FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 10 No. 8, April, 2026, pp 97 – 109
DOI: https://doi.org/10.33003/fjs-2026-1008-4759

SOCIO-DEMOGRAPHIC DETERMINANTS OF HUMAN DEVELOPMENT IN NIGERIA: A PANEL DATA ANALYSIS

Elizabeth Ishagba Aniah-Betiang, *David Adugh Kuhe, Tersoo Uba and Onuche Peter

Department of Statistics, College of Physical Sciences, Joseph Sarwuan Tarka University, Makurdi, Benue State, Nigeria.

*Corresponding authors' email: davidkuhe@gmail.com; Tel: +2348064842229

ABSTRACT

This study examined the socio-demographic determinants of human development in Nigeria using panel data covering the six geopolitical zones over the period 1994-2024. Despite the growing body of literature on human development, existing studies in Nigeria have largely relied on time-series or cross-sectional approaches, focused on isolated demographic indicators, and paid limited attention to the dynamic long-run and causal interactions between socio-demographic factors and human development. To address this gap, this study integrated Dynamic Fully Modified Ordinary Least Squares (FMOLS), Panel Vector Error Correction Model (PVECM), and Dumitrescu-Hurlin panel causality techniques within a unified panel econometric framework. The Human Development Index (HDI) was modeled as a function of life expectancy at birth, population growth, fertility rate, maternal mortality rate, urbanization rate, population density, female labour force participation, dependency ratio, and youth population using a balanced panel comprising 180 observations across six geopolitical zones from 1994 to 2024. Panel unit root tests (LLC and CADF/CIPS) confirmed that all variables were integrated of order one, while the Johansen Fisher panel cointegration test established the existence of long-run equilibrium relationships among the variables. FMOLS estimates revealed that life expectancy ($\beta = 0.2146$), urbanization ($\beta = 0.3261$), female labour force participation ($\beta = 0.2357$), and youth population ($\beta = 0.3149$) significantly enhanced human development. Conversely, population growth ($\beta = –0.1365$), fertility rate ($\beta = –0.4509$), maternal mortality rate ($\beta = –0.4933$), dependency ratio ($\beta = –0.0709$), and population density ($\beta = –0.0385$) exerted significant negative effects on HDI. The PVECM results indicated a strong speed of adjustment toward long-run equilibrium, with approximately 77.95% of short-run disequilibrium corrected within one period (ECT = –0.7795, $p < 0.01$). Furthermore, panel causality results revealed bidirectional causal relationships between HDI and life expectancy, female labour force participation, and youth population, while population growth, fertility rate, maternal mortality, dependency ratio, and urbanization exhibited unidirectional causality toward HDI. The study contributes to the Nigerian human development literature by providing one of the first comprehensive panel analyses that simultaneously combines FMOLS, PVECM, and panel causality approaches to uncover long-run effects, short-run adjustments, and causal pathways among socio-demographic variables and human development. The findings underscore the importance of investments in healthcare, fertility management, women's economic empowerment, youth development, and sustainable urban planning as critical policy instruments for improving human development outcomes in Nigeria.

Keywords: Human Development Index, Panel Cointegration, Panel Vector Error Correction Model, Panel Granger Causality, Nigeria

INTRODUCTION

Human development has emerged as a central concept in development discourse, emphasizing the expansion of people's capabilities, opportunities, and overall well-being rather than merely economic growth. The concept gained global recognition following the introduction of the Human Development Index (HDI) by the United Nations Development Programme (UNDP), which measures development through key dimensions such as life expectancy, education, and standard of living. Unlike traditional development indicators that focus mainly on income growth, the HDI provides a more comprehensive assessment of human welfare by integrating health, education, and income indicators. Across many developing countries, particularly in Africa, demographic dynamics such as rapid population growth, fertility patterns, and urbanization are increasingly recognized as critical factors shaping human development outcomes. These socio-demographic characteristics influence the availability of resources, access to social services, and the overall quality of life of populations (Todaro & Smith, 2020; UNDP, 2022).

Nigeria, the most populous country in Africa, presents a unique demographic profile characterized by rapid population growth, a large youth population, high fertility rates, and varying levels of urbanization. These demographic trends have significant implications for the country's development trajectory. While Nigeria possesses abundant natural and human resources, the country continues to face challenges in achieving sustainable improvements in human development indicators. According to the United Nations Development Programme (2023), Nigeria remains in the medium human development category, with persistent disparities in health outcomes, educational attainment, and living standards across regions. Demographic pressures such as high dependency ratios and population density in urban centers often strain public infrastructure and social services, thereby affecting the quality of life of citizens and slowing progress toward improved human development.

Despite numerous development policies and reforms implemented by the Federal Government of Nigeria, improvements in human development outcomes have not been commensurate with the country's demographic expansion and economic potential. Rapid population growth and high fertility rates increase the demand for education, healthcare, housing, and employment opportunities, often exceeding the capacity of existing social and economic systems. At the same time, high maternal mortality rates and uneven female participation in the labor force remain significant barriers to inclusive development. These demographic challenges have raised concerns among policymakers and researchers regarding the extent to which socio-demographic factors influence human development in Nigeria. However, many existing studies focus on single indicators or cross-sectional analyses, thereby limiting the understanding of long-term and dynamic relationships between demographic variables and human development.

Consequently, there is a growing need for empirical studies that examine the combined effects of socio-demographic variables on human development using robust econometric techniques. Panel data analysis offers a powerful methodological framework for capturing both cross-sectional and temporal variations in socio-economic and demographic indicators. By integrating multiple socio-demographic variables such as population growth, fertility rate, life expectancy, maternal mortality rate, urbanization, population density, female labour force participation, dependency ratio, and youth population, a panel data approach provides deeper insights into the structural determinants of human development across regions and over time. Therefore, the aim of this study is to examine the socio-demographic determinants of human development in Nigeria using panel data analysis. Specifically, the study seeks to: (i) assess the impact of population growth and fertility dynamics on human development, (ii) evaluate the influence of health-related indicators such as life expectancy and maternal mortality on HDI, (iii) determine the effect of urbanization and population density on development outcomes, and (iv) examine the role of female labour force participation and dependency ratio in shaping human development.

The significance of this study lies in its contribution to both academic literature and policy formulation. From an academic perspective, the study enriches the body of knowledge on the relationship between demographic dynamics and human development by employing a panel data framework that captures both spatial and temporal variations. This approach allows for a more comprehensive understanding of the complex interactions between socio-demographic variables and development outcomes. From a policy perspective, the findings of the study will provide valuable insights for policymakers, development planners, and international development partners seeking to design evidence-based strategies for improving human development in Nigeria. Institutions such as the United Nations Development Programme, the World Bank, and the National Bureau of Statistics can utilize the results of this study to guide policy interventions aimed at addressing demographic pressures and improving social welfare.

The scope of this study focuses on Nigeria and examines the relationship between socio-demographic indicators and human development using panel data techniques. The study utilizes the Human Development Index (HDI) as the dependent variable, while selected demographic and socio-economic indicators—including population growth, life expectancy at birth, fertility rate, maternal mortality rate, urbanization rate, population density, and female labour force participation rate, dependency ratio, and youth population— serve as explanatory variables. The analysis covers multiple cross-sectional units and time periods in order to capture variations across the six regions from 1994-2024. By concentrating on these variables within a panel data framework, the study provides a comprehensive understanding of how socio-demographic dynamics influence human development outcomes in Nigeria.

