Advance Journal of Econometrics and Finance
Vol-4, Issue-1, 2026

Advance Journal of Econometrics and Finance

Online ISSN
2959-8990
Print ISSN
2959-8982
https://ajeaf.com/index.php/Journal/About s://ajeaf.

Name of Publisher: SCHOLAR CRAFT EDUCATION & RESEARCH HUB

Review Type: Double Blind Peer Review

Jurnal Frequency: Quarterly Research Journal

Political Periods, Institutional Quality, and Economic Growth in Afghanistan (1996–2024): Historical-Period Associations and Univariate Mean-Break Diagnostics

Dr. Wafiullah Kareemzada[1]

Dr. Wafiullah Kareemzada
PhD in Finance, Institute of Management Studies, University of Peshawar, Pakistan
Email: wafikareemzada@uop.edu.pk
ORCID: https://orcid.org/0009-0009-7397-2509

Abstract

This study evaluates whether historically assigned political periods coincide with statistical changes in Afghan macroeconomic series and whether the estimated period growth association survives reasonable econometric sensitivity checks. The author's 1996–2024 annual archive (N = 29), two near-identical Stata files and an archived Stata output log are cross-checked. The five-period ordinary least squares (OLS) model reproduces the archived joint restriction test, $F(4, 19) = 3.52$, $p = .0261$, although all individual period coefficients are imprecise. Newly calculated heteroskedasticity-consistent Wald tests yield $p = .1203$ (HC1) and $p = .3004$ (HC3), a one lag HAC test yields $p = .0293$, illustrating covariance sensitivity. Removing 2021 gives $p = .1186$; moving the final-period boundary to 2022 gives $p = .2268$. Archived univariate mean-break screens locate tax-revenue breaks near 2001 and 2008, whereas the GDP growth candidate near 2019 is supported at 10% but not 5%. A separate constrained least squares enumeration reproduces the break date minimum residual sum of squares partitions but not the xtbreak test distributions. The results distinguish historical categories, observational conditional contrasts, and univariate changes in means; they do not identify causal political transition effects. Original source vintages and treatment of early zero coded observations remain unverified.

Keywords: Afghanistan; political periods; economic growth; structural breaks; heteroskedasticity-robust inference; influential observations; data provenance.

Advance Journal of Econometrics and Finance
Vol-4, Issue-1, 2026

1. Introduction

Afghanistan's political and economic history presents a difficult question for empirical growth research: do economic series change when political authority changes, or do historical periods merely coincide with other shocks? Between 1996 and 2024, conflict, reconstruction, external financing, fiscal administration and the availability of official statistics changed substantially. An annual political period indicator may summarize those conditions, but it cannot isolate the effect of a transition from the events occurring at the same time.

A second distinction is equally important. Political periods are classified by historical chronology; structural breaks are estimated from economic observations under explicit statistical restrictions. A break in the mean of tax revenue, for example, need not occur in the same year as a change in governing authority, nor does it identify why revenue changed. Conflating these concepts can turn descriptive alignment into an unsupported causal explanation.

The first question is whether the four nonreference period indicators jointly improve a conditional growth regression, and whether this conclusion withstands small-sample covariance and influential-observation checks. The second asks whether statistically estimated mean shifts in output growth, tax revenue and institutional quality follow the same timing as historical categories. A third question concerns whether changing the calendar year treatment of the 2021 transition affects the first finding. All three are observational questions, not causal regime comparisons.

The original Stata log is supplemented by two independently inspected Stata datasets and a model-tracking workbook. This addition permits exact-data reproduction, robust covariance tests, leave-one-year-out checks, and a historically motivated coding sensitivity that were not verified in the previous manuscript. Independent replication is reported separately from the archived Bai–Perron-type critical values and break date intervals.

A chronological partition can be useful for identifying historically meaningful episodes in a macroeconomic record. It permits conditional comparison across intervals shaped by distinct administrative configurations, reconstruction programs and external financing conditions. Yet a period indicator is simultaneously a proxy for all changes that occurred during its assigned years. Changes in conflict intensity, foreign transfers, private investment and the coverage of national accounts are not held fixed merely by including a handful of regressors. The most defensible objective is therefore to assess conditional differences in recorded growth while explicitly acknowledging the unobserved historical bundle associated with each period.

A further complication is that historical and statistical boundaries answer different questions. A political classification is chosen using external historical information. An estimated break date is selected because a statistical objective, such as the fit of a piece wise constant mean, changes at that date. Neither procedure is a substitute for the other. Comparing the two can show whether conventional period labels line up with measured series instability, but an apparent match does not prove a political mechanism. When break dates differ across taxation, governance and GDP growth, the mismatch is a substantive warning against assuming synchronous adjustment of economic and administrative systems (Bai & Perron, 2003).

The article advances a deliberately bounded contribution: it demonstrates the extent to which a conventionally significant joint historical-period test depends on uncertainty estimators and classification choices, while showing why changes in univariate means cannot validate political regime effects. Its new calculations are transparent, reproducible and distinguished from previously reported ARDL–ECM results (Kareemzada et al., 2026).

2. Literature review and conceptual framework

2.1. Political authority, institutions and growth

Institutional economics links growth to enforceable rules, administrative capacity and credible commitments. North (1990) emphasizes how institutions shape exchange and transaction costs; Acemoglu et al. (2001) show why identifying their contribution requires careful empirical design. Neither result turns a historical label into an exogenous treatment in a 29 observation national time series.

Political transitions may change the administration of taxation, cross-border trade, investment expectations and access to external finance. Their observed economic associations depend on inherited conditions, conflict and the timing of assistance or trade disruptions. Afghanistan-specific analyses of revenue institutions and aid-financed statebuilding (Rafi, 2024) reinforce the importance of these overlapping channels. They provide plausible mechanisms, not separately identified coefficients for each period in the present model.

Institutional change can affect economic activity through expectations, enforcement and administrative organization, even when a government does not implement a discrete tax or investment reform. North (1990) views institutions as the constraints governing exchange; Besley and Persson (2009) explain how political incentives affect investments in legal and fiscal capacity. Their theories do not predict that each political transition should coincide with a break in the annual growth mean. Effects can be mediated by administrative continuity, adjustment lags and simultaneous shocks. In country year regressions, the available variation is especially limited because institutional proxies change slowly and may reflect perceptions rather than measurable reforms.

Afghanistan-specific historical work links reconstruction and revenue dependence to external financing and state building choices. Rafi (2024) argues that post-war fiscal arrangements can produce dependence on external assistance, complicating any inference from domestic revenue trajectories. Such work supplies plausible historical mechanisms, not estimated instruments for assigning periods independently of growth. The present study accordingly treats historical chronology as contextual and predetermined, while restricting quantitative claims to the statistical objects actually observed.

2.2. Conflict, external finance and data uncertainty

Conflict affects output through displacement, capital losses, investment uncertainty and disrupted services. Hameed et al. (2022) examine asymmetric conflict growth relationships in Afghanistan and motivate attention to nonlinear adjustment. Saenger et al. (2024) employ satellite nightlights to study the post-2021 economy, demonstrating that assessments can depend on which activity indicator remains observable. Nightlights are not used to replace or reconstruct observations in the present GDP growth regression.

