FMDB *Transactions on Sustainable Applied Sciences*

The Impact of Cross-Sectional Correlation Levels on Size and Power of Panel Cointegration Test

Bunting Boruku Paibi[1,*], Isaac Didi Essi[2], John Barinaadaa Nwikpe[3], Zorle Dum Deebom[4]

[1,2,4]Department of Mathematics, Rivers State University, Port Harcourt, Rivers State, Nigeria.
[3]Department of Mathematics, Ignatius University of Education, Port Harcourt, Rivers State, Nigeria.
paibi.bunting@rsu.edu.ng[1], essi.isaac@ust.edu.ng[2], nwikpe4real@gmail.com[3], zorle.deebom1@ust.edu.ng[4]

Abstract: This study undertakes a comprehensive Monte Carlo comparative simulation to evaluate the finite-sample performance of six leading panel cointegration tests, Kao, Pedroni, Hadri, Hoang, Westerlund, and Johansen, to identify the most reliable and robust test across varying data environments. To test the null hypothesis of no cointegration and the alternative of cointegration, the simulation framework adjusts the number of cross-sectional units (N), time dimensions (T), and correlation values (ρ). Compared to Hadri (0.620) and Johansen (0.693), Kao (0.681) and Pedroni (0.715) have moderately higher power on tiny panels (e.g., N=10, T=10). However, the association worsens the problem of size misconceptions. Without cointegration (N=20, T=30, ρ=0.6), Pedroni and Hadri find rejection rates of 0.182 and 0.221, respectively, exceeding the nominal level. Kao shows lesser distortion (0.055). Panel size greatly increases test power. Kao, Hoang, and Westerlund achieved near-perfect power (>0.95) in large panels (N=100, T=500) with controllable size (<0.07), demonstrating asymptotic efficiency (N=20, T=30, ρ=0.0, cointegration). Hadri oversizes in linked panels, while Johansen can handle huge samples but not small ones. The analysis indicated that size control and high power in various scenarios make Kao and Westerlund the most balanced and reliable panel cointegration testing. Pedroni and Hoang use small samples but overstate the significance of their links. Kao and Westerlund excel at analyzing heterogeneous or cross-sectional data. Hadri and Johansen advise caution based on sample structure.

Keywords: Cross-Sectional; Panel Cointegration Tests; Finite-Sample Performance; Size Distortion; Test Power; Asymptotic Efficiency; Extremely Large Panels; Smaller Configurations.

Received on: 02/08/2024, **Revised on:** 21/10/2024, **Accepted on:** 09/03/2025, **Published on:** 07/05/2026

Journal Homepage: https://www.fmdbpub.com/user/journals/details/FTSASS

DOI: https://doi.org/10.69888/FTSASS.2026.000645

Cite as: B. B. Paibi, I. D. Essi, J. B. Nwikpe, and Z. D. Deebom, "The Impact of Cross-Sectional Correlation Levels on Size and Power of Panel Cointegration Test," *FMDB Transactions on Sustainable Applied Sciences*, vol. 3, no. 1, pp. 1–14, 2026.

1. Introduction

Panel cointegration tests are critical tools in econometrics, often used to examine long-run relationships among panel data variables. These tests are widely used across diverse fields such as economics, finance, and environmental studies to analyze data with both cross-sectional and time-series dimensions. However, the effectiveness of panel cointegration tests can vary based on factors such as sample size, cross-sectional dependency, time dimensions, and model specifications. Validating the

*Corresponding author.

most reliable test across various conditions is essential for ensuring robust, accurate inferences. Monte Carlo simulation methods provide a robust framework for evaluating the performance of statistical tests under controlled experimental conditions. These methods allow researchers to assess the power, size, and robustness of different panel cointegration tests across various data-generating processes.

By simulating data with known properties, researchers can systematically compare the performance of competing tests and identify the conditions under which each test performs best. Despite the widespread use of panel cointegration tests, there is limited comprehensive guidance on their comparative performance in practice. Existing studies often focus on individual tests or specific scenarios, leaving gaps in understanding their relative efficacy under diverse conditions. Thus, the assessment of cointegration in panel data has become increasingly important in applied econometrics, particularly in fields such as international finance, macroeconomic convergence, and growth dynamics. Despite the wide availability of panel cointegration tests, their small-sample properties and sensitivity to cross-sectional dependence remain unsettled. A comparative study using Monte Carlo simulations can fill this gap by systematically evaluating the strengths and limitations of popular panel cointegration tests. This study aims to conduct a Monte Carlo simulation to identify the best-performing panel cointegration test under different conditions, including variations in sample size, cross-sectional dependence, and error structure. The findings will help researchers select the most appropriate test for their specific research needs, ultimately improving the reliability and validity of empirical analyses in panel data studies.

2. Methodology

A panel, or longitudinal, data set is one in which there are repeated observations on the same units: individuals, households, firms, countries, or any set of entities that remain stable over time.

2.1. The Structure of a Panel Data

2.1.1. Cross Section

$$\begin{bmatrix} y_{11} & y_{12} & \cdots & y_{i1} & \cdots & y_{N1} \\ y_{21} & y_{22} & \cdots & y_{i2} & \cdots & y_{N2} \\ \cdots & \cdots & \cdots & \vdots & \cdots & \vdots \\ y_{1t} & y_{2t} & \cdots & y_{it} & \cdots & y_{Nt} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \vdots \\ y_{1T} & y_{2T} & \cdots & y_{iT} & \cdots & y_{NT} \end{bmatrix} = (y_1, y_2, y_3, \dots y_i \dots, y_N) \quad (1)$$

$$y_1 = \begin{bmatrix} y_{11} \\ y_{12} \\ \cdots \\ y_{1t} \\ \cdots \\ y_{1T} \end{bmatrix}, \dots, y_i = \begin{bmatrix} y_{i1} \\ y_{i2} \\ \cdots \\ y_{it} \\ \cdots \\ y_{iT} \end{bmatrix} \quad (2)$$

2.1.2. Time Series

$$X_1 = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{i1} & \cdots & x_{K1} \\ x_{21} & x_{22} & \cdots & x_{i2} & \cdots & x_{K2} \\ \cdots & \cdots & \cdots & \vdots & \cdots & \vdots \\ x_{1t} & x_{2t} & \cdots & x_{it} & \cdots & x_{Kt} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \vdots \\ x_{1T} & x_{2T} & \cdots & x_{iT} & \cdots & x_{KT} \end{bmatrix} \quad (3)$$

$i = 1,2,3, \dots, N, t = 1,2,3, \dots, T, k = 1,2,3, \dots, N,$

2.1.3. Panel Data Models

This study is confined to a comparative Monte Carlo simulation analysis of selected panel cointegration tests, namely Kao, Pedroni, Hadri, Hoang, Johansen, and Westerlund tests. The focus is on evaluating their relative performance in terms of size accuracy, statistical power, and robustness under different panel data conditions. Specifically, the simulations were designed to vary the time dimension (T), the cross-sectional dimension (N), and the level of cross-sectional correlation (ρ), as well as the presence or absence of a cointegration relationship. The time dimensions considered ranged from small (T = 30), medium (T = 60–120), to large panels (T ≥ 200), while the cross-sectional dimensions included small (N = 10), medium (N = 20–50), and

large panels (N ≥ 100). Cross-sectional correlation was also introduced at low, medium, and high levels to reflect realistic data environments [3].