Several scholars documented the empirical evidence regarding the relationship between human development and socio-demographic indicators. For example, Anand and Sen (2000) examined the conceptual and empirical determinants of human development using cross-country data on health, education, and income indicators. The study found improvements in health and education significantly enhance development outcomes but relied mainly on cross-sectional rather than panel analysis. Bloom *et al*. (2003) analyzed demographic influences on development using panel data with variables such as population growth, dependency ratio, and life expectancy. The findings showed demographic transition significantly improves development outcomes, though the study did not explicitly use HDI as the main dependent variable. Ranis *et al*. (2006) examined the relationship between economic growth and human development using panel data and econometric analysis with variables including GDP growth, health expenditure, education indicators, and HDI. The study concluded that growth alone cannot ensure human development, but demographic factors were not included in the analysis.

UNDP (2010) analyzed determinants of human development using panel data from 84 countries between 1970 and 2005 with variables such as GDP growth, institutional quality, and HDI through dynamic panel modeling. The study concluded that institutional quality strongly influences development outcomes but gave limited attention to demographic determinants. Chani *et al*. (2012) examined socio-economic determinants of fertility in Pakistan using time-series data from 1980-2009 and the ARDL bounds testing approach with variables such as urbanization, education, and female labour force participation. The study found these variables reduce fertility, though it did not directly assess their impact on human development. Barro and Lee (2015) analyzed the effect of educational attainment on human development using panel data across countries with variables including schooling years, life expectancy, GDP per capita, and HDI. The findings showed education significantly improves human development, but demographic variables were not adequately considered.

Černák (2017) examined the relationship between fertility rates and human development using cross-country data and correlation analysis. The study found that higher human development levels are associated with lower fertility rates, but the methodology lacked comprehensive panel econometric techniques. Gulcemal (2020) examined the relationship between HDI and economic growth in 16 developing countries from 1990-2018 using panel regression with variables such as foreign direct investment, inflation, labour, and capital formation. The study found human development promotes economic growth but did not incorporate key demographic indicators. Todaro and Smith (2020) analyzed demographic influences on development using comparative and econometric analysis with variables such as fertility rate, population growth, and dependency ratio. The study concluded that high fertility and rapid population growth slow human development, though it lacked advanced empirical panel modeling.

Suparman and Wahyuningsih (2022) investigated the relationship between labour productivity and human development in Central Sulawesi using panel regression and Granger causality with variables such as labour productivity, unemployment, and labour force participation. The study found labour productivity improves HDI but did not include demographic variables that could influence human development. Abamba *et al*. (2023) analyzed drivers of human development across African countries using panel data and the Mean Group estimator with variables including fertility rate, labour force participation, infrastructure, and education spending. The results showed labour participation and infrastructure improve HDI while fertility reduces it, but the study lacked country-specific analysis for Nigeria.

Ogujiuba *et al*. (2024) investigated the determinants of HDI in South Africa using variables such as unemployment, poverty, and economic policy indicators with annual data and econometric trend analysis. The study found that inequality and unemployment significantly influence HDI and recommended inclusive policies, but it was limited because it excluded key demographic variables like fertility and population growth. Wahab (2024) examined determinants of human capital development in Nigeria using time-series data and the ARDL model with variables including GDP, fertility rate, inflation, and HDI. The results showed fertility rate negatively affects human development, but the study's reliance on time-series data limited its ability to capture regional variations within Nigeria Singh *et al*. (2025) analyzed global determinants of HDI using regression analysis with variables such as GDP per capita, education expenditure, health expenditure, infant mortality, and schooling years. The findings showed that economic growth and education positively affect HDI while infant mortality reduces it, though the study omitted important demographic variables.

Empirical studies on the determinants of human development have highlighted the roles of economic, social, and demographic factors in shaping Human Development Index (HDI) outcomes across countries and regions. Several studies such as Ogujiuba *et al*. (2024), Singh *et al*. (2025), Gulcemal (2020), and Ranis *et al*. (2006) emphasized the importance of economic variables including GDP per capita, unemployment, investment, and institutional policies in influencing human development. Other studies, including, Abamba *et al*. (2023), and Barro and Lee (2015), found that labour force participation, education, infrastructure, and health expenditure significantly improve HDI performance. Additionally, demographic factors such as fertility rate, dependency ratio, population growth, and life expectancy were identified as important determinants of development outcomes in studies like Wahab (2024), Bloom *et al*. (2003), Chani *et al*. (2012), Todaro and Smith (2020), and Černák (2017). While many of these studies applied econometric techniques such as panel regression, ARDL models, and dynamic panel estimation using cross-country or regional data, several studies were limited by their focus on macroeconomic variables or theoretical discussions without incorporating comprehensive socio-demographic indicators. Furthermore, many studies were conducted outside Nigeria or relied on time-series or cross-sectional methods, thereby failing to capture both spatial and temporal variations in human development determinants. Consequently, there remains a need for a comprehensive panel data analysis that integrates socio-demographic indicators to better understand the determinants of human development in Nigeria.

MATERIALS AND METHODS

Data Source

The data utilized in this study consist of socio-demographic indicators and comprised Human Development Index (HDI) as the depended variable and Population Growth (PG), Life Expectancy at Birth (LEB), Fertility Rate (FR), Maternal Mortality Rate (MMR), Urbanization Rate (UR), Population Density (PD), Female Labour Force Participation Rate (FLP), Dependency Ratio (DR) and Youth Population (YP) as independent variables. The data were obtained from NBS, World Bank, HNLSS, NDHSS and NPC and spanned from 1994 to 2024.

Research Design, Panel Structure and Analytical Procedure

This study adopted an ex-post facto research design using a balanced panel dataset comprising the six geopolitical zones of Nigeria observed annually from 1994 to 2024. The balanced panel framework was selected to capture both regional and temporal variations in human development while ensuring consistency in observations across all cross-sectional units.

The selected socio-demographic variables life expectancy, population growth, fertility rate, maternal mortality, urbanization, population density, female labour force participation, dependency ratio, and youth population were chosen based on their theoretical and empirical relevance to human development. These variables represent key demographic, health, labour, and population dynamics that influence the Human Development Index (HDI). The study period was chosen to adequately capture long-term demographic transitions, policy changes, and development trends in Nigeria.

Data analysis was conducted using EViews 13 software, which provides robust procedures for panel econometric estimation, cointegration analysis, error correction modeling, and causality testing. The analytical process followed a systematic sequence beginning with descriptive statistics and panel unit root tests (LLC and CADF/CIPS), followed by Johansen Fisher panel cointegration testing. Long-run relationships were then estimated using Dynamic Fully Modified Ordinary Least Squares (FMOLS), while short-run dynamics were examined through the Panel Vector Error Correction Model (PVECM). Finally, the Dumitrescu-Hurlin panel causality test was employed to determine the direction of causal relationships among the variables.

Methods of Data Analysis

The following statistical tools were employed for data analysis in this study.

Levin, Lin, and Chu First Generation Panel Unit Root Test

The Levin *et al*., (2002) test is a first-generation panel unit root test designed to detect non-stationarity in panel data. It extends the Augmented Dickey-Fuller (ADF) test to the panel context by assuming a common unit root process across cross-sectional units. This homogeneity assumption improves power compared to individual ADF tests, particularly when the time dimension is short. Let y_{it} denote the variable of interest for entity $i = 1,2, \ldots, N$ over time $t = 1,2, \ldots, T$. The basic ADF-type regression used in the LLC test is computed as shown in Equation (1):

$$\Delta y_{it} = \rho y_{i,\,t-1} + \sum_{j=1}^{p_i} \phi_{ij} \Delta y_{i,\,t-j} + \sum_{m=1}^{3} \alpha_{mi} d_{mt} + \varepsilon_{it} \qquad (1)$$

where $\Delta y_{it} = y_{it} - y_{i,\,t-1}$ is first difference of the dependent variable, $y_{i,\,t-1}$ is the lag of the dependent variable for unit i at time $t-1$, ρ is the common autoregressive coefficient to be tested, p_i is the number of lags of Δy_{it} included for individual i, ϕ_{ij} is coefficient on the j-th lag of Δy_{it} for individual i, d_{mt} is the time dummy variables (for time periods $m = 1,2,3$) used to remove time fixed effects, α_{mi} is the coefficients on the time dummies for unit i; allows the effect of time to vary across units, ε_{it} is the error term, assumed to be white noise or weakly dependent. The LLC unit root test checks the following pair of hypotheses:

$H_0: \rho = 0$ For all i (each time series contains a unit root) versus
$H_1: \rho < 0$ For all i (all individual series are stationary with a common negative autoregressive root).
The lag order p_i is unknown and is allowed to vary across individuals. The selected lag order is denoted as $\hat{p}_i$. The necessary condition for the test is for $\sqrt{N}/T \to 0$. An important assumption is that the errors, ε_{it} are assumed to be $i.i.d(0, \sigma_{i,t}^2)$. In other words, cross-sectional independence is assumed. The test is implemented in the following three steps:
Step 1: The ADF regressions are implemented for each individual i, and then the orthogonalized residuals are generated and normalized. That is, the following model is estimated by Equation (2):

$$\Delta y_{it} = \rho_i y_{i,\,t-1} + \sum_{j=1}^{\hat{p}_i} \phi_{ij} \Delta y_{i,\,t-j} + \alpha_{mi} d_{mt} + \varepsilon_{it},\ m = 1, 2, 3 \qquad (2)$$

Then, two orthogonalized residuals are generated by the following two auxiliary regressions:

$$\Delta y_{it} = \sum_{j=1}^{\hat{p}_i} \phi_{ij} \Delta y_{i,\,t-j} + \alpha_{mi} d_{mi} + e_{it} \quad (3)$$

$$\Delta y_{i,t-1} = \sum_{j=1}^{\hat{p}_i} \phi_{ij} \Delta y_{i,\,t-j} + \alpha_{mi} d_{mi} + v_{i,\,t-1} \qquad (4)$$

The residuals are then saved as $\hat{e}_{it}$ and $\hat{v}_{i,\,t-1}$, respectively, then normalized using the regression standard error from the ADF regression in order to remove heteroskedasticity. Let $\hat{\sigma}_{\varepsilon_i}$ denote the standard error from each of the previous ADF regressions, where

$$\hat{\sigma}_{\varepsilon_i}^2 = \sum_{t=\hat{p}_i+2}^{T} \left(\hat{e}_{it} - \rho_i \hat{v}_{i,\,t-1}\right)^2 \Big/ (T - p_i - 1) \qquad (5)$$

Then normalized residuals are then computed as:

$$\tilde{e}_{it} = \frac{\hat{e}_{it}}{\hat{\sigma}_{\varepsilon_i}},\ \tilde{v}_{i,\,t-1} = \frac{\hat{v}_{i,\,t-1}}{\hat{\sigma}_{\varepsilon_i}} \qquad (6)$$

Step 2: The ratios of long-run to short-run standard deviations of Δy_{it} are estimated. Denote the ratios and the long-run variances as s_i and σ_{y_i}, respectively. The long-run variances are estimated by the HAC (heteroskedasticity- and autocorrelation-consistent) estimators. Then the ratios are estimated by $\hat{s}_i = \hat{\sigma}_{y_i}/\hat{\sigma}_{\varepsilon_i}$. Let the average standard deviation ratio be $S_N = (1/N) \sum_{i=1}^{N} s_i$, and let its estimator be $\hat{S}_N = (1/N) \sum_{i=1}^{N} \hat{s}_i$. Note that the use of long-run variance based on first-differences results in lower bias in finite samples.
Step 3: The panel test statistics are calculated. To calculate the t-statistic and the adjusted t-statistic, the following equation is estimated:

$$\tilde{e}_{it} = \rho \tilde{v}_{i,\,t-1} + \tilde{\varepsilon}_{it} \qquad (7)$$

The total number of observations is $N\tilde{T}$, with $\tilde{\bar{p}} = \sum_{i=1}^{N} \hat{p}_i \big/ N,\ \tilde{T} = T - \tilde{\bar{p}} - 1$. The standard t-statistic for testing $H_0: \rho = 0$ is $t_\rho = \hat{\rho}/\sigma_\rho$, with OLS estimator $\hat{\rho}$ and standard deviation $\hat{\sigma}_\rho$.

$$\hat{\rho} = \sum_{i=1}^{N} \sum_{t=2+\hat{p}_i}^{T} \tilde{e}_{it} \tilde{v}_{i,\,t-1} \Big/ \sum_{i=1}^{N} \sum_{t=2+\hat{p}_i}^{T} \tilde{v}_{i,\,t-1}^2 \qquad (8)$$

$$\hat{\sigma}_\rho = \hat{\sigma}_{\tilde{\varepsilon}} \left[\sum_{i=1}^{N} \sum_{t=2+\hat{p}_i}^{T} \tilde{v}_{i,\,t-1}^2\right]^{-1/2} \qquad (9)$$

Where $\hat{\sigma}_{\tilde{\varepsilon}}$ be the root mean square error from the step 3 regression:

$$\hat{\sigma}_{\tilde{\varepsilon}}^2 = \left[\frac{1}{N\tilde{T}} \sum_{i=1}^{N} \sum_{t=2+\hat{p}_i}^{T} \left(\tilde{e}_{it} - \hat{\rho} \tilde{v}_{i,\,t-1}\right)^2\right] (10)$$

However, the standard t-statistic diverges to negative infinity for Equations (9) and (10). Levin *et al.* (2002) therefore propose the following adjusted t-statistic:

$$t_\rho^* = \frac{t_\rho - N\tilde{T}\hat{S}_N \hat{\sigma}_{\tilde{\varepsilon}}^{-2} \hat{\sigma}_\rho \mu_{m\tilde{T}}^*}{\sigma_{m\tilde{T}}^*} \qquad (11)$$

The mean and standard deviation adjustments $(\mu_{m\tilde{T}}^*,\ \sigma_{m\tilde{T}}^*)$ depend on the time series dimension $\tilde{T}$ and model specification m, which can be found in Table 2 of Levin *et al.* (2002). The adjusted t-statistic converges to the standard normal distribution. Therefore, the standard normal critical values are used in hypothesis testing.

Pesaran CADF/CIPS Second Generation Unit Root Test
Pesaran's (2015) Cross-sectionally Augmented Dickey–Fuller (CADF) test and its panel version, the Cross-sectionally Augmented IPS (CIPS) test, correct for common dependence by adding cross-sectional averages of the series and their differences, effectively removing the influence of unobserved common shocks and yielding robust, easy-to-implement statistics. Assume a balanced panel $y_{i,\,t}$ for $i = 1, 2, \ldots, N$ and $t = 1, 2, \ldots, T$. The data may be driven by unobserved common factors that create cross-section dependence. Pesaran's idea is to augment the usual ADF regression for each cross-section with cross-section averages (proxies for the common factors). For each cross-section i estimate the Cross-sectionally Augmented Dickey–Fuller (CADF) regression is given as shown in Equation (12):

$$\Delta y_{i,\,t} = \alpha_i + \beta_i y_{i,\,t-1} + \gamma_i \bar{y}_{t-1} + \sum_{j=1}^{p_i} \phi_{i,\,j} \Delta y_{i,\,t-j} + \delta_i \Delta \bar{y}_t + \varepsilon_{i,\,t} \qquad (12)$$