External finance offers another competing explanation for measured growth differences. Assistance and reconstruction spending may sustain demand and imports when domestic revenue is limited; subsequent reductions may depress activity regardless of domestic administrative changes. Trade restrictions, payment system disruptions and changing statistical coverage may operate at the same time. Because reliable annual controls for all these channels are not included, the period indicators remain composite descriptors rather than estimates of discrete transition shocks.

Hameed et al. (2022) use nonlinear time series methods to study the conflict growth nexus, they do not estimate the five period specification presented here. Saenger et al. (2024) instead use nightlights and a synthetic control comparison to analyze post-2021 output disruption. These studies motivate concern about concurrent conflict and changes in observed activity but cannot serve as identifying instruments or diagnostic tests for the current OLS estimates.

Barro (1990) clarifies the theoretical trade off between productive government expenditure and tax-financing distortions, underscoring why the direction of a tax growth relationship is not predetermined. Domestic revenue, imports and output are jointly determined. Import growth may increase customs receipts; an output contraction can mechanically change GDP-denominated tax and trade ratios; and expenditure financed externally may support demand without a commensurate change in domestic taxation. If

these simultaneous developments are incompletely recorded, a political period coefficient absorbs their combined conditional association. Measurement caveats are therefore central to, rather than ancillary to, fiscal and institutional interpretation (Jerven, 2013; Akitoby et al., 2020).

2.3. Institutional quality and moderation

Institutional quality may condition how political authority and fiscal policy relate to economic performance. Administrative effectiveness, regulatory predictability and legal enforcement can affect the use of public resources and investment decisions. In a regression, moderation is represented by an interaction term: the estimated difference between political periods varies with the observed institutional quality score (Brambor et al., 2006). It is distinct from mediation and does not identify an indirect causal mechanism.

The five component Worldwide Governance Indicators (WGI) index used here varies within a narrow, entirely negative observed range. Its values are governance estimates rather than percentile ranks or 0–100 governance scores. The literature motivates testing interactions, but variation and covariance information must support them. The corrected source record does not contain a fully reconciled categorical D2–D5 × centred-IQ coefficient and covariance table; consequently, such moderation cannot be claimed as an empirical finding.

The interaction-model literature distinguishes between estimating a constituent slope and establishing that slopes differ across groups. Brambor et al. (2006) show why conditional effects require explicit marginal-effect formulas and the corresponding coefficient covariance. For this study, a statement that institutional quality moderated the political period association would require a consistently coded interaction model, adequate within-period support and interpretable uncertainty. The archived record does not satisfy those requirements. Consequently, the question remains a well-motivated extension, not a result that can be claimed from the joint period F test.

Kaufmann et al. (2010) emphasize that WGI observations are governance estimates built from multiple source inputs rather than direct administrative audit outcomes. The five component index also compresses different administrative constructs into one average. Perceptions of corruption control, administrative effectiveness, regulatory quality and rule enforcement may evolve differently and may be influenced by the composition of data sources. Interpreting the composite as a measured policy intervention is unwarranted. The 2026 World Bank WGI release recalculated historical estimates, making the distinction between a frozen archival vintage and a newly downloaded governance series essential to reproducibility (World Bank, 2026).

2.4. Structural change and historical periodization

Three inferential distinctions define the analysis. First, D2–D5 coefficients represent conditional annual growth differences relative to D1, not causal transition effects. Unequal period lengths, exceptional observations, and coincident macroeconomic changes constrain interpretation.

Second, historical political periods are predetermined classifications, whereas Bai–Perron-type breaks are estimated separately from economic series under specified trimming and model restrictions. Differences between growth, revenue, and IQ break dates neither invalidate the method nor establish delayed political effects (Bai & Perron, 2003).

Third, additive period models and governance-interaction models address different hypotheses. Joint significance of D2–D5 does not establish heterogeneous IQ slopes, and six-category coefficients cannot be transferred to the five-period specification. Valid marginal effects require consistent coding, observed IQ support, an explicit reference category, and the complete covariance matrix (Brambor et al., 2006).

Accordingly, the study evaluates the joint explanatory contribution of historical periods, compares their boundaries with independently estimated economic breaks, and identifies the limits of institutional-moderation inference. Its contribution is a transparent, source-conscious assessment of political period associations rather than causal attribution or unverified robustness claims.

A Bai–Perron-type procedure permits the dates of specified mean or regression-parameter shifts to be estimated from observations, subject to minimum segment length and assumptions about disturbances (Bai & Perron, 2003). The reported exercises here concern a constant mean within each segment of a single series. They are not tests of changes in all coefficients of the political period OLS model. Neither are they break-adjusted cointegration tests. The difference is conceptually important because a shift in the tax-revenue mean need not alter the slope relating GDP growth to taxes, and a shift in GDP-growth volatility need not imply a change in its mean.

The selected historical periods differ greatly in length. A three year transitional category and a four-year post-2021 category offer limited information compared with a ten year middle period. Moreover, all four comparisons share the same six-year reference group, which includes early observations flagged for uncertain FDI and trade coverage. This creates a common vulnerability: decisions about the earliest data and the choice of omitted period can shape the interpretation of all reported contrasts, even when the regression is calculated correctly.

2.5. Political-Period Classification and Observed Economic Growth

This study distinguishes historically defined political periods from the economic and statistical processes underlying observed differences in GDP growth. Institutional economics suggests that political and administrative changes may influence fiscal capacity, rule enforcement, investment expectations, and economic coordination (North, 1990; Besley & Persson, 2009). In Afghanistan, these mechanisms coincide with variations in conflict, external assistance, trade access, migration, and statistical coverage. Political-period indicators therefore capture historically overlapping conditions rather than independently identifiable institutional interventions.

The empirical framework examines annual real GDP growth across five historical periods (D1–D5) during 1996–2024, with D1 (1996–2001) as the reference category. The additive OLS specification estimates conditional growth differences for D2–D5 relative to D1, controlling for tax revenue, institutional quality, foreign direct investment, trade openness, and population growth. Because these covariates may themselves respond to contemporaneous economic and political conditions, the period coefficients represent conditional associations, not causal transition effects. Their joint significance test evaluates whether the period coefficients are jointly zero within the specified model.

Institutional quality enters the additive specification through an author-constructed, five-component governance index. Although institutional conditions may theoretically modify period-specific growth associations, establishing such moderation requires explicit period-by-institutional-quality interactions and covariance-based conditional-effect inference (Brambor et al., 2006). The available evidence does not provide a fully reconciled categorical interaction model. Accordingly, the analysis does not interpret additive period differences as evidence of institutional moderation.

The framework further distinguishes historical periodization from statistically estimated structural change. Political boundaries are assigned chronologically, whereas univariate mean-break dates are identified from observed series under specified statistical assumptions (Bai & Perron, 2003). Such breaks indicate potential changes in individual series means, not instability in the multivariate growth coefficients or causal effects of political transitions. The analysis therefore evaluates historical-period contrasts and series-specific mean shifts as separate forms of evidence. Alternative coding of 2021 and covariance-sensitive joint tests assess the robustness of the descriptive findings without establishing causal identification.

2.6. Research Expectations

H1 evaluates the joint null hypothesis $\delta_2 = \delta_3 = \delta_4 = \delta_5 = 0$ in the historically defined five-period growth model. Because the annual sample is small (N = 29) and leverage may be substantial, inference is assessed using conventional OLS covariance and, separately, HC1 and HC3 heteroskedasticity-consistent covariance estimators (MacKinnon & White, 1985; Long & Ervin, 2000). A one-lag HAC result is reported only as an additional, tuning-dependent robustness check.

