Vol. 10 No. 2 (2026): Journal of Contemporary Social Research
Journal Link: https://aa.unij.co.uk/index.php/jcsr / Pub. Link: https://aa.unij.co.uk/index.php/jcsr/issue/view/12

Fiscal Discipline and Budget Credibility in Nigeria

Measurement, Forecast Error, and the Long-Run Determinants of Budget Execution

[1]Alfred Okoh, [2]Ojonimi Umaru,
[3]Nuhu Muhammad, [4]Allen Ogedengbe

Alfred Okoh: Department of Management, School of Business Management, Baze University, Abuja, Nigeria..

Corresponding author: Alfred Okoh.

DOI: https://doi.org/10.5281/zenodo.23180541

Abstract

This paper treats measurement and source consistency as first-order empirical questions in assessing fiscal discipline and budget credibility in Nigeria. Using a source-audited Central Bank of Nigeria budget-comparator series for 2010-2022, we find that aggregate expenditure execution averaged 88.1 per cent, that twelve of thirteen years fell below the comparator, and that a one-sample test rejects full credibility ($t(12) = -5.59$, $p < 0.001$; 95% CI 83.4-92.7), confirmed by a Wilcoxon test and bootstrap interval, with no significant time trend. The planned deficit fixed at enactment has averaged about 39 per cent of approved federal spending over 2021-2026 and is highly unstable. Using a single-vintage 2002-2024 dataset, an ARDL error-correction model shows a cointegrating relationship (bounds $F = 4.5$) with a highly significant error-correction term (-0.99, $p = 0.003$): execution reverts almost completely to equilibrium within the year. No individual fiscal or governance variable, including the exchange rate, is a significant long-run driver. The paper contributes a transparent measurement architecture, a single-vintage determinants dataset, and a reproducible finding that Nigerian budget execution is fast-reverting but unanchored by any single measured fundamental, with implications for forecast realism and fiscal transparency.

Keywords: fiscal discipline; budget credibility; budget execution; fiscal forecasting; ARDL; error correction; public financial management; Nigeria

JEL classification: H61; H62; H63; H68; E62; C22; F31

1 Introduction

A budget is an economic contract written before the fiscal year is known, translating macroeconomic assumptions, revenue expectations, borrowing capacity and policy priorities into an authorised plan. That plan is credible when institutions can reasonably expect resources and spending to remain close to what was approved, or expect deviations to be explained transparently. This is stricter than budget preparation: a government can execute 100 per cent of

a revised budget while departing materially from the original plan, and can spend more than the original budget and score an execution rate above 100 per cent even though the deviation signals weak reliability (PEFA Secretariat, 2016; International Budget Partnership & UN DESA, 2023).

Nigeria is a difficult but important setting for this distinction. Oil-related receipts are historically volatile, the naira cost of external obligations is exchange-rate sensitive, debt service competes with discretionary spending, and capital implementation depends on procurement, cash releases and ministry capacity (Baunsgaard, 2003; Basdevant & Imamoglu, 2021; van der Ploeg, 2011). The 2019 PEFA assessment identified low budget credibility as a material weakness, citing poor expenditure and revenue outturn, weak forecasting realism and weak commitment control (PEFA Secretariat, 2019). Khalid (2025) found revenue forecasts exceeded outturn in every year but 2020 over 2011-2023, with an average error of about 1.8 per cent of GDP (roughly 36 per cent of actual revenue), large capital-execution gaps, and understated debt-service forecasts. The pattern is regional: across 39 sub-Saharan African economies over 2021-2024, optimistic revenue forecasts, spending overruns and capital under-execution drive deficit slippage (Lopez Murphy, Sever, Simione, & Zhang, 2026; Kiendrebeogo & Gyesaw, 2026).

This paper places measurement discipline at the centre of a determinants analysis of budget execution, asking: how should credibility be measured when execution may exceed as well as fall short of plan; what can be established from source-audited evidence by formal inference alone; do execution and the fiscal-governance fundamentals share a long-run relationship once every series sits on one documented vintage; and what design prevents result-driven model selection? ARDL is treated as an estimator suited to a particular integration environment, not an identification strategy for simultaneity (Pesaran, Shin, & Smith, 2001).

Inference operates at two levels: thirteen source-audited comparator observations (2010-2022) test whether mean execution differs from full credibility, trends over time, or scales with budget size, and an ARDL/error-correction model on a single-vintage 2002-2024 dataset tests the determinants. The two layers agree that Nigerian expenditure runs persistently below plan, and the dynamic model adds that execution is a fast mean-reverting process; the individual long-run determinants' insignificance is reported plainly. Section 2 gives institutional context; Section 3, the conceptual framework; Section 4, contributions and hypotheses; Section 5, the datasets; Section 6, methodology; Section 7, results; Section 8, robustness; Sections 9-10, interpretation and policy; Section 11, limitations; Section 12, conclusion; Section 13, data and code.

2 Institutional and Fiscal Context

Nigeria's fiscal framework rests on annual Appropriation Acts, a Medium-Term Expenditure Framework, the Debt Management Office, Budget Implementation Reports, and the Fiscal Responsibility Act 2007 (Federal Republic of Nigeria, 2007). The executive prepares a budget on oil-price, oil-production and exchange-rate benchmarks; the National Assembly appropriates; the Budget Office records releases and outturn; the Auditor-General reports on execution. This is comprehensive on paper; the relevant question is whether outturn tracks the plan predictably. Three pressures matter: a large revenue share is oil-linked; debt service is a growing, rigid claim

that displaces discretionary spending when interest costs are under-forecast; and capital expenditure, dependent on procurement and MDA capacity, is the most compressible category when revenue disappoints. Credibility failures rarely trace to one point but arise from the interaction of formulation realism, financing and execution capacity.

Three definitional choices determine whether any credibility statistic is meaningful. The comparator: PEFA evaluates outturn against the amount originally approved, not a later administrative target (PEFA Secretariat, 2016); conflating a revised appropriation with the original flatters credibility precisely when the original plan failed most sharply. The perimeter: federal expenditure, general-government spending, federally collected revenue, retained federal revenue and total public debt are not interchangeable; a ratio mixing perimeters can be numerically precise yet economically meaningless. The vintage: the World Bank revised the WGI methodology in 2025 and recalculated values back to 1996, and Nigeria's GDP was rebased in 2025, so governance or GDP-scaled series must come from one documented vintage (World Bank, 2025b). If a denominator changes meaning across years, an estimated coefficient mixes fiscal behaviour with definitional change. The analysis reconciles the comparator, fixes the perimeter and holds each series to one vintage before estimation.

3 Analytical literature and conceptual framework

Fiscal discipline has two meanings: macro-level control of deficits, debt and fiscal risk, and budget-process discipline, whether financing plans are respected in execution. Institutions matter because budget choices face a common-pool problem in which spending actors gain from programmes whose cost is dispersed across taxpayers or future budgets; rules and transparent scrutiny can change those incentives (Alesina & Perotti, 1996; Alesina, Hausmann, Hommes, & Stein, 1999; Poterba & von Hagen, 1999; Hallerberg, Strauch, & von Hagen, 2007). Budget credibility is narrower: the relationship between an ex-ante plan and ex-post outcomes; PEFA anchors comparison to the original budget (PEFA Secretariat, 2016), and the International Budget Partnership treats credibility as the extent to which targets are met (International Budget Partnership, 2023, 2024; de Renzio & Cho, 2020). A dataset must preserve the temporal identity of the budget compared: original, supplementary/amended, revised and final outturn are not interchangeable.