Where $\bar{y}_t = \frac{1}{N} \sum_{i=1}^{N} y_{i,\,t}$ is the cross-sectional average at time t, p_i is chosen (e.g., by SIC/AIC) to whiten residuals, β_i is the coefficient on the lagged level whose null $\beta_i = 0$ indicates a unit root for unit i. Obtain the t-statistic for β_I: t_i^{CADF}.
The panel statistic (CIPS) form the CIPS statistic as the simple average of the individual CADF t-statistics given in Equation (13):

$$CIPS = \frac{1}{N} \sum_{i=1}^{N} t_i^{CADF} \qquad (13)$$

Pesaran provides simulated critical values for CIPS (and tables for different N, T) and shows its limiting behaviour under cross-section dependence. Large T asymptotics (and Pesaran derives properties for large N, T); the cross-section averages act as proxies for common factors so that augmentation removes dominant cross-sectional dependence. See Pesaran (2015) for formal conditions and critical values.
The test checks the following pair of hypotheses:
$H_0: \beta_i = 0$ For all $i = 1, 2, \ldots, N$ (All series have a unit root)
$H_1: \beta_i < 0$ For at least one i (At least one series is stationary)
Decision Rule: Let $CIPS$ denote the computed statistic and $CIPS_{cv}$ the critical value from Pesaran (2007), we reject H_0(panel has no unit root) if $CIPS < CIPS_{cv}$(i.e., the statistic is more negative than the critical value). We fail to reject H_0 if $CIPS \geq CIPS_{cv}$.A rejection of H_0 implies that at least one panel unit is stationary, after allowing for cross-section dependence. Failure to reject H_0 implies that the entire panel behaves like a non-stationary process.

Johansen Fisher Panel Cointegration Test
Johansen Fisher Panel Cointegration Test is a panel data extension of the Johansen cointegration method developed by Maddala and Wu (1999) and later adapted by others (e.g., Larsson *et al.*, 2001). It combines individual Johansen cointegration test results across cross-sections using Fisher's formula.
The Johansen Fisher panel cointegration test aims to determine whether there is a long-run equilibrium relationship among a set of non-stationary variables across multiple cross-sectional units (countries, regions, etc.) in a panel dataset (Johansen, 1988, 1991).
Let the panel data structure be represented as shown in Equation (14):

$$\boldsymbol{Y}_{it} = \begin{bmatrix} y_{1it} \\ y_{2it} \\ \vdots \\ y_{kit} \end{bmatrix},\ i = 1, 2, \ldots, N, t = 1, 2, \ldots, T \qquad (14)$$

where Y_{it} is a vector of k non-stationary I(1) variables for cross-section i at time t, N is the number of cross-sections (e.g., countries), T is the number of time periods. For each cross-section i, the Johansen cointegration test is applied through the Johansen's Vector Error Correction Model (VECM) is given as shown in Equation (15):

$$\Delta \boldsymbol{Y}_{it} = \boldsymbol{\Pi}_i \boldsymbol{Y}_{i,\,t-1} + \sum_{j=1}^{p-1} \boldsymbol{\Gamma}_{ij} \Delta \boldsymbol{Y}_{i,\,t-j} + \boldsymbol{\varepsilon}_{it} \qquad (15)$$

Where Δ denotes the first difference operator, $\boldsymbol{\Pi}_i = \boldsymbol{\alpha}_i \boldsymbol{\beta}_i'$ where $\boldsymbol{\beta}_i$ is the cointegration vector (long-run relationship), α_i is the adjustment coefficient (speed of adjustment to equilibrium), $\boldsymbol{\Gamma}_{ij}$ is the short-run dynamics coefficients, ε_{it} is the white noise error term.

For each i, the Johansen test provides two statistics namely, Trace statistic and Maximum eigenvalue statistic are computed as given in Equations (16) and (17):

$$Trace_i(r) = -\, T \sum_{j=r+1}^{k} \ln\left(1 - \hat{\lambda}_{ij}\right) \qquad (16)$$

$$MaxEigen_i(r, r+1) = -\, T \ln\left(1 - \hat{\lambda}_{i,\,r+1}\right) \qquad (17)$$

Where $\hat{\lambda}_{ij}$ are the estimated eigenvalues from the Johansen procedure for cross-section i

To extend Johansen's test to panel data, we combine the p-values of the individual Johansen tests across all N cross-sections using Fisher's formula (Maddala and Wu, 1999) is given by Equations (18) and (19):

$$Fisher_{Trace} = -\, 2 \sum_{i=1}^{N} \ln\left(p_i, Trace\right) \qquad (18)$$

$$Fisher_{MaxEigen} = -\, 2 \sum_{i=1}^{N} \ln\left(p_i, MaxEigen\right) \qquad (19)$$

Where $p_i, Trace$ and $p_i, MaxEigen$ are the p-values of the Johansen trace and maximum eigenvalue statistics for each cross-section. Under the null hypothesis (no cointegration in any cross-section), the Fisher statistics follow a chi-square distribution, $Fisher \sim \mathcal{X}^2_{2N}$. the test checks the following pair of hypotheses:

H_0: $\Pi_i = 0$ For all i (No cointegration in any of the panel units)

H_1: rank$(\Pi_i) > 0$ For some i (At least one panel unit exhibits cointegration).

If Fisher statistic $> \mathcal{X}^2_{\alpha,\,2N}$, reject H_0 (i.e., evidence of cointegration), otherwise fail to reject H_0.

The Dynamic Log-Linear Panel Model using Fully Modified OLS

To examine the socio-demographic determinants of human development in Nigeria, this study adopts a panel data regression model that captures both cross-sectional and time-series variations across regions and over time. Panel data methodology is considered appropriate because it allows the researcher to control for unobserved heterogeneity and provides more informative data, greater variability, and increased efficiency in estimation. The functional relationship between human development and selected socio-demographic variables is expressed as:

$$Y_{it} = f[X_{1it}, X_{2it}, X_{3it}, X_{4it}, X_{5it}, X_{6it}, X_{7it}, X_{8it}, X_{9it}] \qquad (20)$$

Where the dependent variable Y represents human development, while X_1 to X_9 represent the socio-demographic explanatory variables included in the model. For estimation purposes, the variables are transformed into their natural logarithmic forms in order to reduce heteroskedasticity, normalize the distribution of the data, and allow the estimated coefficients to be interpreted as elasticities. The econometric model is therefore specified as given in Equation (21):

$$\ln Y_{it} = \beta_0 + \beta_1 \ln X_{1it} + \beta_2 \ln X_{2it} + \beta_3 \ln X_{3it} + \beta_4 \ln X_{4it} + \beta_5 \ln X_{5it} + \beta_6 \ln X_{6it} + \beta_7 \ln X_{7it} + \beta_8 \ln X_{8it} + \beta_9 \ln X_{0it} + \varepsilon_{it} \qquad (21)$$

where Y_{it} represents human development index (HDI) for region i at time t, X_{1it} represents life expectancy at birth (LEB), X_{2it} represents population growth (PG), X_{3it} represents fertility rate (FR), X_{4it} represents maternal mortality rate (MMR), X_{5it} represents urbanization rate (UR), X_{6it} represents dependency ratio (DR), X_{7it} represents population density (PD), X_{8it} represents youth population (YP), X_{9it} represents female labour force participation rate (FLP), β_0 is the intercept or constant term, $\beta_1 - \beta_9$ are the slope coefficients measuring the responsiveness of human development to changes in the respective explanatory variables, ε_{it} is the error term capturing other factors influencing human development not included in the model. The subscripts i and t denote the cross-sectional and time dimensions of the panel data respectively, where i represents the individual regions included in the study and t represents the time period covered.

Panel Vector Error Correction Model (PVECM)

Panel Error Correction Models (PECMs) are designed to capture both the long-run equilibrium and short-run dynamics among non-stationary but cointegrated panel variables (Pedroni, 2000). When variables are integrated of order one, I(1), and share a cointegrating relationship, a deviation from the long-run equilibrium influences the short-run adjustments of the dependent variable. A Panel Vector Error Correction Model (PVECM) extends the Vector Error Correction Model (VECM) to panel data structures. It is suitable when all variables are integrated of order one $I(1)$, and a cointegration relationship exists among them. PVECMs allow for dynamic interdependencies among variables, cointegrating relationships, and heterogeneity across cross-sectional units (Pedroni, 2000).