H2 asks whether estimated mean-break dates align with the historically defined political boundaries. No formal causal null is attached to simple calendar alignment; the purpose is descriptive comparison rather than causal identification.

H3 examines a transparent alternative coding in which 2021 is assigned to the preceding period, reflecting the fact that the transition occurred partway through that year. All of these checks are treated as exploratory sensitivity analyses, not as preregistered confirmatory tests.

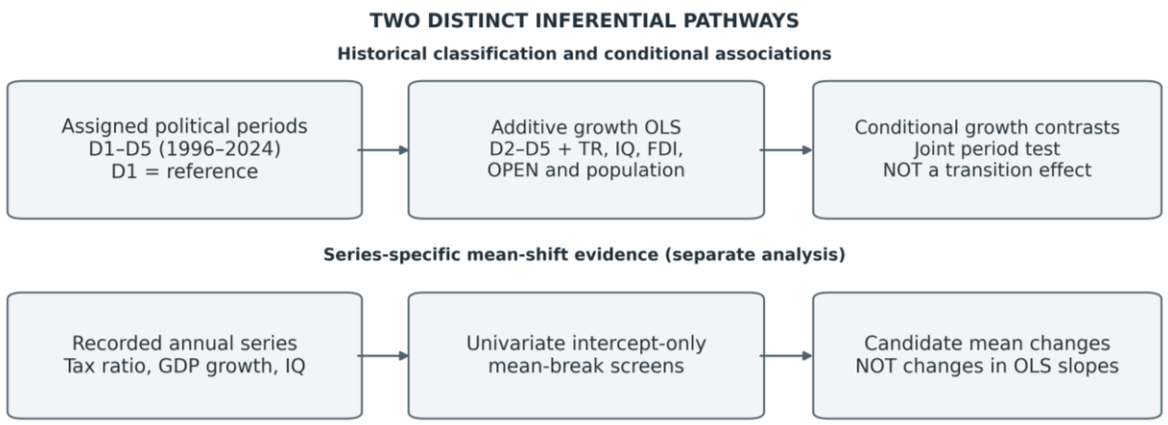


Figure 1. Historical Political Categories, Conditional Period Contrasts, and Separate Series-Mean Screens

3. Data and empirical strategy

3.1. Annual archive, units and vintage discipline

The analysis uses two author-supplied Stata files with 29 annual Afghanistan records, 1996–2024. Independent reading reveals substantively equal outcome and explanatory data across the files, with differences only in floating point rounding, year format and capitalization of unused generated column names. The newer dataset file is the controlling reproduction input. The reported regression statistics were cross-checked against the supplied analysis files, but its upstream source vintage and individual administrative adjustments are not documented in the data files. Real GDP growth is the annual percentage rate, not a welfare or sustainability outcome.

Institutional quality (IQ) is the arithmetic mean of five WGI governance estimates: Control of Corruption, Rule of Law, Government Effectiveness, Voice and Accountability, and Regulatory Quality. Political Stability and Absence of Violence/Terrorism is excluded. Independent row wise checking of all 29 observations confirms the existing iq column agrees with the five component average to within approximately 5×10^{-8} index units. The index is author-constructed and is not an official World Bank composite. Its exact historical source vintage remains unspecified, and it must not be replaced with a 2026 reissued series without re estimation (World Bank, 2026).

$$IQ_t = (CC_t + RL_t + GE_t + VA_t + RQ_t) / 5 \quad (1)$$

The Political Stability and Absence of Violence/Terrorism dimension is excluded by the documented intended formula. IQ uses the WGI governance-estimate scale, not its newer absolute 0–100 scale. The archived WGI release is not dated at the observation level in the log. The World Bank (2026) reports recalculation of historical WGI estimates; a newer series must therefore not be substituted while preserving old coefficients. This is an unresolved reproducibility condition, not evidence that the IQ values are erroneous.

Source inspection reveals four FDI values equal to zero for 1996–1999 and six trade-openness zeros for 1996–2001. Because the input files do not include an observation-level code book distinguishing measured zero, estimated proxy and missing-value placeholder, these zeros are retained exactly for the main reproduction. They cluster in the D1 reference group and may therefore matter substantially to comparisons with later periods. Treating them as missing would eliminate most or all usable D1 observations, changing the estimand rather than providing a clean alternative estimate. Genuine source documentation is required before replacing them.

3.2. Pre-specified historical coding

Five chronological categories are independently reconstructed from calendar years (Table 1 and Figure 1). D1 (1996–2001) is the comparison category; D2 (2002–2004), D3 (2005–2014), D4 (2015–2020) and D5 (2021–2024) represent analytical groupings. The original .dta files also contain six historical indicators, d1–d6, which do not correspond to the five analytical categories. Using those raw dummies to reproduce P01 would estimate a different model. The 2021 calendar year is heterogeneous because Afghanistan's change in authority occurred in August; an alternative boundary starting D5 in 2022 is therefore examined.

Table 1

Five non-overlapping political periods in the annual archive (source defined classification)

Code	Period label	Years	n	Analytical role
D1	Taliban I	1996–2001	6	Reference
D2	Transitional Authority	2002–2004	3	Contrast
D3	Karzai administration	2005–2014	10	Contrast
D4	Ghani administration	2015–2020	6	Contrast
D5	Taliban II	2021–2024	4	Exploratory contrast

Source. Independently reconstructed five-period grouping, n = 6, 3, 10, 6 and 4. The archived six-indicator coding uses different year mappings and is not interchangeable with the five-category classification.

Each observation enters one category only. The coefficients for D2–D5 compare each period with D1 after adjusting for the five recorded macroeconomic controls; they are not comparisons with the preceding period. The sparse D2 and D5 groups have three and four years respectively. A supplementary recoding shifts the D4/D5 boundary from January 2021 to January 2022 without changing the data rows, controls or number of fitted parameters.

3.3. Static model, estimand and conventional inference

Advance Journal of Econometrics and Finance
Vol-4, Issue-1, 2026

$g_t = \alpha + \beta_1 TR_t + \beta_2 IQ_t + \beta_3 FDI_t + \beta_4 TRADE_t + \beta_5 POPGR_t + \sum_{k=2}^{5} \delta_k D_{kt} + u_t$ (2)

Here TR and TRADE are percentages of GDP, FDI is net inflows as a percentage of GDP, POPGR is annual population growth and IQ is the five component index. The joint null is H_0: $\delta_2 = \delta_3 = \delta_4 = \delta_5 = 0$. The OLS specification is reproduced with an intercept and four nonreference-period indicators. With 29 observations, nine slope parameters and an intercept, the error degrees of freedom equal 19. Conventional confidence intervals use t_{19} reference values. The model includes no temporal lag, is not an ARDL, and does not estimate causal interventions.

Inference is explicitly examined under four covariance conventions: conventional OLS, HC1, HC3 (which adjusts for individual leverage and is particularly relevant with this sample size), and a one-lag Bartlett-kernel HAC calculation. HC and HAC changes affect the covariance estimate without changing the fitted coefficient vector; the HAC tuning parameter is not optimized with N = 29. Separately, each country-year is deleted in turn and the joint F test is recalculated under conventional OLS; this influence screen is not an objective rule for dropping observations. A second specification assigns 2021 to D4 and begins D5 in 2022. These are exploratory robustness contrasts, not instruments for regime effects (Long & Ervin, 2000; MacKinnon & White, 1985).