Let EA_t denote actual aggregate expenditure in year t and EB_t the relevant budget comparator. The execution rate answers the directional question:

$$ER_t = 100 \times (EA_t / EB_t) \quad (1)$$

The symmetric absolute deviation from plan answers the magnitude question:

$$EDEV_t = 100 \times |EA_t - EB_t| / EB_t \quad (2)$$

and a proximity-to-plan credibility score follows:

$$ECRED_t = 100 - EDEV_t \quad (3)$$

Credibility does not begin at execution: an optimistic revenue envelope is not financeable at the assumed level. Following Khalid (2025), the revenue forecast error scaled to output is

$$RFE_t = 100 \times (RB_t - RA_t) / Y_t \quad (4)$$

Because Nigerian budgets are frequently amended, the original-to-revised wedge is itself a credibility signal:

$$W_t = 100 \times (EB,orig_t - EB,rev_t) / EB,orig_t \quad (5)$$

A related signal is the planned deficit share, PDS = 100 x (planned deficit / approved spending), measuring how much of the plan is contingent on borrowing at enactment. Equations (1)-(3) separate direction from magnitude; equations (4)-(5), with PDS, locate three upstream sources of credibility failure invisible to a pure execution measure: an unfinanceable revenue forecast, an unstable comparator, and heavy dependence on borrowing.

The framework links planning to execution through five channels: revenue realism, since forecasts display optimism bias and Nigeria's oil-production assumptions are a major error source (Frankel, 2011; ElBerry & Goeminne, 2021; Hameed, 2005; Khalid, 2025); debt-service rigidity; exchange-rate exposure, since depreciation raises the naira cost of foreign-currency debt and imports while affecting oil revenue; administrative capacity (Pattanayak, 2016; Rajaram, Le, Biletska, & Brumby, 2010; Eltokhy et al., 2025); and fiscal institutions, whose effectiveness depends on design and enforceability (Allen, 2009; Allen, Chaponda, Fisher, & Ray, 2017; Eyraud, Debrun, Hodge, Lledo, & Pattillo, 2018; Alonso-Albarran et al., 2025).

Fragmented, opaque procedures contribute to deficit bias, while hierarchical, transparent procedures associate with greater discipline (Alesina & Perotti, 1995, 1996; Alesina et al., 1999; Poterba & von Hagen, 1999; Hallerberg et al., 2007), though the template cannot simply be imported to low-income economies, where institution quality associates with performance heterogeneously and durable institutions evolve slowly (Dabla-Norris et al., 2010; Allen, 2009), and rules anchor policy only if monitored and enforceable (Eyraud et al., 2018; Lledo & Dudine, 2018; Alonso-Albarran et al., 2025). Forecast optimism is well documented (Frankel, 2011), motivating independent fiscal institutions (Beetsma & Debrun, 2016; OECD, 2014), and resource dependence adds procyclicality risk (Talvi & Vegh, 2005; Kaminsky, Reinhart, & Vegh, 2004); plans are less implementable where adjustments are large and institutions weak, a pattern that persists regionally (Lledo & Poplawski-Ribeiro, 2013; Lopez Murphy et al., 2026; Kiendrebeogo & Gyesaw, 2026).

Nigeria's formal architecture is not equivalent to realised credibility. The 2019 PEFA assessment reported outturn far below budget and weak revenue forecasting, alongside progress in tax administration and GIFMIS integration (PEFA Secretariat, 2019); Khalid (2025) adds a forecast-error lens. The gap is a reproducible budget-credibility series with a consistent perimeter, an original-versus-revised distinction, and every regressor on one documented vintage. This paper addresses that gap first and subordinates the determinants model to it.

4 Contribution and hypotheses

Three contributions remain useful whatever the determinants regression shows: a transparent Nigerian budget-credibility architecture storing original, revised and actual values separately; a formal distinction between a directional execution rate and a symmetric deviation-from-plan measure that prevents over-execution from scoring as higher credibility; and a single-vintage determinants dataset for 2002-2024 with a pre-specified modelling and replication protocol. These are measurement, data and design contributions rather than claims of econometric novelty.

The paper advances four testable propositions. H1 and H2 use the source-audited 2010-2022 comparator series; H3 and H4 use the single-vintage 2002-2024 determinants dataset.

H1. Aggregate federal expenditure execution is materially and statistically below the budget comparator over 2010-2022 (mean ER < 100).

H2. The measured shortfall is persistent rather than transitory: there is no significant improving or worsening trend in execution across the sample.

H3. Budget execution and the fiscal-governance set share a long-run (cointegrating) relationship, and deviations from that equilibrium are corrected over time (a significant, well-signed error-correction term).

H4. Among the fiscal and governance fundamentals, exchange-rate depreciation is a statistically significant long-run driver of budget execution.

The data support H1, H2 and the cointegration-and-correction half of H3, but reject H4: the exchange rate is not a significant long-run driver, and neither is any other single fundamental.

5 Data and variable construction

The paper uses two datasets, each traceable to documented sources. The first is a source-audited Central Bank of Nigeria (CBN) budget-comparator series for 2010-2022, used for the descriptive audit and small-sample inference. The second is a single-vintage determinants dataset for 2002-2024, assembled so that every regressor sits on one measurement vintage and the ARDL/error-correction model is estimated without definitional or dating inconsistency. Both are governed by the same discipline: every annual value requires a source document, a unit, a government perimeter, a budget-vintage status and a transformation field, and no observation from a later or alternative dataset is mixed into a series without explicit reconciliation.

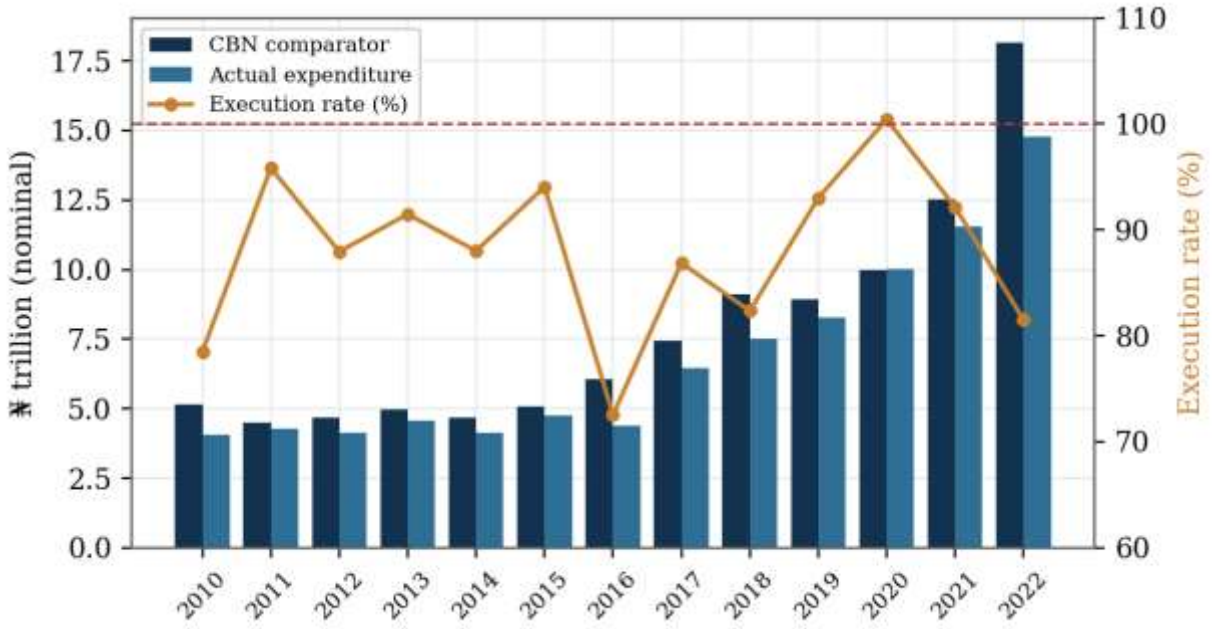
The CBN "Conduct of Fiscal Policy" history permits a defensible audit-level series for 2010-2022 (Table 1). The series is deliberately labelled a comparator rather than a final budget-credibility index: 2010 combines an amended budget with supplementary budgets, 2011 uses an amended budget, 2020 uses the revised COVID-19 budget, and 2012 and 2022 present more than one aggregate figure requiring reconciliation. Other years appear closer to an annual approved aggregate but still require matching to the relevant Appropriation Act and Budget Implementation Report before the series can be promoted to a PEFA-consistent measure.