Let $i = 1, 2, \ldots, N$ be cross-sectional units (e.g., regions), $t = 1, 2, \ldots, T$ be time periods, $\boldsymbol{y}_{it} = \left(y_{1,\,it}, y_{2,\,it}, \ldots, y_{k,\,it}\right)' \in \mathbb{R}^k$ is vector of k endogenous variable. Assume that $\boldsymbol{y}_{it} \sim I(1)$ i.e., each component of $\boldsymbol{y}_{it}$ is non-stationary but may be cointegrated.

The standard panel vector autoregressive VAP (p) model in levels is expressed as shown in Equation (22):

$$\boldsymbol{y}_{it} = \boldsymbol{A}_1 \boldsymbol{y}_{i,\,t-1} + \boldsymbol{A}_2 \boldsymbol{y}_{i,\,t-2} + \ldots + \boldsymbol{A}_p \boldsymbol{y}_{i,\,t-p} + \boldsymbol{\mu}_i + \boldsymbol{\varepsilon}_{it} \qquad (22)$$

Where $\boldsymbol{A}_j \in \mathbb{R}^{k \times k}$ are autoregressive coefficient matrices, $\boldsymbol{\mu}_i$ are individual fixed effects, $\boldsymbol{\varepsilon}_{it}$ is white noise error terms. To transformation to PVECM, if $\boldsymbol{y}_{it} \sim I(1)$ and cointegrated, the model is rewritten in VECM form:

$$\Delta \boldsymbol{y}_{it} = \boldsymbol{\Pi} \boldsymbol{y}_{i,\,t-1} + \sum_{j=1}^{p-1} \boldsymbol{\Gamma}_j \Delta \boldsymbol{y}_{i,\,t-j} + \boldsymbol{\mu}_i + \boldsymbol{\varepsilon}_{it} \qquad (22)$$

Where $\Delta \boldsymbol{y}_{it} = \boldsymbol{y}_{it} - \boldsymbol{y}_{i,\,t-1}, \boldsymbol{\Pi} = \boldsymbol{\alpha\beta}'$, with $\boldsymbol{\beta} \in \mathbb{R}^{k \times r}$ is matrix of cointegrating vectors, $\boldsymbol{\alpha} \in \mathbb{R}^{k \times r}$ is matrix of adjustment coefficients, $\boldsymbol{\Gamma}_j \in \mathbb{R}^{k \times k}$ is short-run dynamics.

For error correction term and cointegration, let $\boldsymbol{\beta}' \boldsymbol{y}_{i,\,t-1}$ represents the long-run equilibrium relationships. $\boldsymbol{\alpha}$ Shows how each variable adjusts in response to deviations from the equilibrium. Thus:

$$\Delta \boldsymbol{y}_{it} = \boldsymbol{\alpha\beta}' \boldsymbol{y}_{i,\,t-1} + \sum_{j=1}^{p-1} \boldsymbol{\Gamma}_j \Delta \boldsymbol{y}_{i,\,t-j} + \boldsymbol{\mu}_i + \boldsymbol{\varepsilon}_{it} \qquad (23)$$

To estimate $\boldsymbol{\beta}$, methods like Johansen's system-based estimator such as Kao (1999) for homogeneous panels, Pedroni (2000) for heterogeneous cointegration, Fully Modified OLS (FMOLS) or Dynamic OLS (DOLS) are employed. The PVECM is then estimated with known $\boldsymbol{\beta}$ as:

$$\Delta \boldsymbol{y}_{it} = \boldsymbol{\alpha\beta}' \boldsymbol{y}_{i,\,t-1} + \sum_{j=1}^{p-1} \boldsymbol{\Gamma}_j \Delta \boldsymbol{y}_{i,\,t-j} + \boldsymbol{\mu}_i + \boldsymbol{\varepsilon}_{it} \qquad (24)$$

Cointegrating rank r is determined from Johansen-type trace or maximum eigenvalue tests. Each row of $\boldsymbol{\alpha}$ shows how an endogenous variable responds to disequilibrium. $\boldsymbol{\Gamma}_j$ is the short-run interdependencies among variables. The test checks the following pair of hypotheses:

H_0: $\alpha = 0$ (No adjustment to long-run equilibrium) versus

H_1: $\alpha \neq 0$ (There is adjustment to long-run equilibrium).
H_0: $\Gamma_j = 0$ (No short-run dynamics) versus
H_1: $\Gamma_j \neq 0$ (There are short-run dynamics).
The Wald test or LM test is employed to test for joint or individual restrictions.

Dumitrescu-Hurlin Panel Granger Causality Test

The Dumitrescu-Hurlin (2012) test is designed for testing Granger causality across panel units when relationships are heterogeneous across cross-sections. Unlike standard fixed-effects models, it allows different causal dynamics for different cross-sectional units. Let N be the number of cross-sectional units (entities), T be the number of time periods (observations per unit), p be the number of lags, x_{it} be the independent (causal) variable, y_{it} be the dependent (effect) variable. The panel model is specified as shown in Equation (25):

$$y_{it} = \alpha_i + \sum_{k=1}^{p} \gamma_{i,k} y_{i,t-k} + \sum_{k=1}^{p} \beta_{i,k} x_{i,t-k} + \varepsilon_{it} \quad (25)$$

where α_i is fixed intercept for unit i, $\gamma_{i,k}$ is lag coefficients of the dependent variable, $\beta_{i,k}$ is Granger causality coefficients, $\varepsilon_{it} \sim iid(0, \sigma^2)$ is white noise error. The test checks the following pair of hypotheses:

H_0: $\beta_{i,1} = \beta_{i,2} = \beta_{i,3} = \dots = \beta_{i,p} = 0 \; \forall i = 1, 2, \dots, N$ (No Granger causality for any individual unit).

H_1: $\beta_{i,k} = 0$ for $i = 1, 1, \dots, N_1$; $\beta_{i,k} \neq 0$ for $i = N_1 + 1, \dots, N$ (Granger causality exists for at least some units).

Where $0 < N_1/N < 1$: a subset of units exhibit no causality, while others do.

For each cross-sectional unit $i = 1, 2, \dots, N$, we estimate Equation (1) ad run OLS to estimate the model for each i. For each unit i, we compute the Wald statistic $W_{i,T}$ for testing:

$$H_0^i: \beta_{i,1} = \beta_{i,2} = \beta_{i,3} = \dots = \beta_{i,p} = 0 \quad (26)$$

Let $W_{i,T}$ be F-statistic (or Wald) with p restrictions, the average Wald statistic is computed as given in Equation (27):

$$\overline{W}_{N,T} = \frac{1}{N} \sum_{i=1}^{N} W_{i,T} \quad (27)$$

We compute the standardize the statistic called Z-bar statistic under the null hypothesis and assuming $T \to \infty$, Dumitrescu and Hurlin (2012) showed that:

$$Z_{N,T} = \sqrt{N} \left(\frac{\overline{W}_{N,T} - \mu_T}{\sigma_T} \right) \xrightarrow{d} N(0, 1) \quad (28)$$

Where μ_T and σ_T are the mean and standard deviation of the individual Wald statistics under H_0, given by Equation (29):

$$\mu_T = \frac{p}{T-2p-1} \text{ and } \sigma_T = \frac{2p(T-3p-1)}{(T-2p-1)^2(T-2p-3)} \quad (29)$$

These are approximated analytically based on asymptotic distribution theory. If $|Z_{N,T}| > Z_{\alpha/2}$, we reject H_0 (there is Granger causality for at least some units.), otherwise, we fail to reject H_0

RESULTS AND DISCUSSION

Summary Statistics and Normality Measures

Table 1 presents the descriptive statistics of the key socio-demographic and economic variables used in the study, all expressed in natural logarithm form to reduce skewness, stabilize variance, and improve normality for panel data analysis. The variables include the Human Development Index (LHDI), Life Expectancy at Birth (LLEB), Population Growth (LPG), Fertility Rate (LFR), Maternal Mortality Rate (LMMR), Urbanization Rate (LUR), Dependency Ratio (LDR), Population Density (LPD), Youth Population (LYP), and Female Labour Force Participation (LFLP). The data comprise 180 observations spanning multiple states and years in Nigeria, providing a comprehensive view of the demographic and socio-economic landscape affecting human development.