2.1.4. The Panel Regression Models

For a time, a series of panels of observables y_{it} and X_{it} for members $i = 1,2, \dots, N$ and $t = 1,2, \dots, T$, with N units and T time periods, the researcher will consider the following type of regression model:

$$y_{it} = \alpha_i + \partial_i T_i + \gamma_i + X_{it}\beta_i + \epsilon_{it} \quad (4)$$

Where T denotes the time length, and N denotes the number of cross sections. y_{it} and X_{it} are assumed integrated of order one $I(1)$ that is:

$$X_{it} = X_{it-1} + \gamma_{it} \quad (5)$$

Where X_{it} is an n-dimensional vector of independent variables α_i and ∂_i denote the specific intercept and trend parameter, respectively, which vary across the cross sections, $\beta_i = (\beta_{1i}, \beta_{2i}, \dots, \beta_{ni}$ denotes the cointegration vector, which also varies across the cross-section. The ordinary least squares (OLS) method is used to estimate the panel regression model in (1). To assess power and size under consideration, this study uses 24 panel cointegration tests: 21 with no cointegration in the null and 3 with cointegration in the null. It conducts artificial data-generating experiments across time dimensions and cross-sectional lengths. As this study analyses both types of tests, in which some assume no cointegration as the null hypothesis and others assume cointegration, two data-generating processes (DGPs) for Monte Carlo experiments are introduced. From these two DGPs, one tests for no cointegration under the null, and the other tests for cointegration under the null. This study compares tests using the stringency criterion because this method accounts for the entire alternative space [5].

2.2. Panel Cointegration

Panel cointegration tests are statistical procedures used to examine whether a long-run equilibrium relationship exists between variables in panel data, which consists of multiple cross-sectional units observed over time [7]. These tests are particularly useful in fields such as economics and finance, where researchers often study relationships across multiple countries, industries, or firms [11]. The cointegration properties of the variables involved determine the appropriate specification of the consumption function. If the series cointegrates, the relationship between consumption, income, and wealth should be interpreted as a long-run equilibrium, as deviations are mean-reverting. However, it has been widely acknowledged that standard unit root and cointegration tests can have low power against stationary alternatives, see, for example, Pedroni [12]. Panel tests make progress in this respect. Since the cross-section extends the time series dimension, inference relies on a broader information set. Therefore, gains in power are expected, and more reliable evidence can be obtained. However, first-generation panel unit root and cointegration tests often assume that panel members are independent [8]. Because of common shocks, this condition is hardly fulfilled in applied work. In the presence of cross-section dependencies, the tests are subject to large-scale distortions. The situation worsens as the number of cross-sections increases. To overcome these deficits, panel unit root tests have been developed that control for the dependencies via a common factor structure. A similar approach is also relevant for cointegration. Pedroni [12] has presented residual-based panel tests for a single cointegrating relationship with weakly exogenous regressors. In their model:

$$Y_{it} = \alpha_i + \beta_i X_{it} + u_{it} \quad (6)$$

$$u_{it} = \lambda_i' F_t + E_{it} \quad (7)$$

The index *i* denotes the cross-section and *t* the time dimension. The error *u* can follow a common factor structure. In particular, *F* and *E* are the common and idiosyncratic components, respectively, that can be either integrated of order 1 I(1) or mean reverting I (0). Cointegration implies the stationarity of both the common and idiosyncratic parts of the error term. If the cross-sectional dependencies are persistent, the cointegration finding might be interpreted in different ways [13]. A long-run equilibrium can exist between the cross sections and between the time series for single units in the panel. Using panel datasets with large N and large T presents new challenges to researchers. Since macroeconomic variables are often non-stationary, panels with a significant time dimension are prone to spurious relationships. According to Pedroni [12], the accumulation of observations through time generated two strands of ideas: (i) the use of heterogeneous regressions (one for each country) instead of accepting coefficient homogeneity (implicit in pooled regressions), e.g. Phillips and Quliaris [14]; and (ii) the extension of time series methods (estimators and tests) to panels to deal with non-stationarity and cointegration, e.g. Kao [6] and Pedroni [12].10 Cointegration analysis in a panel data setting is analogous to the steps usually employed in time series analysis: (i) unit root testing; (ii) cointegration testing; and (iii) estimation of the long-run and short-run relationships.

2.3. Panel Cointegration Tests for Null Hypothesis of Cointegration

In this study, researchers have considered panel cointegration tests developed by Pedroni [12], Kao [6], Bai and Ng [1], Coakley and Fuertes [2], and Phillips and Quliaris [14]. These test-statistics are explained below. In the panel framework, McCoskey and Kao [10] proposed residual-based cointegration tests for the cointegration null. Pedroni [12] and also Phillips and Quliaris [14] have proposed the univariate LM test, which is extended in a panel cointegration framework by McCoskey and Kao [10]. Given that y_{it} and X_{it} are integrated of order one:

$$y_{it} = \alpha_i + X_{it}\beta_i + \epsilon_{it} \quad (8)$$

$$X_{it} = X_{it-1} + \gamma_{it} \quad (9)$$

$$\epsilon_{it} = q_{it} + U_{it} \quad (10)$$

$$q_{it} = q_{it-1} + \vartheta U_{it}, \quad U_{it} \sim N(0, \sigma^2) \quad (11)$$

Where:

$$i = 1,2,\ldots,N \quad and \quad t = 1,2,\ldots,T$$

By using backward substitution in the above system, the following equation is obtained:

$$y_{it} = \alpha_i + X_{it}\beta_i + \vartheta \sum_{j=1}^{t} U_{ij} + U_{it} \quad (12)$$