3.4. Mean-break screens and evidentiary separation

$x_t = \mu_j + \varepsilon_t, \quad \tau_{j-1} < t \leq \tau_j, \quad j = 1, \ldots, m + 1$ (3)

The historical log reports break constant trimming (0.15), each with N = 29. Their tests concern univariate, piecewise-constant means, not regression slope breaks. As a computational cross-check, distinct from the archived xtbreak critical values, the annual series is partitioned into candidate contiguous segments containing at least five observations each; for fixed numbers of breaks, the sum of within-segment squared deviations is exhaustively minimized. This reproduces tax cut-years 2001 and 2008 for two estimated breaks and GDP growth 2019 for one break. Break-number inference and confidence intervals remain those in the original Stata log (Bai & Perron, 2003; Ditzen et al., 2025).

The dating exercise maps mean changes against the preassigned political intervals, rather than selecting periods to match a desired narrative. The 2001 tax break lies at the first grouping boundary, the 2008 tax break lies within D3, and the weak 2019 GDP-growth candidate lies within D4. The governance test reports irregular sequential choices and cannot establish a unique, precise institutional break chronology. Location of a univariate mean change does not identify the historical cause of that change.

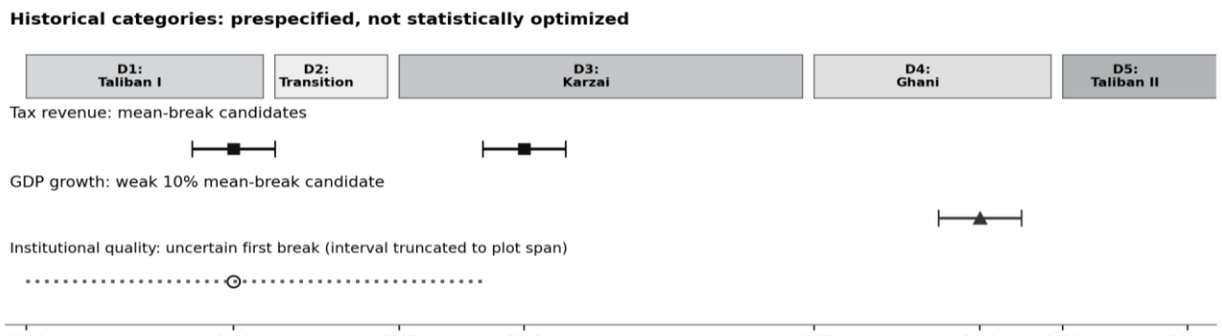


Figure 2: Historical categories and archived series-specific mean-break candidates

Source. Boxes show prespecified periods; markers show univariate mean-break dates, not political effects. GDP 2019 is supported only at 10%; IQ timing is uncertain.

3.5. Prospective governance moderation is not reported as a result

$g_t = \alpha + \gamma' X_t + \eta q_t + \sum_{k=2}^{5} [\delta_k D_{kt} + \kappa_k D_{kt} q_t] + u_t, \quad q_t = IQ_t - \text{mean}(IQ)$ (4)

Equation (4) specifies a possible extension, with X containing the recorded fiscal and macroeconomic controls. The conditional contrast for period k at governance score q is $\delta_k + \kappa_k q$, and its variance requires $\text{var}(\delta_k)$, $\text{var}(\kappa_k)$ and their covariance. The corrected archive does not provide a matched five-period interaction coefficient/covariance table. A separate six-category exploratory regression uses a different reference definition and cannot be relabeled as this model. No categorical marginal effect band or claim of confirmed period specific moderation follows from Equation (4).

3.6. Log verification, source provenance and model boundaries

The full historical Stata log contains the five-period command, 29-observation coefficient table, period-restriction F test, diagnostics. The reported model coefficients, joint restriction and diagnostics are cross-checked against the archived output. Distinct older six-indicator and four category regressions are retained only as model-history evidence, not merged with the five-period model. The original tax growth ARDL(1,0,0,1,0), 1997–2020 (N = 24), contains no IQ term and is used as contextual evidence rather than a newly estimated result.

4. Results

4.1. Descriptive characteristics and cross-variable dependence

Independent data level replication confirms that real GDP growth averages 4.1417% (SD = 9.1791), ranging from −20.7388% to 28.6000%. The recorded tax-to-GDP ratio averages 6.7888% (SD = 2.1652), and IQ averages −1.5312 (SD = 0.1802). Those are descriptive moments of the supplied harmonized data not direct independent confirmation of their original international or Afghan administrative sources.

The pairwise correlations also indicate shared movements across several regressors: tax revenue with trade openness (r = 0.8717), IQ with trade openness (r = 0.8308), and tax revenue with IQ (r = 0.6052). The simple contemporaneous correlation between growth and tax revenue is close to zero (r = −0.0002; p = 0.9993). None of these pairwise associations isolates a structural transmission mechanism.

Table 2

Archived annual variables, 1996–2024 (descriptive summary M00)

Variable	n	Mean	SD	Range
GDP growth (%)	29	4.1417	9.1791	−20.7388 to 28.6000
Tax revenue (% GDP)	29	6.7888	2.1652	3.2000 to 9.8985
IQ (WGI estimate)	29	−1.5312	0.1802	−1.8676 to −1.2810
FDI (% GDP)	29	0.7593	1.1598	0.0000 to 4.3687
Trade openness (% GDP)	29	45.7445	25.3785	0.0000 to 73.3000

Advance Journal of Econometrics and Finance
Vol-4, Issue-1, 2026

Variable	n	Mean	SD	Range
Population growth (%)	29	3.1582	1.1245	0.7620 to 6.1452

Source. Descriptive calculation M00, N = 29, 1996–2024.

In this sample, the standard deviation of annual growth is more than twice its arithmetic mean. That dispersion creates a demanding setting for segment comparisons: a single extreme growth year may be influential in a category with only three or four years. The ratio based covariates are likewise strongly associated, with r(TR, TRADE) = .8717. Both share GDP in the denominator, so the correlation alone cannot establish a trade-tax transmission path. The nearly zero full sample growth tax correlation belongs to a different estimand from the multivariate period coefficients and should not be taken as a replication failure.

4.2. Conditional political period contrasts

The exact-data reproduction matches P01, including overall F(9, 19) = 3.3892 (p = .0120), R^2 = .6162 and adjusted R^2 = .4344. The four-period conventional restriction gives F(4, 19) = 3.5153 (p = .02613), matching the printed Stata F = 3.52, p = .0261. This result describes additional in-sample explanatory fit under one variance assumption, not a causal political effect. The older raw six-dummy regression corresponds to a different partition and returns F(5, 18) = 1.41 (p = .2680), so its p-value is neither an estimate from P01 nor a test that the two partitions are equal.

Individual conventional period estimates are D2 = 7.8960 (SE 13.5151, p = .566), D3 = −1.7866 (SE 14.3906, p = .903), D4 = −6.8288 (SE 16.6527, p = .686) and D5 = −21.2914 (SE 16.0976, p = .202). Every corresponding conventional 95% interval includes zero, and the HC3 intervals are wider. Differences between joint and individual tests arise because their null hypotheses differ; they do not establish relative performance of governments. All estimates are conditional and potentially confounded by time trends and shocks.