Table 1. Source-audited CBN budget comparator versus actual federal expenditure, 2010-2022

Year	Comparator (Ntn)	Actual (Ntn)	ER (%)	EDEV (%)
2010	5.160	4.050	78.49	21.51
2011	4.486	4.299	95.84	4.16
2012	4.697	4.131	87.95	12.05
2013	4.986	4.561	91.47	8.53
2014	4.695	4.132	88.01	11.99
2015	5.068	4.767	94.07	5.93
2016	6.060	4.396	72.54	27.46
2017	7.441	6.464	86.86	13.14
2018	9.120	7.511	82.36	17.64
2019	8.917	8.293	93.00	7.00
2020	9.974	10.017	100.43	0.43
2021	12.512	11.530	92.15	7.85
2022	18.139	14.786	81.52	18.48

Note. CBN, Conduct of Fiscal Policy. ER=100xActual/Comparator; EDEV=100x|Actual-Comparator|/Comparator. 2010, 2011 and 2020 comparators are amended/revised budgets.

Figure 1. CBN budget comparator and federal expenditure outturn, 2010-2022



N trillion, bars; execution rate on right axis. Source: CBN fiscal history.

To estimate the determinants model on internally consistent data, each regressor is drawn from a single documented vintage (Table 3). Fiscal-balance and revenue ratios are IMF general-government revenue and net lending/borrowing as a share of GDP, published through the Sub-Saharan Africa Regional Economic Outlook, keeping numerator and GDP denominator on one internally consistent basis and avoiding the vintage break from Nigeria's 2025 GDP rebasing. The governance measure is the 2025-revision WGI government-effectiveness score, observed continuously from 2002, which sets the sample start. The exchange rate is the official naira/US$ annual average, entered in logs. Budget execution is the aggregate execution ratio (BCI), and the debt-service burden is debt service as a share of retained revenue. The result is a 23-year dataset, 2002-2024, in which every regressor is drawn from one documented source. Aggregate execution fluctuates around a mean below full execution, with pronounced under-execution in the mid-2000s and a partial recovery after 2012, never settling at 100 per cent for long - the descriptive signature of a persistent, mean-reverting credibility gap that Sections 7.1 and 7.6 formalise (Figures 1-3).

Table 3. Single-vintage determinants dataset, 2002-2024 (selected years; full annual series in the replication data)

Year	BCI (%)	FISGDR (%)	DSR (%)	REVGDP (%)	GE	FX (N/$)
2002	86.79	-0.95	9.16	14.72	21.7	120.97
2005	79.56	-3.49	13.45	16.13	30.2	131.27
2008	74.48	-4.05	34.2	14.26	31.6	118.57
2011	78.32	-0.31	44.56	12.59	29.5	153.86
2014	94.8	1.73	51.2	7.77	30.8	158.55
2015	93.5	2.7	70.95	5.15	33.6	193.28
2016	105.1	3.3	64.6	3.63	33.5	253.49
2017	80.5	3.85	75.65	4.67	34.6	305.79
2018	87.4	3.06	82.35	6.04	34.9	306.08
2019	90.3	3.31	40.8	5.56	32.5	306.92
2020	88.2	4.03	53.2	4.66	32.2	358.81
2021	92.15	3.97	71.77	5.12	30.9	409.42
2022	81.54	4	48.45	6.19	30.8	418.79
2023	92.8	3.13	54.97	6.79	34.5	646.37
2024	98.37	1.57	58.91	9.89	31.8	1486.57

Note. BCI=execution rate; FISGDR=-(IMF net lending/borrowing,%GDP); DSR=debt service/retained revenue; REVGDP=IMF revenue (%GDP); GE=WGI 2025 government-effectiveness; LEXR=ln(FX). Sources: IMF SSA REO; World Bank WGI 2025; World Bank/CBN. Full 23-year series used in estimation.

Figure 2. Aggregate budget execution (BCI), 2002-2024



Fluctuates below full execution, with a partial recovery after 2012.

6 Econometric methodology

Functional Form

$$BCI_t = f(FISGDR, DEBTSEVREV, REVGDP, GE, lnEXR)$$

Where:

- BCI = Budget Credibility Index in year ttt
- FISGDR = Fiscal deficit-to-GDP ratio
- DEBTSEVREV = Debt-service-to-revenue ratio
- REVGDP = Government revenue-to-GDP ratio
- GE = Government Effectiveness percentile rank
- lnEXRt = Natural logarithm of the naira per US dollar exchange rate
- εt\varepsilon_tεt = Error term

Mathematical Form

$$BCI_t = \beta_0 + \beta_1 FISGDR_t + \beta_2 DEBTSEVREV_t + \beta_3 REVGDP_t + \beta_4 GE_t + \beta_5 REVGDP_t + \beta_6 LEXR_t + \varepsilon_t \quad -----(2)$$

$\beta_0\beta_1\beta_2\beta_3\beta_4\ \beta_5\ \beta_6$ are coefficient represents the elasticity , α is the intercept. ϵt is the error term.

Econometric Form

$$\Delta BCI_t = \alpha_0 + \sum_{i=1}^{q-1} \alpha_1 \Delta \mathrm{BCI}_{t-1} + \sum_{i=0}^{qi-1} \beta_2 \Delta \mathrm{FISGDR}_{t-1} + \sum_{i=0}^{q2-1} \gamma_3 \Delta \mathrm{DEBTSEVREV}_{t-1}$$
$$+ \sum_{i=0}^{q3-1} \delta_4 \Delta \mathrm{REVGDP}_{t-1} + \sum_{i=0}^{q4-1} \varphi_5 \Delta \mathrm{GE}_{t-1} + \sum_{i=0}^{q5-1} \gamma_6 \Delta \mathrm{EXR}_{t-l} + \delta_7 \mathrm{ECT}_{t-1}$$
$$+ \lambda_1 \mathrm{BCI}_{t-1} + \lambda_2 \mathrm{FISGDR}_{t-1} + \lambda_3 \mathrm{DEBTSEVREV}_{t-1} + \lambda_4 \mathrm{REVGDP}_{t-1}$$
$$+ \lambda_5 \mathrm{GE}_{t-1} + \lambda_6 \mathrm{LEXR}_{t-l} + \mu_t$$

. Where α_0 and μ_t are the constant parameter and error term respectively. The expression with the signs of summation in the equation is error correction dynamics. The parameter coefficient denotes the short run effects and lambda (λ) is the corresponding relationship in the long run.

While ECT is the error correction term

7 Results

7.1 Expenditure execution is materially and significantly below the comparator

Across the thirteen comparator years the mean execution rate is 88.05 per cent (median 88.01), with a standard deviation of 7.70 points and a coefficient of variation of 8.7 per cent. The lowest rate is 72.54 per cent (2016) and the highest 100.43 per cent (2020). Pooling the naira amounts before ratioing gives 87.83 per cent, and twelve of thirteen years fall below the comparator, so the result is not driven by a few late-sample years (Table 4).