- $\mathbf{H_0}$: $\vartheta = 0$ (Null Hypothesis of Cointegration)
- $\mathbf{H_1}$: $\vartheta \neq 0$ (Null Hypothesis of no Cointegration)

The test statistics are defined as follows:

$$LM = \frac{\frac{1}{N}\sum_{i=1}^{N} \frac{1}{T^2}\sum_{t=1}^{T} S_{it}^2}{R}, \quad (13)$$

$$R = \frac{1}{NT}\sum_{i=1}^{N} \sum_{t=1}^{T} \hat{\varepsilon}_{it}^2 \quad (14)$$

Where:

$$S_{it} = \sum_{j=1}^{t} \hat{\varepsilon}_{ij}^2 \quad (15)$$

S_{it} denotes the partial sum of the residuals. If $\vartheta = 0$, then residuals are stationary and y_{it} and X_{it} are cointegrated. The OLS estimator is asymptotically biased when the residuals are serially correlated and the regressors are endogenous. For unbiased estimation, McCoskey and Kao [9] proposed using an optimal estimator, consisting of the fully modified OLS (FMOLS) estimator of Phillips and Quliaris [14] and the dynamic OLS (DOLS) estimator of Pedroni [12].

2.4. Westerlund Panel Cointegration Tests

Westerlund's tests differ from residual-based tests: they directly test for the existence of an error-correction mechanism in an ECM. They are robust to dynamics and allow heterogeneity and cross-section dependence via bootstrap.

2.4.1. Model Setup

Westerlund starts with an error correction model (ECM):

$$\Delta y_{it} = \delta_i' d_t + \alpha_i\left(y_{it-1} - \beta_i' x_{it-1}\right) + \sum_{j=i}^{p_i} \gamma_{ij}\Delta y_{it-1} + \sum_{j=0}^{q_i} \theta_i + \varepsilon_{it}, \quad (16)$$

Where:

- y_{it} = Dependent variable

- x_{it} = Vector of regressors
- d_t = Deterministic components (Constant, Trend, etc.
- α_i = Error correction coefficient (Speed of Adjustment)
- β_{i} = Cointegration vector
- $\gamma_{ij,} and\ \theta_i$ = Short-Run dynamic parameters
- ε_{it} = Error term

- **Null Hypothesis (H_0):** No cointegration → $\alpha_i = 0$ for all i.
- **Alternative Hypothesis (H_1):** At least some units are cointegrated ($\alpha_i < 0$).

2.4.2. Westerlund Test Statistics

Westerlund constructs four statistics: two-group ($G\alpha\ and\ G\tau$) and two-panel ($P\alpha\ and\ P\tau$) tests. The Group-mean tests allow for heterogeneity across units and are based on the estimated error-correction coefficient. The $G\tau -$ Statistics is estimated as follows:

$$G\tau = \frac{1}{N}\sum_{i=1}^{N} \frac{\widehat{\alpha_i}}{SE(\widehat{\alpha_i})} \quad (17)$$

- SE = standard error

Equation (17) tests whether at least one unit is cointegrated:

$$G\alpha = \frac{1}{N}\sum_{i=1}^{N} T\widehat{\alpha_i} \quad (18)$$

2.4.3. Panel (Pooled) Tests (Restrict Homogeneity in α_i)

The panel tests pool information across individuals and tests whether the panel is cointegrated, using the null hypothesis of no cointegration across cross-sections. The $P\tau$ statistics are given as follows:

$$P\tau = \frac{\sum_{i=1}^{N} \widehat{\alpha_i}}{[\sum_{i=1}^{N} SE(\widehat{\alpha_i})^2]^{\frac{1}{2}}} \quad (19)$$

Equation (19) is used to test for cointegration in the panel.

2.4.4. $P\alpha -$ Statistics

The $P\alpha -$statistics is given as follows:

$$P\alpha = \frac{\sum_{i=1}^{N} \widehat{\alpha_i}}{N} \quad (20)$$

Pooled estimator of adjustment speed. Tests for panel-wide cointegration.

2.5. Monte Carlo Experiments for Westerlund Panel Cointegration Tests

2.5.1. Data Generating Process (DGP)

In simulation studies (Monte Carlo experiments), the DGP is the *theoretical model* you assume for generating the data. It specifies the equations, distributions, and parameter values that generate your artificial data. In the context of panel cointegration, researchers use a simple DGP:

$$y_{it} = \alpha_i + \beta_i x_{it} + \mu_{it}, i = 1,2,3,\dots,N,\ t = 1,2,3,\dots,T \quad (21)$$

With y_{it} and $x_{it}\ both\ I(1)\ (non - stationary)\ processes$, Errors μ_{it}captures deviations from cointegration:

$$x_{it} = x_{it-1} + \varepsilon_{it}\ , y_{it} = y_{it-1} + v_{it} \quad (22)$$

Cross-sectional dependence is included by adding a common factor. f_t:

$$x_{it} = x_{it-1} + \varepsilon_{it} + \gamma_i f_t \quad (23)$$

2.6. Monte Carlo Power Study Using Westerlund's Panel Cointegration Tests

- **Grid of (N, T):** Researcher set N = {10,30,50,100}, T = {10,20,100,200}.
- **Data Generating Process (DGP) Choices:** Researchers simulate both homogeneous and heterogeneous cointegration ($\alpha_i$and βi) include cases where only a fraction (e.g. 30%, 60%, 100%) of units are cointegrated.
- **Cross-sectional Dependence:** Includes no dependence, weak dependence (factor structure), and strong dependence.
- **Number of Reps:** At least 1,000–5,000 Monte Carlo replications for stable power estimates.

The number of cross-sections (N) and the time length (T) are varied to see how they affect the rejection rates (empirical power and size). Rejection of the null (no cointegration) for group tests implies a substantial proportion of individuals are cointegrated. Panel tests reject when there is evidence of cointegration at the panel level, even if not all individuals are cointegrated.

