Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A Computational Diagnostic Framework for Detecting Structural Break Under Small-Sample Conditions

Vilasa Chandran

vilaschandran1@gmail.com

Govt. Hamidia Arts and Commerce College

Sharad Tiwari

Govt. Hamidia Arts and Commerce College

Research Article

Keywords:

Posted Date: November 2nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10032015/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

A Computational Diagnostic Framework for Detecting Structural Break Under Small-Sample Conditions

Vilasa Chandran[1*] and Sharad Tiwari[2]

[1*]Department of Economics, Govt Hamidia Arts and Commerce College, Bhopal, 462001, Madhya Pradesh,India.

[2]Department of Economics, Govt Hamidia Arts and Commerce College, Bhopal, 462001, Madhya Pradesh, India.

*Corresponding author(s). E-mail(s): vilaschandran1@gmail.com;
Contributing authors: tiwarisharad042@gmail.com;

Abstract

Developing countries are often marked by changes in base years, periodic revisions in key macroeconomic indicators and unavailability of consistent back series for many indices. Consequently, researchers are frequently left with a short but homogeneous time series. Applying standard econometric break tests in such environments remains unreliable due to low-statistical power, while splicing together series from heterogeneous sources to satisfy statistical requirements mathematically introduces bias to the analysis and alters the underlying data-generation process. To overcome this limitation, this paper proposes a comprehensive computational diagnostic framework engineered specifically for small-sample settings. The framework consists of a core engine: the Ensemble Break Diagnostic (EBD) which operates as a location extractor by integrating the Zivot-Andrews unit root test with Maximum Entropy Bootstrap to map the frequency distribution of break-dates. The framework relies on a multi-layered validation structure combining an Algorithmic Stability Diagnostic based on Jensen-Shannon Divergence, a Null Monte-Carlo stress test, an Artificial Break Diagnostic, and a simulation-derived Threshold Derivation logic designed to control Type I and Type II errors. Applied to India's macro-fiscal time series (2011-2024), the framework successfully diagnoses a structural break in Real Revenue Expenditure in 2019, while Real GDP was diagnosed with a weak structural break/structural instability in 2019, and no evidence for a structural break was found in Real Total Expenditure and Real Capital Expenditure. A 20-country out-of-sample cross-validation confirms the framework's general applicability, providing the first purpose-built computational toolkit designed for detecting parameter shift in a small-sample time series with serial dependence.

Keywords: Structural Break Detection, Small-Sample Time Series, Jensen-Shannon Divergence, Maximum Entropy Bootstrap, Change Point Detection, Zivot-Andrews Unit Root Test, Ensemble Learning

1 Introduction

Policy interpretation relies critically on the strength and robustness of the estimation and inference drawn from the macro-fiscal time series. Nelson and Plosser (1982) demonstrate that many of the macro-economic time-series are non-stationary, implying that a shock may have a permanent effect and the series may not revert to

the trend line. Perron (1989) subsequently indicated that the unit-root behaviour may appear due to the break in the parameter (intercept, slope and trend), as the data-generation process pre and post break may diverge due to a shift in the parameters. Consequently, the validity and robustness of estimation and inference derived from such a series is conditional on the correct identification of exogenous and regime-specific data-generation process since pooling observations together under two distinct data-generation processes may yield misleading conclusions.

However, Perron (1989) also cautions against exogenously imposed break dates, highlighting the risk of data mining and spurious inference. To address this concern, Zivot and Andrews (1992) propose a unit root test with an endogenously determined single break, selected through repeated testing across every candidate for the break-date. While this approach mitigates data mining concern associated with exogenous break selection, this method also has a well known limitation in connection to the low power in small observational time-series and restriction to a single break point which minimizes the t-statistic Although the Bai and Perron (1998) test is powerful in a moderate to large sample and when breaks are well separated, it is less reliable in small-sample time series. Accordingly, given the short sample and the objective of diagnosing structural break date and uncertainty surrounding its occurrence, this paper develops a computational diagnostic framework by integrating elements from the standard econometrics, ensemble learning, information theory, computational resampling and Wald's minimax decision theory. Furthermore, our study takes a methodological pivot by utilising the Zivot-Andrews unit root test as a diagnostic tool for break-date identification rather than relying on the conventional hypothesis testing approach.

Following this, India presents a particularly challenging empirical environment for macroeconomic and fiscal analysis. The country is characterised by periodic revisions of key macroeconomic indicators, changes in base years and unavailability of consistent back series for newly introduced measurement indices. As a result, researchers may unintentionally violate standard assumptions underlying the classical econometric inference by imposing linearity and parameter stability on data generated under multiple regimes. These challenges are especially consequential in studies involving public expenditure and deflation to real values. In particular, the absence of consistent back series for the CPI NEW (urban+rural) prior to 2011 complicates the construction of real fiscal aggregates.

Additionally, India's high-frequency data exhibit pronounced "last-quarter rush" effects which artificially inflate volatility and may mimic regime shifts. This renders quarterly or monthly series noisy and unsuitable for diagnosing genuine structural breaks. Although recent institutional reforms, including the SNA (2021) and SNA SPARSH (2023) based disbursement mechanism, were intended to smooth fiscal execution, they simultaneously alter the underlying data-generation process.

Consequently, researchers face a fundamental Bias-Variance Trade-off. Extending the observations through splicing or index substitution may reduce variance but risks introducing bias by pooling observations generated under two distinct regimes. As cautioned by Lucas (1976), such regime changes can undermine the validity and robustness of the policy inference. In this setting, standard econometric tests may misclassify measurement changes or institutional reforms as a genuine structural break. Accordingly, restricting the observations to a shorter but internally homogeneous period preserves the integrity of the data-generation process and yields more reliable policy diagnostics than mechanically extending the sample at the cost of increased bias.

As a result, this paper adopts a diagnostic rather than a causal perspective, with the objective of diagnosing break and uncertainty around break-date. Consequently, allowing for precise policy interpretation under conditions of regime instability, the main analysis focuses on annual Real Revenue Expenditure, Real Capital Expenditure, Real Total Expenditure, and Real GDP of India over the period 2011–2024, with a secondary cross-country validation in Appendix C applying the framework to Real GDP measured in local currency for twenty distinct economies to prove the general applicability of the proposed framework. This study engineers a computational diagnostic framework by embedding the Zivot-and-Andrews unit root test within the

Maximum Entropy Bootstrap procedure by Vinod (2006)—hereafter referred to as the Ensemble Break Diagnostic (EBD)–to map the frequency distribution of break–dates. To assess the stability of this algorithm, the Ensemble Break Diagnostic is independently executed a second time, and the Jensen-Shannon Divergence is employed to quantify the divergence between the initial reference run and an entirely independent replication. In addition, the framework subjects the Ensemble Break Diagnostic to a Null Monte Carlo stress test. A deterministic linear time series, generated to contain no structural break, is processed through the Ensemble Break Diagnostic to generate a null break-date frequency distribution. The Jensen-Shannon Divergence is then used to quantify the divergence between the empirical and the regular null frequency distribution, providing a quantitative mechanism for distinguishing a genuine structural break from the regular null dispersion of frequency.

To evaluate sensitivity, distinct directional artificial breaks are introduced into two controlled synthetic time series (upward-rising and downward-falling) and processed through the Ensemble Break Diagnostic to examine whether the procedure successfully detects the imposed break. Finally, to objectively quantify a structural break and isolate it from bootstrap-induced dispersion, the Signal to Noise Ratio (SNR) and Dominance Share (%) calculations are embedded directly within the artificial break algorithm. This integrated engine is then subjected to a computationally intensive 100-iteration nested loop simulation involving approximately 1598400 distinct regressions to derive the thresholds for both the Signal-to-Noise Ratio (SNR) and Dominance Share (%). Within the proposed framework, a structural break is diagnosed only when an empirical series exceeds both computationally derived thresholds.

Our results indicate a pronounced clustering of break-date frequencies for Real GDP and Real Revenue Expenditure around 2019. However, Real GDP exceeded the threshold for dominant share (%) but could not cross the threshold for the Signal-to-Noise Ratio (SNR), suggesting evidence for structural instability/weak structural break in Real GDP in 2019. Meanwhile, Real Revenue Expenditure exceeded the thresholds for both Signal-to-Noise Ratio (SNR) and Dominance Share (%), providing strong evidence for a structural break in 2019. In contrast, Real Capital Expenditure and Real Total expenditure did not exceed the thresholds for the Signal-to-Noise Ratio (SNR) and Dominance Share (%), implying absence of a meaningful structural break. These findings underscore the need for caution in policy interpretation when data continuity is limited and the assumptions underlying standard econometric inference are potentially violated.

2 The Empirical Context: Diagnostic Gaps and Data Constraints

The empirical literature on public expenditure and GDP in India is extensive and has yielded important insights for policy makers, much of the work re-lies upon the application of standard time-series econometric framework, often under the assumption of parameter stability and data continuity.

However, in an applied macro-economic environment characterised by frequent policy reforms, base year revisions and changes in statistical measurement frameworks, these assumptions are difficult to maintain . Several influential studies (eg. Bose and Bhanumurthy (2013.), Srinivasan (2013), Tulsidharan (2006)) therefore rely on nominal expenditure series and assume parameter stability across different policy reforms, external shocks over long-period of time. This choice is methodologically pragmatic, as it preserves sample length and avoid technical challenges associated with dealing with the structural break and splicing heterogeneous price indices across different regime.

Against this backdrop, this study complements the existing literature by adopting a diagnostic perspective to examine structural instability under data constraints, with particular emphasis on preserving regime homogeneity and characterizing expenditure reorientation rather than imposing parametric stability.

3 Data

This study employs annual data on India's real gross domestic product (GDP), Total Expenditure (Real), Revenue Expenditure (Real), Capital Expenditure(Real) over the period 2011-2024. Expenditure data are obtained from the Handbook of Statistics on the Indian Economy published by the Reserve Bank of India (2025), while Real GDP data are obtained from the Economic Survey of India (2025) and the Handbook of Statistics (2025). Expenditure series are deflated using the Consumer Price Index (Combined Urban + Rural) base year = 2012

$$RealValue = \frac{NominalValue(CurrentYear)}{CPI(CurrentYear)} \times 100 \quad (1)$$

4 Research Methodology

Motivated by severe data constraints and the overarching objective to detect instability in macroeconomic time series, this paper adopts a diagnostic perspective by constructing an overarching Computational Diagnostic Framework to detect structural parameter shifts in small-sample settings. The framework systematically consists of four integrated computational layers: the Ensemble Break Diagnostic (EBD), the Algorithmic Stability Diagnostic, the Artificial Break Diagnostic, and the Threshold Derivation Architecture. All estimation and simulation procedures are conducted under a fixed random seed to ensure exact computational reproducibility across alternative data environments.

4.1 Mathematical Formulation of the Core Engine: Ensemble Break Diagnostic (EBD)

Successfully detecting a structural break in a small-sample macroeconomic series imposes a significant challenge for researchers, as the standard break detection procedure suffers from "low statistical power", especially in environments where observational strength is significantly low. To resolve this predicament, we rely on a bootstrap procedure especially designed to generate resampled perturbations of the underlying data-generation process while preserving the strict serial dependence structure. The generated ensemble will closely resemble the underlying data-generation process, and any shift in the parameters will certainly reflect on the resampled series.

Following this logic, this study combines the entropy-preserving resampling method with the endogenized break detection test to operationalize our core computational engine.