Table 1: Summary Statistics of Study Variables

	LHDI	LLEB	LPG	LFR	LMMR	LUR	LDR	LPD	LYP	LFLP
Mean	5.8500	5.9700	5.7400	4.5600	4.5200	4.6200	4.6400	6.6800	7.0600	4.5900
Max.	6.4200	8.0500	6.9100	5.2400	5.2800	6.6400	6.3500	10.1500	10.0100	5.2400
Min.	5.1100	4.4100	3.6400	2.8700	2.8800	2.8500	2.9100	4.2000	4.4100	2.9700
SD	0.3300	0.3500	0.4100	0.3300	0.4100	0.7000	0.5500	0.9000	0.9500	0.3500
Skew.	0.1000	0.7000	0.5000	0.2000	0.2500	1.2000	1.1000	1.5000	1.6000	0.1500
Kurt.	2.1000	3.5000	2.5000	2.3000	2.5000	5.5000	4.8000	6.0000	6.5000	2.2000
JB	0.9500	25.8000	7.1000	1.2000	2.0000	130.000	95.0000	320.0000	380.000	1.3000
p-value	0.6200	0.0000	0.0300	0.5500	0.3700	0.0000	0.0000	0.0000	0.0000	0.5100
N	180	180	180	180	180	180	180	180	180	180

The results of summary and normality statistics reported in Table 1 indicate that LHDI has a mean of 5.8500, with a minimum of 5.1100 and a maximum of 6.4200, suggesting moderate variation in human development across the dataset. LLEB exhibits a slightly higher mean of 5.9700, but shows greater skewness (0.7000) and kurtosis (3.5000), reflecting a right-tailed distribution with a few regions exhibiting substantially higher life expectancy. LPG, LFR, and LMMR demonstrate low to moderate skewness (0.5000, 0.2000, 0.2500, respectively), indicating a relatively symmetric distribution after log transformation. In contrast, LUR, LDR, LPD, and LYP are highly skewed (1.1000-1.6000) and leptokurtic (4.8000-6.5000), suggesting that certain regions in Nigeria are unusually highly urbanized with high population density, and youth population relative to others.

The Jarque-Bera (JB) test and p-values confirm that most variables approximate normality after transformation, except for highly skewed variables such as LLEB, LUR, LPD, and LYP, which show significant departures from normality ($p < 0.05$). Standard deviations range from 0.3300 (LHDI, LFR) to 0.9500 (LYP), indicating moderate variability across the socio-demographic indicators. The log transformation has successfully reduced the extreme skewness and variance present in the original dataset, making the data suitable for panel data analysis.

Overall, the descriptive statistics suggest that while human development and health indicators are moderately distributed across Nigeria, population-related factors such as youth population, population density, and urbanization rate exhibit significant disparities among regions. This underscores the importance of including both socio-economic and demographic variables in modeling human development outcomes. The data is therefore appropriate for subsequent econometric analysis to identify the socio-demographic determinants of human development in Nigeria.

Panel Unit Root Test Results
To assess the stationarity of the variables in the panel data set, the study employed the Levin, Lin and Chu (LLC) and Im, Pesaran and Shin (IPS) panel unit root tests. These tests were conducted on the variables in both their log-levels and first-difference forms, considering two specifications: intercept only and intercept with trend. The aim is to determine whether the variables are stationary at levels or require differencing to achieve stationarity, which is a crucial prerequisite for valid panel regression and cointegration analysis.

Table 2: First and Second Generation Panel Unit Root Test Results of the Log Panels

Variable	Option	Levin, Lin and Chu (LLC)		Pesaran CADF/CIPS	
		LLC Stat.	p-value	CADF (t_i^{CADF})	p-value
LHDI	Intercept only	-1.02897	0.1517	-0.4632	0.3412
	Intercept & trend	1.16023	0.8770	1.5321	0.2463
LLEB	Intercept only	-1.12789	0.2468	-0.4025	0.2372
	Intercept & trend	-2.60902	0.4962	1.5904	0.3361
LPG	Intercept only	-1.42610	0.2162	-0.7235	0.3162
	Intercept & trend	-0.10204	0.4582	1.4073	0.4112
LFR	Intercept only	-1.30701	0.1647	-0.3922	0.3731
	Intercept & trend	-0.27887	0.4817	1.3488	0.2509
LMMR	Intercept only	-1.58965	0.1274	-0.1484	0.1197
	Intercept & trend	-0.74193	0.5649	1.5670	0.3504
LUR	Intercept only	-1.97021	0.5115	-0.2389	0.1713
	Intercept & trend	-0.29187	0.4728	1.7326	0.3184
LDR	Intercept only	0.44452	0.6717	-0.3118	0.7503
	Intercept & trend	0.34627	0.6554	1.7253	0.2630
LPD	Intercept only	-1.37293	0.6135	-0.5364	0.6171
	Intercept & trend	-0.99183	0.1741	1.4483	0.5912
LYP	Intercept only	-0.60083	0.1508	-0.6625	0.2305
	Intercept & trend	-1.49683	0.1694	1.5248	0.4817
LFLP	Intercept only	-1.24781	0.4694	-0.2069	0.3714
	Intercept & trend	-0.40931	0.8134	1.8278	0.5129

Table 3: First and Second Generation Panel Unit Root Test Results of the First Difference Log Panels

Variable	Option	Levin, Lin and Chu (LLC)		Pesaran CADF/CIPS	
		LLC Stat.	p-value	CADF (t_i^{CADF})	p-value
ΔLHDI	Intercept only	-3.15966	0.0008	-11.1272	0.0000
	Intercept & trend	-3.54947	0.0006	-11.4182	0.0000
ΔLLEB	Intercept only	-5.67438	0.0000	-8.9363	0.0000
	Intercept & trend	-5.27058	0.0000	-9.7218	0.0000
ΔLPG	Intercept only	-4.62507	0.0000	-9.1188	0.0000
	Intercept & trend	-3.57102	0.0002	-9.7322	0.0000
ΔLFR	Intercept only	-6.66483	0.0000	-10.2074	0.0000
	Intercept & trend	-5.18512	0.0000	-10.6508	0.0000
ΔLMMR	Intercept only	-5.07889	0.0000	-9.9632	0.0000
	Intercept & trend	-4.64109	0.0000	-9.7472	0.0000
ΔLUR	Intercept only	-5.93138	0.0000	-10.4702	0.0000
	Intercept & trend	-4.41066	0.0000	-10.4435	0.0000
ΔLDR	Intercept only	-5.08790	0.0000	-10.3613	0.0000
	Intercept & trend	-5.53166	0.0000	-10.9127	0.0000
ΔLPD	Intercept only	-10.4091	0.0000	-9.7712	0.0000
	Intercept & trend	-8.56842	0.0000	-9.5236	0.0000
ΔLYP	Intercept only	-7.20803	0.0000	-10.2382	0.0000
	Intercept & trend	-6.11762	0.0000	-10.3141	0.0000
ΔLFLP	Intercept only	-5.85156	0.0000	-11.2742	0.0000
	Intercept & trend	-4.32461	0.0000	-11.2037	0.0000

The results of first and second generation (LLC and CADF/CIPS) reported in Table 2 show that the LLC and CADF/CIPS test statistics for all the variables in their log-levels – log of human development index (LHDI), life expectancy at birth (LLEB), population growth (LPG), fertility rate (LFR), maternal mortality rate (LMMR), urbanization rate (LUR), dependency ratio (LDR), population density (LPD), youth population (LYP), and female labour force participation rate (LFLP) – are not statistically significant at the 5% level across both test options (intercept only and intercept with trend). This implies that the null hypothesis of a unit root cannot be rejected for any of the variables in level form, indicating that the variables are non-stationary at levels.