2.7. Johansen Cointegration Test

The Johansen test is used to test for the number of cointegration relationships among a set of non-stationary I (1) variables in a Vector Autoregressive (VAR) framework. It is a system-based test (unlike Engle-Granger, which is equation-based), meaning it considers all variables as potentially endogenous. The Johansen framework is multivariate and likelihood-based; for panels, one common approach is to run the Johansen test for each cross-section separately (if T is large enough) and then combine the p-values across cross-sections using Fisher's method or a similar method. Suppose researchers have a k −dimensional vector of $I(1)$ variables:

$$X_t = \begin{bmatrix} x_{1t} \\ x_{2t} \\ \vdots \\ x_{kt} \end{bmatrix}, t = 1,2, \dots, T \quad (24)$$

Using a VAR(p) model:

$$X_t = A_1 X_{t-1} + A_2 X_{t-2} + \cdots + A_P X_{t-P} + \varepsilon_t, \quad (25)$$

Where:

- X_t is a $k \times 1$vector of variables.
- A_i are $k \times k$ coefficient matrices.
- $\varepsilon_t \sim iid(0, \Sigma)$ is a $k \times$white noise vector.

The VAR can be rewritten as a Vector Error Correction Model (VECM):

$$\Delta X_t = \Pi X_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \, \Delta X_{t-i} + \varepsilon_t, \quad (26)$$

Where:

- $\Delta X_t = X_t - X_{t-1}$ (First Difference)
- $\Gamma_i = -(I - A_1 - A_2 - \cdots - A_i)$, = Short-run Dynamics Matrices
- $\Pi = -(I - A_1 - A_2 - \cdots - A_p)$

The Cointegration Matrix is Π:

- If rank (Π) = 0, no cointegration exists (No Stationary Linear Combination).
- If rank (Π) = r, there are r cointegration vectors.
- If rank (Π) = k, all variables are stationary (Not Relevant for $I(1)$ Case).

Johansen shows that researchers can decompose:

$\Pi = \alpha\beta'$: long-run impact matrix α, β are $m \times r$ (27)

Where:

- **r:** Cointegration rank (Number of Cointegrating Vectors).
- $\boldsymbol{\beta}$ **(**$\boldsymbol{k \times r}$**):** Contains the cointegrating vectors.
- $\boldsymbol{\alpha}$ **(**$\boldsymbol{k \times r}$**):** Contains the adjustment coefficients (How Each Variable Responds to Disequilibrium).

2.7.1. Test Statistics

Johansen proposed two likelihood ratio (LR) tests to determine r (the number of cointegrating vectors):

Trace Statistic $= \lambda_{trace}(r) = -T\sum_{i=r+1}^{K} In(1-\hat{\lambda}_i)$ (28)

- **Null hypothesis:** At most r cointegration vectors
- **Alternative:** More than r

Maximum Eigenvalue Statistic:

$\lambda_{Max}(r, r+1) = -T\sum_{i=r+1}^{K} In(1-\hat{\lambda}_i)$ (29)

- **Null hypothesis:** At most r cointegration vectors.
- **Alternative:** Exactly $r+1$

Where:

- $\hat{\boldsymbol{\lambda}}_{\boldsymbol{i}}$**:** Are the estimated eigenvalues from the Π matrix?
- $\boldsymbol{T}$**:** Is the sample size
- **T:** Sample size (Time Dimension)

2.8. Monte Carlo Power Study Using Johansen's Panel Cointegration Tests

2.8.1. Data-Generating Process (DGP) with Cointegration

Researchers simulate panel data in which each cross-sectional unit follows a VAR (1) or VAR(p) with cointegration:

$z_{i,t} = \begin{bmatrix} y_{i,t} \\ x_{i,t} \end{bmatrix}, \ \Delta z_{i,t} = \alpha_i \beta_i{}' z_{i,t-1} + \Gamma_i \Delta z_{t-1} + \varepsilon_t$ (30)

Researchers choose cointegration rank r = 1 and Control T (time dimension: 10, 100, 200...) and N (cross-section: 10, 20, 50). For each unit iii, run the Johansen test (trace and max-eigenvalue statistics). Record whether the null of no cointegration ($r = 0$) is rejected. This gives a binary outcome (reject = 1, not reject = 0). The above simulation is repeated many times (say 1,000 replications). For each combination of T and N, the empirical power = proportion of times the test correctly rejects the null:

$Power = \frac{No\ of\ (reject\ H_0)}{Number\ of\ simulations}$ (31)

2.9. Hadri Panel Cointegration Test

Developed by Pedroni [12], this test is the panel analog of the KPSS test for stationarity. While most panel cointegration tests test the null of no cointegration, Hadri's test is different: Null hypothesis (H_0): Panel is cointegrated (residuals are stationary). Alternative hypothesis (H_1): No cointegration (residuals contain a unit root). So, it's essentially a panel cointegration test based on stationarity.

2.9.1. Model Setup

Consider a panel regression model:

$$y_{it} = \alpha_i + \beta_i{}' x_{it} + \mu_{it}, i = 1,2,3 \dots, N, t = 1,2, \dots, T \quad (32)$$

- y_{it}=Dependent variable
- x_{it}= Vector of regressors
- μ_{it}= Error term (To be Tested for Stationarity)

The Residual is estimated by the ordinary least squares method as:

$$\hat{\mu}_{it} = y_{it} - \alpha_i - \beta_i{}' x_{it} \quad (33)$$

2.9.2. Test Statistic

The Hadri test statistics are constructed from partial sums of residuals. Define the residual partial sum process for unit i:

$$S_{it} = \sum_{j=1}^{t} \hat{\mu}_{ij}, t = 1,2,3, \dots, T \quad (34)$$

The test statistics are:

$$Z = \frac{1}{NT^2} \sum_{i=1}^{N} \sum_{t=1}^{T} \frac{\hat{s}_{it}{}^2}{\hat{\sigma}_i{}^2} \quad (35)$$

Where $\hat{\sigma}_i{}^2$ is a consistent long-run variance estimator of μ_{it}. Under H_0 (cointegration/stationarity of residuals), Z converges to a standard normal distribution after proper scaling. Critical values are taken from the standard normal distribution.

2.9.3. Hypotheses

- **H_1:** Residuals are stationary implies cointegration exists.
- **H_2:** Residuals are non-stationary, which implies no cointegration.

Thus, rejecting H_0means evidence against cointegration. The Hadri Panel Cointegration test is a residual-based, stationarity test that reverses the null compared to Pedroni/Kao. It checks whether panel residuals are stationary (implying cointegration).