4.1.1 Zivot-Andrews Break-Date Diagnostic Procedure

This study utilizes the Zivot-Andrews unit-root test as a primary constituent part in the formulation of the core engine. The Zivot-Andrews procedure extends the Augmented Dickey-Fuller (ADF) framework by allowing for a single unknown structural break in the intercept, the slope, or both parameters of the time series. The test evaluates all possible break-dates within the series and settles on the specific point that yields the least favorable result for the null unit-root hypothesis. Consequently, the break-point is selected by the procedure which minimizes the one-sided t-statistic.

Critically, while the test is mathematically designed to always report the most adverse observation as a potential break-point, this date should not be interpreted as a definitive structural break in isolation. Given the short span of the observations ($n = 14$), the Zivot-Andrews test is not employed here for traditional hypothesis confirmation. Instead, it serves as a diagnostic tool within the core engine to identify the points which yield the most minimal one-sided t-statistics across the generated ensemble. This allows the framework to map the distribution of break-dates rather than relying on a single, potentially biased point estimate.

Model Specification and Degree of Freedom Consideration:

To maximize diagnostic strength under limited observations, the analysis adopts Model C that allows for both intercept and slope shift. This specification is consistent with Indian data-specific environment.

To preserve the degree of freedom, the lag length is restricted to 1, while this choice reduces

flexibility, it ensures tractability and avoids overfitting in a small observation environment. The effective loss of observation at the end points of each series is explicitly acknowledged and treated as an unavoidable consequence of the diagnostic framework.

Regression Equation for Model C:

$$y_t = \mu^c + \theta^c DU_t(\lambda) + \beta^c t + \gamma^c DT_t^*(\lambda) + \alpha^c y_{t-1} + \sum_{j=1}^{k} c_j^c \Delta y_{t-j} + e_t \tag{2}$$

Where,

1. $DU_t(\lambda) = 1$ if $t > T\lambda$ and 0 otherwise; (intercept shift)
2. $DT_t(\lambda) = (t - T\lambda)$ if $t > T\lambda$ and 0 otherwise; (trend shift)
3. $T\lambda =$ candidate break date
4. $k = 1$ is the selected lag length

4.1.2 Maximum Entropy Bootstrap Framework

Given the limited number of observations and inherent non-stationarity of Indian macro- fiscal time series, conventional resampling techniques are not appropriate. In particular, i.i.d bootstrap of Efron (1979) is invalid in the presence of serial dependence, as it fails to preserve the temporal correlation structure of the data (Singh, 1981). Block bootstrap methods partially address this limitation by preserving serial dependence through block resampling however, they typically rely on assumption of weak stationarity and require long observations to form representative blocks.

To address these constraints while preserving the underlying dependence structure and trend of macro fiscal time series. This study employs Maximum Entropy Bootstrap proposed by Vinod (2006). The method develops 999 sets of plausible series from the original series using the principle of maximum entropy, subject to mean-preserving constraints. As a result, the grand mean of all bootstrap generated ensembles is equal to the mean of the original series.

Unlike conventional bootstrap approaches that rely on detrending or differencing to achieve stationarity, the Maximum Entropy Bootstrap avoids explicit transformation of the data. Instead, the generated ensembles retain basic shape, trend behavior ,and time-dependence properties of the original series, including autocorrelation and partial autocorrelation structures. While this strict preservation of temporal dependence has been identified as a limitation of the Maximum Entropy Bootstrap for generating null distributions in classical hypothesis testing (Bergamelli et al., 2015), it is particularly what makes the procedure advantageous for this specific diagnostic setting. This feature is particularly important in small–sample data environment where pre-filtering of time series may distort the underlying data generation process necessary for detecting shift in the parameters accurately.

As a result, the study adopts the Maximum Entropy Bootstrap as the second constituent component for the formulation of the core engine, ensuring that the generated ensembles strictly preserve the serial dependence and non-linear trend of the original series. The algorithmic construction of the MEB procedure follows the step outlined in Vinod (2006) and is summarized in Appendix A.

4.1.3 Integration and Engine Synthesis: Ensemble Break Diagnostic (EBD)

To complete the core engine, the Zivot-Andrews unit-root test is applied to each bootstrap replication generated by the Maximum Entropy Bootstrap. This allows the framework to assess the sensitivity of break-date identification to the uncertainty generated by data-consistent bootstrap perturbations of the original series. The resulting distribution of break-dates across the bootstrap ensemble is utilized as a diagnostic measure of regime instability and uncertainty.

The Ensemble Break Diagnostic (EBD) is mathematically formulated as follows:

$$L_j^* = index_t[min(\mathcal{T}_{ZA}(y_j^*))] \quad j \in 1, 2, \ldots, 999 \tag{3}$$

$$f(L_t) = \sum_{j=1}^{J} \mathbb{I}(L_j^* = t) \quad t \in 1, 2, \ldots, 14 \tag{4}$$

$$\mathbb{I}(L_j^* = t) = \begin{cases} 1 & if \quad L_j^* = t \\ 0 & otherwise \end{cases} \tag{5}$$

Where, the variables are defined as:

1. L_j^*: The location of the minimum-most one-sided t-statistic identified in perturbation j

2 .y_x: The empirical macroeconomic series where $x = \{1, 2, \ldots, n\}$ denotes the sequence of observations.

3. $n = 14$: The total number of observations in the sample.

4.L_t: The discrete location variable, where $t \in x$ and $L_t \in \{1, 2, \ldots, 14\}$.

5. index$_t$: The operator that extracts the specific serial observation number $t \in x$.

6. min: The operator identifying the minimum-most one-sided t-statistic within a single perturbation.

7. $\mathcal{T}_{ZA}$: Model C of the Zivot-Andrews unit-root test specification and lag=1.

8. $J = 999$: The total ensemble size generated through the Maximum Entropy Bootstrap.

9. $\sum \mathbb{I}$: The aggregation (frequency counting).

10. j: j^th series of the 999 generated series.

11. $f(L_t)$: denotes the frequency count of break-date at observation t.

4.2 Algorithmic Stability Diagnostic

The Maximum Entropy Bootstrap procedure relies on independently drawn random number from the [0,1] uniform intervals to generate an ensemble of plausible series while preserving the time dependence structure of the original series. Accordingly, it is essential to assess whether the identified break-date distribution is an intrinsic property of the underlying series and is stable across the independent runs of the algorithm.

This study employs Jensen-Shannon Divergence by Lin (1991) to compare the break-date frequency distribution from the initial (reference) run of the Ensemble Break Diagnostic (EBD) with that obtained from an entirely independent run of the same EBD. Given that Jensen-Shannon Divergence is symmetric, bounded and equals zero if two distributions are identical, making it particularly handy to measure divergence between two probability distributions and suitable to compare frequency distributions generated by the framework.

Following common practice in applied settings such as in signal processing, information theory and ensemble learning in machine learning, etc. Jensen- Shannon Divergence values below a small tolerance level are interpreted as indicating negligible divergence between frequency distributions . In this study, Jensen-Shannon Divergence value below 0.01 is interpreted as indicating stability across independent algorithmic runs and two distributions being practically indistinguishable.

Jensen-Shannon Divergence (JSD) between two probability distributions P and Q with equal weight $\pi_1 = \pi_2 = \frac{1}{2}$, is defined as

$$JSD(P||Q) = 0.5 \times (KL(P||R) + KL(Q||R)) \quad (6)$$

$R = 0.5 \times (P + Q)$ denotes the midpoint of the probability vectors P and Q and $KL(P||R)$, $KL(Q||R)$ denotes the Kullback-Leibler(KL) Divergence.

4.3 Null Monte-Carlo Stress Test

To assess whether the generated break-date frequency distribution is an artifact of the bootstrap procedure or reflects inherent structural instability in the empirical time series, the framework is subjected to a rigorous null stress test. A deterministic null placebo series, constructed to contain no structural break , is generated to serve as the null data-generation process.

$$y_t = \alpha + \beta t + \epsilon_t \quad (7)$$

1. $\epsilon_t \sim N(0, 1)$,Normal Distribution
2. $\alpha = 100$, constant for all t
3. $\beta = 2$, constant for all t
4. No shift in Intercept and Slope
5. $t = 14$

To generate a perfect null environment, the study selects a standard normal distribution $N(0, 1)$ for the residuals. This choice introduces necessary stochastic variation into the series without distorting the deterministic path of the placebo series. Following the fundamental assumptions of Gaussian innovations, keeping the mean at 0 and the standard deviation at 1 ensures that the residuals do not mimic a structural break or unintentionally influence the Ensemble Break Diagnostic (EBD). As a result, the generated placebo series strictly preserves the deterministic trend throughout all observations. This allows the framework to operate in an environment where the placebo series is purely linear, containing no shifts in either slope or intercept, thereby providing a

rigorous baseline for measuring empirical divergence. The Ensemble Break Diagnostic (EBD) is then applied to the null placebo series to obtain a null frequency distribution. The divergence between the empirical and null frequency distribution is then quantified using the Jensen-Shannon Divergence (JSD), defined in 4.2.

Let,

1. f_{emp} denotes the empirical break-date frequency distribution

2. f_{null} denotes the null break-date frequency distribution

$$JSD(f_{emp}||f_{null}) = 0.5(KL(f_{emp}||R) + KL(f_{null}||R)) \quad (8)$$

$KL(f_{emp}||R)$ and $KL(f_{null}||R)$ denote the Kullback-Leibler(KL) Divergence.

To provide scale for interpretation, the study employs a Signal to Noise Ratio (SNR) based on Algorithmic Stability Diagnostic (Jensen-Shannon Divergence). The baseline noise level is defined as the maximum observed Jensen-Shannon Divergence value obtained from the initial (reference) and independent run , specified in 4.2.

$$SNR = \frac{JSD(f_{emp}||f_{null})}{JSD_{BASELINE}} \quad (9)$$

$JSD_{BASELINE} = max_{ij} JSD(f_i, f_j)$

(where f_i and f_j are frequency distribution from initial (reference) and independent run detailed in the 4.2)

The ratio serves as a quantitative benchmark for evaluating whether the detected clustering represents a genuine structural signal for break rather than dispersion induced by the bootstrap procedure. The SNR combined with Dominance Share (%) (percentage of concentration for the highest frequency break year), is utilized as a "double-lock condition" to diagnose and objectively confirm structural break. Rather than relying arbitrary heuristics, the objective thresholds for both Signal to Noise Ratio (SNR) and Dominance Share (%) are computationally derived through the rigorous simulation framework detailed in section 4.5.

4.4 Artificial Break Diagnostic

To evaluate the sensitivity of the proposed break-detection framework, the study generates two distinct controlled synthetic time series:(i) an upward-rising time series and (ii) a downward-falling time series. Artificial structural breaks under four different conditions are then subsequently introduced within the parameters(intercept & slope) of each controlled synthetic time series. The objective is to examine whether the Ensemble Break Diagnostic (EBD) generates a prominent break-date clustering when a known shift in intercept and slope is embedded in the data-generating process.

Model Specification for Controlled Synthetic Time Series:

(i) Upward-Rising Time Series:

$$y_{ut} = \alpha_u + \beta_u t + \epsilon_t \quad (10)$$

(ii) Downward-Falling Time Series:

$$y_{dt} = \alpha_d + \beta_d t + \epsilon_t \quad (11)$$

Where the parameters are established as:

1.$\alpha_u = 100$ (Upward Rising Series Intercept)

2.$\alpha_d = 1500$ (Downward falling series Intercept)

3.$\beta_u = 20$ (Upward Rising Series Slope)

4.$\beta_d = -15$ (Downward Rising Series Slope)

5.$\epsilon_t \sim N(0, 1)$ Normal Distribution

6.$t = 14$

The residuals of the controlled synthetic time series follow a standard normal distribution, $\epsilon_t \sim N(0, 1)$, to introduce the stochastic variation while preventing residuals to influence Ensemble Break Diagnostic (EBD).