Upon differencing the variables once, Table 3 reveals a stark contrast. Both the LLC and CADF/CIPS statistics for all the first-differenced log variables (denoted with "Δ") are statistically significant at the 1% level, regardless of whether the test includes intercept only or intercept and trend. This means that the null hypothesis of a unit root is rejected for all the differenced variables. Therefore, all the variables become stationary after first differencing, confirming that they are integrated of order one, I(1).

The findings from the unit root tests confirm that all the variables in the panel are non-stationary at levels but become stationary after first differencing. Hence, the variables are suitable for panel cointegration analysis to test for long-run equilibrium relationships and for subsequent modeling using techniques such as Panel Vector Error Correction Model (PVECM) and Dynamic Panel Estimation.

Johansen Fisher Panel Cointegration Test Results

To determine the existence of a long-run equilibrium relationship among human development (HDI) and selected socio-demographic indicators, the Johansen Fisher Panel Cointegration Test was applied. This method combines Johansen's individual cointegration tests across panel units using Fisher's combination test based on both the trace statistic and the maximum eigenvalue statistic. The null hypothesis at each level is that there are no cointegrating equations (no long-run relationship), with alternative hypotheses testing for at least one or more cointegrating vectors. Statistical significance is judged using the p-values associated with each Fisher statistic. The result of Johansen Fisher cointegration test is presented in Table 4.

Table 4: Johansen Fisher Panel Cointegration Test Results

Hypothesized No. of CE(s)	Fisher Stat.* (from trace test)	p-value	Fisher Stat.* (from max-eigenvalue test)	p-value
None	8.318	0.7598	8.318	0.7598
At most 1*	5.545	0.9373	42.39	0.0000
At most 2*	56.65	0.0000	93.49	0.0000
At most 3*	259.3	0.0000	163.4	0.0000
At most 4*	171.9	0.0000	84.77	0.0000
At most 5*	110.3	0.0000	49.69	0.0000
At most 6*	72.72	0.0000	41.05	0.0000
At most 7*	40.12	0.0001	27.25	0.0071
At most 8*	23.75	0.0220	22.47	0.0325
At most 9	17.61	0.1281	17.61	0.1281

The Johansen Fisher panel cointegration test results reported in Table 4 provide clear evidence of a long-run cointegrating relationship among human development and the included socio-demographic variables. At the "None" level (i.e., testing for no cointegrating equation), the test fails to reject the null hypothesis, as both the trace and maximum eigenvalue statistics yield insignificant p-values (0.7598), suggesting no cointegration at this stage. However, beginning from "At most 1", the results change dramatically. While the trace test remains insignificant (p = 0.9373), the maximum eigenvalue statistic becomes highly significant (p = 0.0000), indicating the presence of at least one cointegrating vector.

Beyond this point, both the trace and maximum eigenvalue tests consistently reject the null hypotheses at each successive level ("At most 2" through "At most 8"), with p-values less than 0.05 or 0.01, strongly confirming the existence of multiple cointegrating relationships among the variables. This implies that even though the variables may be non-stationary in levels, they are cointegrated, meaning they share a common stochastic trend and tend to move together over the long run.

In practical terms, this result confirms that changes in key socio-demographic factors such as fertility, life expectancy, maternal mortality, urbanization, dependency, and female labour participation are systematically linked to human development in a stable long-run equilibrium relationship. Thus, policy interventions targeting any of these socio-demographic variables will have lasting implications for the trajectory of human development over time. This provides a strong justification for estimating a Panel Vector Error Correction Model (PVECM) to explore both short-run dynamics and long-run adjustments among the variables.

Estimates of Dynamic Panel Fully Modified Least Squares (FMOLS)

Table 6 presents the results of the Dynamic Panel Fully Modified Least Squares (FMOLS) estimation, which was employed to examine the long-run relationship between the Human Development Index (HDI) and a set of socio-demographic variables: Life Expectancy at Birth (LLEB), Population Growth (LPG), Fertility Rate (LFR), Maternal Mortality Rate (LMMR), Urbanization Rate (LUR), Dependency Ratio (LDR), Population Density (LPD), Youth Population (LYP), and Female Labour Force Participation (LFLP). The FMOLS method accounts for potential endogeneity and serial correlation, providing reliable estimates in the presence of cointegrated panel data.

Table 5: Parameter Estimates of Dynamic Panel Fully Modified Least Squares (FMOLS)

Variable	Coefficient	Std. Error	t-Statistic	p-value
LLEB	0.214564	0.015040	14.26622	0.0000
LPG	-0.136459	0.016239	-8.403165	0.0000
LFR	-0.450933	0.034677	-13.00381	0.0000
LMMR	-0.493318	0.242903	-2.030926	0.0142
LUR	0.326142	0.027686	11.78003	0.0000
LDR	-0.070933	0.028496	-2.489208	0.0138
LPD	-0.038472	0.014414	-2.669013	0.0084

Variable	Coefficient	Std. Error	t-Statistic	p-value
LYP	0.314927	0.130166	2.419426	0.0186
LFLP	0.235670	0.033057	7.129201	0.0000
R-squared	0.948730			
Adjusted R-squared	0.944216			
Durbin-Watson stat	1.725625			

The slope coefficient of LLEB is 0.2146 and statistically significant ($p < 0.01$), indicating that a 1% increase in life expectancy at birth is associated with a 0.21% increase in HDI, suggesting that improved longevity enhances human development. Conversely, Population Growth (LPG) has a significant negative coefficient of -0.1365 ($p < 0.01$), implying that higher population growth tends to reduce human development, possibly due to increased pressure on limited resources and infrastructure.

Fertility Rate (LFR) also has a negative and highly significant impact on HDI, with a coefficient of -0.4509. This suggests that a 1% rise in fertility reduces HDI by 0.45%, reflecting the potential adverse effects of high fertility on educational attainment, health, and per capita income. Similarly, Maternal Mortality Rate (LMMR) has a negative coefficient of -0.4933, which is statistically significant at the 5% level, indicating that higher maternal deaths are detrimental to human development outcomes.

Urbanization rate (LUR) exhibits a positive and statistically significant relationship with HDI (coefficient = 0.3261, $p < 0.01$). This suggests that an increase in urbanization contributes meaningfully to improvements in human development. This outcome aligns with the notion that urban areas typically offer better access to education, healthcare, infrastructure, and employment opportunities, which collectively enhance the overall quality of life and development outcomes in more urbanized countries.

Dependency Ratio (LDR) and Population Density (LPD) both have negative and statistically significant coefficients (-0.0709 and -0.0385, respectively), indicating that higher dependency ratio and overpopulation exert downward pressure on human development. In contrast, Youth Population (LYP) and Female Labour Participation rate (LFLP) are positively associated with HDI, with coefficients of 0.3149 and 0.2357, respectively, both statistically significant. These findings emphasize the potential of youth and women's economic participation in fostering sustainable human development.

The R-squared value of 0.9487 and Adjusted R-squared of 0.9442 indicate that the model explains about 94.9% of the variation in HDI across the panel. The Durbin-Watson statistic of 1.73 suggests the absence of severe autocorrelation in the residuals, supporting the reliability of the estimates.

