from 22 in March 2015 to 13 in August 2016 to 11 in September 2017, while 12 countries moved from Low or Moderate to High or In Distress by 2019 (International Monetary Fund, 2026). Countries exiting Low or Moderate Risk ratings during this window had spent the preceding two to four years accessing IDA credits rather than grants, at the moment of maximum fiscal deterioration. Throughout this window, the commodity stress test derived its projection from a price signal that was zero by construction during a common bust.

7.5 Zambia: The Quantified Case Study

Zambia's trajectory between 2012 and 2017 allows the common-bust mechanism to be traced quantitatively through successive DSA reviews. The 2012 DSA rated Zambia at Low Risk, with the present value of public and publicly guaranteed (PPG) external debt at 9.1 percent of GDP and total public debt at 28 percent; projections showed all indicators below thresholds through 2032 (International Monetary Fund, 2012). The 2015 DSA upgraded Zambia to Moderate Risk, with PV external debt rising to 22.7 percent of GDP; the baseline assumed copper prices would stabilize at US$2.60 per pound (International Monetary Fund, 2015). Copper reached US$2.00 per pound by late 2015 and continued to decline. The 2017 DSA upgraded Zambia to High Risk, with PV external debt at 46.2 percent of GDP, breaching

the 40 percent Moderate-tier threshold, and total public debt at 61 percent of GDP (International Monetary Fund, 2017). Total public debt had more than doubled, from 28 percent to 61 percent, in four years while the framework assigned a rating of Low Risk, then Moderate Risk. The DSF issued its High rating in September 2017, after the commodity bust had run its course and the fiscal deterioration had become irreversible: a delay spanning two review cycles.

The 2015 Zambia DSA also modified the export-shock stress test downward for country-specific reasons. The standard B2 scenario assumes that losses in export proceeds are compensated one-to-one by additional public borrowing. The 2015 DSA noted, however, that "the exporting sector is dominated by foreign-owned mining companies[...] a shortfall in export proceeds will be partly offset by a decline in mining income outflows[...] requiring less additional government borrowing than the stress test assumes" (International Monetary Fund, 2015). IMF staff thus calibrated the single-leg export stress to show a smaller response for Zambia specifically, making an already-insensitive test less sensitive still in the country with the highest copper-revenue share in the Africa-20 sample.

7.6 Kenya: Testing OI-1 and OI-2

Debt-service payments as a share of export revenues rose from 10.7 percent (2013) to 20.9 percent (2016) for Ethiopia, from 5.2 percent to 26.0 percent for Ghana, from 3.0 percent to 10.4 percent for Zambia, and from 4.7 percent to 11.2 percent for Kenya (World Bank, 2025b). In each country, the debt-service burden at least doubled during the years the LIC-DSF was certifying Low or Moderate Risk.

Kenya illustrates how normalization of deviance (OI-1, Section 2.3) and constitutive authority (OI-2, Section 2.3) operated simultaneously within a single country. The common-bust episode passed each LIC-DSF review cycle without being classified as a distinct risk category, not because staff failed to identify it, but because the framework's procedural structure, specifically the absence of a joint price-decline scenario in the DSA template, prevented its formal expression. The institutional certification of safety reshaped financing flows during the years fiscal damage was accumulating.

Commodity exposure. During the decade of its Low Risk rating (2008–2018), Kenya derived 65–70 percent of its export revenue from primary commodities, with agriculture alone accounting for 57 percent of total exports (UNCTAD, 2026). Tea, coffee, and horticulture accounted for the largest share of export revenue. On the import side, petroleum was the single largest input cost. This bilateral dependence places Kenya in the common-bust exposure class: a global commodity price downturn compresses both legs of the terms-of-trade simultaneously, leaving the ratio P_x/P_m near its pre-shock level even as both absolute price levels fall.

The 2014–16 episode and the DSA's interpretation. Global commodity prices fell between mid-2014 and early 2016: oil declined 70 percent from its pre-bust peak and agricultural commodity indices contracted in parallel, compressing both legs of the terms-of-trade simultaneously. By construction, the IMF calibrates LIC-DSF stress tests to the *ratio* of export to import prices; when both legs fall proportionally, the ratio attenuates the measured shock even as the absolute fiscal shock is real and accumulating.

The IMF's January 2017 Article IV consultation for Kenya illustrates the interpretive consequence. The executive summary concluded: "Debt sustainability analysis continues to show a low risk of debt distress, although vulnerabilities have increased" (International Monetary Fund, 2017). The fall in oil prices during 2014–16 was framed as net-favorable for the current account: the current account deficit "fell by 3 percentage points of GDP to 6.8 percent in 2015, primarily due to lower oil prices and investment-related imports that more than offset weaker exports" (International Monetary Fund, 2017). The same document acknowledged that "the vulnerability to export shocks has increased," yet the Low rating was retained.

IMF staff entered the oil price decline as an import-bill saving that *offset* weaker export performance, the inferential procedure the ratio-based stress test architecture systematically induces by construction. IMF staff distributed the fiscal damage from compressed export revenues across line items rather than recording it as a net fiscal loss.

The debt trajectory during the Low-rating window. Kenya held a rating of Low Risk of debt distress from 2008 through September 2018. During this window public debt rose from 48.6 percent of GDP in FY2014/15 to 60.7 percent of GDP by June 2018, a 12 percentage point increase while the framework continued to certify Low Risk (International Monetary Fund, 2017). The 6.1 percentage point common-bust estimate is consistent with, though it cannot be attributed solely to, the trajectory observed in Kenya over this interval.

7.7 Kenya: Constitutive Authority in Operation

The mechanical override provides the clearest evidence. IMF staff acknowledged that "the mechanical external and overall risk of debt distress is 'high', as it was in the October 2018 DSA" (International Monetary Fund, 2020). The mechanical model had signaled High Risk in October 2018, within the Low-rating decade, but staff judgment was applied to hold the rating at Moderate at that review. When Kenya was ultimately assigned a High rating in 2020, the attribution was "purely a consequence of the impact of COVID-19" (International Monetary Fund, 2020).

The PRGT disbursement record reflects the rating signal with a lag that compounds the assessment gap. Kenya made net repayments to the PRGT throughout the common-bust years, SDR 38.6 million in 2014, SDR 47.1 million in 2015, and SDR 48.7 million in 2016, years in which the framework detected no emergency financing need. When COVID-19 arrived in 2020, SDR 542.8 million was disbursed under the Rapid Credit Facility.

The DSA rating does not merely summarize Kenya's debt outlook; it constitutively shaped how creditors, donors, and co-financiers assessed sovereign risk during the rating period (Barnett and Finnemore, 2004). The Low rating constituted a certification of safety in each successive review between 2008 and 2018, while public debt rose at a rate the mechanical model would later classify as High Risk. The common-bust scenario was not invisible; it appeared in every Article IV consultation. What the framework lacked was the aggregation structure needed to recognize co-movement between export and import prices as a distinct fiscal hazard rather than two offsetting line items: a failure of the procedural template to express what staff perceived, not a failure of staff perception.

7.8 Coverage Within the LIC-DSF Universe

Twelve of the twenty Africa-20 economies appear on the World Bank LIC-DSF country list.[1] The findings are not confined to the full sample. When the sample is restricted to those twelve countries, the structural export coefficient at $h = 4$ is -5.92 (SE 1.99) and the symmetry restriction is rejected at $p = 0.003$ (Appendix Table 15). The common-bust direction is consistent across the full sample and the LIC-DSF-eligible subsample (Appendix Table 15). The same framework applies to 70 countries globally, of which 58–87 percent are commodity-dependent by the 50-percent export-share criterion (UNCTAD, 2026).

The LIC-DSF's 50-percent export-share trigger creates a two-layer coverage problem that compounds the stress-test architecture gap. Above the threshold, the commodity stress test is triggered but applies only the export-price component. Below the threshold, covering one-third of LIC-DSF countries, including the most food- and fuel-import-dependent economies in South Asia and East Africa, the test is not triggered, and countries receive only a generic B3 export-volume shock. The coverage gap runs directly opposite to where the stress falls: common-bust income effects are larger and more precisely estimated in the below-threshold group, because for those economies (where export revenues do not rise with global commodity prices) a global commodity bust delivers only the import-cost damage with no export-revenue windfall to offset it, so the coverage gap is largest where the fiscal exposure is greatest.

A simulation against each country's own 2026 LIC-DSF threshold (Weak/Medium/Strong, face-value debt-to-GDP) illustrates what the gap costs in practice (Appendix Table 13). Four countries (Ghana, Malawi, Sudan, and Chad) already breach their threshold under the composition-aware common-bust scenario at the time of the simulation. The ToT-only projection registers zero commodity stress in a common bust, so none of the four breaches would be detected. Among the eight non-breaching countries, Kenya's remaining buffer above its threshold is only 0.20 times the size of the compression error, the discrepancy between the composition-aware projection and the ToT-only projection (1.8 pp buffer against a 9.1 pp compression gap), and Niger's is 0.23. A shock exceeding the benchmark scenario calibration, or country-specific commodity exposure differing from the panel average, would be sufficient to change the rating for those two countries as well. At the other end, Madagascar has a buffer 1.81 times the compression error. Under IDA20+ rules (in effect since July 2022), a country that should be rated High Risk but is assessed as Moderate receives zero grant share on its entire IDA allocation, when the correct assessment would deliver 100 percent grants. For a country receiving the median IDA allocation in the Africa-20 sample, the difference between zero and 100 percent grant share is equivalent to the full present-value face amount of that cycle's IDA credits. Ghana and Malawi, both currently rated High Risk or In Distress (the 100 percent grant tier), are the countries where a common-bust episode in the period before their rating crossed the threshold would have been invisible to the framework. Sudan and Chad, also breaching in the simulation, have been at High Risk or In Distress for most of the sample period; the framework's inability to project common-bust stress means any reassessment toward Moderate during a temporary price recovery would understate the fragility of that recovery.

[1] The twelve: Burkina Faso, Cameroon, Chad, Côte d'Ivoire, Democratic Republic of Congo, Ghana, Kenya, Madagascar, Malawi, Niger, Senegal, Sudan.