4.4.1 Injection of Artificial Break Under Different Conditions

To fully analyze how the Ensemble Break Diagnostic (EBD) behaves under different structural breaks, the study engineers four distinct scenarios:(i) Surge and Surge, (ii) Rise and Crash, (iii) Crash and Surge, and (iv) Crash and Fall, where the intercept and slope shift distinctly. Subsequently, the four structural break scenarios are embedded within the two controlled synthetic time series to fully simulate structural shifts.

1.**Surge and Surge:** Under this scenario, both the intercept and slope shift upward post-break.

2.**Rise and Crash:** Under this scenario, the intercept shifts upward immediately after the break, but the slope shifts downward following the break.

3.**Crash and Surge:** Under this scenario, the intercept shifts downward post-break, followed by an upward shift in the slope.

4.**Crash and Fall:** This scenario follows a condition in which both the intercept and slope shift downward, post-break.

Model Specification for Injection of Structural Break:

(i) Upward-Rising Series with Break:

$$y_{ut} = \begin{cases} \alpha_u + \beta_u t + \epsilon_t & t < T_B \\ (\alpha_u + \delta_\alpha) + (\beta_u + \delta_\beta)t + \epsilon_t & t \geq T_B \end{cases} \tag{12}$$

(ii) Downward-Falling Series with Break:

$$y_{dt} = \begin{cases} \alpha_d + \beta_d t + \epsilon_t & t < T_B \\ (\alpha_d + \delta_\alpha) + (\beta_d + \delta_\beta)t + \epsilon_t & t \geq T_B \end{cases} \tag{13}$$

Parameterization of Artificial Breaks:

To ensure exact computational reproducibility, the specific magnitudes of the injected structural shifts were explicitly programmed into the simulation environment. The breaks were introduced at observation six($T_B = 6$), with the magnitude and directional signs of the intercept shift(δ_α) and slope shift (δ_β) calibrated to test the diagnostic limits of Ensemble Break Diagnostic (EBD).

The exact parameterization for each artificial break permutations(four scenarios applied across two controlled synthetic time series) are given in Table: (1). The generated synthetic series are subsequently processed through the Ensemble Break Diagnostic (EBD) to map the frequency distributions of the break-dates.

4.4.2 Algorithmic Stability of the Synthetic Series with Artificial Structural Breaks

To assess the internal stability of the algorithm when processing massive structural shocks, the study utilizes the Jensen-Shannon Divergence (JSD) by Lin (1991). This metric compares the frequency distribution obtained from the initial application of the Ensemble Break Diagnostic (EBD) on the 8 distinct controlled synthetic time series against an entirely independent run of the same series. As established in section 4.2, the study adopts a Jensen-Shannon Divergence value of below 0.01 is taken as indicative of minimal stochastic variation between independent algorithmic runs.

The internal algorithmic stability noise floors for each scenario are denoted as follows:

Upward-Rising Controlled Series:

1.JSD_1: Represents the calculated JSD value for the Surge and Surge scenario.

2.JSD_2: Represents the calculated JSD value for the Rise and Crash scenario.

3.JSD_3: Represents the calculated JSD value for the Crash and Surge scenario.

4.JSD_4: Represents the calculated JSD value for the Crash and Fall scenario.

Downward-Falling Controlled Series:

1.JSD_5: Represents the calculated JSD value for the Surge and Surge scenario.

2.JSD_6: Represents the calculated JSD value for the Rise and Crash scenario.

3.JSD_7: Represents the calculated JSD value for the Crash and Surge scenario.

4.JSD_8: Represents the calculated JSD value for the Crash and Fall scenario.

4.4.3 Null Baseline Series Generation

To accurately interpret the structural break signals, the study generates two continuous null baseline series representing the uninterrupted data-generating processes:(i) an upward-rising series and (ii) a downward-falling series, both containing no structural breaks. These dynamic nulls provide a robust, directionally unbiased baseline against which to measure the magnitude of the structural ruptures artificially induced within each controlled time series. Ultimately, quantifying the divergence between the artificial break scenarios and these dynamic nulls directly facilitates the derivation of computationally verified, objective thresholds for both the Signal-to-Noise Ratio (SNR) and the Dominance Share(%).

Table 1 Parameterization of Artificial break conditions $T_B = 6$

Structural Scenario	**Baseline DGP**	**Intercept Shift**(δ_α)	**Slope Shift**(δ_β)	**Post-Break Dynamics**
Surge & Surge	Upward-Rising ($\beta_u > 0$)	+100	+45	Accelerated Growth
	Downward-Falling ($\beta_d) < 0$	+110	+45	Trend Reversal(Upward)
Rise & Crash	Upward-Rising ($\beta_u > 0$)	+100	−34	Immediate Spike,Subsequent Decay
	Downward-Falling ($\beta_d < 0$)	+150	−30	Immediate Recovery, Accelerated Decay
Crash & Surge	Upward-Rising ($\beta_u > 0$)	−100	+35	Immediate Collapse, Accelerated Recovery
	Downward-Falling ($\beta_d < 0$)	−100	+45	Deepened Collapse, Trend Reversal
Crash & Fall	Upward-Rising ($\beta_u > 0$)	−90	−30	Structural Collapse
	Downward-Falling ($\beta_d < 0$)	−200	−25	Accelerated Structural Collapse

[1]Notes: All simulated scenarios utilize a sample size of $t = 14$ with standard normal gaussian noise $\epsilon_t \sim N(0, 1)$,applied continuously across all observations

Model Specification for the Baseline Null:

To establish the diagnostic control group, the study specifies the data-generating process for two dynamic null baselines. These baselines represent the pure trajectories without any parameteric shifts.

The models are defined as:

$$null_{up} = \alpha_{up} + \beta_{up}t + \epsilon_t \tag{14}$$

$$null_{down} = \alpha_{do} + \beta_{do}t + \epsilon_t \tag{15}$$

1.$\alpha_{up} = 100$
2.$\beta_{up} = 2$
3.$\alpha_{do} = 1500$
4.$\beta_{do} = -20$
5.$\epsilon_t \sim N(0, 1)$
6.$t = 14$

These continuous null baselines are then subjected to Ensemble Break Diagnostic (EBD) to generate their respective break-date frequency distributions. let $f_{null_{up}}$ denotes the frequency distribution generated for the upward-rising null baseline series, and $f_{null_{down}}$ denote the frequency distribution generated for the downward-falling null baseline series.

4.4.4 Quantifying Sensitivity via Jensen-Shannon Divergence

To interpret the structural break signal and test the sensitivity of the framework, the study utilizes the Jensen-Shannon Divergence, as defined in Section 4.2, to measure the divergence between each controlled synthetic time series engineered to hold a parameter shift and its respective continuous baseline. The upward-rising controlled synthetic time series are measured strictly against the upward-rising baseline, and vice-versa.

The resulting divergence value, combined with the chronological clustering around the artificially

imposed break date, indicates whether the framework is responsive to genuine structural shifts rather than random bootstrap dispersion.

Let the frequency distributions of the simulated break dates obtained from applying Ensemble Break Diagnostic on the respective controlled scenarios be defined as follows:

Upward-Rising Controlled Series:

1.f_{art_1}: Frequency distribution obtained from the Surge and Surge scenario.

2.f_{art_2}: Frequency distribution obtained from the Rise and Crash scenario.

3.f_{art_3}: Frequency distribution obtained from the Crash and Surge scenario.

4.f_{art_4}: Frequency distribution obtained from the Crash and Fall scenario.

Downward-Falling Controlled Series:

1.f_{art_5}: Frequency distribution obtained from the Surge and Surge scenario.

2.f_{art_6}: Frequency distribution obtained from the Rise and Crash scenario.

3.f_{art_7}: Frequency distribution obtained from the Crash and Surge scenario.

4.f_{art_8}: Frequency distribution obtained from the Crash and Fall scenario.

The divergence between the artificial break signals and the null baselines is formally expressed as:

(i)Upward-Rising Series Divergence:

$$JSD(f_{art_i}||f_{null_{up}}) = \frac{1}{2}\big[KL(f_{art_i}||R) + KL(f_{null_{up}}||R)\big] \quad i \in 1,2,3,4 \tag{16}$$

(ii)Downward-Falling Series Divergence:

$$JSD(f_{art_i}||f_{null_{down}}) = \frac{1}{2}\big[KL(f_{art_i}||R) + KL(f_{null_{down}}||R)\big] \quad i \in 5,6,7,8 \tag{17}$$

4.5 Threshold Derivation Through Simulation

To derive the thresholds for the Signal-to-Noise Ratio (SNR) and Dominance Share (%), the study relies on a computationally intensive simulation by embedding the SNR calculation directly within the artificial break diagnostic algorithm.

To calculate the SNR, the study first isolates the maximum observed Jensen-Shannon Divergence values from the internal algorithmic stability checks of both the upward-rising and downward-falling controlled synthetic time series (as detailed in Section 4.4.5). These maximum internal divergence values serve as the two distinct algorithmic noise baselines. The SNR is then computed as the ratio of the structural divergence—measured between the controlled synthetic time series and their respective continuous nulls. (detailed in Section 4.4.4)

The respective Signal-to-Noise Ratio (SNR) are formalized as follows:

(i) Upward-Rising Controlled Series (SNR):

$$SNR_{art_i} = \frac{JSD(f_{art_i}||f_{null_{up}})}{JSD_{BASELINE_1}} \quad i \in 1,2,3,4 \tag{18}$$

where,

$JSD_{BASELINE_1 = max(JSD_1, JSD_2, JSD_3, JSD_4)}$

(ii) Downward-Falling Controlled Series (SNR):

$$SNR_{art_i} = \frac{JSD(f_{art_i}||f_{null_{down}})}{JSD_{BASELINE_2}} \quad i \in 5,6,7,8 \tag{19}$$

where,

$JSD_{BASELINE_2 = max(JSD_5, JSD_6, JSD_7, JSD_8)}$

To calculate the Dominance Share (%), the frequency distributions obtained from applying the Ensemble Break Diagnostic on the simulated controlled scenarios (detailed in Section 4.4.4) are utilized. The study isolates the maximum observed frequency peak from each distribution to calculate the dominance share(%).

The respective Dominance Shares are formalized as follows:

(iii)Upward-Rising Controlled Series:

$$DominanceShare(DS)(\%)_{art_i} = \frac{max(f_{art_i})}{999} \quad i \in 1,2,3,4 \tag{20}$$

(iv)Downward-Falling Controlled Series:

$$DominanceShare(DS)(\%)_{art_i} = \frac{max(f_{art_i})}{999} \quad i \in 5,6,7,8 \tag{21}$$

This core algorithm is subsequently subjected to a 100-iteration Monte Carlo simulation, dynamically subjecting the associated variables to stochastic variation across each run. The nested simulation architecture involves approximately 1,598,400 distinct regressions, requiring roughly 25 hours of continuous computation on a standard personal computer. Due to this severe computational intensity, the study restricts the Monte Carlo simulation to 100 runs, which is mathematically sufficient to exhaust the stochastic probability space for the $t = 14$ sample size and stabilize the significance thresholds.

4.5.1 Threshold Calculation via Boundary Derivation

In small-sample environments, relying on standard asymptotic p-values often leads to spurious conclusions due to low statistical power. Therefore, this study adopts the Minimax decision theory (Wald, 1950) and the framework of the Minimum Detectable Signal (MDS) to computationally derive statistically significant thresholds for both the Signal-to-Noise Ratio (SNR) and the Dominance Share (%).