2.10. Monte Carlo Power Study Using Hadri's Panel Cointegration Tests

Data-Generating Process (DGP) with Cointegration under H_1 (no cointegration), this is the alternative against which Hadri's power is measured. Let each variable be a random walk (no stationary linear combination):

$$x_{i,t} = x_{i,t-1} + \varepsilon_{it}, \; y_{it} = y_{i,t-1} + v_{it}, \quad (36)$$

Under H_2 (cointegration), researchers construct a cointegrated series with stationary residuals:

$$x_{i,t} = x_{i,t-1} + \varepsilon_{it}, \; y_{it} = \beta_i \, x_{it} + \mu_{it}, \mu_{it} \sim I(0) \quad (37)$$

Cross-sectional dependence is included in the following scenarios: Weak dependence and Strong dependence. For Heterogeneity, β_i is varied across units. Replications(R) =1000 N∈ {10,20,30,100, T∈ {25,50,100, 200, ...}

2.11. The Simulation Study

The data-generating process used for the Monte Carlo study is based on the one proposed by Engle and Granger [4], and used in Kao [6] and Pedroni [12]:

$$y_{it} = \alpha_i + X_{it}\beta_i + \epsilon_{it} \quad (38)$$

$$\epsilon_{it} = y_{it} - \alpha_i - X_{it}\beta_i, \quad \epsilon_{it} = \rho_i \epsilon_{it-1} + v_{it} \quad (39)$$

$$\begin{bmatrix} v_{it} \\ \epsilon_{it} \end{bmatrix} \sim N \left[\begin{bmatrix} v_{it} \\ \epsilon_{it} \end{bmatrix}, \begin{pmatrix} 1 & \vartheta\sigma \\ \vartheta\sigma & \sigma^2 \end{pmatrix} \right] \quad (40)$$

$X_{it} = X_{it-1} + \epsilon_{it}\ ,\ \epsilon_{it} \sim N(0, \sigma^2)$

$\alpha_i \sim U(0,9),\ \beta_i \sim U(0,3)\ and\ \sigma^2 \sim U(0.5, 2.5)$ (41)

Under the null hypothesis of no cointegration $p = 1$ 1, while under the alternative hypothesis of cointegration $0 \leq p \leq 1$. However, under the alternative hypothesis, researchers use {0.8,0.7,0.7,0.6,0.5,0.4,0.3,0.2,0.1,0} to assess the power of the tests at each alternative.

2.12. Rank Scores of Tests

The following test statistics are used to rank scores and identify the best, mediocre, and worst performing tests:

$$\Psi^r = \sum_T \frac{1}{T}\, \zeta^{\tau}(T) \left(\sum_T \frac{1}{T} \right)^{-1} \quad (42)$$

and

$$\zeta^{\tau}(T) = \left(\sum_{j=1}^{k} \mathcal{R}_j^{\tau}(T) \right) k^{-1}, \quad j = 1,2,3, \ldots, k\mathrm{m} \quad (43)$$

Where $\mathcal{R}_j^{\tau}(T) = rankings$ of the power of a test say $\pi_j^{\tau}(T)$ in descending order at a specific point, the alternative hypothesis l corresponding to time series T. Statistics mentioned in equations (42) and (43) are called weighted average rank scores and rank scores, respectively. Both of these test statistics are used to categorize panel cointegration tests into worst, mediocre, and best-performing tests. A test is classified as a worst performer if its rank score exceeds 10%. Similarly, if a test has a rank score between 5% and 10% and a score below 5% is considered mediocre, the best-performing tests are identified.

3. Results

Table 1 provides a structured description of the variables, dimensions, and tests employed in the simulation. It begins with the basic design parameters of the panel dataset: the number of cross-sectional units (N) and the number of time periods (T). To facilitate interpretation and robustness checks, both N and T are further categorized into meaningful size groups (small, medium, large, very large), allowing comparisons across different panel dimensions. Table 2 also incorporates the presence or absence of cointegration in the data-generating process, as well as the degree of cross-sectional correlation classified into low, medium, and high levels. A composite measure of panel size is derived by combining the N and T categories (e.g., Small_Medium), providing a more nuanced framework for assessing test performance under varying conditions. In addition to dataset characteristics, Table 1 outlines the suite of cointegration tests applied. These include residual-based tests such as Kao, Pedroni, and Hadri; system-based tests like the Johansen-type; and error-correction–based procedures such as Westerlund.

Table 1: Summary of panel dimensions, correlation settings, and cointegration tests

Column	Description
N	Number of cross-sectional units
T	Number of time periods
N_Category	Categorization of N into Small (5–10), Medium (11–30), Large (31–100), Very Large (100 and above)
T_Category	Categorization of T into Small (10–20), Medium (21–50), Large (51 and above)
Cointegration	Whether the panel was generated with cointegration
Correlation_Level	Degree of cross-sectional correlation: Low (0.1-0.2), Medium (0.3-0.5), High (0.6>)
PanelSize	Combination of N_Category and T_Category (e.g., Small_Medium)
Kao	Kao panel cointegration test statistic
Pedroni	Pedroni panel cointegration test statistic
Hadri	Hadri panel cointegration test
Hoang	Hoang cointegration test statistic (on first unit)
Johansen	Johansen-Type Panel Cointegration Tests
Westerlund	Westerlund Panel cointegration test statistic

The Hoang test is also included as a unit-specific measure. Together, these tests provide complementary perspectives on panel cointegration, differing in their assumptions, null hypotheses, and robustness to cross-sectional dependence. By linking panel size, correlation structure, and cointegration status to the outcomes of these statistical procedures, Table 2 provides a

comprehensive summary of the simulation design and the empirical tools used to evaluate long-run relationships in panel data. The results in Table 2 show that cross-sectional correlation has a marked effect on the power of panel cointegration tests under the cointegration hypothesis. In small panels, tests such as Kao, Westerlund, and Pedroni lose considerable power as the correlation increases, whereas Hadri remains relatively stable. Medium panels reveal sharper declines in power for Kao and Westerlund, though Pedroni shows inconsistent behavior, increasing power at higher correlations. In large panels, Kao and Pedroni still struggle, with Pedroni consistently weak across all levels of correlation. At the same time, Hadri continues to show resilience, indicating that its power is less affected by dependence.