The shortfall is statistically significant: a one-sample t-test against the null of full credibility rejects decisively, t(12) = -5.59, $p < 0.001$, 95 per cent CI 83.40-92.71. This is confirmed by a Wilcoxon signed-rank test ($p < 0.001$) and a bootstrap interval of 83.86-91.93 (Table 5). All three methods agree: expenditure was reliably and significantly below the comparator over 2010-2022, supporting H1.

Table 4. Descriptive statistics of the comparator series, 2010-2022 (n = 13)

Statistic	Value	Statistic	Value
Mean execution rate (%)	88.05	Pooled execution rate (%)	87.83
Median execution rate (%)	88.01	Mean symmetric alignment (%)	87.99
Std. deviation (pp)	7.70	Mean absolute deviation (%)	12.01
Coefficient of variation (%)	8.75	Years below / above comparator	12 / 1
Minimum (2016) (%)	72.54	Maximum (2020) (%)	100.43
Restricted 2013-2019 mean (%)	86.90	Observations	13

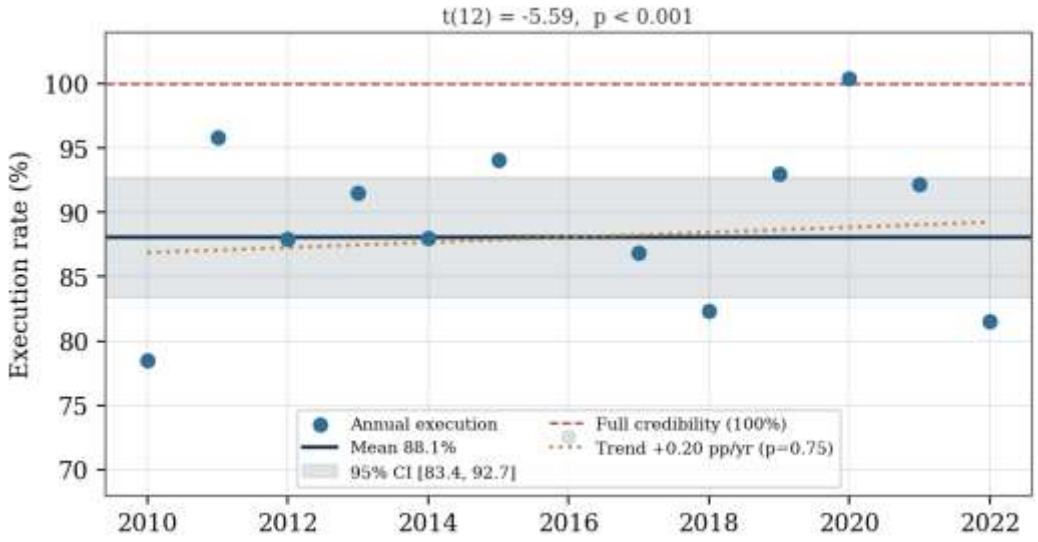
Note. Computed from Table 1. The 2013-2019 window uses years closest to an annual approved aggregate.

Table 5. Formal inference on the comparator series

Test / quantity	Statistic	95% interval	p-value
Mean execution vs 100 (t-test, df 12)	t = -5.59	[83.40, 92.71]	< 0.001
Mean execution vs 100 (Wilcoxon)	W = 1.0	n/a	< 0.001
Mean execution (bootstrap, 10,000)	mean = 88.05	[83.86, 91.93]	n/a
Time trend in execution (OLS slope)	+0.20 pp/yr (SE 0.59)	n/a	0.747
Execution vs log budget size (Pearson r)	r = -0.03	n/a	0.924
R-squared of the trend regression	0.010	n/a	n/a

Note. All three methods reject full credibility; trend and size correlations are insignificant.

Figure 3. Execution rate: mean, 95% confidence interval and estimated trend



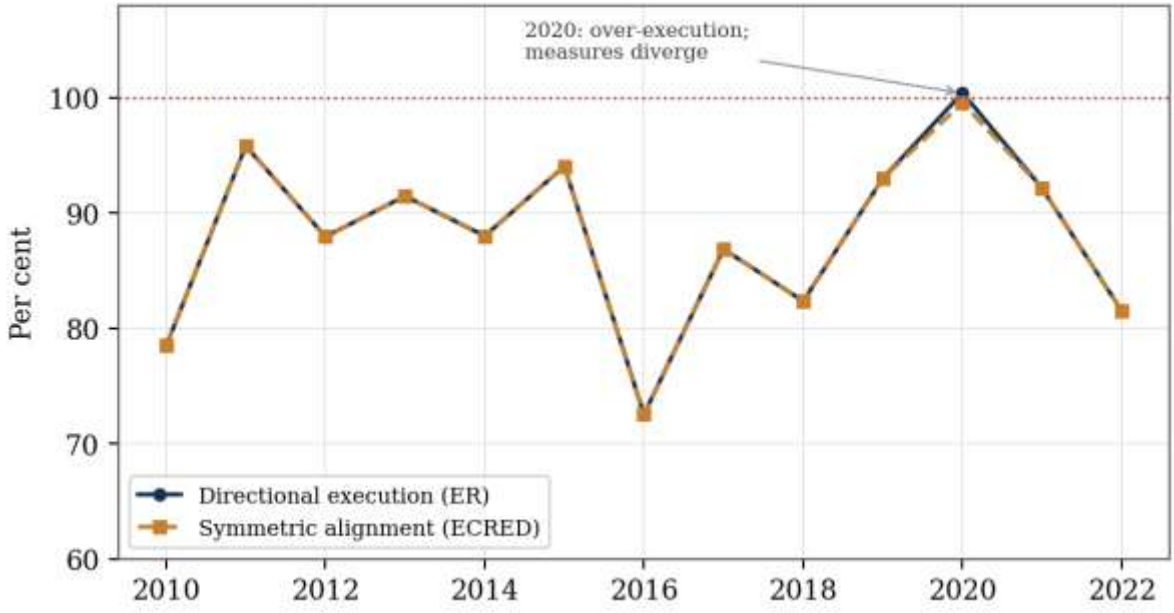
t(12) = -5.59, p < 0.001; no significant time trend.

7.2 The shortfall is persistent, and the measurement correction matters conceptually

Regressing execution on a centred time trend yields an insignificant slope of +0.20 points/year (SE 0.59; p = 0.75; R-squared = 0.01), and its correlation with the log of the nominal comparator is essentially zero (r = -0.03; p = 0.92). The shortfall is persistent rather than improving or worsening and does not scale with budget size, supporting H2 and indicating a structural rather than transitional gap.

The measurement distinction in equations (1)-(3) matters conceptually though it changes few sample values: mean symmetric alignment is 87.99 per cent, only 0.07 points below mean execution, since twelve of thirteen observations lie below 100. The measures diverge where it matters: in 2020 execution is 100.43 per cent but symmetric alignment 99.57 per cent, so an execution-only index would score overshoot above perfect credibility. Only two observations lie within 5 per cent of the comparator, six within 10, nine within 15, with a mean absolute deviation of 12.01 per cent - high execution is not tight alignment.