The threshold is derived by establishing two absolute mathematical boundaries: the maximum observable signal from the worst-case momentum scenario (controlling false positive detection: Type I error), and the minimum observable signal from the best-case structural rupture scenario (reducing false negative detection: Type II error). The exact midpoint between these two boundaries serves as the baseline diagnostic threshold. Because this midpoint is subject to stochastic variation due to the nested loop architecture of the of the artificial break diagnostic algorithm, it is evaluated over multiple runs of the algorithm to achieve stabilization.

The Ensemble Break Diagnostic (EBD) is expected to produce significant clustering of break dates when the intercept and/or slope break in a direction opposite to the pre-break trend (e.g., an upward-rising series experiencing a crash). Consequently, the study utilizes the minimum observable SNR and Dominance Share from these opposing-break scenarios $(SNR_2, ..., SNR_7)$ and $(DS_{art_2}, ..., DS_{art_7})$ to define the upper boundary of the threshold.

Conversely, for the lower boundary, the study utilizes the maximum observable SNR and Dominance Share from the (SNR_1, SNR_8) and (DS_{art_1}, DS_{art_8}) scenarios. In these scenarios, both the intercept and slope break in the exact same direction as the pre-break trajectory of the controlled series (e.g., Surge and Surge on an upward-rising series). These represent severe trend accelerations rather than genuine trend inversions.

The discrete threshold for a single iteration is defined as the midpoint between these boundaries:

(i) Threshold For Signal to Noise Ratio:

$$Threshold_{SNR} = \frac{1}{2}\Big[\min(SNR_2, SNR_3, SNR_4, SNR_5, SNR_6, SNR_7) + \max(SNR_1, SNR_8)\Big] \tag{22}$$

(ii) Threshold For Dominance Share (%):

$$Threshold_{DS} = \frac{1}{2}\Big[\min(DS_{art_2}, DS_{art_3}, DS_{art_4}, DS_{art_5}, DS_{art_6}, DS_{art_7}) + \max(DS_{art_1}, DS_{art_8})\Big] \tag{23}$$

To derive the final, objective thresholds applicable to empirical data, the study takes the arithmetic mean of these calculated thresholds across the 100-iteration Monte Carlo simulation. This strictly stabilizes the threshold against any residual bootstrap-induced dispersion(Additional results demonstrating stability of 100 iteration Monte Carlo Simulation, is detailed in Appendix B):

(iii) Final Threshold for Signal to Noise Ratio:

$$FinalThreshold(SNR) = \frac{1}{100}\sum_{i=1}^{100}(Threshold_{SNR})_i \quad (24)$$

(iv) Final Threshold for Dominance Share(%):

$$FinalThreshold(DominanceShare) = \frac{1}{100}\sum_{i=1}^{100}(Threshold_{DS})_i \quad (25)$$

5 Results

5.1 Algorithmic Stability Diagnostic Results

Table 2 Algorithmic Stability Diagnostic(JSD Values)

Series	JSD Value[1]	Interpretation
Real GDP	0.005831868	High Algorithmic Stability
Real Revenue Expenditure	0.005904571	High Algorithmic Stability
Real Capital Expenditure	0.003860585	High Algorithmic Stability
Real Total Expenditure	0.004582688	High Algorithmic Stability

Note: Jensen-Shannon Divergence value less than 0.01 indicate high algorithmic stability.

[1] Jensen-Shannon Divergence between initial (reference) and independent run refer section 4.2

The Jensen-Shannon Divergence values reported in Table: 2 for all series, are uniformly small and well below the adopted tolerance level of 0.01. This indicates the break-date frequency distributions generated by Ensemble Break Diagnostic are highly stable across independent algorithmic runs. Consequently, the identified concentration and dispersion patterns reflect intrinsic features of the underlying data rather than artifacts of simulation.

5.2 Ensemble Break Diagnostic (EBD) Break-Dates Frequency Distribution

5.2.1 Real GDP

The break-date frequency distribution for Real GDP (Fig: 1(a)) exhibits a pronounced clustering around 2019, accounting for 39.23% of the 999 bootstrap replications. The remaining years record comparatively smaller break-date frequencies, indicating a concentration of break-date frequencies around a single dominant year.

5.2.2 Real Revenue Expenditure

The frequency distribution for Real Revenue Expenditure (Fig: 1(b)) shows a strong clustering around 2019, which accounts for 60.26%of the total bootstrap replications. Other years contribute relatively modest frequencies ,suggesting a clear dominance of one single year within the break-date frequency distribution.

5.2.3 Real Capital Expenditure

The generated break-date frequency distribution for Real Capital Expenditure (Fig: 1)(c)) appears more dispersed across 2019-2020, with no single year overwhelmingly dominating break-date frequency distribution

5.2.4 Real Total Expenditure

The frequency distribution of break dates for Real Total Expenditure (Fig: 1(d)) is comparatively more diffused. The bootstrap replications are spread across multiple years, with no single year exhibiting dominant clustering in the frequency distribution.

5.3 Null Monte-Carlo Stress Test Results

The empirical findings reported in Table: 3 indicate substantial divergence between the break-date frequency distributions obtained for Real GDP, Real Revenue Expenditure, and Real Capital Expenditure and break-date frequency distribution generated from the null placebo series constructed to contain no break in either intercept or slope. In contrast, Real Total Expenditure exhibits minimal divergence with its frequency

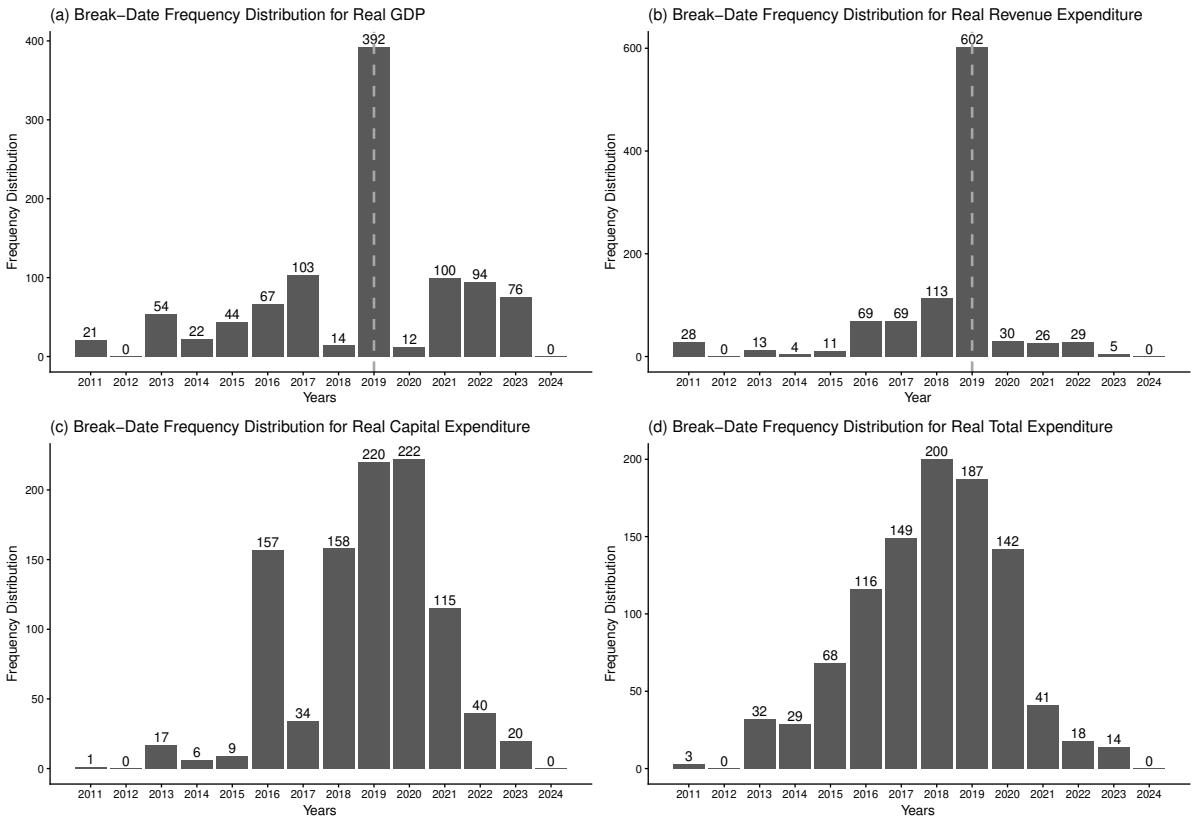


Fig. 1 Application of Ensemble Break Diagnostic (EBD) on Real Expenditure Components and Real GDP of India. Panels illustrate results for:(a) Real GDP, (b) Real Revenue Expenditure,(c) Real Capital Expenditure, (d) Real Total Expenditure. Notes: Year 2012 and 2024 omitted due to loss of observation

Table 3 Null Monte-Carlo Stress Test

SERIES	JSD VALUE[1]	SNR[2]	HIGHEST FREQUENCY BREAK YEAR	DOMINANCE SHARE[3](%)
Real GDP	0.1681126	28.47159	2019	39.23%
Real Revenue Expenditure	0.1986975	33.65147	2019	60.26%
Real Capital Expenditure	0.1340756	22.70709	2020	22.22%
Real Total Expenditure	0.0231613	3.922661	2018	20.02%

JSD VALUE: represents jensen-shannon divergence value between the empirical break date frequency distribution(f_{emp}) and null break date frequency distribution(f_{null}). SNR (SIGNAL-TO-NOISE RATIO): is calculated as the ratio between JSD($f_{emp}||f_{null}$) and $JSD_{BASELINE} = 0.005831868$. DOMINANCE SHARE(%): refers to concentration of highest frequency break year acrross 999 bootstrap replications.

distribution closely aligning with the null placebo series.

This divergence is reflected in the Jensen-Shannon Divergence values computed between

the empirical and null frequency distributions. Real GDP (0.1681126), Real Revenue Expenditure (0.1986975), and Real Capital Expenditure (0.1340756) exhibit elevated divergence, while Real Total Expenditure shows minimal divergence (0.0231613), indicating close similarity to the null frequency distribution.

The Signal-to-Noise Ratio (SNR), constructed to quantify break signals relative to algorithmic baseline instability, further reinforce this pattern. Real Revenue Expenditure records the highest SNR (33.65), followed by Real GDP (28.47) and Real Capital Expenditure (22.70). In contrast, Real Total Expenditure displays a low SNR (3.92), consistent with the absence of meaningful structural break signal.

5.4 Artificial Break Diagnostic Results

Figures 2 and 3 present the results from the application of the Ensemble Break Diagnostic (EBD) on the artificially engineered structural break scenarios. The results indicate visible and dominant clustering of break-date frequencies around the imposed shift point when the intercept and/or slope break in a direction opposite to the pre-break trend. This precise identification is highly evident in the (a) Rise & crash (Fig: 3(b)),(b) Surge & Surge (Fig: 3(c)), and (c) Crash & Surge (Fig: 3(d)) scenario for the downward-falling controlled synthetic series (Fig: 3) as well as the (a) Crash & Fall (Fig: 2(a)), (b) Crash & Surge (Fig: 2(b)) , and (c) Rise & Crash (Fig: 2(d)) scenarios for the upward-rising controlled synthetic series (Fig: 2).

Conversely, in scenarios where both the intercept and slope break in the exact same direction as the pre-break trend (e.g., Surge & Surge (Fig: 2(c)) scenario for the upward-rising controlled synthetic series (Fig: 2) and the Crash & Fall (Fig: 3(a)) scenario for the downward-falling controlled synthetic series (Fig: 3)) the EBD does not register a dominant clustering around the imposed shift point. This divergence in clustering pattern suggests that when a series is merely experiencing severe trend-acceleration, the framework may not register such shift as a genuine structural break.