8 Design Gaps and Proposed Corrections

8.1 Why the Gap Persists

Time-horizon mismatch. A stress-test calibration anchored to short-run revenue and GDP responses captures the first-order export-revenue effect but may not fully incorporate the longer-run debt accumulation. Collier and Goderis (2012), using panel cointegration on a global sample of commodity exporters from 1963 to 2008, find that commodity booms have positive short-term effects on output but adverse long-term effects, with the damage operating through real exchange-rate appreciation and increases in public consumption. The windfall expands government expenditure commitments in the short run, but expenditure commitments tend not to reverse when commodity prices fall. Annex V of the 2026 reform validates its stress calibration against transitory GDP and revenue responses that mean-revert by the third year following a bust, the horizon over which the short-run output expansion that Collier and Goderis document operates. The debt response does not mean-revert at that horizon; it is still rising at $h = 4$. A framework that validates its stress calibration against transitory quantities may understate the accumulation phase even if it correctly specified the price components.

Political economy of review cycles. The persistence of these design choices across three review cycles (2012, 2017–18, 2026) reflects IO rule-making under principal-agent constraints. Copelovitch (2010) develops a common-agency model of the IMF in which agency slack, the latitude available to staff relative to shareholder control, expands when principal preferences are heterogeneous or weakly aligned. Under low shareholder salience over the methodological details of LIC debt-stress calibration, functional departments retain wide discretion over technical design. Agency slack does not guarantee optimal rules; it guarantees the rules preferred by the department with drafting authority. Dreher, Sturm, and Vreeland (2015) find that even a closely scrutinized dimension of IMF program architecture, the number and stringency of conditionality clauses, reflects political rather than purely technical determination: temporary UN Security Council members receive 30 percent fewer conditions on IMF arrangements than otherwise comparable non-members, across 314 programs from 1992 to 2008. If conditionality itself is measurably distorted by political salience, the stress-test architecture faces the same pressures without the same external oversight. Reinsberg, Stubbs, and Kentikelenis (2022) analyze 668 IMF arrangements from 1980 to 2009 and find that intra-agency bargaining is a robust predictor of program non-completion, with functional departments appending conditions reflecting their own priorities without weighing implementation constraints. The LIC-DSF commodity stress module is a product of the same bargaining structure: designed by the department with expertise in export-price diagnostics, it is technically coherent within its own specification while systematically excluding the import-price channel. Stone (2011)'s theory of informal governance at international organizations explains why these conventions persist through successive review cycles: formal rules govern routine decisions, while informal influence, exercised through agenda control and pre-board-meeting deference, allows leading member states to shape what reaches a formal vote. A diagnostic reform that would reweight commodity stress testing toward import prices, making low-income borrowers appear more vulnerable, faces systematic resistance at every review cycle, without any explicit veto being required.

8.2 Epistemic Persistence and Constitutive Scope

The primary mechanism operates at the level of institutional knowledge rather than political economy. Three review cycles of procedural fidelity to the Px/Pm stress test have normalized that procedure as technically valid (OI-1), and each "Low Risk" or "Moderate Risk" classification during a common-bust episode has reshaped creditor, donor, and co-financier behavior, compounding the fiscal deterioration the rating understates (OI-2). This constitutive authority is most consequential where no private-market alternative exists. Of 54 African sovereign borrowers, 31–57 percent carry no sovereign credit rating from Moody's, S&P, or Fitch (Countryeconomy.com, 2026). For those countries, the LIC-DSF is not one signal among several; it is the only systematic external debt risk assessment available to creditors, donors, and co-financiers. The remaining 23 rated African sovereigns show that private agency signals did not correct the common-bust diagnostic gap during 2014–16: S&P held Ethiopia at B Stable from its May 2014 debut through 2019, consistent with the DSF's Moderate classification and equally insensitive to the commodity bust; S&P and Fitch held Kenya at B+ throughout the bust period, aligned with the DSF's Low rating. The common-bust assessment gap was not specific to the DSF; private agencies working from the same price-ratio framework produced the same incorrect low-risk classifications. What distinguishes the DSF is not that it had less information but that its "Low" and "Moderate" classifications carried binding consequences (grant eligibility, PRGT access, NCB ceilings) that private ratings do not activate. Because DSA ratings directly determine grant eligibility, PRGT access, and NCB ceilings, correcting the stress-test architecture would do more than improve a forecast; it would alter the institutional acts that currently compound the fiscal deterioration it understates.

The cross-institutional pattern confirms OI-3. I find that none of the seventeen sovereign debt assessment frameworks surveyed (covering the African, Asian, and Inter-American Development Banks; the European Stability Mechanism and European Commission; the IMF's External Balance Assessment and Market-Access Country DSA; the sovereign credit rating methodologies of Moody's, S&P Global, and Fitch Ratings; and UNCTAD, UN DESA, and the G20 Common Framework) uses a bilateral export-import price decomposition in its stress-test architecture. The IMF's External Balance Assessment is the closest exception: it explicitly constructs both a commodity export-price index P_x and a commodity import-price index P_m across 42 commodity categories, but then uses only their ratio in estimation, so a simultaneous proportional decline in both legs registers as a zero net shock, the same compression failure expressed in a different mathematical form. The only operational tool with an explicit import-price component is the Asian Development Bank's Countercyclical Support Facility (2026), which disburses when "sharp rises in fuel and food import bills" are documented for Pacific Island LICs. This is a financing trigger, not a debt sustainability assessment; the import-price component has entered operational design in one regional instrument but has not been integrated into any framework's DSA methodology. The cross-institutional pattern thus reflects a shared epistemic tradition in which export revenue is the primary channel of commodity-price transmission and import-price co-movement is outside the scope of debt sustainability assessment.

8.3 Ethiopia and Ghana: OI-1 and OI-2 in Two Episodes

The same normalization dynamic appeared in the 2013 Ethiopia consultation. IMF Country Report 13/308 states that "the terms-of-trade shock is the most extreme potential risk for the sustainability of the external debt," yet concludes that "even under the possibility of a large deterioration in the terms-of-trade, no indicators would breach the indicative thresholds during the period covered by the DSA" (International Monetary Fund, 2013). Ethiopia was exporting coffee, gold, and oilseeds while importing oil and manufactures; all commodity prices were declining simultaneously in 2013–15. The DSA named the terms-of-trade shock as the dominant risk and simultaneously certified that it produced no threshold breach, because the ratio P_x/P_m changes less when both price components fall together than either price level alone.

Subsequent Article IV reports show the same pattern during the bust period. IMF Country Report 16/322, the 2016 Article IV consultation, recorded: "Export revenue stagnated as pronounced falls in merchandise export prices (by 11 percent) were only partially offset by an expansion in volumes. . . Savings from lower fuel import prices were more than compensated by large drought-related food imports" (International Monetary Fund, 2016). Staff were documenting the joint movement of export and import prices in the common-bust configuration, in a consultation that simultaneously rated Ethiopia at Moderate Risk. The same cycle recurred in 2015, in 2016, and again in 2026: staff documented co-movement risk that the framework's scenario architecture did not translate into a threshold breach.

IMF Country Report 26/20, covering Ethiopia's Sixth Review under its Extended Credit Facility arrangement, states that "adverse terms of trade shocks, falls in gold and coffee prices, rises in oil prices, could widen trade and fiscal deficits and make achievement of reserve accumulation targets more difficult" (International Monetary Fund, 2026). Staff articulate the simultaneous export-fall and import-rise structure as a named downside risk. The same report simultaneously assesses Ethiopia's prospective debt at Moderate Risk using a DSA template that does not generate a stress projection for that named scenario. Staff awareness of the common-bust risk does not propagate into the framework's formal scenario architecture, not because staff lack the insight but because the template prevents its expression. This is the Barnett and Finnemore "autonomous authority" mechanism (Barnett and Finnemore, 2004): procedural rules about what counts as a legitimate scenario shape the output regardless of what staff perceive.

Ghana's 2026 Fifth Review illustrates the opposite directional failure. IMF Country Report 25/343 records that "although all DSA debt indicators are below respective thresholds under the baseline, staff applies judgment to maintain Ghana's classification at high risk of debt distress, as a premature reclassification would be unwarranted given significant uncertainties surrounding the commodity price" outlook (International Monetary Fund, 2025). The mechanical model produces a safe rating; staff judgment overrides it upward because commodity price uncertainty the model cannot quantify is understood to be material. This contrasts with the Kenya episode of 2018, in which the mechanical model produced a High rating and judgment held it at Moderate, an override in the opposite direction for a closely related reason. In both cases, the override reflects a failure of the diagnostic architecture the staff are working within, rather than a failure of staff capacity itself: when the commodity stress architecture cannot capture commodity price uncertainty, judgment substitution in both directions becomes the operational response.

8.4 The Required Correction and the 2026 Reform's Response

The correction does not require a redesign of the LIC-DSF. It requires adding one scenario to the existing stress-test battery: a *joint export-import price stress* in which both prices decline simultaneously by the standard commodity-shock magnitude.

The existing commodity stress test projects a debt path under a one-standard-deviation export shock. The new scenario would project the debt path under a simultaneous one-standard-deviation decline in both the export and import price indices. Where a country's import basket contains the same commodity as its export basket (the condition that currently triggers the netting adjustment), the joint scenario is relevant: those are the economies most likely to experience common busts.

The debt response to the joint scenario is not the sum of the two single-leg responses; it must be constructed from the impact matrix. The required inputs (the estimated debt response to each structural shock and the cross-price impact coefficients) are derivable from country-specific price data and a standard local-projection estimation. No new data collection is required.

Half of the six-percentage-point common-bust debt increase operates through increased issuance (Table 11), the GFN channel the revised framework now monitors. A framework that adds a GFN threshold without the common-bust stress scenario that generates that pressure will detect the distress only after it appears in financing accounts, not in the price signal that caused it.