The IQ coefficient is negative in this particular conditional regression (−54.2263; SE = 19.1869; p = 0.011). Its sign must not be interpreted as proof that improved governance lowers growth. The index has limited within-sample range and correlates with trade openness; endogeneity, shared trends and measurement features remain plausible explanations. Likewise, the trade coefficient (0.5071; p = 0.085) does not meet a 5% threshold.

Table 3

Full-sample conditional OLS model of real GDP growth (P01)

Regressor	b	SE	p	95% CI†
Tax revenue	-0.2123	1.6548	0.899	[-3.68, 3.25]
Institutional quality	-54.2263	19.1869	0.011	[-94.38, -14.07]
FDI	0.9465	1.5783	0.556	[-2.36, 4.25]
Trade openness	0.5071	0.2787	0.085	[-0.08, 1.09]
Population growth	1.5119	1.6135	0.360	[-1.87, 4.89]
D2 (2002–04)	7.8960	13.5151	0.566	[-20.39, 36.18]
D3 (2005–14)	-1.7866	14.3906	0.903	[-31.91, 28.33]
D4 (2015–20)	-6.8288	16.6527	0.686	[-41.68, 28.03]
D5 (2021–24)	-21.2914	16.0976	0.202	[-54.98, 12.40]
Intercept	-101.9931	34.4041	0.008	[-174.00, -29.98]

Source. Model P01: OLS with intercept, TR, IQ, FDI, TRADE, POPGR and D2–D5, 1996–2024, N = 29,

The importance of uncertainty becomes clearer with small-sample leverage adjustment. For D5, the classical estimate is negative but imprecise; HC3 yields a substantially wider interval, and its p-value remains above .05. Thus neither the conventional nor leverage-adjusted estimate identifies a period-specific contraction. The substantive question is not which period coefficient has the largest magnitude but whether these coefficients remain informative after allowing for variance misspecification and influential observations.

The IQ coefficient is −54.2263 (classical SE 19.1869, p = .011) within P01. Under HC3 its p value rises to .0514. This illustrates model-specific instability rather than proof that governance improvements reduce growth. The five WGI measures are perception-based, and the index is correlated with economic variables and historical periods. Neither the association nor its alternate p values resolves reverse causation or omitted contemporaneous conditions.

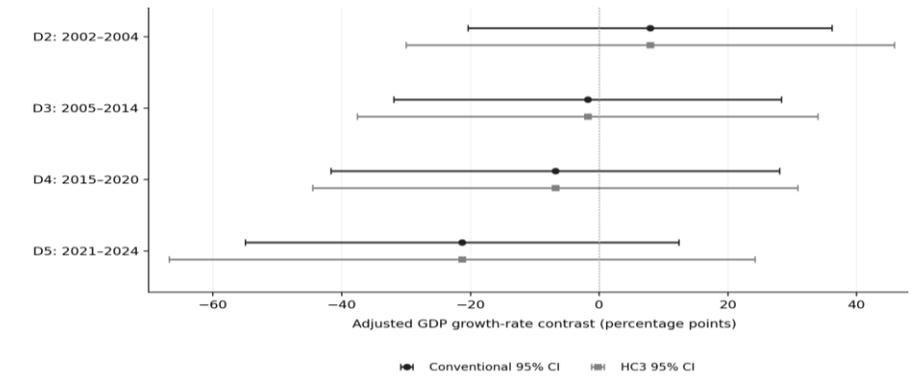


Figure 3: Individual period contrasts: conventional and HC3 confidence intervals

Note. P01, Afghanistan annual data 1996–2024, N = 29, OLS coefficients and two t_{19}-based 95% covariance-specific intervals. Black circles show conventional confidence intervals (archived A, independently reproduced C); gray squares show HC3 confidence intervals (new independent C). The chart describes uncertainty in historical period indicators relative to D1, not causal regime effects.

The institutional quality main-effect estimate in P01 also depends on the covariance convention. Its conventional slope is −54.2263 (SE = 19.1869; p = .0108; 95% CI [−94.3849, −14.0677]), but the HC3 estimate has SE = 26.0774, p = .0514, and 95% CI [−108.8068, 0.3543]. This is a sensitivity of the conditional main-effect estimate, not evidence that governance moderates political period effects or that a governance intervention causes growth. The governance scale and shared trends further limit structural interpretation.

4.3. Independent covariance, influence, and calendar-boundary sensitivity

The reproduced P01 model exhibits covariance-sensitive joint inference (Table 4). Under the conventional OLS variance assumption, p = .0261; using HC1 gives p = .1203 and HC3 gives p = .3004. The one-lag HAC p-value is .0293, but its short-sample tuning cannot be treated as decisive. Figure 3 shows substantially wider individual HC3 confidence intervals. The effect of the 2021 boundary and the leave-one-year-out checks is reported separately from the archived model, since those calculations change the sample or categories.

Table 4

Joint period restrictions under alternative covariance and coding assumptions

Advance Journal of Econometrics and Finance
Vol-4, Issue-1, 2026

Model ID	Years / N	Estimator / covariance	Joint F	df	p
P01 – conventional	1996–2024; N = 29	OLS, classical	3.5153	4, 19	.0261
P01 – HC1	1996–2024; N = 29	OLS, HC1	2.1031	4, 19	.1203
P01 – HC3	1996–2024; N = 29	OLS, HC3	1.3143	4, 19	.3004
P01 – HAC(1)	1996–2024; N = 29	OLS, HAC lag 1	3.4029	4, 19	.0293
P01–L2021	1996–2024, excluding 2021; N = 28	OLS, classical	2.1324	4, 18	.1186
P01–B2022	1996–2024; N = 29	OLS, classical	1.5552	4, 19	.2268
P01–B2022–HC3	1996–2024; N = 29	OLS, HC3	1.0885	4, 19	.3903

Source. The joint null is D2 = D3 = D4 = D5 = 0. No coefficient-level SE or CI belongs to a four-restriction F test; classical coefficient SEs and CIs appear in Table 3 and HC3 comparisons in Figure 3.

Influence diagnostics reveal Cook's distances of 0.4756 (2004), 0.3199 (2021), and 0.2920 (2002), compared with the conventional heuristic 4/29 ≈ 0.138. The joint classical p-value is below .05 in 26 of 29 leave-one-year-out fits, but excluding 2021 increases it to .1186. This is evidence of sensitivity, not a reason to delete 2021. Reassigning 2021 to the 2015–2021 group and using 2022–2024 for D5 changes the classical p to .2268 (HC3 p = .3903). These exploratory alternatives reinforce the dependence of conventional inference on a short calendar-year partition.

4.4. Archived residual tests and specification limitations

Table 5

Residual and specification checks for full-sample political period model P01

Diagnostic	Test statistic	p	Interpretive status
Breusch–Godfrey, lag 1	$\chi^2 = 3.518$	.0607	Near 10% rejection
Breusch–Godfrey, lag 2	$\chi^2 = 5.022$	.0812	Near 10% rejection
Breusch–Pagan	$\chi^2 = 2.87$	.0904	Near 10% rejection
Ramsey RESET	$F(3,16) = 2.84$	.0706	Near 10% rejection
Skewness–kurtosis	$\chi^2(2) = 0.66$	.7193	Not rejected
Shapiro–Wilk	W = 0.97128	.5949	Not rejected

Source. Model P01, static OLS (lag 0), 1996–2024, N = 29.