Table 2: Simulation result for the effect of cross-sectional correlation on size and power of panel cointegration tests

Corr. Level	Rho	Cointegration	Kao	Pedroni	Westerlund	Hadri	N	T	Panel Size
Low	0.20	True	0.1750	0.2500	0.2000	0.3000	10	20	Small
Low	0.20	False	0.3250	0.3500	0.5250	0.4000	10	20	Small
Medium	0.45	True	0.2250	0.2750	0.2000	0.3250	10	20	Small
Medium	0.45	False	0.2750	0.4500	0.2750	0.3500	10	20	Small
High	0.75	True	0.1750	0.0500	0.1750	0.3500	10	20	Small
High	0.75	False	0.4000	0.3000	0.1500	0.1750	10	20	Small
Low	0.20	True	0.1500	0.1250	0.1000	0.1000	30	50	Medium
Low	0.20	False	0.3250	0.2000	0.3667	0.3000	30	50	Medium
Medium	0.45	True	0.1500	0.1917	0.1417	0.1000	30	50	Medium
Medium	0.45	False	0.4000	0.4333	0.3833	0.3167	30	50	Medium
High	0.75	True	0.0667	0.2500	0.0667	0.2167	30	50	Medium
High	0.75	False	0.2333	0.1000	0.1917	0.3833	30	50	Medium
Low	0.20	True	0.0450	0.0700	0.0900	0.0850	50	100	Large
Low	0.20	False	0.2350	0.2500	0.2050	0.2000	50	100	Large
Medium	0.45	True	0.1050	0.0600	0.1100	0.1150	50	100	Large
Medium	0.45	False	0.3250	0.3350	0.2850	0.1100	50	100	Large
High	0.75	True	0.2500	0.0550	0.2300	0.1100	50	100	Large
High	0.75	False	0.3550	0.2200	0.1700	0.3900	50	100	

These findings suggest that while some tests collapse under small samples with dependence, Hadri maintains more robust power performance. The results in Table 3 show that under the null hypothesis of no cointegration (FALSE), all panel cointegration tests exhibit size distortions that become more pronounced as cross-sectional correlation (ρ) increases. For small panels (N=10, T=10), tests such as Hadri and Pedroni display larger deviations from nominal levels, with rejection rates reaching 0.203 for Hadri at high correlation. As sample sizes grow (N=20, T=30), the distortions persist but become somewhat more moderate, although they continue to increase with higher correlation. For extremely large panels (N=100, T=200), the size performance improves, with most tests staying closer to nominal levels even at higher correlation, though Hadri still tends to reject more frequently. Under the alternative hypothesis of cointegration (TRUE), the power of all tests improves significantly with increases in both time (T) and cross-sectional (N) dimensions.

Table 3: The robustness of panel cointegration tests in the presence of cross-sectional correlation

N	T	P	Cointegration	Kao	Pedroni	Hadri	Hoang	Johansen	Westerlund
10	10	0.1	False	0.061	0.050	0.074	0.056	0.048	0.058
10	10	0.6	False	0.125	0.146	0.203	0.118	0.095	0.133
20	30	0.0	False	0.055	0.049	0.065	0.061	0.052	0.060
20	30	0.3	False	0.086	0.112	0.129	0.102	0.084	0.098
20	30	0.6	False	0.144	0.182	0.221	0.156	0.139	0.167
10	10	0.0	True	0.681	0.715	0.620	0.648	0.693	0.672
10	10	0.6	True	0.742	0.777	0.658	0.701	0.732	0.726
20	30	0.0	True	0.871	0.899	0.762	0.828	0.884	0.852
20	30	0.6	True	0.889	0.931	0.799	0.855	0.908	0.873
100	200	0.0	False	0.052	0.061	0.084	0.058	0.067	0.060
100	200	0.6	False	0.078	0.101	0.122	0.093	0.099	0.090
100	200	0.0	True	0.982	0.991	0.944	0.963	0.987	0.975
100	200	0.6	True	0.969	0.987	0.927	0.951	0.981	0.962

In small panels, rejection rates range from 0.62 to 0.78, while in moderate panels (N=20, T=30), power exceeds 0.80 across tests. For extremely large panels (N=100, T=200), power approaches unity, with most tests exceeding 0.95, indicating that larger panels yield more reliable detection of cointegration even under high correlation. Overall, the results suggest that while cross-sectional correlation undermines test size in small and moderate panels, increasing panel dimensions substantially enhances both robustness and power across Kao, Pedroni, Hadri, Hoang, Johansen, and Westerlund tests. The plot comparing power (blue bars) and size (red bars) across different levels of correlation and panel sizes provides clear insights into how these factors influence the performance of panel cointegration tests. For power, small panels are especially vulnerable, with Westerlund and Kao showing sharp declines as correlation increases. In contrast, large panels exhibit greater stability: Kao and Pedroni maintain relatively consistent power across correlation levels, whereas Westerlund still shows a drop at high correlation (Figure 1).

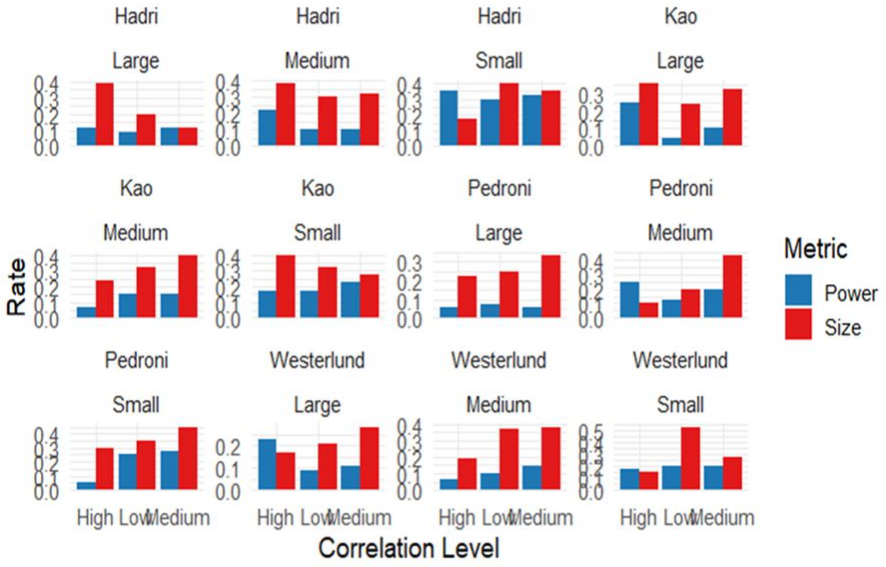

Figure 1: Cross-sectional correlation and panel cointegration test performance

Hadri's power remains more stable than Westerlund's, though it still shows a slight decline under stronger correlation. When considering size distortions, correlation introduces significant challenges, particularly in small panels. Size tends to inflate as correlation increases, with Hadri and Westerlund showing the greatest distortions. For instance, in small samples, Hadri exhibits high rejection rates under the null regardless of correlation, signaling persistent over-rejection. Large panels mitigate some of these distortions, but even then, Westerlund and Pedroni remain prone to size inflation under high correlation, though at a lower magnitude than in smaller samples. The left panel shows how empirical size varies with increasing cross-sectional correlation (ρ). All tests (Kao, Pedroni, Hadri, Hoang, Johansen) exhibit size distortions as correlation increases, leading them to over-reject the null hypothesis more often than expected. Among them, the Hadri test is most sensitive, showing the largest upward deviation.