8.5 The 2026 Reform's Retained Architecture

The revised framework applies scenario-based structural identification to climate shocks: direct physical damage functions, indirect productivity effects, and a 30-year debt-path local projection are combined to trace how projected future warming translates into sovereign debt trajectories. The logic is identical to the two-leg decomposition: separate channels, magnitudes from historical data, and a traceable path to the DSA rating. The methodological asymmetry is pronounced: the framework applies scenario-based structural identification with 30-year debt-path projections to climate shocks, while retaining an HP-filter export-price index multiplied by a fixed revenue elasticity for commodity price shocks that are observable in real time and have been estimated in panel data covering 38–51 countries over 30 years.

The same estimator appears in the reform's internal validation of the commodity shock. Annex V reports local projection estimates of real GDP growth and government revenue responses to commodity price bust episodes, the same estimator used for debt responses (IMF and World Bank, 2026, Annex V, para. 8). The reform found that GDP growth and revenue losses were transitory, mean-reverting by the third year following a bust, which it used to justify retaining the existing calibration (IMF and World Bank, 2026, Annex V, paras. 9–10). The reform's LP analysis, however, extended only to revenues and GDP; it did not estimate local projections for sovereign debt itself. The debt response does not mean-revert: it reaches 6.1 percentage points at $h = 4$, after the revenue and GDP paths have returned toward their pre-bust levels. A framework that validates its calibration against transitory revenue and GDP effects therefore captures the first-order short-run responses but may understate the longer-run debt accumulation that persists after those transitory

effects have reversed.

The September 2026 review identified concerns about the calibration and adequacy of the single-leg commodity shock but did not address the architecture. The recalibrated export shock corresponds to 0.6 standard deviations relative to post-COVID realized outcomes (IMF and World Bank, 2026, Annex V, Table V.1). The review responded by recalibrating the shock magnitude and adding the GFN threshold. The architecture of the stress test, centered on one price and treating the other as a buffer, was unchanged.

Henning (2026) documents that under the current framework, DSAs "raise false alarms (type I errors) and miss crises (type II errors), each in about one-third of cases." That figure reflects the baseline classification problem, before the stress-test architecture is considered. The stress test creates an additional diagnostic gap: it assigns zero fiscal consequence to common busts, the dominant adverse price shock in commodity-dependent economies, accounting for 57 percent of severe price episodes in the Africa-20 data, because the terms-of-trade ratio suppresses the price-shock measure when both price components fall together. A framework that misclassifies one-third of episodes at baseline and applies a stress scenario uninformative for the dominant shock type thereby compounds both error types simultaneously.

The IMF-World Bank review dropped the Present Value of external public debt-to-exports ratio as a debt-burden indicator, citing evidence that it increased false positives (para. 22). This was the only debt-burden indicator (DBI) that contained an export-denominator link to commodity price movements; its removal makes the architecture's insensitivity to commodity-price compression more complete rather than less.

8.6 Baseline Optimism and Architecture Gaps: A Compound Challenge

The LIC-DSF's commodity stress-test failure is not an accident of inadequate data or technical oversight, nor is it the same problem that existing evaluations have identified. Gaudin, Harnoys Vannier, and Kessler (2024) document that LIC-DSF forecast errors trace mainly to optimism in baseline growth and deficit projections rather than stress-test undercalibration.

Berg, Berkes, and Pattillo (2014) provide the first empirical evaluation of the LIC-DSF's two core components, the debt threshold approach (DTA), which applies separate thresholds to five debt-burden ratios, and the worst-case aggregator (WCA), which classifies a country as high-risk whenever any single threshold is breached. Berg, Berkes, and Pattillo find that neither component had been subject to empirical testing prior to their study. On accuracy, the DTA loses information relative to simpler alternatives, and the WCA is too conservative; it predicts crises too often relative to the loss function the DSF itself specifies, and is dominated in predictive accuracy by simple alternative aggregators. The stress-test architecture was thus embedded in a framework whose risk-classification machinery was deployed without testing whether it could identify distress accurately.

Mooney and de Soyres (2017) assess LIC-DSF projection performance over 2005–2015 and document a systematic tendency toward optimism in both public and external debt projections, largest for countries rated at Moderate Risk of debt distress, the group for which accurate projections matter most for classification decisions. Projection errors persist after controlling for unanticipated shocks, inconsistent with the hypothesis that forecast errors reflect unforeseeable events. The baseline from which the commodity stress test is applied is therefore systematically biased upward, so the stress-test's architecture gap is compounded

before it begins.

The 2026 reform's own back-testing confirms this. Figure 7 reports "some optimism bias in revenue and export forecasts over longer horizons" (IMF and World Bank, 2026) and identifies as an "Emerging Challenge" that "effectiveness of realism tools in ensuring forecast accuracy could be strained as uncertainty around forecasts increases with frequency of shocks." The reform's retrospective documentation acknowledges this as its primary diagnostic limitation: the framework is least reliable in common-bust episodes. That problem is distinct from the optimism-bias problem: the stress-test scenario itself is incompletely specified for common-bust episodes, not merely insufficiently activated. A baseline that is calibrated to the correct distribution and a stress scenario that relies on a net-price aggregator would still understate common-bust pressure. The correction operates on the stress architecture, not the baseline parameterization. An optimistic baseline combined with a stress scenario that does not separately model both price legs compounds the assessment gap at every common-bust episode.

The compounding is most severe in commodity-dependent LICs: Annex IV of the reform's own backward-looking forecast analysis finds that optimism bias in export and revenue forecasts is "more pronounced in commodity exporters" (IMF and World Bank, 2026, Annex IV, n. 4). Commodity-dependent LICs simultaneously face the largest baseline forecast optimism and the largest stress-test architecture gap. Both failures appear in the reform's own documentation, yet neither is connected there to their shared origin in the single-leg commodity architecture.

8.7 The Cost of Misclassification: Financing Flows During the 2014–16 Bust

The scale of the classification errors appears in IDA allocation data. African countries rated Low or Moderate Risk at the March 2015 DSA snapshot received IDA credit allocations of SDR 4,249 million in FY2015 and SDR 4,286 million in FY2016. Under IDA17 rules then in effect, High or In Distress ratings carried 100 percent grant financing; Low ratings carried zero. If the LIC-DSF had produced risk ratings that fully reflected the separate export- and import-price pressure—rather than netting import exposure against the export shock—and assigned High Risk ratings accordingly, the full allocations would have been grant-financed. An illustrative calculation of the two-year grant shortfall, confined to African LIC-DSF members receiving credits while rated Low or Moderate Risk during the peak bust years, yields SDR 8,535 million ($11.8 billion at prevailing 2015–16 SDR valuation rates)[2] (World Bank, 2025). Ethiopia and Nigeria alone account for $4.6 billion of this figure. These resources do not represent grants forgone; they constitute concessional loans bearing debt-service obligations that compound the fiscal deterioration the common-bust episode produced.

Across all PRGT-eligible economies, net PRGT flows turned negative in 2016, SDR 676 million disbursed against SDR 769 million in repayments, for a net transfer of negative SDR 93 million, during the deepest trough of the 2014–16 commodity bust (International Monetary Fund, 2024). The same facility disbursed SDR 6,828 million in 2020, a 10.1-fold increase. The asymmetry reflects the diagnostic rating: the commodity bust of 2014–16 generated no

[2]SDR valuation rates are published monthly in the IMF's *International Financial Statistics*.

emergency-risk DSA classification for most affected countries, because common busts leave the terms-of-trade near their pre-shock level, while the COVID-19 episode generated consistently High Risk and In Distress ratings. The framework's classification error shapes whether the IMF's concessional facility responds to a fiscal shock or classifies it as an ordinary credit event.

Emergency financing through the Rapid Credit Facility (RCF) requires no DSA clearance and is available without a Fund-supported program. The Extended Credit Facility (ECF), which requires a macroeconomic framework consistent with debt sustainability, was the sole PRGT facility active for the primary common-bust-affected economies during 2014–16; no RCF was activated for any of those countries during the bust period. The SDR 6,828 million disbursed in 2020 was predominantly RCF, emergency financing, unconditional on DSA. The DSA was therefore the binding gatekeeper when the commodity bust was deepest, and absent as a gatekeeper when COVID-19 produced unambiguous distress indicators.

The net-negative PRGT response during 2014–16 was not offset by the African Development Bank. For Zambia, which carries the highest copper-revenue share in the Africa-20 sample, AfDB ODA disbursements peaked at $127.3 million in 2015, then contracted to $69.2 million in 2016, $47.5 million in 2017, and $39.0 million in 2018: a 69 percent cumulative decline over three years (OECD and African Development Bank, 2024). Total AfDB ODA disbursements to Sub-Saharan Africa fell 16.1 percent in 2014 (from $4.31 billion to $3.62 billion), recovered partially in 2015, then contracted again in 2016 ($3.65 billion) and 2018 ($3.43 billion). The multilateral development system did not serve as a countercyclical buffer during the commodity bust; it contracted alongside the IMF's concessional facility, tracking the same commodity price deterioration.

8.8 Structural Costs: Haircut Persistence and Expenditure Lock-In

Cruces and Trebesch (2013) compile a database of 180 sovereign debt restructurings with foreign creditors from 1970 to 2010 and find that the size of the creditor haircut at resolution is a strong predictor of both subsequent market access costs and the duration of capital market exclusion. Haircuts above 50 percent are followed by spreads 250 basis points higher and exclusion periods three years longer than low-haircut restructurings. A framework that understates debt accumulation during a common bust permits debt to accumulate beyond the level at which earlier corrective intervention would have been triggered, exposing the borrower to restructuring with a larger stock and higher required haircut.

A related dynamic operates on the expenditure side. Adler and Magud (2015) document forty-two commodity terms-of-trade boom episodes from 1970 to 2012 and find that governments increase expenditure commitments during commodity booms at a rate that exceeds saving. The marginal saving rate across these episodes was 45 percent of the windfall, meaning that 55 percent of every unit of boom revenue was committed as recurring expenditure. When commodity prices fall, these expenditure commitments do not reverse at the same rate. For African commodity exporters, this pattern implies that common-bust episodes begin with expenditure bases calibrated to boom conditions, requiring new domestic borrowing to finance the shortfall, the quantity that local projections identify as rising at $h = 4$. A stress test that projects no net fiscal consequence from a common bust provides no early warning of this expenditure pattern.