Several Stata residual diagnostic probabilities are between .06 and .10, and all must be interpreted with low sample power. Breusch–Godfrey serial-correlation tests yield p = .0607 (lag 1) and p = .0812 (lag 2); the original fitted-value Breusch–Pagan and RESET procedures likewise approach 10% significance. This ambiguity supports the additional variance and leverage sensitivity tests, but the HC3 calculation does not correct omitted variables, measurement problems, persistence, or causal confounding. No alternative-model diagnostic result is silently borrowed from P01.

The log's level unit-root screens also disagree for growth (ADF p = .0829; PP p = .0001), and its revenue, trade and IQ level results do not reject unit roots at 5%. These screens, possible breaks and small samples make conventional static inference fragile; regression on a growth rate and persistent level regressors is not automatically free of spurious association. A fully redesigned dynamic model should be pre-specified and distinguished from these replication tests, not reported here without estimation.

4.5. Independent mean-break calculations and historical crosswalk

Archived xtbreak results identify revenue mean-break dates 2001 and 2008 under the reported criteria, a GDP growth candidate near 2019 only at the 10% criterion, and an unsettled governance break count. A separate exhaustive within-segment least-squares calculation using five-year minimum segments reproduces the 2001/2008 tax cut-years with residual sum of squares 21.5540 (two breaks) and the GDP 2019 cut-year with residual sum of squares 1870.4083 (one break). Those new minima confirm the partition arithmetic, not the Bai–Perron critical values, asymptotic test distribution or intervals.

Table 6

Reported univariate mean-break screens and comparability to the political periods

Series	UDmax	Reported date/interval	Evidentiary status
Real GDP growth	7.62	2019 [2018, 2020]	10% candidate; not 5%; interior to D4
Tax revenue	67.74	2001 [2000, 2002]; 2008 [2007, 2009]	Two stronger shifts; 2001 near D1/D2, 2008 within D3
Institutional quality	74.98	First 2001 [1995, 2007]	Count 1–4 uncertain; interval extends outside dataset

Source. B01 growth, B02 tax revenue and B03 IQ are univariate intercept-break screens from the original Stata log, N = 29, 1996–2024, trimming = .15, no external regressors or lags. The displayed UDmax and reported break-date intervals are original-log evidence (A), whereas the separate exhaustive least-squares date matching is new computed evidence (C). No p-value, SE or CI for a regression slope is implied. The GDP candidate does not pass the 5% criterion; governance sequential break counts are nonmonotonic. See Bai and Perron (2003) and Ditzen et al. (2025).

The alignment map shows no automatic correspondence between political and statistical breaks. The 2001 revenue mean shift is near the start of D2; 2008 is inside D3, not at a category boundary. The weaker 2019 output-growth candidate falls inside D4, preceding the 2021 annual recoding. Independent least-squares segmentation matches these archival locations but does not establish that any elected or unelected authority caused the change. Fiscal definitions, denominator movements, aid and shocks remain potential alternative explanations.

Because the first governance interval reportedly extends outside the sample and the selected number of breaks is not stable, it cannot be represented as a precision estimate of when institutional performance changed. It is safer to report the available information exactly, flag the inconsistency, and require the original command, settings and standard-error assumptions for re-estimation. Equally, tax mean breaks need not measure tax reforms: administrative coverage, national accounts revisions and shifts in the ratio denominator can affect the recorded series.

4.6. Institutional moderation and unestimated causal pathways

The log includes a six-dummy interaction design based on d2–d6, not the correctly reconstructed five period variable regime5. That exploratory regression has 13 residual degrees of freedom and different coefficient meanings. It cannot supply a valid five-category IQ-by-period interaction table. A separate linear ordered period index by IQ regression imposes

an arbitrary numerical ordering and also does not establish moderation. P01 includes IQ only as a main effect control. The IQ main-effect coefficient is not evidence of period-specific moderation, irrespective of its covariance-dependent significance.

The distinct tax-revenue × IQ interaction is also statistically imprecise (−2.1304; SE = 5.5997; p = 0.707). It tests a tax-related slope hypothesis rather than the historical period interactions. No governance mediation, threshold, or period specific marginal effect plot is therefore reported. Valid categorical moderation would require a documented re-estimation, the observed IQ support by period, and covariance based marginal effect intervals.

An additional model reported in the archive imposes an ordered numerical index on political periods and interacts that artificial scale with centered IQ. Its interaction is not statistically distinguishable from zero and does not reproduce the categorical hypothesis tested here. Giving five politically distinctive periods equally spaced numerical values assumes a constant marginal transition distance without an empirical basis. In consequence, neither this exploratory scalar interaction nor the incompatible six-category interaction provides validated marginal effects for the four D2–D5 periods. The manuscript does not draw an unsupported Johnson–Neyman region or pronounce a threshold.

5. Discussion

5.1. Political contrasts are composite historical associations

Reproducing the baseline does not establish a historically causal growth ranking. The original joint period test rejects under conventional OLS covariance but not under HC1 or HC3; all individual period intervals include zero. The sensitivity is statistically intelligible with a small, highly adjusted model and uneven 3- and 4-year categories. D5 also starts in January 2021 despite the shift in political authority occurring in August, which makes a calendar-boundary sensitivity substantively warranted. The finding is about conditional statistical specification, not relative government competence.

The negative estimated IQ association is similarly model-specific. A contemporaneous growth regression with non-stationarity or persistent macroeconomic regressors does not resolve whether governance affects growth, growth influences governance assessments, or both reflect a third shock. Institutional theory remains a useful basis for future hypotheses, but its plausibility cannot substitute for identification.

The result depends on the research design used to represent political chronology. In 29 observations, a five-period categorical regression with five controls and an intercept leaves 19 error degrees of freedom. There is substantial collinearity-particularly with trade openness and periods that co-move over time and highly leveraged years. The HC3 interval expansion and the 2021 deletion show that one marginal p-value should not be mistaken for structural evidence. Alternative governance-vintage values or corrected zero codes would require separate, documented recalculation.

The estimated 2021–2024 contrast deserves additional caution. The period is historically distinct, but its four observations cannot separately identify financial isolation, changes in international assistance, household demand, drought, statistical discontinuity and political decision-making. Satellite-based studies provide an independent view of economic activity after 2021 (Saenger et al., 2024), yet they differ in outcome construction and counterfactual design. Their estimates should not be imported to validate the conventional period OLS coefficient, particularly when GDP growth and nightlight-derived activity are not interchangeable.

5.2. Why non-coincident breaks are the central contribution

The revenue series exhibits reported mean breaks near 2001 and 2008, whereas the GDP-growth break candidate is near 2019 and weakly supported. That difference is substantively informative even without causal attribution: fiscal records, administrative systems and aggregate output need not change on an identical timetable. Their measurement conventions may also change unevenly. A Bai–Perron-type break test identifies an estimated statistical partition under its assumptions; it does not validate the historical explanation for a date (Bai & Perron, 2003).

For fragile-state applications, keeping the historical event map distinct from the series-specific statistical break map is a methodological contribution. It prevents alignment chosen after observing outcomes from being reported as an independently confirmed political effect. Equally, a failure to find a common break chronology does not imply that transitions are economically irrelevant.