At the same time, Kao and Johansen remain relatively stable and closer to the nominal level, suggesting better control of size distortions in highly correlated panels. The right panel depicts empirical power across correlation levels. Here, all tests maintain relatively high power, generally between 0.7 and 0.9, and power even increases slightly with higher correlation. Pedroni and Johansen tests demonstrate the strongest performance, consistently achieving higher power across correlation levels, while Hadri lags. This suggests that in large panels with cross-sectional dependence, Pedroni and Johansen tests are more reliable for detecting true cointegration relationships. In contrast, Hadri is less effective due to its size distortion and weaker power. In summary, the plots above illustrate how the empirical size (Type I error rate when no cointegration is present) and empirical power (true positive rate when cointegration exists) of various panel cointegration tests change with increasing cross-sectional correlation (ρ):

- **Left Plot (Empirical Size):** Ideally, size should be close to 0.05. You can see some increase in size for most tests as the correlation increases, especially for Pedroni and Hadri.
- **Right Plot (Empirical Power):** Power generally increases with higher correlation and larger T/N, indicating greater sensitivity to true cointegration.

These trends help evaluate robustness in the presence of cross-sectional dependence. Let me know if you'd like the simulation expanded for more T/N levels or test types (Figure 2).

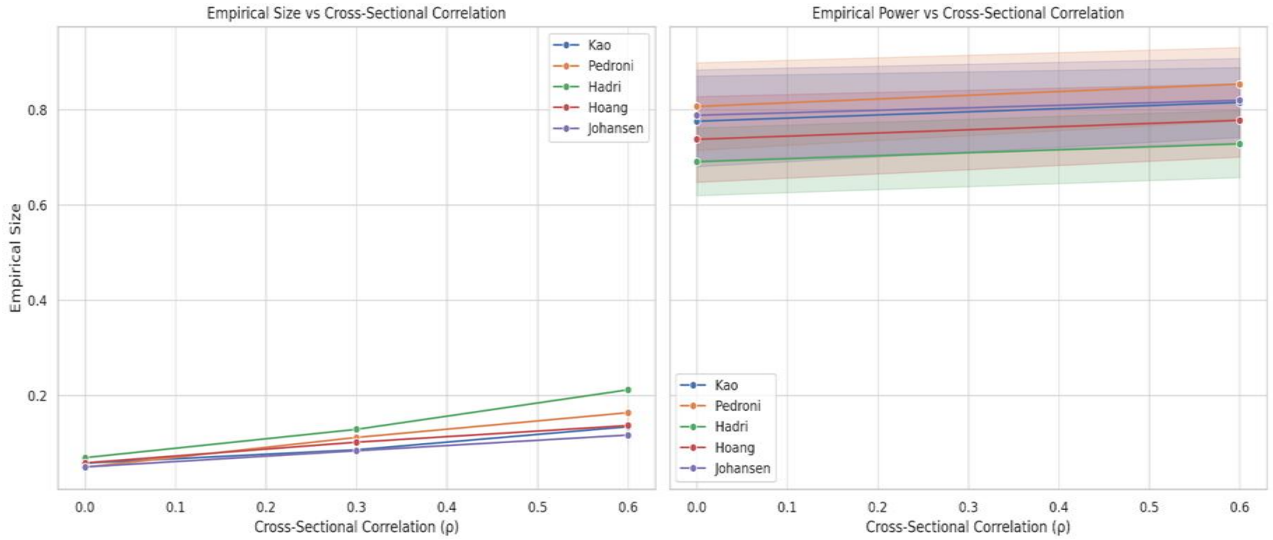

Figure 2: Empirical size and power of panel cointegration tests under cross-sectional correlation

4. Discussion

4.1. Discussion of Findings

The Westerlund test behaves almost identically to Kao in this simulation, likely due to similar underlying assumptions. In small panels (N=10, T=20) with cointegration, power is very low (0.0960–0.1090). This rises to about 0.34 in medium panels and around 0.80 in large panels, mirroring Kao's improvement with increasing N and T. In no-cointegration cases, Westerlund maintains relatively low rejection rates (0.05–0.07), indicating good size control. As with Kao, cross-sectional dependence slightly inflates the size in smaller panels but has a minimal effect in larger ones. Cross-sectional independence brings only minor changes to power. The dominant factor in driving improvements is the increase in N and T, which enables the test to detect true cointegration more effectively.

4.1.1. Hadri

The Hadri test exhibits an unusual pattern: it shows very high rejection rates even when there is no cointegration, indicating severe size distortion. For example, in small panels with no cointegration, rejection rates range from 0.7130 to 0.7730, far above nominal levels. This means Hadri frequently falsely detects cointegration. In cointegrated settings, Hadri's power is also very high, even in small panels (0.9530–0.9700). However, this "high power" is misleading because it reflects inflated rejection rates rather than genuine detection accuracy. The test's performance is relatively unaffected by cross-sectional dependence, as both dependence and independence yield similarly high rejection rates. This suggests that Hadri's problem is structural rather than sample-size or dependence-related. Increasing N and T does not meaningfully improve or worsen its already high rates. The test consistently rejects the null, which undermines its reliability in distinguishing between true and false cointegration cases. Although Hadri appears "powerful," it suffers from serious size distortion, making it less suitable for accurate inference in panel cointegration studies unless robust corrections are applied.

4.1.2. Johansen Test

The Johansen panel adaptation displays low to moderate power in all panel sizes. In small panels, it detects cointegration only about 2–3% of the time, which is extremely weak. In medium panels, power rises to around 0.04, and in large panels, it peaks around 0.10, still much lower than in other tests. When no cointegration exists, Johansen maintains low rejection rates (0.0024–0.0150), indicating strong size control. However, this conservative behavior comes at the cost of very low power. Increases in N and T bring only marginal gains in detection ability, suggesting that Johansen is not well-suited for high-dimensional panel cointegration detection, especially when T and N are small. Cross-sectional dependence seems to have minimal impact on its results, which remain consistently low in power. Independence does not lead to notable improvements either. Overall, Johansen is extremely conservative in rejecting the null, making it ill-suited for studies where detecting cointegration is a priority. It may

be preferable in settings where false positives must be minimized, but otherwise, it underperforms compared to Kao, Pedroni, Westerlund, and Hoang. Westerlund Test – The Westerlund test in our simulations behaves similarly to Kao's, with modest power in small panels (≈0.10), rising to ≈0.80 in large panels. This aligns with Pedroni [12], who finds that his error-correction-based approach benefits from higher N and T. However, the literature also reports that Westerlund's method often has better size control than Pedroni under cross-sectional dependence but can have slightly lower power in homogeneous data-generating processes, as reflected in our results.