8.9 Compounding Gaps: Hidden Debt and Domestic Markets

The calibration problem is compounded by incomplete debt stocks. Horn, Reinhart, and Trebesch (2021) find that 50 percent of Chinese loan commitments to developing countries are absent from the World Bank's Debtor Reporting System (DRS), an average gap of USD 200 billion per year over the decade to 2016. Because LIC-DSF baseline projections depend on DRS-recorded debt stocks, DSAs in countries with substantial unreported Chinese exposure begin from an understated baseline. A commodity stress test applied to an understated baseline will project a smaller response than actual exposure warrants, even if the stress-test architecture were otherwise correct.

This gap has grown with the development of African domestic debt markets. Manger et al. (2025) introduce the African Debt Database, covering more than 50,000 individual debt instruments issued by 54 African countries between 2000 and 2024, amounting to USD 6.3 trillion in total, and document that domestic sovereign debt markets have more than tripled in size since 2010, rising from USD 150 billion to USD 500 billion. African governments now raise more than half of their annual financing domestically, typically at maturities far shorter than the external loans that anchor LIC-DSF stress scenarios and at interest rates that reset rapidly in the event of credit deterioration. This maturity and currency structure means that a commodity bust that forces new domestic borrowing is also repricing the rollover risk of the existing domestic stock, a dynamic that the LIC-DSF commodity stress scenario, calibrated to external debt dynamics, does not capture.

8.10 Why the Reform's New Metrics Do Not Close the Gap

The 2026 reform's addition of the GFN threshold makes the case for the common-bust scenario stronger, not weaker. The estimates indicate that 50 percent of the common-bust debt accumulation operates through increased domestic issuance (the numerator of GFN; Table 11). A country in a common bust would see its GFN pressure rise, but the framework would not project any commodity stress that would trigger the scenario. The new threshold adds detection capacity for a pressure the framework does not yet model.

The reform's own retrospective analysis documents the consequence of this diagnostic gap. Annex IV finds that forecast errors for overall public debt widened during and after COVID-19, "driven mainly by larger-than-anticipated domestic borrowing, higher interest costs, and the emergence of previously unrecorded liabilities" (IMF and World Bank, 2026, Annex IV, para. 3). Larger-than-anticipated domestic borrowing is the pattern that common-bust local projections identify, reaching 3 percentage points of GDP at $h = 4$ (Table 11). The framework has repeatedly observed this consequence without building the scenario architecture that would have anticipated it.

The reform also created a new domestic debt module (Annex II) designed to identify countries facing elevated domestic debt risks. The module's empirical foundation (Box II.2) identifies the primary indicator preceding domestic debt stress: a rise in net domestic debt issuance-to-GDP in the year before the onset of stress, with the median ratio doubling from non-stress to pre-stress periods across LIC-DSAs over 2015–24 (IMF and World Bank, 2026, Box II.2) — the same issuance surge that common-bust local projections identify. The domestic-issuance component of the debt response (the "remainder" in the stock-identity decomposition) accounts for 50 percent of the six-percentage-point common-bust debt accumu-

lation (Table 11). The reform built a module to detect the issuance surge that common-bust episodes generate, but the domestic debt module is activated by financial-sector indicators (credit cycles, banking stress), not by commodity price co-movement. A common bust that generates fiscal pressure and forces domestic borrowing would produce the issuance surge the module is calibrated to identify, yet the commodity stress test produces no indicator that would activate it.

Finally, the reform recalibrated the granularity of the Moderate Risk rating, the mechanism that distinguishes countries with "limited," "some," and "substantial" space to absorb shocks before a downgrade to High Risk. Back-testing in Annex VIII shows that 71.8 percent of subsequent High Risk downgrades originated from Moderate Risk countries previously classified as having limited space (IMF and World Bank, 2026, Annex VIII, Table VIII.1). The granularity exercise calibrates these space boundaries from historical stressed debt paths. If the stressed path systematically understates the debt response in common-bust episodes, countries near the boundary between "some" and "limited" space would be misclassified, and the high-risk-transition predictability that Annex VIII validates would deteriorate in the commodity price cycles where it matters most.

The 2026 reform's own new metrics are compromised by the same compression error. Structural local projections on revenue/GDP, exports/GDP, and interest/revenue, the three objects the revised framework monitors under its new "realism" and "liquidity" pillars, all reject symmetry at $h = 4$ ($p = 0.013$, $p = 0.007$, and $p = 0.035$ respectively), with the export leg 2–3 times larger than the import leg across all three outcomes (Appendix Table 14). The 2026 reform added new thresholds but retained the ToT-based stress scenario; the compression error therefore propagates into every new metric simultaneously. The simultaneous decline of fossil fuel export prices and the increase in food import prices have recurred in African data since 2019. The correction is technically tractable: it adds a scenario, not a new rating scale, and requires no threshold renegotiation. By December 2025, the reform's own Figure 5 shows 48 percent of LIC-DSF users carrying a High Risk or In Distress external debt rating, the tier in which a one-category misclassification fully eliminates IDA grant access under IDA20+ rules. At this scale, the stress-test architecture is not a technical detail; it is the primary input into whether 48 percent of the LIC-DSF universe receives resource transfers or debt obligations from the international system.

8.11 Limitations

The analysis has three limitations.

First, the structural identification requires DJP's bilateral price data, which are available for only 38 developing economies. The common-bust estimate is therefore derived from a sample that does not cover the full LIC-DSF country universe. The shift-share corroboration, which requires no DJP structure and extends to 51 countries, replicates the qualitative finding, but the six-point estimate should be read as specific to the African commodity-export structure.

Second, the analysis identifies fiscal effects conditional on the structure of price shocks as of 2019. The post-pandemic commodity cycle (2020–2023) and the geopolitically driven price realignments since 2022 may alter the relative magnitudes of the export and import responses, particularly for food-importing economies where the import price volatility has increased relative to the pre-2019 period.

Third, this is a diagnostic of the stress-test architecture, not a calibrated structural model. The proposed correction specifies what additional scenario should be run; it does not derive the welfare-optimal threshold against which the scenario's projection should be compared.

9 Conclusion

Sovereign debt vulnerability in commodity-dependent economies is shaped not only by the shocks countries face, but by whether the frameworks used to assess those risks can distinguish the economic forces driving them. When institutions measure commodity exposure through a net-price representation, the class of price episodes most damaging to low-income sovereign debt — simultaneous declines in export and import prices — registers as a near-zero fiscal event. The six-percentage-point common-bust estimate documented in Section 6 is the empirical content of that gap.

The LIC-DSF provides the principal case study. Its commodity stress test has persisted across three review cycles without incorporating import-price co-movement. The 2026 review added new thresholds, recalibrated shock magnitudes, and introduced domestic debt monitoring, while leaving unchanged the architectural element that registers near-zero fiscal impact from common busts. That survival is not an oversight; it reflects the normalization dynamic Barnett and Finnemore predict. Procedural fidelity across three cycles has embedded the terms-of-trade ratio as the legitimate instrument, displacing the question of whether it correctly identifies the relevant stress channel.

The constitutive consequence is most fully documented for Kenya, where a decade of Low Risk certification shaped PRGT net repayments and creditor pricing throughout the twelve-percentage-point public-debt increase. Zambia's rating moved from Low to Moderate to High in four years while the commodity bust that drove the deterioration produced a stress-test measure of approximately zero. During the 2014–16 bust's deepest trough, the PRGT recorded net outflows, an outcome that did not reflect forecast error but constitutive authority in operation: the rating simultaneously certified debt sustainability and, through the constitutive mechanism documented in Section 2, shaped the creditor and donor responses that accompanied the fiscal deterioration it failed to detect.

The IO literature asks when evidence produces institutional revision and when an IO's authority over its own assessment prevents such revision. Over three review cycles, across seventeen frameworks, and with six percentage points of common-bust debt accumulation that the stress test could not project, the evidence has accumulated for twenty years without producing an architectural revision. The Barnett-Finnemore mechanisms account for this trajectory better than agency-slack, measurement-constraint, or genuine-uncertainty alternatives, each of which is inconsistent with the evidence presented in Section 2.

The required intervention is not additional documentation but a concrete correction that requires no new data and no threshold renegotiation: adding a joint export-import price stress scenario to the existing DSA template. That addition would enable the framework to formalize what staff observations document when both price components decline simultaneously. Future reform cycles will provide evidence on whether normalization of deviance can be overcome by formal empirical documentation. A calibrated model of sovereign borrowing with commodity-dependent fiscal accounts is the natural extension of this analysis; deriving the welfare-optimal threshold against which the joint scenario's projection should be compared remains an open question for future research.

References

Adler, Gustavo, and Nicolas E. Magud, 2015, "Four Decades of Terms-of-Trade Booms: A Metric of Income Windfall," *Journal of International Money and Finance* 55: 162–192.

Ahir, Hites, Nicholas Bloom, and Davide Furceri, 2022, "The World Uncertainty Index," NBER Working Paper No. 29763 (Cambridge, MA: National Bureau of Economic Research).

Bal Gündüz, Yasemin, 2017, "Debt Sustainability in Low-Income Countries: Policies, Institutions, or Shocks?" IMF Working Paper WP/17/114 (Washington, DC: International Monetary Fund).

Berg, Andrew, Enrico G. Berkes, Catherine A. Pattillo, Andrea F. Presbitero, and Yorbol Yakhshilikov, 2014, "Assessing Bias and Accuracy in the World Bank–IMF's Debt Sustainability Framework for Low-Income Countries," IMF Working Paper WP/14/048 (Washington, DC: International Monetary Fund).

Byrne, Joseph P., Giorgio Fazio, and Norbert Fiess, 2013, "Primary Commodity Prices: Co-Movements, Common Factors and Fundamentals," *Journal of Development Economics* 101: 16–26.

Barnett, Michael, and Martha Finnemore, 1999, "The Politics, Power, and Pathologies of International Organizations," *International Organization* 53 (4): 699–732.

Barnett, Michael, and Martha Finnemore, 2004, *Rules for the World: International Organizations in Global Politics* (Ithaca, NY: Cornell University Press).