Table 4 Sensitivity of the Framework to Parameter Shift

Upward-Rising Controlled Synthetic Series		
Scenario	Dominance Share (%)	Highest Frequency Observation
Surge & Surge	17.51%	7th
Rise & Crash	93.29 %	6th
Crash & Surge	42.74%	6th
Crash & Fall	77.37%	6th
Downward-Falling Controlled Synthetic Series		
Scenario	Dominance Share (%)	Highest Frequency Observation
Surge & Surge	50.75%	6th
Rise & Crash	83.08%	6th
Crash & Surge	86.28%	6th
Crash & Fall	20.02%	6th

Note: Dominance Share denotes to the share of maximum frequency distribution to the total frequency($n = 999$)

5.4.1 Stability of The Artificial Break Algorithm

Table 5 Stability Diagnostic

Upward-Rising Controlled Synthetic Series		
Scenario	JSD Value[1]	Interpretation
Surge & Surge	0.003669292	High Algorithmic Stability
Rise & Crash	0.000621027	High Algorithmic Stability
Crash & Surge	0.005427287	High Algorithmic Stability
Crash & Fall	0.002929807	High Algorithmic Stability
Downward-Falling Controlled Synthetic Series		
Scenario	JSD Value[1]	Interpretation
Surge & Surge	0.006036794	High Algorithmic Stability
Rise & Crash	0.002273323	High Algorithmic Stability
Crash & Surge	0.004685045	High Algorithmic Stability
Crash & Fall	0.005372539	High Algorithmic Stability

Jensen-Shannon Divergence value less than 0.01 indicate high algorithmic stability.

[1]Jensen-Shannon Divergence between initial (reference) and independent run refer section 4.4.2

The Jensen-Shannon Divergence values reported in Table 5 for all scenarios are well below the adopted threshold level of 0.01. Consequently, this indicates that the frequency distributions generated by the Ensemble Break Diagnostic (EBD) reflect the innate features of the engineered structural break scenarios rather than stochastic noise. This confirms that the artificial break diagnostic algorithm remains highly stable across multiple independent runs.

5.4.2 Quantifying Sensitivity via Jensen-Shannon Divergence

Table 6 Quantifying Sensitivity of the Framework to Artificially Induced Parameter Shift

Upward-Rising Controlled Synthetic Series		
Scenario	JSD Value[1]	Interpretation
Surge & Surge	0.02500345	Minimum Divergence
Rise & Crash	0.6142217	Divergence
Crash & Surge	0.2736684	Divergence
Crash & Fall	0.434734	Divergence
Downward-Falling Controlled Synthetic Series		
Scenario	JSD Value[1]	Interpretation
Surge & Surge	0.2963486	Divergence
Rise & Crash	0.4840979	Divergence
Crash & Surge	0.525285	Divergence
Crash & Fall	0.07565687	Minimum Divergence

[1]JSD value measures divergence between the respective null baseline(($f_{null_{up}}$) and ($f_{null_{down}}$)) with its respective corresponding controlled synthetic time series (section 4.4.4)

JSD Value=0 represent virtually 0 divergence with null and JSD Value=1 represent complete divergence with their respective null

Table 6 presents the Jensen-Shannon Divergence (JSD) values quantifying the divergence between the artificial break scenarios and their respective continuous null baselines. The reported metrics provide compelling evidence supporting the framework's sensitivity to artificially induced structural breaks at observation 6. Specifically, scenarios in which the intercept and/or slope break in a direction opposite to the pre-break trend display substantially high JSD values. This includes the Rise & Crash, Crash & Surge, and Crash & Fall scenarios for the upward-rising controlled synthetic series, as well as the Surge & Surge, Rise & Crash, and Crash & Surge scenarios for the downward-falling controlled synthetic series.

Conversely, in scenarios where the intercept and slope break in the exact same direction as the pre-break trend—symbolizing a severe trend-acceleration, the calculated Jensen-Shannon Divergence values remain notably low. This indicates that these trend driven scenarios remain highly similar to their respective no-break null baselines.

Taking the divergence evidence in Table 6 together with the concentration data reported in Table 4 confirms that the diagnostic framework is highly responsive to genuine structural ruptures, successfully clustering break-date distributions around the precise, artificially induced shift point. Crucially, the data proves that the framework does not misclassify trend-acceleration scenarios as structural breaks.

Taken together, the algorithmic stability diagnostic, the Null Monte Carlo stress test, and the artificial break diagnostic establish a coherent and rigorous computational validation structure, highlighting that the proposed framework is robust to stochastic noise, immune to trend-acceleration misclassification, and yet highly responsive to genuine structural break.

5.5 Decision Based on Computationally Derived Threshold

After executing the computationally intensive(1.5 million regression) 100-iteration Monte-Carlo simulation of the algorithm(as detailed in section 4.6), to stabilize stochastic variation to derive boundaries of structural break. The final "double-lock" thresholds for the signal-to-noise ratio (SNR) and Dominance Share(%) are calculated as follows:

1. $FinalThreshold(SNR) = 29.45913$
2. $FinalThreshold(DominanceShare) = 31.62563\%$

To be successfully diagnosed with a genuine structural break, an empirical time series must strictly exceed both of these computationally derived thresholds. The year 2019 emerges as a critical point of shift for both Real Revenue Expenditure and Real GDP. As reported in Table

Table 7 Decision Table

Series	SNR[1]	Highest Frequency year	Dominance Share(%)[2]	Decision[3]
Real GDP	28.47159	2019	39.23%	weak structural break/structural instability in 2019
Real Revenue Expenditure	33.65147	2019	60.26%	structural break in 2019
Real Capital Expenditure	22.70709	2020	22.22%	no structural break
Real Total Expenditure	3.922661	2018	20.02%	no structural break

[1]**SNR (Signal-to-Noise Ratio)**: is calculated as the ratio between $JSD(f_{emp}||f_{null})$ and $JSD_{BASELINE} = 0.005831868$ [2]**Dominance Share**: refers to share of highest frequency break year across 999 bootstrap replication, [3]**Decision Rule**: structural break is confirmed only after double lock condition:$SNR > FinalThreshold(SNR)$ and Dominance Share> $FinalThreshold(DominanceShare)$

7, the calculated Signal-to-Noise Ratio (33.65147) and Dominance Share (60.26%) for Real Revenue Expenditure both strictly exceed the computationally derived thresholds. Consequently, this series satisfies the framework's "double-lock condition," providing strong evidence consistent with a structural break in 2019.

In contrast, Real GDP failed to satisfy the strict double-lock condition. While its Dominance Share (39.23%) successfully crossed the final threshold, its Signal-to-Noise Ratio (28.47159) missed the critical diagnostic boundary (29.46) by a narrow margin of approximately one point. Accordingly, the study finds that while Real GDP did not experience a complete a structural break, it is diagnosed with strong volatility and severe structural instability in 2019. This diagnosis of a "weak structural break" is confirmed by the near-miss of the SNR threshold coupled with a dominance share that remained well above the final threshold for dominance share.

Finally, the study found no evidence of a structural break in either Real Total Expenditure or Real Capital Expenditure, as both series failed to cross both the final thresholds.Furthermore, to verify that this strict double lock condition successfully isolate structural break from the stochastic noise across diverse data generating processes, a comprehensive 20-country global cross validation of this diagnostic framework is provided in Appendix C.

6 Policy and Economic Interpretation

6.1 The 2019 Point of Shift

The year 2019 emerges as a point of rupture in the data generation process for Real GDP and Real Revenue Expenditure. The diagnostic framework detects evidence consistent with a structural break in the Real Revenue Expenditure around 2019-20, as the series successfully satisfies the computationally derived double-lock condition. However, Real GDP failed to cross both required thresholds (28.47159 <29.45913 and 39.23% >31.62563%). Panels (a) and (b) of Figure: 4 illustrate visible change in intercept and slope, marked by the dashed vertical line between 2019 and 2020.

For Real GDP(Fig 4(a)), the post 2019 period reflects an initial downward shift in the intercept followed by recovery characterized by a steeper growth trajectory. In contrast Real Revenue Expenditure (Fig: 4(b)) exhibits an upward intercept shift, accompanied by increased variability post 2019 period.

Importantly, the identified point of shift in 2019 precedes the full macroeconomic impact of COVID-19 pandemic in 2020, during which India experienced a marked contraction in GDP alongside expanded public expenditure. The temporal alignment between the detected point of shift and

the pandemic period is consistent with the interpretation that COVID-19 constituted a major exogenous shock affecting the underlying data generation process. However, given the diagnostic nature of the diagnostic framework, this interpretation reflects temporal correspondence rather than formal causal estimation.

6.2 Trend Acceleration in Real Capital Expenditure

Real Capital Expenditure did not satisfy the strict double lock condition established to diagnose structural break. However, the Signal-to-Noise Ratio (22.70709) calculated for the series highlights a notable divergence from the null placebo series, even though it did not exceed the computationally derived diagnostic threshold. This case therefor represents empirical evidence of a "trend-acceleration" present in the data generating process of Real Capital Expenditure.

Panel (c) of the Figure 4 illustrates this trend acceleration. The series follows a rapid growth trajectory, characterised by a steeper slope post 2019. This accelerated momentum aligns fundamentally with the government's push for infrastructure during the period. Real Capital Expenditure displays a transition from a mild upward drift into a higher growth regime post 2019, reflecting amplified policy momentum rather than a structural break.

6.3 Real Total Expenditure

The study finds no evidence of a structural break in Real Total Expenditure, as the results of the null Monte-Carlo Stress Test indicate close alignment with the null placebo series . Panel (d) of Figure 4 illustrates this relative stability in the series. Although Real Total Expenditure exhibits a pronounced increase in 2020, the subsequent trajectory remains aligned with the broader pre-existing trend, underscoring the relative stability of the underlying data-generating process.

7 Methodological Implications

This study presents several methodological implications for structural break analysis in a short-sample and data-constrained environment. These implications result from a cross-disciplinary synthesis that integrates information-theoretic divergence measures, entropy-based resampling, ensemble learning principles, signal processing-based diagnostics, and minimax decision theory under uncertainty. Each component is drawn from an established discipline outside conventional statistics and econometrics to address specific limitation of existing procedure under small sample conditions. Taken together, this integration establishes a unified computational architecture whose constituent parts may independently offer solutions to related problems in applied time-series analysis.

First, the analysis adopts a diagnostic perspective by moving beyond reliance on a single point estimate of the structural break date. In small-sample settings, conventional structural break procedures may lack sufficient statistical power to reliably identify the presence and precise timing of a structural break. To address this limitation, the study maps the distribution of break dates across resampled series generated using Maximum Entropy Bootstrap procedure. This approach preserves the inherent serial dependence and trend characteristics of the original data- generating process while embedding the Zivot-Andrews Unit Root Test within the resampling framework to formulate Ensemble Break Diagnostic (EBD). Under this setup, a genuine structural break is expected to produce a pronounced clustering of break-date frequencies around a particular year rather than appearing as a dispersed frequency distribution.

Second, to ensure the stability of the diagnostic framework , the study draws on concepts from the information theory by adopting Jensen-Shannon Divergence to measure divergence between the independent runs of the diagnostic algorithm. A tolerance threshold of 0.01 is adopted to verify that repeated executions of the Ensemble Break Diagnostic (EBD) produce break-date frequency distribution consistent with the underlying data generation process, thereby ensuring algorithmic robustness.