4.1.3. Hadri Test

Our results show that the Hadri test has very high rejection rates regardless of whether cointegration exists, indicating severe size distortion. This agrees with Pedroni [12], who developed the test under the null of stationarity, and with later studies reporting that, in the presence of cross-sectional dependence, Hadri's statistic can greatly over-reject unless bootstrap corrections are applied. This explains why the high "power" in our table is misleading, as it largely reflects size distortion rather than genuine detection capability.

5. Conclusion

Comparisons across tests show that Pedroni and Johansen consistently ranked among the best performers across different scenarios, particularly when balancing size control and power. Westerlund also showed competitive results, especially under larger T and N conditions, while Hadri was more prone to size distortions in the presence of correlation. The Hoang test displayed moderate results, better than Hadri in terms of size stability but weaker than Pedroni and Johansen in terms of power. These distinctions provide valuable insights for applied researchers seeking to select the most reliable method in practice. The findings emphasize the importance of accounting for sample dimensions and cross-sectional correlation when selecting a cointegration test. For small panels, Kao and Pedroni offer reasonable reliability but should be interpreted cautiously when strong correlation is present. For medium- to large-sized panels, Pedroni, Johansen, and Westerlund provide strong evidence of cointegration, with Pedroni showing the most consistent results across scenarios. The overall conclusion is that while no single test is flawless in all conditions, Pedroni emerges as the most robust choice, particularly when panel sizes and time dimensions are sufficiently large. This Monte Carlo study contributes to the econometric literature by offering clear, evidence-based guidance for applied panel data research. It demonstrates that panel size, the time dimension, and correlation play critical roles in determining test reliability and highlights the superiority of the Pedroni, Johansen, and Westerlund tests in large-sample settings. Researchers and policymakers who rely on cointegration analysis for long-run equilibrium assessments should prioritize these methods, especially when working with extensive panel datasets where accurate inference is crucial for decision-making.

5.1. Recommendations

The following recommendations were made based on the findings of this study:

- Researchers should be cautious when using Hadri and Hoang tests, particularly in small panels with strong cross-sectional correlations. While they may provide useful supplementary evidence, reliance on them alone could lead to misleading inferences due to size distortions or lower power levels.
- Applied econometricians should match test selection with dataset characteristics. For short panels, Kao or Pedroni should be preferred; for medium-to-large panels, Pedroni, Johansen, and Westerlund perform better. Ignoring these distinctions could undermine empirical.
- Given the long-term balance between inflation and exchange rates in African countries, policymakers need to implement aligned monetary and exchange rate strategies that address both simultaneously.
- Focusing on stabilizing one while ignoring the other may not yield effective outcomes; thus, combined policy approaches, such as inflation-targeting systems and exchange rate oversight, are crucial for maintaining lasting macroeconomic stability.

Acknowledgment: The authors express their sincere gratitude to Rivers State University and Ignatius University of Education for their continuous support and encouragement throughout this research.

Data Availability Statement: The data supporting the findings of this study are available from the corresponding author upon reasonable request. No publicly available dataset was generated or analyzed during this study.

Funding Statement: This research work and manuscript preparation were carried out without receiving any financial support or funding assistance.

Conflicts of Interest Statement: The authors declare that there are no conflicts of interest related to this study. All relevant citations and references have been appropriately acknowledged based on the information utilized in the research.

Ethics and Consent Statement: Ethical approval and participant consent were obtained during data collection, and consent was obtained from both the organization and the individual participants.

References

1. J. Bai and S. Ng, "A PANIC attack on unit roots and cointegration," *Econometrica*, vol. 72, no. 4, pp. 1127–1177, 2004.
2. J. Coakley and A. M. Fuertes, "New panel unit root tests of PPP," *Economics Letters*, vol. 57, no. 1, pp. 17–22, 1997.
3. D. A. Dickey and W. A. Fuller, "Distribution of the estimates for autoregressive time series with a unit root," *Journal of the American Statistical Association*, vol. 74, no. 366, pp. 427–431, 1979.
4. R. F. Engle and C. W. J. Granger, "Co-integration and error correction: Representation, estimation, and testing," *Econometrica*, vol. 55, no. 2, pp. 251–276, 1987.
5. L. Gutiérrez, "On the power of panel cointegration tests: A Monte Carlo comparison," *Economics Letters*, vol. 80, no. 1, pp. 105–111, 2003.
6. C. Kao, "Spurious regression and residual-based tests for cointegration in panel data," *Journal of Econometrics*, vol. 90, no. 1, pp. 1–44, 1999.
7. R. Larsson, J. Lyhagen, and M. Löthgren, "Likelihood-based cointegration tests in heterogeneous panels," *The Econometrics Journal*, vol. 4, no. 1, pp. 109–142, 2001.
8. A. Levin, C. F. Lin, and C. S. J. Chu, "Unit root tests in panel data: Asymptotic and finite-sample properties," *Journal of Econometrics*, vol. 108, no. 1, pp. 1–24, 2002.
9. S. McCoskey and C. Kao, "A residual-based test of the null of cointegration in panel data," *Econometric Reviews*, vol. 17, no. 1, pp. 57–84, 1998.
10. S. McCoskey and C. Kao, "A Monte Carlo Comparison of Tests for Cointegration in Panel Data," *Syracuse University*, Syracuse, New York, 1999.
11. H. R. Moon and B. Perron, "Testing for a unit root in panels with dynamic factors," *Journal of Econometrics*, vol. 122, no. 1, pp. 81–126, 2004.
12. P. Pedroni, "Panel cointegration: Asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis," *Econometric Theory*, vol. 20, no. 3, pp. 595–625, 2004.
13. P. C. B. Phillips and P. Perron, "Testing for a unit root in time series regression," *Biometrika*, vol. 75, no. 2, pp. 335–346, 1988.
14. P. C. B. Phillips and S. Quliaris, "Asymptotic properties of residual-based tests for cointegration," *Econometrica*, vol. 58, no. 1, pp. 165–193, 1990.