Bachas, Pierre, Matthew Fisher-Post, Anders Jensen, and Gabriel Zucman, 2026, "Globalization, Capital Taxation, and Development: Evidence from a Macrohistorical Database," *American Economic Journal: Applied Economics* 18 (4): 509–549.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel, 2022, "Quasi-Experimental Shift-Share Research Designs," *Review of Economic Studies* 89 (1): 181–213.

Cashin, Paul, C. John McDermott, and Alasdair Scott, 2002, "Booms and Slumps in World Commodity Prices," *Journal of Development Economics* 69 (1): 277–296.

Collier, Paul, and Benedikt Goderis, 2012, "Commodity Prices and Growth: An Empirical Investigation," *European Economic Review* 56 (6): 1241–1260.

Copelovitch, Mark S., 2010, *The International Monetary Fund in the Global Economy: Banks, Bonds, and Bailouts* (Cambridge: Cambridge University Press).

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

Chwieroth, Jeffrey M., 2010, *Capital Ideas: The IMF and the Rise of Financial Liberalization* (Princeton, NJ: Princeton University Press).

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller, 2008, "Bootstrap-Based Improvements for Inference with Clustered Errors," *Review of Economics and Statistics* 90 (3): 414–427.

Dreher, Axel, Jan-Egbert Sturm, and James Raymond Vreeland, 2015, "Politics and IMF Conditionality," *Journal of Conflict Resolution* 59 (1): 120–148.

Di Pace, Federico, Luciana Juvenal, and Ivan Petrella, 2025, "Terms-of-Trade Shocks Are Not All Alike," *American Economic Journal: Macroeconomics* 17 (2): 24–64.

Driscoll, John C., and Aart C. Kraay, 1998, "Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data," *Review of Economics and Statistics* 80 (4): 549–560.

Henning, C. Randall, 2026, *Getting Debt Sustainability Analysis Right: Eight Reforms for the Framework for Low-Income Countries*, Carnegie Endowment for International Peace (Washington, DC: Carnegie Endowment for International Peace), February.

Horn, Sebastian, Carmen M. Reinhart, and Christoph Trebesch, 2021, "China's Overseas Lending," *Journal of International Economics* 133: 103539.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift, 2020, "Bartik Instruments: What, When, Why, and How," *American Economic Review* 110 (8): 2586–2624.

Haas, Peter M., 1992, "Introduction: Epistemic Communities and International Policy Coordination," *International Organization* 46 (1): 1–35.

Gruss, Bertrand, and Suhaib Kebhaj, 2019, "Commodity Terms of Trade: A New Database," IMF Working Paper WP/19/21 (Washington, DC: International Monetary Fund).

International Monetary Fund, 2026, *List of LIC DSAs for PRGT-Eligible Countries*, as of March 31, 2026 (Washington, DC: International Monetary Fund).

International Monetary Fund and World Bank, 2018, *Guidance Note on the Bank-Fund Debt Sustainability Framework for Low Income Countries*, IMF Policy Paper (Washington, DC: International Monetary Fund).

International Monetary Fund and World Bank, 2026, *Review of the Bank-Fund Debt Sustainability Framework for Low-Income Countries, Proposed Reforms*, IMF Policy Paper No. PP/26/034 (Washington, DC: International Monetary Fund), September.

International Monetary Fund, 2025, *World Economic Outlook* database, April 2025 vintage (Washington, DC: International Monetary Fund).

Pindyck, Robert S., and Julio J. Rotemberg, 1990, "The Excess Co-Movement of Commodity Prices," *Economic Journal* 100 (403): 1173–1189.

Prebisch, Raúl, 1950, *The Economic Development of Latin America and Its Principal Problems* (New York: United Nations Economic Commission for Latin America).

Reinsberg, Bernhard, Thomas Stubbs, and Alexander Kentikelenis, 2022, "Unimplementable by Design? Understanding (Non-)Compliance with International Monetary Fund Policy Conditionality," *Governance* 35 (3): 689–715.

Schmitt-Grohé, Stephanie, and Martín Uribe, 2018, "How Important Are Terms-of-Trade Shocks?" *International Economic Review* 59 (1): 85–111.

Levy, Antoine, Luca Antonio Ricci, and Alejandro Mariano Werner, 2020, "The Sources of Fiscal Fluctuations," IMF Working Paper No. WP/20/220 (Washington, DC: International Monetary Fund).

Singer, Hans W., 1950, "The Distribution of Gains between Investing and Borrowing Countries," *American Economic Review* 40 (2): 473–485.

Spatafora, Nikola, and Issouf Samaké, 2012, "Commodity Price Shocks and Fiscal Outcomes," IMF Working Paper No. 12/112 (Washington, DC: International Monetary Fund).

Stone, Randall W., 2011, *Controlling Institutions: International Organizations and the Global Economy* (Cambridge: Cambridge University Press).

Teulings, Coen N., and Nick Zubanov, 2014, "Is Economic Recovery a Myth? Robust Esti-

mation of Impulse Responses," *Journal of Applied Econometrics* 29 (3): 497–514.

Jordà, Òscar, 2005, "Estimation and Inference of Impulse Responses by Local Projections," *American Economic Review* 95 (1): 161–182.

Jordà, Òscar, and Alan M. Taylor, 2025, "Local Projections," *Journal of Economic Literature* 63 (1): 59–110.

Känzig, Diego R., 2021, "The Macroeconomic Effects of Oil Supply News: Evidence from OPEC Announcements," *American Economic Review* 111 (4): 1092–1125.

Gaudin, Antoine, Brendan Harnoys Vannier, and Martin Kessler, 2024, "A Critical Analysis of DSA Projections," *Development* 67 (3): 233–247.

Manger, Mark S., David Mihalyi, Ugo Panizza, Niccolò Rescia, Christoph Trebesch, and Ka Lok Wong, 2025, "Africa's Domestic Debt Boom: Evidence from the African Debt Database," Graduate Institute Working Paper HEIDWP15-2025 (Geneva: Graduate Institute of International and Development Studies).

Montiel Olea, José Luis, and Mikkel Plagborg-Møller, 2021, "Local Projection Inference Is Simpler and More Robust Than You Think," *Econometrica* 89 (4): 1789–1823.

Mooney, Henry, and Constance de Soyres, 2017, "Debt Sustainability Analyses for Low-Income Countries: An Assessment of Projection Performance," IMF Working Paper WP/17/220 (Washington, DC: International Monetary Fund).

Müller, Karsten, Chenzi Xu, Mohamed Lehbib, and Ziliang Chen, 2025, "The Global Macro Database: A New International Macroeconomic Dataset," NBER Working Paper No. 33714 (Cambridge, MA: National Bureau of Economic Research).

International Monetary Fund, 2017, *Kenya: 2017 Article IV Consultation, Press Release; Staff Report; and Statement by the Executive Director for Kenya*, IMF Country Report No. 17/25 (Washington, DC: International Monetary Fund).

International Monetary Fund, 2020, *Kenya: Request for Disbursement Under the Rapid Credit Facility, Press Release; Staff Report; and Statement by the Executive Director for Kenya*, IMF Country Report No. 20/3 (Washington, DC: International Monetary Fund).

International Monetary Fund, 2026, *Ethiopia: Sixth Review Under the Extended Credit Facility Arrangement, Press Release; Staff Report; and Statement by the Executive Director for Ethiopia*, IMF Country Report No. 26/20 (Washington, DC: International Monetary Fund).

International Monetary Fund, 2025, *Ghana: Fifth Review Under the Extended Credit Facility Arrangement, Press Release; Staff Report; and Statement by the Executive Director for Ghana*, IMF Country Report No. 25/343 (Washington, DC: International Monetary Fund).

International Monetary Fund, 2024, *IMF Financial Data by Country: Poverty Reduction and Growth Trust Disbursements and Repayments*, IMF Financial Data Query Tool (Washington, DC: International Monetary Fund), accessed October 2026.

United Nations Conference on Trade and Development, 2026, *State of Commodity Dependence 2024* (Geneva: UNCTAD).

International Monetary Fund, 2012, *Zambia: Article IV Consultation and First Review Under the Three-Year Extended Credit Facility, Staff Report*, IMF Country Report No. 12/200 (Washington, DC: International Monetary Fund).

International Monetary Fund, 2015, *Zambia: 2015 Article IV Consultation, Staff Report; Press Release; and Statement by the Executive Director for Zambia*, IMF Country Report No. 15/152 (Washington, DC: International Monetary Fund).

International Monetary Fund, 2017, *Zambia: 2017 Article IV Consultation, Press Release; Staff Report; and Statement by the Executive Director for Zambia*, IMF Country Report No. 17/327 (Washington, DC: International Monetary Fund).

International Monetary Fund, 2013, *Ethiopia: 2013 Article IV Consultation, Staff Report; Press Release; and Statement by the Executive Director for Ethiopia*, IMF Country Report No. 13/308 (Washington, DC: International Monetary Fund).

International Monetary Fund, 2016, *Ethiopia: 2016 Article IV Consultation, Staff Report; Press Release; and Statement by the Executive Director for Ethiopia*, IMF Country Report No. 16/322 (Washington, DC: International Monetary Fund).

Countryeconomy.com, 2026, *Sovereign Credit Ratings: Africa*, historical rating actions compiled from Moody's, S&P, and Fitch (countryeconomy.com), accessed October 2026.

OECD and African Development Bank, 2024, *DAC2a: Aid by Donor and Recipient*, OECD International Development Statistics, CRS Creditor Reporting System (Paris: OECD), accessed October 2026.

World Bank, 2025, *World Development Indicators*: DT.TDS.DECT.EX.ZS (Debt Service on External Debt, Total, % of Exports of Goods, Services, and Primary Income), International Debt Statistics, accessed October 2026.

World Bank, 2025, *IDA Grant and Credit Allocations by Country, FY2012, FY2025*, compiled from IDA Annual Replenishment Allocation Tables (Washington, DC: World Bank Group), accessed October 2026.