Third, the analysis constructs a deterministic series with no structural break in either intercept and/or slope to serve as a benchmark for comparison. Break-date frequency distributions generated by this series are compared with empirical frequency distribution using Jensen-Shannon Divergence to quantify the strength of the structural break signal. The study further introduces

a Signal-to-Noise Ratio as an interpretative scale, which is combined with the dominance share (the percentage of concentration of break-date frequencies for the highest frequency break year) to diagnose the presence of structural breaks.

Fourth, a sensitivity experiment is conducted by artificially generating two controlled synthetic time series and subsequently introducing distinct structural break scenarios (e.g., Surge & Surge, Rise & Crash, Crash & Surge, Crash & Fall) into the intercept and slope of each controlled synthetic series. These series are processed through the Ensemble Break Diagnostic (EBD) to generate break-date frequency distribution. The study then generates two baseline null without structural breaks, which are utilized as benchmarks for the Jensen-Shannon Divergence to measure the divergence between the two null baseline and distinct break scenarios. The resulting break-date frequency distributions exhibits distinct clustering around the artificially induced break points. Coupled with high Jensen-Shannon Divergence values, this confirms that the framework is highly sensitive to structural break within the data-generating process.

Fifth, the study generates precise thresholds for the Signal-to-Noise Ratio (SNR) and Dominance Share (%) by embedding these calculation directly within the artificial break diagnostic algorithm. To establish these boundaries, the analysis utilizes Wald's Minimax Decision Theory, extracting the maximum value from the worst-case scenario and minimum value from the best-case scenario to mathematically derive a midpoint that minimizes the error. The study then subjects this algorithmic logic to a computationally intensive, 100 iteration Monte-Carlo simulation encompassing over 1.5 million distinct regression. This extensive iteration process stabilizes stochastic variation, allowing the study to computationally derive a stable and objective thresholds for SNR and Dominance Share (%).

Taken together, these methodological implications open new avenues for structural break analysis in short-sample macroeconomic time series.

8 Limitations and Future Research

To address the severe limitations imposed by relatively small sample sizes, this study developed a computational diagnostic framework to examine structural breaks in small time series. While the proposed framework successfully detects shifts in the parameters of the underlying data-generating process, the study acknowledges several limitations that provide critical avenues for future research.

First, the framework is designed primarily as a diagnostic tool, it does not provide formal causal attribution for the identified structural break. Second, because the study integrates the Zivot-Andrews unit root test with the Maximum Entropy Bootstrap to diagnose structural shifts within the empirical time series, the current version framework is restricted to identifying a single structural break.

Third, the analysis derives its thresholds through a computationally intensive Monte Carlo simulation. Consequently, the derived thresholds are applicable specifically for an $n = 14$ observational point series using standard Gaussian noise structures, even though the underlying theoretical framework remains uniformly applicable.

Finally, the deriving and recalibration of these thresholds is constrained by the severe computational intensity of the algorithm. The nested simulation setup required to derive these thresholds encompassed more than 1.5 million distinct regressions, demanding more than 25 hours of continuous computation on a standard personal computer for just 100 iterations. Consequently, testing the framework under alternative observational lengths, distinct stochastic noise structures (e.g., non-Gaussian residuals), or more expansive calibrated directional parameter shifts presents a significant bottleneck for standard computational solutions.

Future research should utilize High-Performance Computing (HPC) solutions to expand the simulation space. Enhanced computational strength would allow researchers to recalibrate and publish a generalized table of statistically significant Signal-to-Noise Ratio (SNR) and Dominance Share (%)thresholds for diverse, small-sample sizes (e.g., $n = 10$ to $n = 50$).

Furthermore, extending the diagnostic framework to accommodate multiple break identification remains a highly promising frontier for applied time-series econometrics.

9 Conclusion

The proposed computational diagnostic framework provides researchers working with relatively small sample sizes with a tailored diagnostic tool for detecting structural breaks without violating standard econometric assumptions. The study adopts a principled approach that prioritizes the homogeneity of the data-generating process over the splicing of series belonging to different data-generating regimes. The main contribution of this analysis lies in pivoting from causal attribution towards a diagnostic examination of structural breaks, while integrating elements from information theory, standard econometric procedures, resampling methods, and Wald's Minimax Decision Theory within a unified framework.

The study maps the frequency distribution of break dates by embedding the Zivot-Andrews Unit Root Test within a Maximum Entropy Bootstrap procedure. To guarantee the reliability of the framework, a rigorous multi-layered computational validation structure is designed. This involves first constructing an Algorithmic Stability Diagnostic utilizing the Jensen-Shannon Divergence. The second layer consists of a Null Monte Carlo Stress Test, benchmarking empirical time series against a null series generated to contain no break in either intercept or slope. By applying a derived Signal-to-Noise Ratio (SNR), the procedure successfully distinguishes genuine structural breaks from bootstrap-induced frequency dispersion. The third and final layer consists of an Artificial Break Diagnostic that confirms the framework is highly responsive to true parameter shifts in the data-generating process.

The application of this computational diagnostic framework to India's macro-fiscal time series (2011–2024) yields precise insights. The results indicate strong evidence for a definitive structural break around 2019 for Real Revenue Expenditure, meanwhile diagnosing a weak structural break or structural instability—in Real GDP during the same period. In contrast, Real Capital Expenditure reveals characteristics in conformity with trend acceleration. Ultimately, both Real Capital Expenditure and Real Total Expenditure failed to satisfy the strict double-lock conditions established by the framework to be formally diagnosed with a structural break.

Ultimately, this paper provides a reproducible toolkit for applied researchers navigating the fundamental bias-variance trade-off in empirical macroeconomics. While future comprehensive studies utilizing High-Performance Computing (HPC) solutions will universally formalize the framework—by recalibrating the parameters under different stochastic noise structures and widening the simulations to include a broad range of observational lengths—the current framework stands as a powerful proof of concept and a robust diagnostic alternative. It allows policymakers and researchers to systematically evaluate structural breaks, instability, and the reorientation of macro-fiscal series without unintentionally violating the assumptions of standard econometric inference.

10 Computational Framework and Replication

The complete algorithmic framework, including data processing, threshold derivation, multi-layered validation, artificial break diagnostics, and the multi-country analysis is implemented in R and made publicly available at:

https://github.com/vilaschandran1/Computational-Diagnostic-Framework-for-Detecting-Structural-Break.git

The repository contains three primary execution scripts.

File 1 (01_Main_Indian_Macro_Fiscal_Analysis.R) contains the code for the primary Indian macro-fiscal application, artificial break diagnostics, and algorithmic stability checks.

File 2(02_Threshold_Derivation.R) executes the robust Monte Carlo threshold derivation. This simulation is highly computationally intensive and requires an estimated runtime of 25+ hours on standard hardware.

File 3(03_Multi_Country_Analysis.R) extends the framework's application across 20 distinct international economies, allowing researchers to fully reproduce the global out-of-sample results seamlessly.

Declaration

- **Competing Interests:** The authors declare no competing interests
- **Author Contributions: V.C.** contributed to the development and conceptualization of the methodology, implementation of algorithm, threshold derivation, formal analysis and preparation of original draft. **S.T.** provided supervision to the project and contributed in the process of review and editing of the draft. All authors reviewed and approved the final manuscript
- **Funding:** V.C. gratefully acknowledges the financial support provided by the National Backward Classes Finance & Development Corporation (NBCFDC), Govt of India, under Senior Research Fellowship scheme. S.T. declares that no independent funding was received for his contribution to this manuscript.
- **Data Availability:** The data used for the primary Indian macro-fiscal series are publicly available and can be accessed from the official databases of the Reserve Bank of India (RBI) and the Ministry of Finance, Govt of India. The data for the out-of-sample multi-country analysis were extracted from the World Bank's World Development Indicators (WDI) database using the `WDI` package in R Studio.
- **Code availability:** The complete algorithmic framework and execution scripts are publicly available in the project's GitHub repository: https://github.com/vilaschandran1/Computational-Diagnostic-Framework-for-Detecting-Structural-Break.git

Appendix A Algorithmic Steps for Maximum Entropy Bootstrap

This study utilizes the Maximum Entropy Bootstrap (MEB) procedure outlined by Vinod and Lopez-de Lacalle (2009) to generate 999 resampled series. Let x_t be an observed chronological time series of length T. The algorithm proceeds as follows:

1. Step 1: Sort the original chronological x_t in a increasing order to create the sorted order statistics, denoted as y_t. Store the original chronological ordering index vector.

2. Step 2: Compute the intermediate points z_t for the sorted series using the simple average of consecutive order statistics:$z_t = 0.5(y_t + y_{t+1})$ for $t = 1, ..., T-1$

3. Step 3: To establish the lower and upper limits for the intermediate point (z_t), a trimmed mean (m_{trm}) is derived from the deviations of the original chronological series$(|x_t - x_{t-1}|)$. The lower limit is defined as $z_0 = y_t - m_{trm}$ and upper limit as $z_T = y_T + m_{trm}$,which are then alongside the intermediate points to construct a continuous sequence of T half-open intervals: $I_1 = (z_0, z_1], I_2 = (z_1, z_2], ..., (z_{T-1}, z_t]$

4. Step 4: To satisfy the mean-preserving constraint ensuring that mean (m_t) of each interval I_t equals weighted sum of the sorted data-weights from the set{0.25,0.50,0.75} are applied to satisfy the following relations:

$m_1 = 0.75y_1 + 0.25y_2$, for the lowest interval.

$m_k = 0.25y_{k-1} + 0.50y_k + 0.25y_{k+1}$, for $k = 2, ..., T-1$

$m_T = 0.25y_{T-1} + 0.75y_T$, for the highest interval.

5. Step 5: Draw T random numbers from the uniform interval $[0, 1]$ and sort them in ascending order. To compute the sample quantiles for the Maximum Entropy density, apply piecewise linear interpolation across the intervals I_t, rigidly shifting the natural interval bounds to ensure their centers strictly satisfy the mean(m_t) calculated in Step 4.

6. Step 6: Reorder the sorted sample quantiles by using the ordering index vector preserved in Step 1. This process retrieve the time-dependence characteristics and structural shape of the original observed series x_t.

7. Step 7: Repeat Steps 2 through 6 for J replications (e.g.,J=999) to generate the final ensemble of bootstrapped time series.

Appendix B Additional Results from Threshold Simulation

To add robustness to the framework additional stability measures were incorporated into the 100 iteration nested simulation of the algorithm designed specifically for the threshold derivation.

B.1 Hit Rates

To measure whether the algorithm successfully detected the artificially imposed structural break exactly at the imposed shift point across every iteration, the study embedded a process to Hit Rates from the simulation.

The Hit Rate for each scenario (i) is calculated as the percentage of iterations where the maximum peak of the break-date frequency distribution (argmax) perefectly aligned with the imposed structural break (T_B):

$$HitRate_i(\%) = \frac{\sum_{j=1}^{100} I(argmax(f_{art_{i,j}}) = T_B)}{100} \times 100 \quad \text{(B1)}$$

where:

1. $i \in 1, 2, 3, 4, 5, 6, 7, 8$ represent structural break scenario.

2. $j \in 1, 2, 3, ...100$ represents the specific simulation iteration.

3. $T_B = 6$ represents the artificially imposed break point at observation 6.

4. $I(.)$ is an indicator function that equals 1 if the algorithm correctly identifies T_B as the highest frequency break date and 0 otherwise.

This procedure provides further insight into how the diagnostic framework behaved under the stochastic variation of the nested simulation. The results reported in Table B1 highlight the algorithm's precise ability to distinguish between genuine structural breaks and trend accelerations. Across the 100-iteration simulation, the framework successfully identified the parameter shift at the exact artificially imposed breakpoint (100% Hit Rate) for all scenarios where the intercept and/or slope broke in the opposite direction of the pre-break trend.