Figure 4. Directional execution and symmetric comparator alignment, 2010-2022



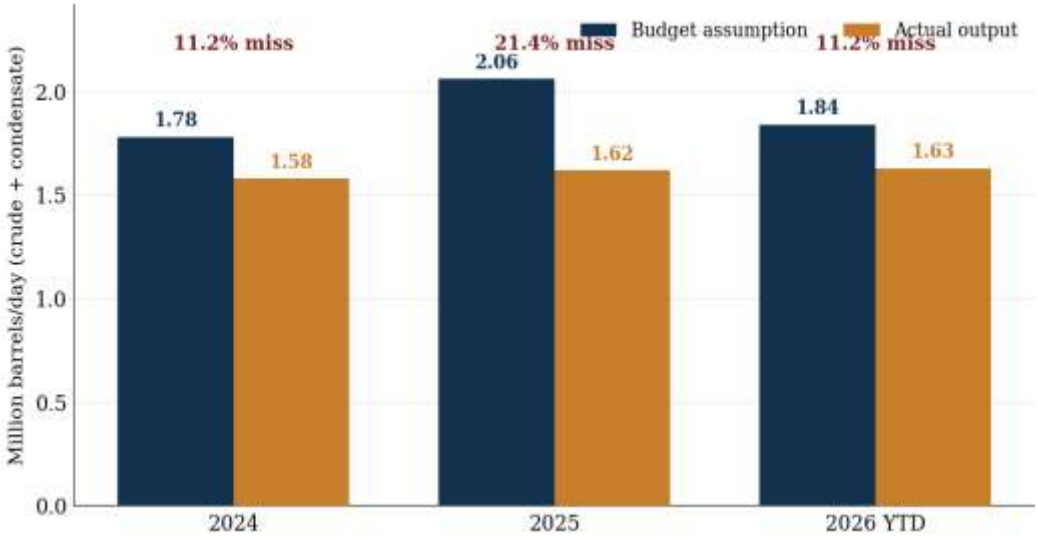
Measures coincide below plan, diverge only when spending exceeds it (2020).

7.3 Independent evidence identifies forecast realism and capital execution as mechanisms

Khalid (2025) finds Nigerian revenue forecasts exceeded outturn in every year but 2020 over 2011-2023, averaging 1.8 per cent of GDP (about 36 per cent of actual revenue), dominated by optimistic oil-production assumptions, supporting the revenue-realism mechanism of equation (4). Capital expenditure is over-budgeted most severely, by more than 70 per cent of actual executed capital spending, while debt-service costs are typically under-budgeted. If revenue fails but spending compresses in response, the eventual deficit can look closer to forecast than either component alone, so the fiscal-balance error understates budget weakness.

Oil-production assumptions illustrate the mechanism directly: budgets assumed 1.78 million barrels/day for 2024 against an outturn of 1.58, 2.06 for 2025 against 1.62, and 1.84 for 2026 against roughly 1.63 to date - shortfalls of 11-21 per cent that translate almost mechanically into revenue optimism and compression of capital spending (Figure 8).

Figure 5. Budget oil-production assumptions versus actual output, 2024-2026



Million barrels/day; budget assumption exceeds outturn by 11-21%. Sources: NUPRC/Reuters.

7.4 The planned deficit is an ex-ante financing stance, not a realised balance

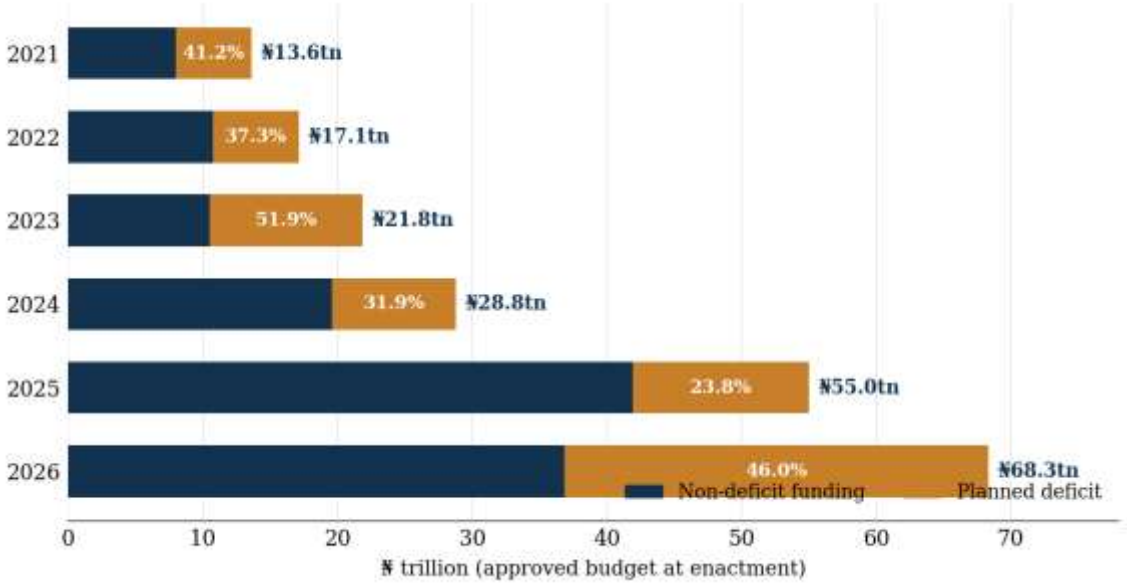
The figure quoted as "the deficit" is the planned financing gap at enactment, not the realised balance. Across the six approved budgets from 2021 to 2026, the planned deficit averaged 38.7 per cent of approved spending, ranging from 23.8 per cent (2025) to 51.9 per cent (2023) (Table 8), rebounding to 46.0 per cent in 2026 (N31.46tn, the largest in nominal terms) - roughly two-fifths of spending is contingent on borrowing at enactment, and the share is not converging.

Table 8. Approved federal budget and planned deficit at enactment, 2021-2026

Fiscal year	Approved spending (Ntn)	Planned deficit (Ntn)	Deficit share (%)
2021	13.59	5.60	41.2
2022	17.13	6.39	37.3
2023	21.83	11.34	51.9
2024	28.77	9.18	31.9
2025	54.99	13.08	23.8
2026	68.32	31.46	46.0
Mean 2021-2026	n/a	n/a	38.7

Note. Approved at enactment; deficit is the planned gap, not the realised balance.

Figure 6. Approved federal budget and planned deficit at enactment, 2021-2026



N trillion; amber segment is the planned financing gap.

This deficit-funded portion is most exposed when financing is constrained, and since capital spending is the compressible margin (7.3), a high deficit share raises the fragility of exactly the components that under-execute most. Borrowing two-fifths of the budget yearly also feeds the debt-service ratio, which rises from about 9 per cent of revenue in 2002 to the 40-75 per cent range recently (Table 3). The 2025 budget illustrates the point: approved at N54.99tn with a N13.08tn deficit, later revised to N59.99tn with a N14.10tn deficit - a nine-per-cent upward revision that is exactly the wedge W of Section 3.

7.5 Institutional and regional evidence reinforce the picture

The 2019 PEFA assessment found discipline undermined by poor outturn against the original budget, unrealistic forecasts and weak release predictability, alongside progress in tax administration and GIFMIS integration - procedural reform is necessary but not sufficient. Regionally, deviations across 39 countries in 2021-2024 are driven by optimistic forecasts, overruns and capital under-execution, smaller where institutions and IMF-supported programmes are stronger (Lopez Murphy et al., 2026; Kiendrebeogo & Gyesaw, 2026). Nigeria fits the pattern, with particular weight on oil-production optimism, a heavy deficit stance and capital compression.

7.6 The determinants model, 2002-2024

With every regressor on one documented vintage, the headline is precise: execution and the fiscal-governance set form a cointegrating system correcting deviations almost completely within the year, but no individual fundamental, exchange rate included, is a significant long-run driver.