A Figures

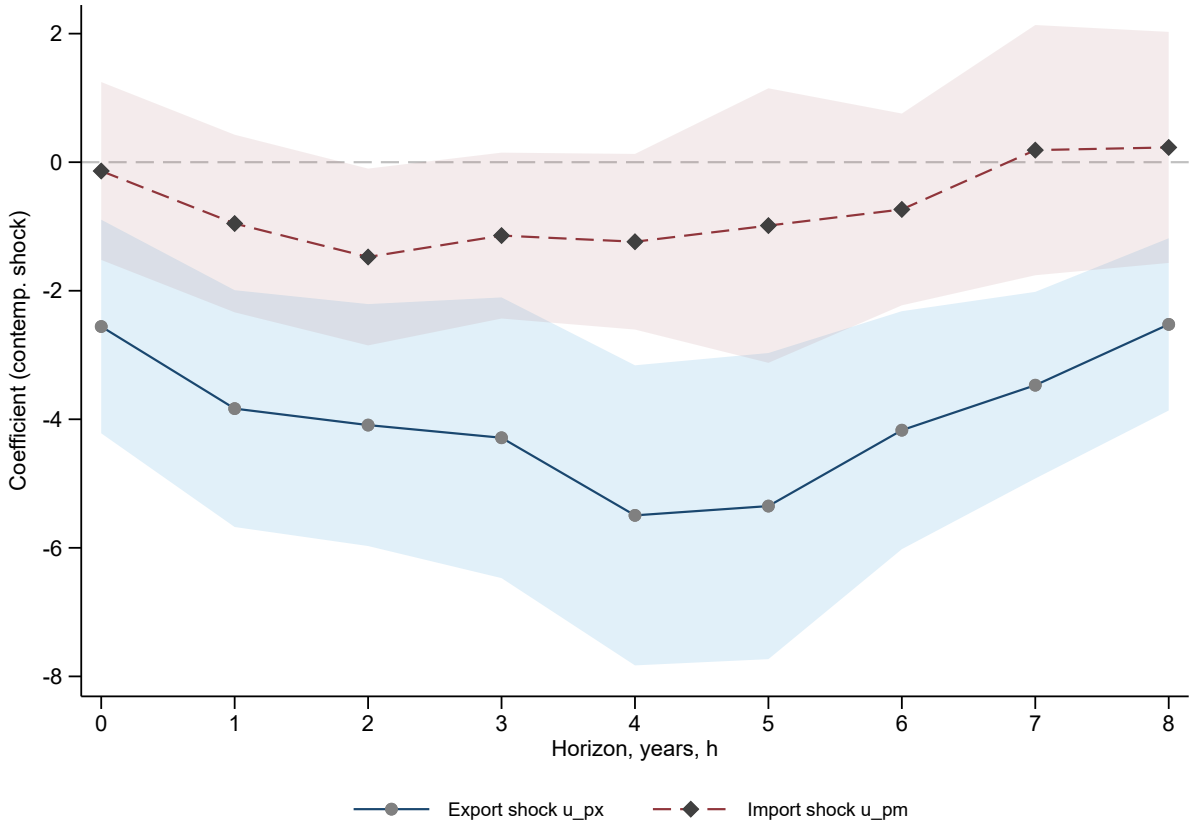


Figure 1: Structural Price Shocks and Gross Government Debt: Impulse Responses

Notes: Structural local projections of gross government debt (% of GDP) on DJP export-price (left panel) and import-price (right panel) shocks, Africa-20, 1990–2019. Long-difference specification, country fixed effects. Shaded bands: 68 and 90 percent confidence intervals from country-clustered standard errors. Horizon in years on the horizontal axis.

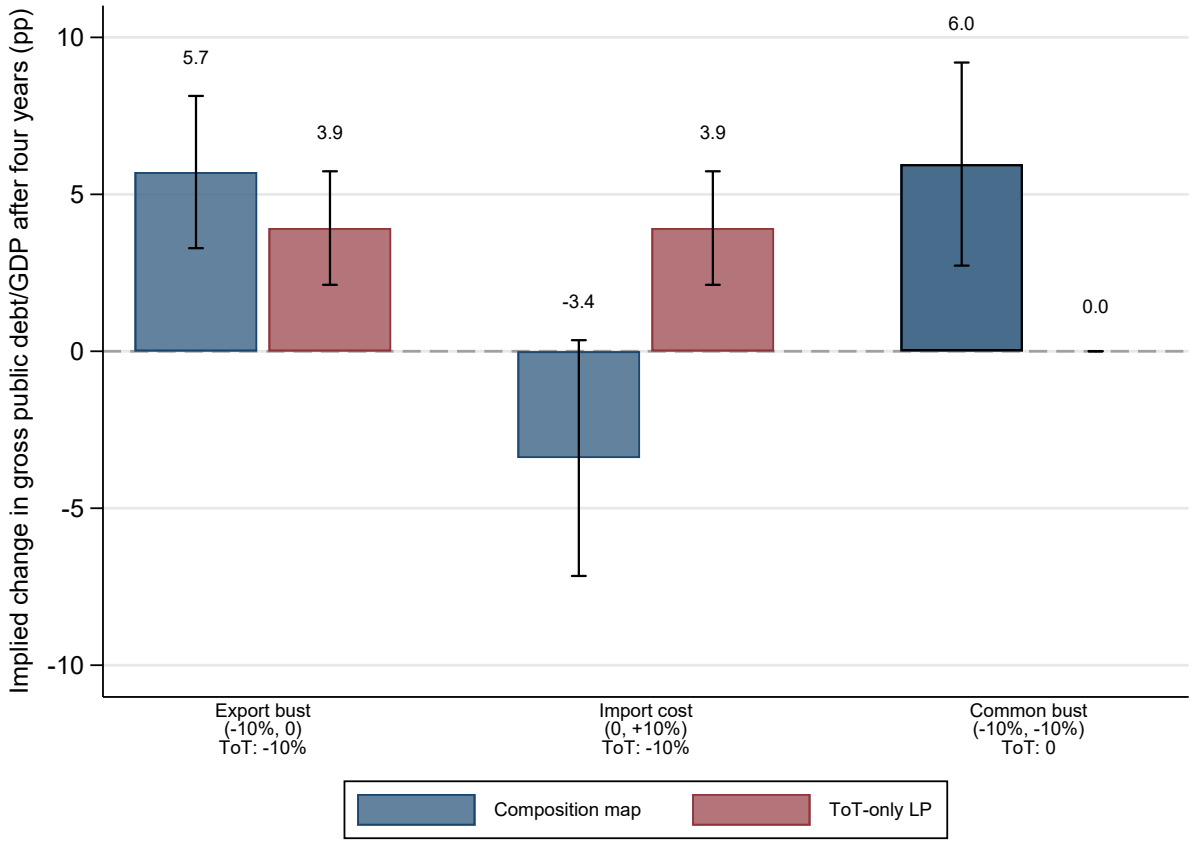


Figure 2: Composition Stress Map: The Two Failure Modes at $h = 4$

Notes: Debt response per 10 percent own-price move at $h = 4$, percentage points of GDP. Each scenario is constructed from the structural impact matrix. "ToT only" is the fitted value from a single-regressor local projection of gross government debt on the terms-of-trade change ($h = 4$, same controls and sample as Table 3). The common-bust scenario estimate (+6.1 pp) exceeds the export-bust scenario estimate (+5.7 pp) yet the terms-of-trade ratio registers zero change in the common-bust scenario.

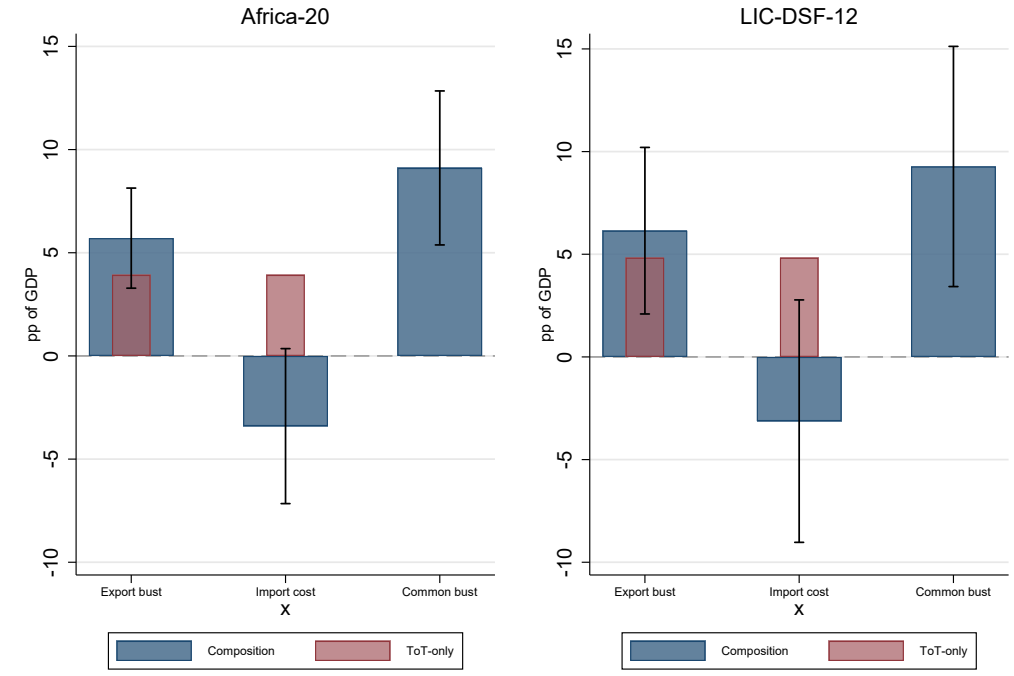


Figure 3: Composition Stress Map: LIC-DSF 12-Country Subsample

Notes: Same construction as Figure 2, restricted to the twelve Africa-20 economies on the World Bank LIC-DSF country list. The pattern is quantitatively similar to the full sample.