Table B1 Hit Rate (%) Results

Upward-Rising Controlled Synthetic Series	
Scenario	Hit Rate (%)
Surge & Surge	42%
Rise & Crash	100%
Crash & Surge	100%
Crash & Fall	100%
Downward-Falling Controlled Synthetic Series	
Scenario	Hit Rate (%)
Surge & Surge	100%
Rise & Crash	100%
Crash & Surge	100%
Crash & Fall	81%

Conversely, in scenarios exhibiting an acceleration of pre-existing trends, the framework did not universally lock onto the imposed shift point. Specifically, the hit rate dropped to 42% for the Surge & Surge scenario within the upward-rising controlled synthetic series, and to 81% for the Crash & Fall scenario within the downward-falling controlled synthetic series. This confirms the diagnostic robustness of the framework, as it actively avoids misclassifying momentum-driven trend accelerations as definitive structural ruptures.

B.2 Relative Variation from the Baseline

In the reference run for the artificial break diagnostic, the scenarios representing mere trend acceleration (e.g., Surge & Surge for the Upward-Rising controlled synthetic series, and Crash & Fall for the Downward-Falling controlled synthetic series) produced rather subdued clustering of break-date frequencies, whereas genuine structural break scenarios generated dominant clustering. The Hit Rate (%) detailed in Table B1 further highlights the framework's ability to distinctly differentiate between trend acceleration and true structural breaks.

To ascertain how the nested simulation affects this frequency distribution and to add an additional layer of validation to the framework's capabilities, the study incorporated the Mean Dominance Share (MDS) within the threshold derivation simulation. This metric represents the expected mean of the maximum dominance share for each artificial break scenario across all 100 iterations. The Mean Dominance Share for each scenario (i) is formalized as:

$$MDS_i = \frac{1}{100} \sum_{j=1}^{100} \left(\frac{max(f_{art_{i,j}})}{999} \times 100 \right) \quad \text{(B2)}$$

Where:

1. $i \in 1, 2, 3, 4, 5, 6, 7, 8$ represents structural break scenario.

2. $j \in 1, 2, 3, ...100$ represents the specific simulation iteration.

3. $\max(f_{art_{i,j}})$ represents the maximum break-date frequency peak observed in iteration j for scenario i (out of the 999 Maximum Entropy Bootstrap replications).

This Mean Dominance Share is subsequently compared against the reference Dominance Share

(DS_{art_i}) initially detailed in Table 4. A close alignment between the simulated mean and the reference baseline confirms the computational robustness of the framework. Crucially, this mathematically validates the underlying logic utilized to establish the maximum worst-case and minimum best-case scenarios used for the formal derivation of the final diagnostic thresholds.

Table B2 Relative Variation from Baseline

Upward-Rising Controlled Synthetic Series		
Scenarios	MDS_i	Dominance Share(DS_i[1])
Surge & Surge	17.67568%	17.51751%
Rise & Crash	92.28529%	93.29329%
Crash & Surge	42.65065%	42.74274%
Crash & Fall	74.70370%	77.37737%
Downward-Falling Controlled Synthetic Series		
Scenarios	MDS_i	Dominance Share(DS_i)[1]
Surge & Surge	53.97698%	50.75075%
Rise & Crash	82.58559%	83.08308%
Crash & Surge	84.98498%	86.28628%
Crash & Fall	20.58358%	20.02002%

[1]Dominance Share denotes to the share of maximum frequency distribution to the total frequency($n = 999$) detailed in Table 4

Table B2 reveals a strict alignment between the simulated Mean Dominance Share (MDS_i) and the corresponding reference Dominance Share (DS_i) across all artificial break scenarios. These results provide critical validation for the proposed framework. Crucially, the stochastic variability introduced during the nested simulation did not destabilize the algorithm; the aggregated means across the 100 iteration simulation space remained highly comparable to the initial independent reference points. Consequently, this alignment confirms the computational robustness of the logic utilized for threshold derivation, definitively proving the framework's capacity to reliably diagnose structural breaks and suppress false-positive trend accelerations.

Appendix C Cross-Country Global Validation of the Diagnostic Framework

To evaluate the external validity and generalizability of the proposed computational diagnostic framework, the algorithm was applied to 20 global economies, evenly divided between 10 Developed (High-Income) and 10 Developing (Upper- and Lower-Middle-Income) nations as classified by the World Bank.

The cross-country validation follows the methodological precedent set by the primary analysis, strictly bounding to an n=14 observational period (2011–2024) utilizing pure Real GDP data measured in Constant Local Currency Units (LCU). The data was extracted directly from the World Bank's World Development Indicators (WDI) database (Arel-Bundock, 2026). Each series was processed through the Ensemble Break Diagnostic (EBD) (999 replications) and evaluated against the computationally derived thresholds established in Section 4.5: a Signal-to-Noise Ratio (SNR) threshold of 29.45913 and a Dominance Share (%) threshold of 31.62563%.

C.1 Discussion of Cross-Country Findings

The results reported in Tables C3 and C4 confirm the validity of the proposed framework and its precise identification of structural breaks. As observed in the primary Indian dataset, the year 2019 emerges as a massive, globally synchronized structural break for 11 of the 20 evaluated economies, preceding the full impact of the 2020 COVID-19 pandemic. This temporal alignment with the pandemic period supports the interpretation that COVID-19 caused a major shock affecting the parameters. However, following the precedent established in the primary study, this interpretation only reflects temporal correspondence rather than formal causal attribution. Economies such as Mexico, South Africa, the UK, France, and Japan have shown high SNR and Dominance Share values, confirming the severity of the parameter shift in their Real GDP.

Table C3 Diagnostic Results for Developed (High-Income) Economies

Country	SNR	Dominance Share (%)	HighestFrequency Break Year	Decision[1]
United States	14.11124	24.62462%	2019	No Structural break
United Kingdom	171.5235	94.59459%	2019	Structural Break in 2019
Germany	115.08	55.15516%	2019	Structural Break in 2019
France	241.1472	83.58358%	2019	Structural Break in 2019
Canada	59.3016	77.07708%	2019	Structural Break in 2019
Japan	262.8055	89.78979%	2019	Structural Break in 2019
Australia	23.20369	23.32332%	2019	No Structural break
South Korea	10.41248	19.71972%	2019	No Structural break
Switzerland	32.00758	44.64464%	2019	Structural Break in 2019
Singapore	72.9429	48.74875%	2019	Structural Break in 2019

[1]Decision Rule requires satisfying the double-lock condition: SNR >29.45913 and Dominance Share(%) >31.62563%

Crucially, the framework successfully avoided uniformly diagnosing 2019 as a universal structural break. Economies characterized by robust internal momentum or distinct policy support mechanisms (such as the United States, South Korea, China, Vietnam, Australia, Turkey, Egypt, and Indonesia) generated frequency distributions closely aligned with the null placebo series detailed in Section 4.3. By strictly failing the computationally derived thresholds, the framework correctly diagnosed these trajectories as maintaining baseline continuity, rather than experiencing a fundamental regime shift.

Furthermore, the algorithm successfully isolated domestic shocks independent of the 2019 global pattern. Notably, the framework diagnosed a distinct structural break for Russia in 2017, aligning with the country's specific macroeconomic adjustments due to the oil price collapse in 2014–2015 and subsequent geopolitical sanctions. This specific detection solidifies the framework's sensitivity to localized structural shifts within the strict $n = 14$ constraint.

Table C4 Diagnostic Results for Developing/Transitional Economies

Country	SNR	Dominance Share (%)	HighestFrequency Break Year	Decision[1]
Argentina	1747.77	95.0951%	2019	Structural Break in 2019
Turkiye	11.88358	16.31632%	2019	No Structural Break
Vietnam	12.60178	15.51552%	2017	No Structural Break
Brazil	113.0623	71.37137%	2019	Structural Break in 2019
Mexico	244.5938	93.59359%	2019	Structural Break in 2019
South Africa	442.0411	99.4995%	2019	Structural Break in 2019
Indonesia	8.379531	20.02002%	2019	No Structural Break
Egypt	2.558569	19.01902%	2018	No Structural Break
Russia	125.7757	37.93794%	2017	Structural Break in 2017
China	10.26777	15.51552%	2017	No Structural Break

[1]Decision Rule requires satisfying the double-lock condition: SNR >29.45913 and Dominance Share(%) >31.62563%

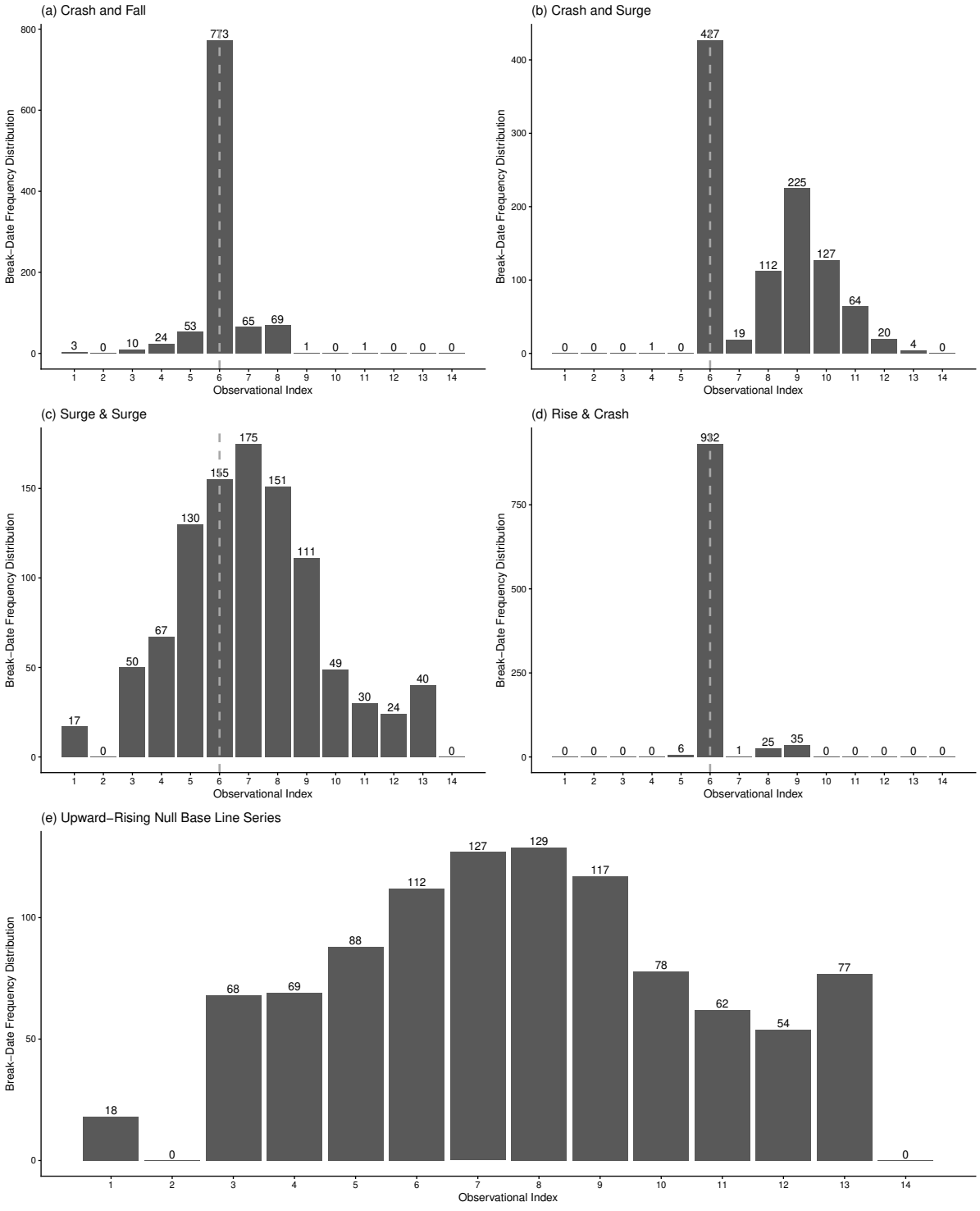


Fig. 2 Artificial Break Diagnostic Results. frequency distribution of break-dates for artificially induced structural break scenarios at observation 6 in a upward-rising controlled synthetic series and upward rising null baseline series, panels represent results for:(a) Crash and Fall, (b) Crash and surge, (c) Surge and Surge, (d) Rise and Crash, (e) upward-rising null baseline series. Notes: 1. year 2012-13 and 2024-25 omitted due to loss of observations,2. Dashed vertical line denotes artificially imposed structural break