Unit-root tests show mixed orders (Table 9): execution and government effectiveness are near-stationary in levels, while the fiscal ratios, debt-service ratio and log exchange rate are first-difference stationary. No series is unambiguously I(2) - the mixed-order environment the bounds procedure is designed for.

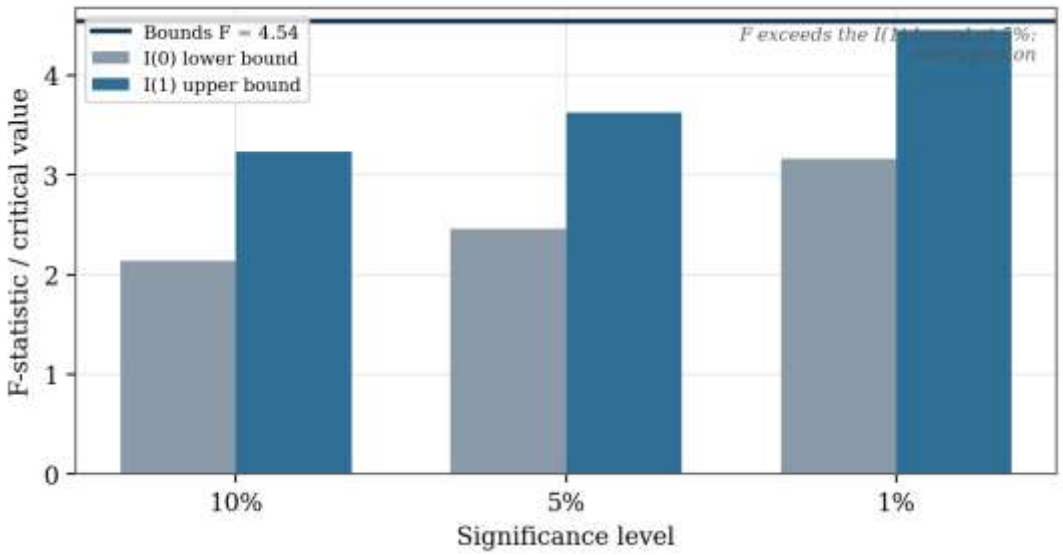
Table 9. Unit-root tests (Augmented Dickey-Fuller p-values), 2002-2024

Variable	Level	First difference	Indicated order
BCI (execution)	0.046	0.401	I(0)
FISGDR (deficit/GDP)	0.993	0.997	I(1) (PP)
DSR (debt service/revenue)	0.116	0.351	I(1)
REVGDP (revenue/GDP)	0.318	1.000	I(1) (PP)
GE (govt effectiveness)	0.179	0.001	I(0)/I(1)
LEXR (ln exchange rate)	1.000	0.157	I(1)

Note. MacKinnon p-values, AIC lag length; no variable is I(2).

The bounds F-test rejects no levels relationship: 4.16 (restricted constant) and 4.54 (unrestricted), both above the 5 per cent I(1) bound of 3.63 (Table 10) - the cointegration half of H3, robust across deterministic cases.

Figure 7. Bounds F-test against the Pesaran-Shin-Smith critical bounds



F-statistic (4.54) exceeds the I(1) upper bound, rejecting no levels relationship.

Table 10. Bounds cointegration test

Specification	F-statistic	5% I(0)-I(1) bound	Decision
Restricted constant (case 2)	4.16	2.28 - 3.29	Cointegration
Unrestricted constant (case 3)	4.54	2.46 - 3.63	Cointegration

Note. Critical values: Pesaran, Shin and Smith (2001). F-statistic exceeds the I(1) bound in both cases.

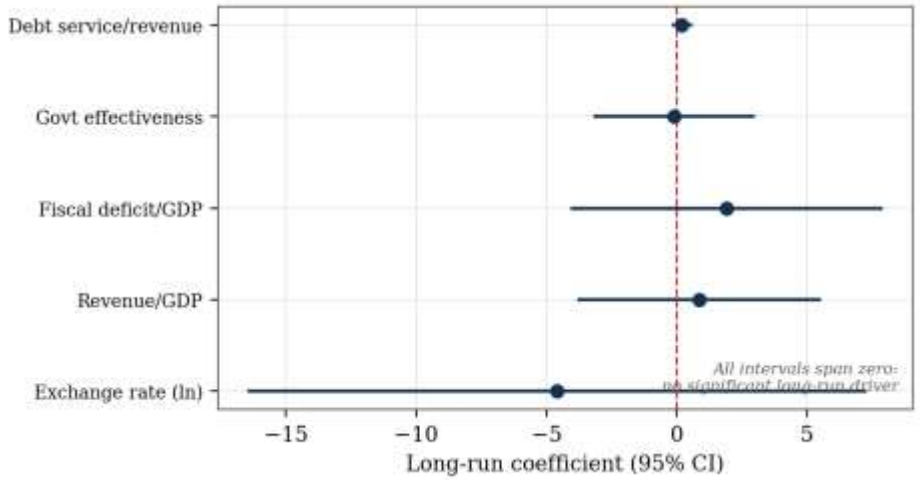
The long-run coefficients (Table 11) are, without exception, insignificant: debt service (0.21, p = 0.20), fiscal deficit (1.92, p = 0.49), revenue (0.88, p = 0.68), government effectiveness (-0.09, p = 0.95), and the exchange rate (-4.59, p = 0.41); every interval spans zero. This is decisive for H4: exchange-rate depreciation carries no detectable long-run effect, and no other fundamental replaces it as an anchor.

Table 11. Long-run coefficients (dependent variable: BCI)

Regressor	Coefficient	Std. error	95% CI	p-value
Debt service / revenue	0.21	0.15	[-0.13, 0.54]	0.197
Fiscal deficit / GDP	1.92	2.66	[-4.01, 7.85]	0.487
Revenue / GDP	0.88	2.07	[-3.73, 5.49]	0.680
Govt effectiveness	-0.09	1.36	[-3.13, 2.95]	0.948
Exchange rate (ln)	-4.59	5.30	[-16.40, 7.23]	0.408

Note. From the ARDL(1,1,1,1,1,1) conditional error-correction model. None is significant.

Figure 8. Long-run coefficients with 95% confidence intervals



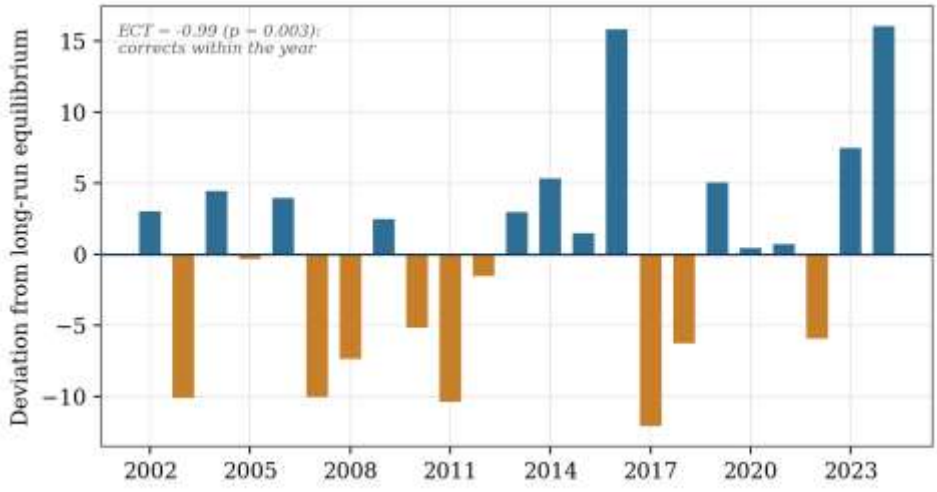
Every interval spans zero: no variable is a significant long-run driver.