The main scholarly value of placing break dates beside historical categories is the refusal to assume that separate variables react on the same schedule. Revenue administration can change through revisions to border collection or fiscal reporting; real activity can adjust through investment, security and imported inputs; governance estimates may respond to perceptions and international information flows. The timing of these mechanisms may differ, and each observed proxy may be revised retrospectively. The divergence between 2008 revenue and 2019 growth dates in the archived record is consistent with asynchronous responses or measurement histories, but it cannot discriminate between them.

The means-break matching is independently reproducible from the shared dataset, strengthening numerical traceability but not adding a break-adjusted regression. Newly independently computed covariance and calendar-boundary checks are a more substantive extension to the earlier manuscript: they change what can be concluded from the historical-period F test. The two publication objects remain sharply distinct from the previously published ARDL–ECM tax growth estimates.

5.3. Model limitations, influence and plausible competing accounts

The new influence checks identify the years 2004 and 2021 among the largest Cook's distances. Although the conventional joint test rejects in 26 of 29 leave-one-year-out fits, its p-value becomes .1186 when 2021 is removed. This asymmetry is consistent with a narrow model whose conditional contrasts are sensitive to particular years and residual leverage. No observation was removed from the reported baseline, and observations were never selected for exclusion based on desired significance. Early FDI and trade zeros remain source-audit priorities, especially because most occur in D1.

A second model checks the historical calendar convention without dropping observations: 2021 is assigned to D4 while D5 begins with 2022. The conventional joint p-value rises to .2268 and its HC3 counterpart is .3903. This is not a political-effect estimate, because the 2021 annual growth outcome blends pre- and post-August months as well as other shocks. It demonstrates that ordinary period labels are consequential model decisions. Full causal or break-adjusted inference would require independent event-timing data, a counterfactual framework and better control of external assistance, conflict and measurement revisions.

5.4. Distinctness from the January 2026 publication

The Stata files are substantively equal apart from rounding and representation; the data level five-period OLS is independently reproduced. The author-supplied workbook includes a procedure checklist and source-data copies rather than year-level WDI/official vintage documentation. Consequently, origin and fiscal-to-calendar alignment of historical measurements remain unverified. Numeric reproducibility of the supplied file is not the same as authentication of its original administrative records.

This study is not a second estimate of the primary tax growth ARDL–ECM relationship in Kareemzada et al. (2026). Its distinctive empirical objects are the joint political period contrast, the uncertainty of individual comparisons, and the independently estimated mean-break chronology. Disclosure of the shared data framework and clear model separation are therefore part of its originality claim.

Kareemzada et al. (2026) investigate taxation, institutional quality, growth, and political transitions in Afghanistan using an overlapping ARDL–ECM archive. The present study treats those dynamic estimates as contextual evidence. Its additional analyses concern covariance sensitivity, influential observations, alternative calendar boundaries, and cross-checks of archived univariate mean-break partitions. These analyses are descriptive and do not identify causal political-transition effects.

This manuscript adds executed sensitivity tests and reports unfavorable as well as favorable results. The editor must still assess whether that incremental contribution is sufficient for the selected journal, and the author must confirm the original coauthors' contribution and data reuse rights.

5.5. Policy implications and external validity

The findings support improved continuity of national accounts, fiscal reporting and archival documentation, particularly during administrative transitions. They do not estimate how much growth a particular regime, governance reform or revenue instrument would generate. More specific policy impacts would require credible treatment variation, better shock measurement or a robust dynamic design.

For statistical agencies and fiscal analysts, the clearest implication is to preserve exact observation definitions and historical release vintages. Changes in institutional classification, fiscal calendar and national-account benchmarks should be documented with machine-readable bridge tables. An annual data system capable of separating tax-type from collection office, budget transfers from recurrent receipts and estimated from observed output would reduce the ambiguity surrounding historical contrasts. These improvements follow from the measurement problem and public-finance research, not from a proven treatment benefit in the present regression.

For research on fragile states, policy lessons should likewise concern research design: pre-specified historical groupings, distinguish breaks in means from slopes, and report imprecise comparisons as imprecise. A small sample can be informative when its limitations are stated clearly; its value decreases when significant tests are used to rank governments or attribute GDP changes to one institution without an identification strategy. The present evidence supports better measurement and reproducibility, not a numerical return to any individual reform.

6. Conclusion

The independently reproduced five-period OLS equation matches the archived conventional joint test, $F(4, 19) = 3.5153$, $p = .0261$. The same joint null is not rejected using HC1 ($p = .1203$) or HC3 ($p = .3004$), whereas a one-lag HAC calculation rejects at $p = .0293$. Omitting 2021 from the original categories produces $p = .1186$; coding the final category from 2022 produces $p = .2268$ (HC3 $p = .3903$). Thus the observed period association is covariance- and calendar-sensitive rather than robustly established, and no individual period effect is statistically precise under the original conventional model.

The original Stata output locates statistically better-supported tax-revenue mean breaks at 2001 and 2008, compared with a weaker GDP-growth break candidate near 2019 and an ambiguous IQ chronology. New constrained segmentation reproduces the associated residual-minimizing tax and growth dates. These are changes in individual series means, not break dates for the growth-regression slopes, formal transition treatment effects or evidence of policy-induced economic gains.

The defensible conclusion is that historical-period coefficients in a short Afghanistan macroeconomic series depend materially on uncertain variance assumptions, particular country-years, and coarse annual political coding. A significant conventional F statistic should therefore be presented with its robustness failures rather than as proof of political or governance effects. The paper contributes a traceable reanalysis with newly executed diagnostics but does not independently validate the historical inputs against their exact WDI, WGI and Afghan reporting vintages.

All supplied Stata data files and the command log have been audited, and the principal OLS findings are reproduced independently. The analysis is limited by the absence of authenticated observation-level source vintages and harmonization records.

References:

Acemoglu, D., Johnson, S., & Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review*, *91*(5), 1369–1401. https://doi.org/10.1257/aer.91.5.1369

Akitoby, B., Honda, J., & Primus, K. (2020). Tax revenues in fragile and conflict-affected states: Why are they low and how can we raise them? (IMF Working Paper No. 2020/143). International Monetary Fund. https://doi.org/10.5089/9781513550848.001

Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, *18*(1), 1–22. https://doi.org/10.1002/jae.659

Barro, R. J. (1990). Government spending in a simple model of endogenous growth. *Journal of Political Economy*, *98*(5, Part 2), S103–S125. https://doi.org/10.1086/261726

Besley, T., & Persson, T. (2009). The origins of state capacity: Property rights, taxation, and politics. *American Economic Review*, *99*(4), 1218–1244. https://doi.org/10.1257/aer.99.4.1218

Brambor, T., Clark, W. R., & Golder, M. (2006). Understanding interaction models: Improving empirical analyses. *Political Analysis*, *14*(1), 63–82. https://doi.org/10.1093/pan/mpi014

Ditzen, J., Karavias, Y., & Westerlund, J. (2025). Testing and estimating structural breaks in time series and panel data in Stata. *The Stata Journal*, *25*(3), 526–560. https://doi.org/10.1177/1536867X251365449

Hameed, M. A., Rahman, M. M., & Khanam, R. (2022). Assessing the asymmetric war-growth nexus: A case of Afghanistan. *PLOS ONE*, *17*(8), e0272670. https://doi.org/10.1371/journal.pone.0272670

Jerven, M. (2013). *Poor numbers: How we are misled by African development statistics and what to do about it*. Cornell University Press.