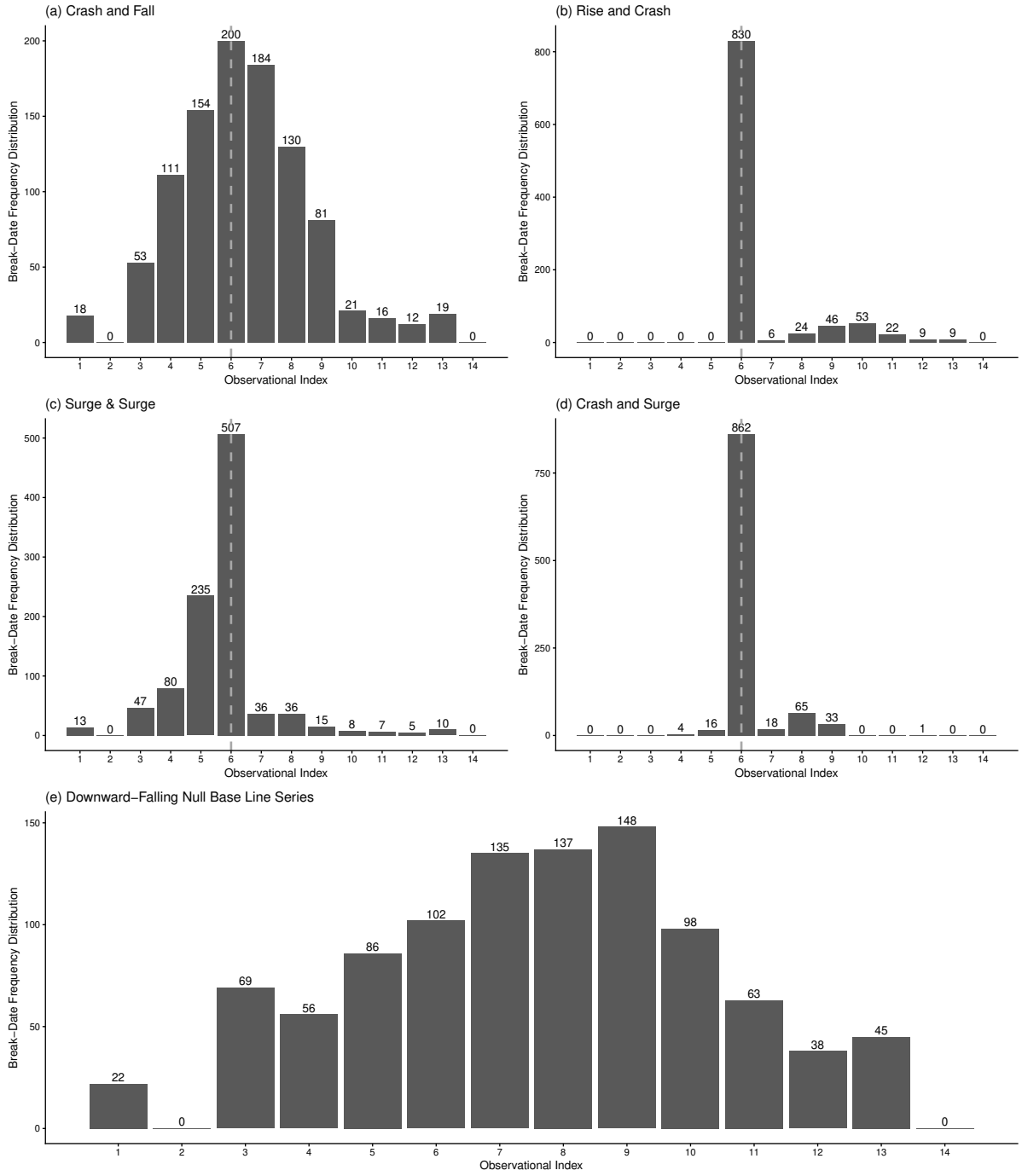


Fig. 3 Artificial Break Diagnostic Results. frequency distribution of break-dates for artificially induced structural break scenarios at observation 6 in a downward-falling controlled synthetic series and downward-falling null baseline series, panels represent results for:(a) Crash and Fall, (b) Rise and Crash, (c) Surge and Surge, (d) Crash and Surge, (e) downward-falling null baseline series. Notes:1. year 2012-13 and 2024-25 omitted due to loss of observations,2. Dashed vertical line denotes artificially imposed structural break

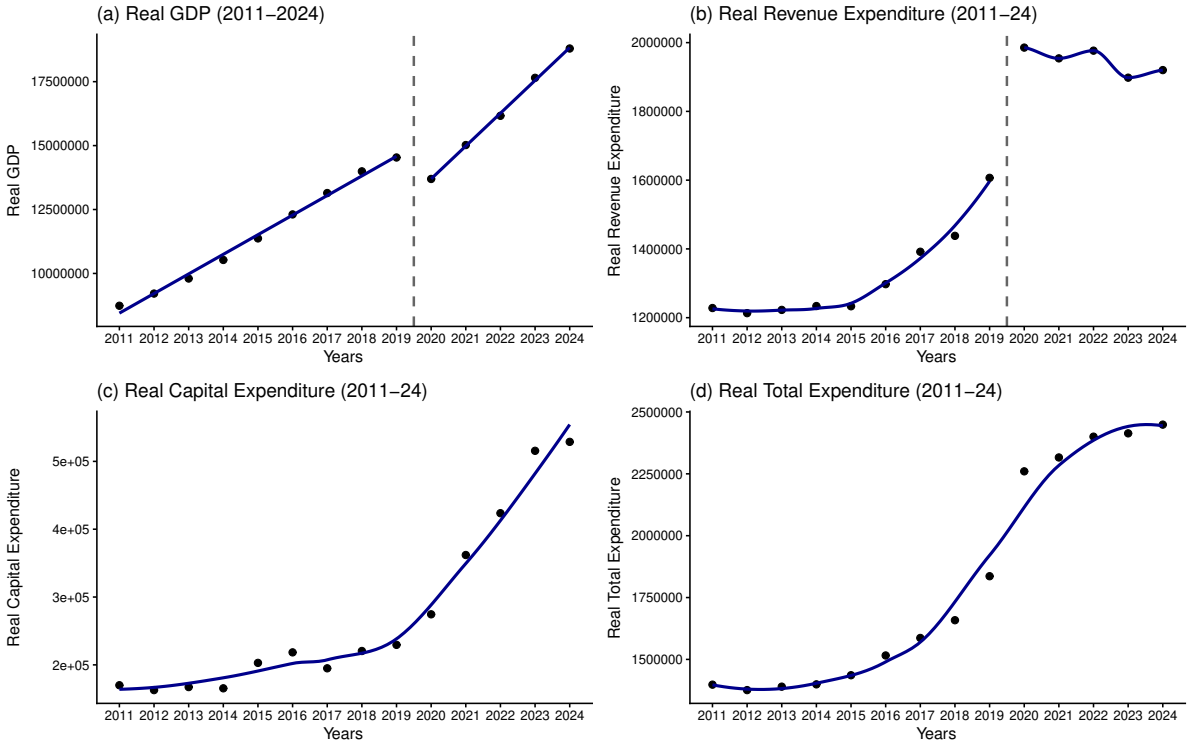


Fig. 4 Real GDP and Real Public Expenditure Components(2011-2024), Panels depicts trajectories for:(a) Real GDP, (b) Real Revenue Expenditure,(c) Real Capital Expenditure, (d) Real Total Expenditure, Dashed Vertical line represent diagnostically detected break for Real Revenue Expenditure and weak structural break for Real GDP

References

Arel-Bundock, V.: WDI: World Development Indicators and Other World Bank Data. R package Version 2.7.10 https://CRAN.R-project.org/package=WDI (2026)

Bose, S., Bhanumurthy, N.R.: Fiscal Multipliers for India. National Institute of Public Finance and Policy. Working Papers 13/125 https://RePEc:npf:wpaper:13/125 (2013)

Bergamelli, M., Novotný, J., Urga, G.: Maximum non-extensive entropy block bootstrap for non-stationary processes. L'Actualité Économique, **91(1-2)**, 115–139 (2015) https://doi.org/10.7202/1036916ar

Bai, J., Perron, P.: Estimating and testing linear models with structural changes. Econometrica **66(1)**, 47–78 (1998) https://doi.org/10.2307/2998540

Efron, B.: Bootstrap methods: Another look at the jackknife. The Annals of Statistics **7(1)**, 1–26 (1979). http://www.jstor.org/stable/2958830

Lin, J.: Divergence measures based on the shannon entropy. IEEE Transactions on Information Theory **37**, 145–151 (1991) https://doi.org/10.1109/18.61115

Lucas, R.E.: Econometric policy evaluation: A critique. Carnegie-Rochester Conference Series on Public Policy **1**, 19–46 (1976) https://doi.org/10.1016/S0167-2231(76)80003-6

Nelson, C.R., Plosser, C.I.: Trends and random walks in macroeconomic time series: Some evidence and implications. Journal of Monetary Economics **10(2)**, 139–162 (1982) https://doi.org/10.1016/0304-3932(82)90012-5

Finance., M.: Economic Survey 2024-25. Government of India, New Delhi (2025)

India, R.B.: Handbook of Statistics on Indian Economy. Reserve Bank of India, Mumbai (2025)

Perron, P.: The great crash, the oil price shock, and the unit root hypothesis. Econometrica **57(6)**, 1361–1401 (1989) https://doi.org/10.2307/1913712

Singh, K.: On the asymptotic accuracy of efron's bootstrap. The Annals of Statistics **9(6)**, 1187–1195 (1981) https://doi.org/10.1214/aos/1176345636

Srinivasan, P.: Causality between public expenditure and economic growth: The indian case. International Journal of Economics and Management **7(2)**, 335–347 (2013). https://ssrn.com/abstract=2376143

Tulsidharan, S.: Government expenditure and economic growth in india (1960-2000). Finance India **20(1)**, 169–179 (2006). https://api.semanticscholar.org/CorpusID:151221600

Vinod, H.D.: Maximum entropy ensembles for time series inference in economics. Journal of Asian Economics **17(6)**, 955–978 (2006) https://doi.org/10.1016/j.asieco.2006.09.001

Vinod, H.D., Lopez-de-Lacalle, J.: Maximum entropy bootstrap for time series: The meboot r package. Journal of Statistical Software **29(5)**, 1–19 (2009) https://doi.org/10.18637/jss.v029.i05

Wald, A.: Statistical Decision Functions. John Wiley & Sons, New York (1950)

Zivot, E., Andrews, D.W.K.: Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. Journal of Business & Economic Statistics **10(3)**, 251–270 (1992) https://doi.org/10.2307/1391541