The dynamics, in contrast, are sharp. The error-correction coefficient is -0.986 (SE 0.253, p = 0.003, t = -3.89), comfortably stable:

DBCI_t = c - 0.986 (BCI_{t-1} - theta'X_{t-1}) + short-run terms, (SE = 0.253, p = 0.003) (10)

A coefficient of -0.99 means essentially the entire deviation from equilibrium in one year is corrected by the next: execution is a fast mean-reverting process, with shocks almost fully unwound within twelve months.

Figure 9. Deviation from long-run equilibrium (cointegrating residual), 2002-2024



Error-correction coefficient -0.99 (p = 0.003): corrected within the year.

Table 12. Specification diagnostics for the determinants model

Diagnostic	Statistic (p)	Decision	Reading
Breusch-Godfrey serial correlation	0.024	Fails	Residual dynamics remain
Breusch-Pagan heteroscedasticity	0.493	Passes	Constant variance
Ramsey RESET (functional form)	0.018	Fails	Possible non-linearity/mis-specification
Jarque-Bera normality	0.895	Passes	Residuals approx. normal
R-squared (change in execution)	0.84	n/a	High explanatory fit

Note. From the conditional error-correction model. Robust content: the cointegration and error-correction result.

The model explains 84 per cent of the variation in the change in execution and passes heteroscedasticity and normality tests, but fails the Breusch-Godfrey ($p = 0.024$) and Ramsey RESET ($p = 0.018$) tests (Table 12). More parsimonious specifications did not clear both tests, and the failures trace to the same source: an annual execution ratio over twenty-three observations carries residual dynamics that five slow-moving fundamentals do not fully capture. The robust content is the cointegration-and-correction result, not any individual coefficient; where the exchange rate or debt-service ratio appear significant in narrower specifications, those specifications fail the diagnostics.

8 Robustness and diagnostics

On measurement, the conclusion is invariant to the choice between the directional execution rate and the symmetric alignment score: their means differ by 0.07 points, both well below 100. It also survives excluding years whose comparators are explicitly amended or revised - the 2013-2019 subsample still averages 86.90 per cent - so the shortfall is not an artefact of the obvious revision cases. Rejection of full credibility holds across the parametric, distribution-free and bootstrap methods, and the two null results (no trend, no scaling with budget size) are themselves evidence the shortfall is structural.

On the determinants model, both robust findings survive every reasonable variation. The cointegration result holds under both deterministic cases, and the error-correction coefficient stays close to -1 and highly significant (-0.93 to -0.99, $p <= 0.01$) with a post-2015 exchange-rate break dummy and with depreciation entered as a rate rather than a level. The insignificance of the long-run drivers is equally robust: no specification places a conventional confidence level on any single fundamental once diagnostics are respected, and apparent exchange-rate significance appears only in diagnostic-failing specifications. Two caveats: the serial-correlation and RESET

failures indicate no fully well-specified ARDL exists at this sample size, and the exchange rate's rapid post-2015 depreciation strains the I(0)/I(1) requirement.

9 Discussion

The main finding has two parts. Once "budget credibility" is fixed - the original plan as denominator, overspending treated as deviation rather than success, and each regressor on one vintage - the Nigerian outturn is significantly and persistently below plan, and execution forms a cointegrating system with the fiscal-governance set that corrects deviations almost completely within the year. Yet no individual fundamental is a significant long-run driver of that system. These are not contradictory: a near-unit error-correction coefficient with insignificant long-run loadings describes a variable strongly tied to its own equilibrium and adjusting quickly, but whose equilibrium is not pinned by any single measured fundamental in a short annual sample. Execution behaves less like the output of a structural fiscal reaction function and more like a fast-reverting process disciplined by an interaction of formulation, financing and capacity that no one aggregate captures.

The exchange-rate result matters because theory gives it a strong prior - debt service and imported capital inputs are foreign-currency denominated and oil revenue is dollar-linked - yet the long-run coefficient is statistically indistinguishable from zero, attaining significance only in short-run, differenced terms, consistent with depreciation operating through fast within-year channels rather than as a stable long-run anchor.

The evidence also changes how under-execution should be read. A low capital-execution rate is easily attributed to weak implementation capacity, and PEFA does identify planning and control weaknesses, but Khalid (2025) shows the budget itself often begins from optimistic revenue assumptions, and capital spending is the most compressible component when expected oil revenue does not arrive. An implementation gap therefore combines over-ambitious formulation, financing shortfall and project capacity, which a single governance coefficient cannot disentangle; debt-service pressure warrants the same caution, since it can rise from falling revenue or exchange-rate moves rather than from spending discipline alone.

A budget borrowing close to two-fifths of its spending, in a share swinging from under a quarter to more than half between adjacent years, is fragile before execution begins, and successive planned deficits drive the debt-service ratio that later competes with discretionary spending. Stronger domestic tax capacity can reduce dependence on volatile resource receipts but requires integrated policy, law and administration rather than rate increases alone (Baer et al., 2025); Nigeria already has formal fiscal-responsibility architecture, and the practical gap is the weak feedback loop between forecast performance and the next budget, not the absence of another rule (Eyraud et al., 2018; Alonso-Albarran et al., 2025). A twenty-three-year single-vintage dataset yielding a defensible cointegration result is more informative than a longer series spliced from incompatible vintages.

10 Policy implications

Nigeria should publish and retain, in machine-readable form, the original approved budget, every amending budget, in-year releases, the financing plan and the final outturn under a stable perimeter, without revisions overwriting the original plan - and reconcile the planned deficit against the realised balance given how high and volatile that stance has become. Forecast-performance reviews should be institutionalised, comparing each budget's assumptions and outturn and decomposing errors in oil price, production, exchange rate and tax bases. Because capital spending is the adjustment margin when revenue underperforms, in-year monitoring should distinguish projects delayed for lack of cash release from those delayed by procurement or capacity failures, since remedies differ (Pattanayak, 2016; Eltokhy et al., 2025). Further, evidence-consistent recommendations: domestic revenue mobilisation can reduce oil dependence if built on realistic capacity (Baer et al., 2025); fiscal-rule design should preserve a medium-term anchor with well-defined shock responses (Lledo & Dudine, 2018; Eyraud et al., 2018); and independent scrutiny of macro-fiscal assumptions can reduce optimistic forecasting (OECD, 2014; Beetsma & Debrun, 2016). This paper does not estimate the causal effect of a fiscal council or a specific debt, exchange-rate or oil-price rule.

11 Conclusion

Nigeria's budget-credibility problem is easy to describe and harder to measure. This paper shows why measurement must precede estimation: the researcher must know which budget is the denominator, which perimeter each variable covers, whether overspending is rewarded, and whether every regressor sits on one vintage.

That discipline delivers substance. On the comparator series, execution averaged 88.05 per cent (pooled 87.83), twelve of thirteen years fell below the comparator, and a one-sample test rejected full credibility ($t(12) = -5.59$, $p < 0.001$) across parametric, non-parametric and bootstrap methods, with no significant trend. On the 2002-2024 determinants dataset, execution and the fiscal-governance set form a cointegrating system (bounds $F = 4.5$) that corrects almost completely within the year (error-correction coefficient -0.99, $p = 0.003$), but no individual fundamental, exchange rate included, is a significant long-run driver.

A result is not made more credible by estimator sophistication when data mix perimeters or vintages. Fixing measurement first turns this into a reproducible programme of Nigerian fiscal analysis whose message is clear even where it is null.

References

Alesina, A., Hausmann, R., Hommes, R., & Stein, E. (1999). Budget institutions and fiscal performance in Latin America. Journal of Development Economics, 59(2), 253-273.