Kareemzada, W., Rashid, M. U., & Shah, S. H. A. (2026). Tax revenue, institutional quality, and sustainable growth in Afghanistan (1996–2024): A dual source ARDL–ECM analysis. *Journal of Management Science Research Review*, *5*(1), 374–404. https://doi.org/10.5281/zenodo.19485920

Kaufmann, D., Kraay, A., & Mastruzzi, M. (2010). The worldwide governance indicators: Methodology and analytical issues (Policy Research Working Paper No. 5430). World Bank. https://doi.org/10.1596/1813-9450-5430

Long, J. S., & Ervin, L. H. (2000). Using heteroscedasticity consistent standard errors in the linear regression model. *The American Statistician*, *54*(3), 217–224. https://doi.org/10.1080/00031305.2000.10474549

MacKinnon, J. G., & White, H. (1985). Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *Journal of Econometrics*, *29*(3), 305–325. https://doi.org/10.1016/0304-4076(85)90158-7

North, D. C. (1990). *Institutions, institutional change and economic performance*. Cambridge University Press. https://doi.org/10.1017/CBO9780511808678

Advance Journal of Econometrics and Finance
Vol-4, Issue-1, 2026

Rafi, K. (2024). Minimalist economic management, deferred revenue regime and aid dependency: Explaining contradictory post-war statebuilding aims. *Global Policy*, *15*(5), 869–885. https://doi.org/10.1111/obes.12377

Saenger, T. R., Kapstein, E. B., & Sircar, R. (2024). Estimating the collapse of Afghanistan's economy using nightlights data. *PLOS ONE*, *19*(12), e0315337. https://doi.org/10.1371/journal.pone.0315337

World Bank. (2026). *Worldwide Governance Indicators: 2026 update* [Data set]. https://www.worldbank.org/en/publication/worldwide-governance-indicators

- **Appendix A. Model registry and source provenance**

Table A1

Distinct archived empirical objects and valid inference

ID	Window / N	Method and variables	Permitted claim
M00	1996–2024 / 29	Descriptives, correlations; growth, TR, IQ, FDI, TRADE, POPGR	Master Stata log; source-coded descriptives only
P01	1996–2024 / 29	OLS; growth, TR, IQ, FDI, TRADE, POPGR, D2–D5; no lags	Log-verified static contrasts and P01 diagnostics
B01–B03	1996–2024 / 29	Separate Bai–Perron-type constant-mean break screens	Log-verified univariate mean breaks; no causal timing
A01	1997–2020 / 24	ARDL(1,0,0,1,0); growth, TR, FDI, TRADE, POPGR	Separate dynamic baseline; not re-estimated here
X01	Unexecuted	Prospective D2–D5 × centered-IQ interaction; no fitted covariance matrix	Class D: unexecuted; no marginal-effects or period-governance conclusion
S01	1996–2024; N = 29 (P01 and sensitivities)	Independently executed HC1, HC3, HAC(1), leave-one-year-out and 2022 boundary sensitivities; static OLS, lag 0	Class C: reported in Table 4 and Appendix B; robust results do not confirm conventional joint rejection
P00	1996–2024 / 29	OLS, four historical periods; 1996–2001 / 2002–2014 / 2015–2020 / 2021–2024	Distinct historical coding; no D2–D5 joint test reported
P02	1996–2024; N = 29; d2 2002–08, d3 2009–13, d4 2014–17, d5 2018–21, d6 2022–24 (d1 1996–01 omitted)	OLS, five legacy dummies and five macro controls, static lag 0; separate joint F test	Independently reproduced F(5,18) = 1.4093, p = .2680; contrasts refer to different categories than P01
S02	Not yet executable	Original source-vintage reconstructions and matched WDI–official R0–R3 source substitutions	Class D: absent original observation-level source ledger; no coefficients or comparative claims

The archived analysis dataset contains 29 annual records, a nine-variable five component IQ cross-check, a single 1996–2024 time panel, and the legacy six-period dummy coding. Data integrity audit: FDI zeros occur in 1996–1999; trade-openness zeros in 1996–2001. IQ equals the mean of CC, RL, GE, VA and RQ to floating-point precision; PS is excluded. Two .dta versions agree substantively, but the workbook's data mirrors have different year representations.

Table A2

Alternative historical codings documented in the original Stata log

Model ID	Coding and years	N; estimator	Reported statistical result	Interpretation
P01 (preferred)	D1 1996–01; D2 2002–04; D3 2005–14; D4 2015–20; D5 2021–24	29; static OLS, lag 0	Period restrictions F(4,19)=3.52; p=.0261	Directly verified; focal P01 model
P00 (alternative)	1996–01; 2002–14; 2015–20; 2021–24	29; static OLS, lag 0	Overall F(8,20)=3.37; p=.0131	Overall fit test; no joint period-only test reported
P02 (legacy)	d1 1996–01; d2 2002–08; d3 2009–13; d4 2014–17; d5 2018–21; d6 2022–24	29; static OLS, lag 0	Independent reproduction: F(5,18)=1.4093; p=.2680	Different period windows than P01; no combined interpretation

Source. A denotes results transcribed from the archived Stata output; C denotes calculations reproduced from the archived annual dataset.

- **Appendix B. Candidate interactions and robustness sequence**

$\Delta_k(q) = \delta_k + \kappa_k q, \quad k \in \{2, 3, 4, 5\}$ (B1)

$\mathrm{Var}[\Delta_k(q)] = \mathrm{Var}(\delta_k) + q^2\mathrm{Var}(\kappa_k) + 2q\mathrm{Cov}(\delta_k, \kappa_k)$ (B2)

Completed independent sensitivity checks (class C): (1) exact-data reproduction of P01; (2) classical, HC1, HC3 and one-lag HAC four-restriction tests; (3) a 29-fit leave-one-year-out analysis with reported minimum/maximum and the 2021 result; (4) D4 through 2021 and D5 beginning 2022; (5) five-year-minimum exhaustive least-squares break-date partitions for growth, revenue and IQ; (6) cross-file numeric consistency and five component IQ validation. Not executed (class D): source-verified substitution of early zero records, a governance-vintage rerun, valid categorical IQ-by-period marginal effects, a formal break-adjusted multivariate growth regression, and causal transition identification. These are substantive limitations, not promises of favorable statistical results.

Table B1

Sensitivity-model definitions and numerical outcomes

Model / sensitivity	Years / N	Independent result	Interpretive boundary
P01 classical / HC1 / HC3 / HAC(1)	29; 1996–2024	p = .0261 / .1203 / .3004 / .0293	Same fitted mean; alternative covariance
P01 LOO 2021	28; 1996–2024 except 2021	p = .1186	Observation influence; separate fit
P01 calendar boundary	29; 1996–2024	p = .2268; HC3 p = .3903	D4 ends 2021; D5 begins 2022
B02 tax SSR, two breaks	29; 1996–2024	2001 and 2008, SSR = 21.5540	Univariate five-year-minimum partition
B01 GDP-growth SSR, one break	29; 1996–2024	2019, SSR = 1870.4083	Univariate five-year-minimum partition