B LIC-DSF Country Coverage in the Africa-20 Sample

Table 2: Africa-20 Countries: LIC-DSF Eligibility and Primary Export

Country	ISO3	Primary export	LIC-DSF
Algeria	DZA	Oil and gas	No
Burkina Faso	BFA	Gold, cotton	Yes
Cameroon	CMR	Oil, cocoa	Yes
Chad	TCD	Oil	Yes
Côte d'Ivoire	CIV	Cocoa, oil	Yes
Dem. Rep. Congo	COD	Minerals	Yes
Egypt	EGY	Oil, gas	No
Equatorial Guinea	GNQ	Oil	No
Gabon	GAB	Oil	No
Ghana	GHA	Gold, cocoa, oil	Yes
Kenya	KEN	Tea, horticulture	Yes
Madagascar	MDG	Vanilla, minerals	Yes
Malawi	MWI	Tobacco, tea	Yes
Mauritius	MUS	Services, sugar	No
Morocco	MAR	Phosphates	No
Niger	NER	Uranium, gold	Yes
Nigeria	NGA	Oil	No
Senegal	SEN	Gold, fish	Yes
South Africa	ZAF	Gold, minerals	No
Sudan	SDN	Oil, gold	Yes

Notes: LIC-DSF eligibility as of end-April 2026, from the World Bank Debt Toolkit. Countries are selected on data availability, not on debt-distress status or DSF eligibility.

C Tables

Table 3: Structural Price Shocks (DJP): Gross Government Debt

	Horizon h (years)								
	0	1	2	3	4	5	6	7	8
$\hat{\beta}_{u_x}$: structural export shock	-2.56^{***}	-3.83^{***}	-4.09^{***}	-4.29^{***}	-5.49^{***}	-5.35^{***}	-4.17^{***}	-3.47^{***}	-2.52^{***}
	(0.85)	(0.94)	(0.96)	(1.11)	(1.19)	(1.21)	(0.94)	(0.74)	(0.68)
$\hat{\beta}_{u_m}$: structural import shock	-0.14	-0.95	-1.48^{**}	-1.14^{*}	-1.24^{*}	-0.99	-0.74	0.19	0.23
	(0.71)	(0.70)	(0.70)	(0.66)	(0.70)	(1.09)	(0.76)	(0.99)	(0.92)
$F(H_0 : \hat{\beta}_{u_x} + \hat{\beta}_{u_m} = 0)$	8.38	18.21	24.90	21.93	33.54	10.17	17.04	7.77	3.10
$p(H_0 : \hat{\beta}_{u_x} + \hat{\beta}_{u_m} = 0)$	0.009	<0.001	<0.001	<0.001	<0.001	0.005	<0.001	0.012	0.094
Observations	441	441	441	441	441	421	401	381	361

Notes: Structural local projections using DJP export- and import-price shocks. Long-difference specification with country fixed effects and baseline controls. Units: percentage points of GDP per one standard deviation of the shock. Standard errors clustered at the country level in parentheses. Symmetry rows report the clustered $F(1, 19)$ statistic and p-value for $H_0 : \hat{\beta}_{u_x} + \hat{\beta}_{u_m} = 0$. Stars: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 4: Structural Gross-Debt Responses per 10 Percent Own-Price Move

	$h = 0$	$h = 2$	$h = 4$
A. Impact first stage (log price)			
$\partial \ln P_x / \partial u_x$	0.096^{***}		
	(0.012)		
$\partial \ln P_m / \partial u_m$	0.036^{***}		
	(0.003)		
B. Debt, per one SD of the shock			
$\hat{\beta}_{u_x}$	-2.56^{***}	-4.09^{***}	-5.49^{***}
	(0.85)	(0.96)	(1.19)
$\hat{\beta}_{u_m}$	-0.14	-1.48^{**}	-1.24^{*}
	(0.71)	(0.70)	(0.70)
C. Debt, per 10 percent own-price move			
Export-price shock	-2.66	-4.25	-5.71
Import-price shock	-0.38	-4.05	-3.40

Notes: Panel C rescales using $\tilde{\beta}_{x,h} = \hat{\beta}_{x,h} \cdot 0.10/\hat{g}_{x,0}$ and $\tilde{\beta}_{m,h} = \hat{\beta}_{m,h} \cdot 0.10/\hat{g}_{m,0}$ with $\hat{g}_{x,0} = 0.0962$ and $\hat{g}_{m,0} = 0.0364$. Panel C reports point estimates only. Stars: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 5: The Two Failure Modes at Every Horizon, $h = 0$–4

	Source failure			Common-shock failure	
h	Export bust	Import cost	ToT-only	Common bust	95% interval
$h = 0$	2.66	-0.38	1.56	1.67	[-1.65, 5.00]
$h = 1$	3.98	-2.62	2.40	4.39	[0.88, 7.90]
$h = 2$	4.25	-4.05	2.26	5.83	[2.36, 9.30]
$h = 3$	4.45	-3.14	3.05	5.10	[1.91, 8.28]
$h = 4$	5.71	-3.40	3.93	6.10	[2.73, 9.20]

Notes: Percentage points of GDP per 10 percent own-price move. Export bust and Import cost are the rescaled structural leg responses (Panel C, Table 4). ToT-only is the fitted value from a regression using the terms-of-trade ratio in place of the two separate prices. Common bust is the debt response to a simultaneous 10 percent decline in both prices, constructed from the impact matrix **A**. Intervals use the estimated covariance of the two leg coefficients.

Table 6: Structural Export-Price Shocks and Sovereign Debt Restructuring

	Horizon h				
	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$
Panel A: Restructuring entry					
$\hat{\beta}_{u_x}$	0.007	0.063**	0.043*	0.038*	0.012
	(0.021)	(0.026)	(0.021)	(0.019)	(0.020)
$\hat{\beta}_{u_m}$	-0.011	0.003	-0.016	-0.015	-0.014
	(0.014)	(0.012)	(0.018)	(0.021)	(0.024)
N	437	437	437	437	437
Panel B: In default					
$\hat{\beta}_{u_x}$	0.002	-0.002	-0.009	-0.010	-0.012
	(0.006)	(0.003)	(0.008)	(0.010)	(0.011)
$\hat{\beta}_{u_m}$	0.006	0.005	0.000	-0.000	0.000
	(0.005)	(0.007)	(0.007)	(0.008)	(0.006)
N	471	471	471	471	471

Notes: Linear probability model local projections of binary sovereign-distress outcomes on DJP structural shocks. Twenty African economies, 1990–2019. Restructuring entry uses Horn, Reinhart, Trebesch bank and bond restructurings combined with Asonuma, Trebesch preemptive and post-default restructurings, identified for 16 of 20 economies. Country fixed effects; standard errors clustered by country ($G = 20$). Stars: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 7: Structural Mechanism Decomposition: Revenue, Expenditure, Primary Balance

Outcome	Leg	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$	N
Revenue	Export	0.34	0.82^{**}	0.53^{**}	0.32^{*}	0.41^{**}	467
		(0.24)	(0.34)	(0.24)	(0.18)	(0.19)	
	Import	0.04	−0.19	0.09	−0.10	0.18	467
		(0.12)	(0.21)	(0.12)	(0.19)	(0.20)	
Expenditure	Export	−1.37	0.68	0.55	−0.19	−0.09	468
		(1.06)	(1.11)	(0.60)	(0.26)	(0.24)	
	Import	1.07	−0.21	−0.26	0.07	0.06	468
		(1.27)	(0.29)	(0.24)	(0.20)	(0.22)	
Primary balance	Export	1.59	0.02	−0.07	0.41^{*}	0.41	467
		(1.11)	(1.03)	(0.39)	(0.21)	(0.24)	
	Import	−1.13	0.08	0.27	−0.21	0.02	467
		(1.26)	(0.30)	(0.21)	(0.18)	(0.17)	

Notes: Structural local projections using DJP shocks. Long-difference specification with country fixed effects and baseline controls. Units: percentage points of GDP per one standard deviation of the shock. Standard errors clustered at the country level ($G = 20$). Stars: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 8: LIC-DSF Commodity Netting Cells: IDA21 Generic and Uganda Workbook

Object	Cell	IDA21 generic	Uganda workbook
Commodity export share (trigger)	G_{29}	56.27	56.27
Fuel price shock	G_{32}	−30.89	−29.16
Fuel export share	G_{33}	14.61	14.61
Fuel import mitigation	G_{49}	5.49	5.49
Adjusted fuel export share	G_{34}	9.12	9.12
Non-fuel price shock	G_{36}	−21.69	−23.12
Non-fuel export share	G_{43}	41.66	41.66
Non-fuel import mitigation	G_{50}	0.00	0.00
Adjusted non-fuel export share	G_{44}	41.66	41.66
Composite commodity shock	G_{46}	−23.34	−24.20

Formulas: $G_{34} = \max(0, G_{33} - G_{49})$; $G_{44} = \max(0, G_{43} - G_{50})$.

Notes: LIC-DSF commodity cells from the IMF-World Bank workbook, sheet "Input 6 , Tailored Tests." Import mitigation is asymmetric ($G_{49} > 0$, $G_{50} = 0$): only fuel imports reduce the export share entering the composite shock. Uganda is not in the LP sample; shares illustrate the workbook algebra.

Table 9: Historical Validation: Realized versus Predicted Debt Response in Common-Bust Episodes

Episode	N	Realized	Linear map	Asymmetric map
2014–16 bust	14	23.5	14.4	23.2
2008–09 GFC	14	4.9	17.5	27.9
Pre-2007 (HIPC/MDRI era)	27	−18.9	9.8	15.5

Notes: Percentage points of GDP, four-year average change. Sample: country-years in the Africa-20 panel where both $\Delta \ln P_x < -0.05$ and $\Delta \ln P_m < -0.05$, grouped by episode. Linear map applies $h = 4$ composition-stress-map coefficients assuming symmetric boom/bust responses. Asymmetric map uses sign-split coefficients (bust costs weighted separately from boom savings). In the 2014–16 episode, the post-MDRI episode in which the export-price mechanism operates without debt-relief confounders, the asymmetric map (23.2 pp) matches the realized average (23.5 pp) within 0.3 pp. The 2008–09 over-prediction reflects that episode's diffuse, non-export-led character. Pre-2007 debt relief dominates realized changes; neither map is expected to replicate observed debt changes.