Alesina, A., & Perotti, R. (1995). The political economy of budget deficits. IMF Staff Papers, 42(1), 1-31.

Alesina, A., & Perotti, R. (1996). Fiscal discipline and the budget process. American Economic Review, 86(2), 401-407.

Allen, R. (2009). The challenge of reforming budgetary institutions in developing countries (IMF Working Paper No. 2009/096). International Monetary Fund.

Allen, R., Chaponda, T., Fisher, L., & Ray, R. (2017). Medium-term budget frameworks in Sub-Saharan African countries (IMF Working Paper No. 2017/203). International Monetary Fund.

Alonso-Albarran, V., Arroyo, C., Aydin, O., Davoodi, H. R., Lam, W. R., Balasundharam, V., Hegab, G., Nguyen, A. D. M., Salazar, N., Sher, G., Solovyeva, A., & Tchelishvili, N. (2025). Fiscal rules and fiscal councils: Recent trends and revisions since the pandemic (IMF Working Paper No. 2025/198). International Monetary Fund.

Baer, K., Bellon, M., Davies, M., de Mooij, R., Gaspar, V., Lemgruber, A., Mansour, M., Sawadogo, F., Takebe, M., & Vellutini, C. (2025). Building tax capacity for growth and development (IMF Departmental Paper No. 2025/007). International Monetary Fund.

Basdevant, O., & Imamoglu, E. (2021). How to design a fiscal strategy in a resource-rich country (IMF How-To Note No. 2021/001). International Monetary Fund.

Baunsgaard, T. (2003). Fiscal policy in Nigeria: Any role for rules? (IMF Working Paper No. 2003/155). International Monetary Fund.

Beetsma, R. M. W. J., & Debrun, X. (2016). Fiscal councils: Rationale and effectiveness (IMF Working Paper No. 2016/086). International Monetary Fund.

Budget Office of the Federation. (2026). Quarterly budget implementation reports. Federal Republic of Nigeria.

BudgIT Foundation. (2025). Analysis of the 2025 Federal Government approved budget. https://budgit.org/post_infographics/2025-fg-approved-budget/

BudgIT Foundation. (2026). Federal Government approved budget analyses, 2021-2026.

Central Bank of Nigeria. (n.d.). Conduct of fiscal policy. https://www.cbn.gov.ng/MonetaryPolicy/Fiscal.html

Central Bank of Nigeria. (2025). Statistical bulletin and annual economic report.

Dabla-Norris, E., Allen, R., Zanna, L.-F., Prakash, T., Kvintradze, E., Lledo, V., Yackovlev, I., & Gollwitzer, S. (2010). Budget institutions and fiscal performance in low-income countries (IMF Working Paper No. 2010/080). International Monetary Fund.

de Renzio, P., & Cho, C. (2020). Exploring the determinants of budget credibility. International Budget Partnership.

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association, 74(366a), 427-431.

ElBerry, N. A., & Goeminne, S. (2021). Fiscal transparency, fiscal forecasting and budget credibility in developing countries. Journal of Forecasting, 40(1), 144-161.

Eltokhy, K., Feruglio, N., Miao, K., Navarro, A., & Tandberg, E. (2025). How to improve public investment management in low-income countries (IMF How-To Note No. 2025/001). International Monetary Fund.

Eyraud, L., Debrun, X., Hodge, A., Lledo, V. D., & Pattillo, C. A. (2018). Second-generation fiscal rules: Balancing simplicity, flexibility, and enforceability (IMF Staff Discussion Note No. 2018/004). International Monetary Fund.

Federal Republic of Nigeria. (2007). Fiscal Responsibility Act, 2007.

Frankel, J. A. (2011). Over-optimism in forecasts by official budget agencies and its implications. Oxford Review of Economic Policy, 27(4), 536-562.

Hallerberg, M., Strauch, R., & von Hagen, J. (2007). The design of fiscal rules and forms of governance in European Union countries. European Journal of Political Economy, 23(2), 338-359.

Hameed, F. (2005). Fiscal transparency and economic outcomes (IMF Working Paper No. 2005/225). International Monetary Fund.

International Budget Partnership. (2023). Strengthening budget credibility for service delivery.

International Budget Partnership. (2024). Strengthening budget credibility: The pocket guide for auditors.

International Budget Partnership & United Nations Department of Economic and Social Affairs. (2023). Strengthening budget credibility through external audits: A handbook for auditors.

International Monetary Fund. (2025). Sub-Saharan Africa Regional Economic Outlook database: General government revenue and net lending/borrowing for Nigeria (Series NGAGGRG01GDPPT, NGAGGXCNLG01GDPPT) [Data set]. FRED, Federal Reserve Bank of St. Louis.

Kaminsky, G. L., Reinhart, C. M., & Vegh, C. A. (2004). When it rains, it pours: Procyclical capital flows and macroeconomic policies. NBER Macroeconomics Annual, 19, 11-53.

Khalid, S. (2025). Fiscal forecasting errors in Nigeria (IMF Selected Issues Paper No. 2025/095). International Monetary Fund.

Kiendrebeogo, Y., & Gyesaw, B. K. (2026). Fiscal forecast errors in East Africa: Nature, drivers, and policy implications (IMF Working Paper No. 2026/055). International Monetary Fund.

Lledo, V. D., & Dudine, P. (2018). How to select fiscal rules: A primer (IMF How-To Note No. 2018/001). International Monetary Fund.

Lledo, V., & Poplawski-Ribeiro, M. (2013). Fiscal policy implementation in Sub-Saharan Africa. World Development, 46, 79-91.

Lopez Murphy, P., Sever, C., Simione, F. F., & Zhang, Q. (2026). Budget credibility in Sub-Saharan Africa (IMF Departmental Paper No. 2026/006). International Monetary Fund.

Nigerian Upstream Petroleum Regulatory Commission. (2026). Nigeria crude oil and condensate production data.

OECD. (2014). Recommendation of the Council on principles for independent fiscal institutions (OECD/LEGAL/0401).

Pattanayak, S. (2016). Expenditure control: Key features, stages, and actors (IMF Technical Notes and Manuals No. 2016/002). International Monetary Fund.

PEFA Secretariat. (2016). PEFA framework for assessing public financial management.

PEFA Secretariat. (2019). Nigeria 2019: Federal Government PEFA assessment.

Pesaran, M. H., Shin, Y., & Smith, R. J. (2001). Bounds testing approaches to the analysis of level relationships. Journal of Applied Econometrics, 16(3), 289-326.

Phillips, P. C. B., & Perron, P. (1988). Testing for a unit root in time series regression. Biometrika, 75(2), 335-346.

Poterba, J. M., & von Hagen, J. (Eds.). (1999). Fiscal institutions and fiscal performance. University of Chicago Press.

Rajaram, A., Le, T. M., Biletska, N., & Brumby, J. (2010). A diagnostic framework for assessing public investment management (World Bank Policy Research Working Paper No. 5397). World Bank.

Reuters. (2026). Nigeria fiscal and oil-production reporting. Thomson Reuters.

Talvi, E., & Vegh, C. A. (2005). Tax base variability and procyclical fiscal policy in developing countries. Journal of Development Economics, 78(1), 156-190.

van der Ploeg, F. (2011). Natural resources: Curse or blessing? Journal of Economic Literature, 49(2), 366-420.

World Bank. (2025b). The worldwide governance indicators: Revised methodology for measuring governance using perception data.

World Bank. (2025c). Official exchange rate (LCU per US$, period average) for Nigeria [Data set] (Series PA.NUS.FCRF). World Development Indicators.