Table 10: Debt Local Projections: Common vs. Relative Shock Direction

h	Semi-partial R^2 (common)	Semi-partial R^2 (relative)	Common share	N
0	0.0137	0.0089	60.6%	441
2	0.0223	0.0040	84.9%	441
4	0.0204	0.0066	75.7%	441

Notes: Within-R^2 attribution from structural debt LPs on $u_{\text{common}} = u_x + u_m$ and $u_{\text{relative}} = u_x - u_m$ plus baseline controls and country fixed effects. Semi-partial R^2 is the drop in within-R^2 when that direction is omitted from the joint model. Common share is common/(common+relative); this is a relative comparison, not a decomposition of total explained variance. The common $[1, 1]$ direction, which the ToT ratio discards entirely, accounts for 61–85 percent of the attributable variation in debt responses. Fiscal analogue of Di Pace, Juvenal, and Petrella (2025, Appendix A.4.3).

Table 11: Structural Stock Identity: Common-Bust Debt Response, Budget versus Remainder

	Full sample			Excl. 2008–09, 2014–16		
h	Debt	Budget	Remainder	Debt	Budget	Remainder
0	1.7	−2.0	3.7	3.1	−1.6	4.7
1	4.4	−1.8	6.2	9.0	−1.7	10.6
2	5.8	−1.2	7.0	11.2	−0.9	12.0
3	5.1	−1.5	6.6	9.3	−1.7	11.0
4	6.0	−1.2	7.2	10.4	−1.6	12.0

Notes: Percentage points of GDP. Common-bust scenario constructed from the joint shock combination solving $\mathbf{As} = \mathbf{p}$ for a $-10\%/-10\%$ price path. Budget = cumulative structural primary-balance IRF, combined across legs; Remainder = Debt − Budget (GDP-denominator effects, valuation, interest/financing, timing residual). The budget share at $h = 4$ is −20 percent (full sample) and −15 percent (excluding crisis episodes): the fiscal channel is a net drag on debt, not a source of the increase. The debt effect is larger when crisis episodes are excluded (10.4 vs 6.0 pp), confirming the pattern is not crisis-driven.

Table 12: Non-DJP Replication: Shift-Share Symmetry Test on a Broader Global Panel

Horizon	$\hat{\beta}_x$	$\hat{\beta}_m$	$p(H_0\text{: } \beta_x + \beta_m = 0)$	N	Countries
Panel A: All 51 commodity-dependent economies					
$h = 0$	−8.80	−4.79	< 0.001	1420	51
$h = 2$	−13.06	−12.82	< 0.001	1318	51
$h = 4$	−16.27	−7.63	< 0.001	1216	51
Panel B: 38 countries, excluding all 13 in the Africa-20 sample (out-of-sample test)					
$h = 0$	−9.75	−4.76	< 0.001	1043	38
$h = 2$	−14.62	−11.31	< 0.001	967	38
$h = 4$	−20.17	−3.75	< 0.001	891	38

Notes: Country fixed-effects local projections of general-government debt/GDP on a BACI/Pink Sheet Bartik shift-share instrument (1995–99 CEPII BACI trade shares interacted with World Bank Pink Sheet commodity price growth). No DJP shocks or prices enter anywhere. Panel A covers all countries clearing a commodity-concentration threshold ($\geq 30\%$ tracked commodity exports) with ≥ 15 years of debt coverage. Panel B excludes the 13 Africa-20 countries in the main sample, a pure out-of-sample test. $H_0 : \beta_x + \beta_m = 0$ is rejected at $p < 0.001$ at every horizon in both panels. Country-clustered SEs in parentheses; horizons $h = 1$ and $h = 3$ omitted for brevity; results are qualitatively unchanged at those horizons.

Table 13: LIC-DSF Sensitivity: Country Distance from Compression-Error Threshold

Country	DCC	Threshold	Stressed PV	Buffer	Sensitivity ratio
SDN	Weak	35	211.9	−176.9	−19.42
MWI	Weak	35	60.1	−25.1	−2.76
GHA	Medium	55	71.1	−16.1	−1.77
TCD	Weak	35	50.1	−15.1	−1.66
KEN	Strong	70	68.2	1.8	0.20
NER	Medium	55	52.9	2.1	0.23
CIV	Medium	55	52.2	2.8	0.31
SEN	Strong	70	63.8	6.2	0.68
COD	Weak	35	25.7	9.3	1.02
BFA	Medium	55	45.1	9.9	1.09
CMR	Medium	55	43.1	11.9	1.30
MDG	Medium	55	38.5	16.5	1.81

Notes: Thresholds: Weak 55, Medium 65, Strong 75 percent of GDP (face-value, 2026 LIC-DSF review). Stressed PV is the baseline plus the 9.1 pp composition-aware compression gap at $h = 4$ (the discrepancy between the composition-aware path and the ToT-only path, which registers zero in a common bust). Buffer is threshold minus stressed PV (negative = already breaching). Sensitivity ratio is buffer divided by the compression gap (9.11 pp for every country). A ratio near zero means the buffer does not materially exceed the gap between the two approaches; a negative ratio marks a country already breaching under the composition-aware path. Four countries breach (negative ratio); Kenya (0.20) and Niger (0.23) are near the threshold.

Table 14: Structural LPs on 2026 LIC-DSF Realism and Liquidity Objects

	$h = 0$			$h = 4$		
	$\hat{\beta}_{u_x}$	$\hat{\beta}_{u_m}$	p(sym)	$\hat{\beta}_{u_x}$	$\hat{\beta}_{u_m}$	p(sym)
Revenue (% GDP)	0.34	0.04	0.150	0.41	0.18	0.013
	(0.24)	(0.12)		(0.19)	(0.20)	
Exports/GDP (pp)	1.17	0.58	< 0.001	0.88	0.32	0.007
	(0.35)	(0.20)		(0.39)	(0.26)	
Interest/revenue (pp)	-0.65	-0.04	0.124	-0.64	-0.24	0.035
	(0.30)	(0.19)		(0.41)	(0.38)	

Notes: DJP structural local projections of the 2026 LIC-DSF revised framework's "realism" and "liquidity" metrics, Africa-20, 1990–2019. Long-difference specification with country fixed effects, as in Table 3. Units: pp of GDP (or named ratio) per 1 s.d. shock. Country-clustered SEs in parentheses. All three objects reject symmetry ($H_0 : \hat{\beta}_{u_x} + \hat{\beta}_{u_m} = 0$) at $h = 4$ at the 5 percent level. The 2026 review added these metrics but retained the ToT-based stress scenario, so the compression error propagates into every new object the revised framework monitors.

Table 15: Gross Debt LPs: Full Africa-20 vs. LIC-DSF 12-Country Subsample

	$h = 0$			$h = 4$		
	$\hat{\beta}_x$	$\hat{\beta}_m$	p(sym)	$\hat{\beta}_x$	$\hat{\beta}_m$	p(sym)
Panel A: Structural shocks (u_{px}, u_{pm})						
Full Africa-20 ($G = 20$, $N = 441$)	-2.56	-0.14	0.009	-5.49	-1.24	< 0.001
	(0.85)	(0.71)		(1.19)	(0.70)	
LIC-DSF-12 ($G = 12$, $N = 255$)	-2.77	-0.23	0.025	-5.92	-1.14	0.003
	(1.01)	(0.91)		(1.99)	(1.10)	
Panel B: Reduced form ($\Delta \ln P_x$, $\Delta \ln P_m$)						
Full Africa-20 ($G = 20$, $N = 441$)	-12.57	-2.51	0.368	-26.25	-55.72	< 0.001
	(5.72)	(19.63)		(9.04)	(17.76)	
LIC-DSF-12 ($G = 12$, $N = 255$)	-8.55	-15.61	0.232	-26.91	-56.53	0.005
	(7.46)	(23.33)		(20.23)	(30.77)	

Notes: Reduced-form and structural local projections of gross government debt for the Africa-20 and the twelve LIC-DSF economies in the sample, 1990–2019. Long-difference specification, country fixed effects, baseline controls. Units: pp of GDP per unit $\Delta \ln P$ or per 1 s.d. shock. Country-clustered SEs in parentheses. LIC-DSF-12: BFA, CMR, TCD, CIV, COD, GHA, KEN, MDG, MWI, NER, SEN, SDN.

Table 16: Synchronized Common Busts in Structural Shock Space

Year	Countries in common bust	Mean u^{common}	Mean u^{ToT}	Direction error
2001	18/20	−2.060	+0.177	Yes, wrong sign
1998	16/20	−1.682	+0.125	Yes, wrong sign
2015	15/20	−1.853	−0.363	No, right sign, 5.1× too small
2019	20/20	−1.839	−0.060	No, right sign, 30.7× too small

Notes: Africa-20 DJP structural shocks. A country-year is a common bust if $u_x < 0$ and $u_m < 0$. $u^{\text{common}} = u_x + u_m$ and $u^{\text{ToT}} = u_x - u_m$. In 2001 and 1998 the panel-mean ToT signal is positive while the common component is strongly negative: the ratio points the wrong direction. In 2015 and 2019 the ToT signal is negative but 5.1 and 30.7 times smaller in absolute value than the common component. Years with ≥ 10 simultaneous common busts also include 2016 (15/20), 1997 (13/20), and 1991 (12/20).

Table 17: Common versus Relative Variation: Prices and Structural Shocks

Object	Common $[1,1]$ share	Relative $[1,-1]$ share
Raw price innovations $(\Delta \ln P_x, \Delta \ln P_m)$	0.682	0.318
Structural shocks (u_x, u_m)	0.560	0.440
Cross-country synchronization (mean pairwise correlation)		
$u^{\text{common}} = u_x + u_m$	0.673	
$u^{\text{ToT}} = u_x - u_m$	0.057	

Notes: Africa-20, 1990–2019 ($N = 600$). Common and relative shares are $\text{Var}(c)/(\text{Var}(c) + \text{Var}(r))$ with $c = x + m$ and $r = x - m$. Structural shocks are the published DJP posterior medians. Mean pairwise correlations are averages of country-pair correlations of the annual series.