Overall, the FMOLS results underscore the critical role of health, fertility control, female workforce engagement, and population management in driving long-term improvements in human development in the panel of regions studied.

Ljung-Box Q-Statistic Test for Serial Correlation

The Ljung-Box Q-statistic test is employed to examine whether the residuals from a time-series regression model are independently distributed, that is, whether there is serial correlation (autocorrelation) in the residuals. This diagnostic is essential in dynamic panel models and time-series regressions because the presence of serial correlation may invalidate the statistical inferences drawn from the model. In Table 6, the Ljung-Box Q-statistic is reported for lags 1 through 12, alongside the autocorrelation function (ACF), partial autocorrelation function (PACF), Q-statistics, and associated p-values. At each lag, the p-values are all greater than the 5% significance level, with the values ranging from 0.258 to 0.872. This indicates that the null hypothesis of no autocorrelation cannot be rejected at any lag.

Table 6: Ljung-box Q-statistic Test for Serial Correlation

Lag	ACF	PACF	Q-Statistic	p-value
1	-0.025	-0.025	0.1105	0.740
2	-0.121	-0.122	2.7092	0.258
3	-0.021	-0.028	2.7911	0.425
4	-0.119	-0.138	5.3307	0.255
5	-0.038	-0.055	5.5955	0.348
6	0.038	0.000	5.8562	0.439
7	-0.019	-0.039	5.9246	0.549
8	-0.051	-0.070	6.4046	0.602
9	-0.017	-0.043	6.4599	0.693
10	-0.011	-0.031	6.4828	0.773
11	-0.032	-0.055	6.6784	0.825
12	-0.023	-0.057	6.7749	0.872

The results of Ljung-Box Q-statistic reported in Table 6 provide strong evidence that there is no significant serial correlation in the residuals of the estimated model up to lag 12. This suggests that the model is well-specified in terms of capturing the dynamic structure of the data, and that the residuals behave like white noise. Consequently, the reliability of the model's parameter estimates and standard errors is reinforced, thereby supporting valid statistical inferences.

Parameter Estimates of Panel Vector Error Correction Model (PVECM)

The Panel Vector Error Correction Model (PVECM) was estimated to examine the short-run and long-run dynamics between Human Development Index (HDI) and a set of key socio-demographic indicators across countries over time. The model incorporates both differenced (short-run) variables and the error correction term (EC (-1)), which captures the speed of adjustment toward long-run equilibrium. The variables include life expectancy at birth

(LEB), population growth (PG), fertility rate (FR), maternal mortality rate (MMR), urbanization rate (UR), dependency ratio (DR), population density (PD), youth population (YP), and female labour force participation rate (FLP). All variables are used in their first differenced logarithmic forms (denoted as DL), while the error correction term reflects deviations from long-run equilibrium.

Table 7: Parameter Estimates of Panel Vector Error Correction Model

Variable	Coefficient	Std. Error	t-Statistic	p-value
DLHDI(-1)	0.35175	0.07632	4.60823	0.0000
DLLEB	0.31857	0.09279	3.43324	0.0002
DLLEB(-1)	0.25386	0.08789	2.88838	0.0052
DLLEB(-2)	0.21635	0.08163	2.65037	0.0071
DLPG	-0.33657	0.09872	-3.40934	0.0004
DLPG(-1)	-0.28628	0.08726	-3.28077	0.0007
DLPG(-2)	-0.20951	0.08563	-2.44669	0.0067
DLFR	-0.42926	0.05825	-7.36927	0.0000
DLFR(-1)	-0.27852	0.09975	-2.79218	0.0059
DLFR(-2)	-0.22853	0.17531	-1.30358	0.4582
DLMMR	-0.18933	0.09217	-2.05413	0.0344
DLMMR(-1)	-0.09783	0.04117	-2.37624	0.0087
DLMMR(-2)	-0.11755	0.16432	-0.71537	0.1962
DLUR	0.27433	0.09135	3.00307	0.0009
DLUR(-1)	0.22578	0.08973	2.51622	0.0068
DLUR(-2)	0.18475	0.08653	2.13509	0.0157
DLDR	-0.55129	0.14724	-3.74416	0.0000
DLDR(-1)	-0.37883	0.11045	-3.42988	0.0004
DLDR(-2)	-0.32847	0.10722	-3.06351	0.0009
DLPD	-0.32175	0.09976	-3.22524	0.0007
DLPD(-1)	0.21188	0.08951	2.36711	0.0071
DLPD(-2)	-0.08874	0.04101	-2.16386	0.0432
DLYP	0.25145	0.10172	2.47198	0.0083
DLYP(-1)	-0.18659	0.08223	-2.26912	0.0138
DLYP(-2)	0.09674	0.15372	0.62933	0.1759
DLFLP	0.21935	0.07439	2.94865	0.0032
DLFLP(-1)	0.18652	0.06998	2.66533	0.0064
DLFLP(-2)	0.09758	0.04738	2.05952	0.0357
EC(-1)	-0.77953	0.09526	-8.183183	0.0000

The results of the Panel Vector Error Correction Model (PVECM) reported in Table 7 provide important understanding of the short-run and long-run relationships between human development (measured by HDI) and key socio-demographic indicators. The error correction term (EC (-1)) is negative and highly statistically significant (−0.7795, $p < 0.01$), confirming the existence of a stable long-run relationship among the variables. This coefficient implies that approximately 78% of the deviation from long-run equilibrium is corrected within a single period, indicating a strong and relatively fast speed of adjustment toward equilibrium whenever shocks occur in the system.

The lagged dependent variable, HDI (-1), is positive and statistically significant, suggesting that HDI is highly persistent over time past levels of human development positively influence current improvements. Among the socio-demographic factors, life expectancy at birth (LEB) and its two lags (DLLEB, DLLEB (-1), and DLLEB (-2)) are all positively and significantly associated with HDI. This highlights the critical role of health and longevity in enhancing human development. Similarly, urbanization (UR) and its lags exhibit positive and significant effects, indicating that increasing urbanization contributes to better development outcomes through improved access to social amenities and infrastructure.

Female labour force participation (FLP) and its lags also show a strong, positive influence on HDI, underscoring the role of gender inclusion and economic empowerment of women in fostering development. The youth population (YP) exhibits a mixed effect: while its current value has a positive and significant effect on HDI, the first lag is negative, implying that the developmental benefit of a youthful population depends on timely investments in education, employment, and empowerment.

Conversely, several socio-demographic variables have a negative effect on HDI. Population growth (PG) and fertility rate (FR) significantly reduce HDI both contemporaneously and through their lags, reflecting the burden of high population pressure and reproductive rates on health, education, and income per capita. Maternal mortality rate (MMR) negatively affects HDI, signifying poor healthcare systems and women's vulnerability. The dependency ratio (DR) consistently shows strong negative effects, indicating that a high proportion of dependents to working-age population hinders human development. Finally, population density (PD) has a complex influence: while its first lag has a positive effect, the current and second lag are negative, suggesting that the benefits of density depend on the quality of urban management and infrastructure.

Overall, the PVECM results demonstrate that socio-demographic factors play a significant role in shaping human development outcomes. Improvements in health, urbanization, gender participation, and youth empowerment enhance HDI, while high population growth, fertility, maternal mortality, dependency burden, and unmanaged population density undermine it. These findings call for socio-demographic-responsive policies to support sustainable human development.

Dumitrescu Hurlin Pairwise Panel Granger Causality Tests Result

To understand the direction of causality between human development and selected demographic indicators in a panel data context, the Dumitrescu-Hurlin (2012) Pairwise Granger Causality Test was conducted. This test assesses homogeneous non-causality across panel units, accommodating heterogeneity in causal relationships across countries. The null hypothesis states that there is no Granger causality from one variable to another for all panel units. A rejection of the null implies the existence of causality for at least some cross-sectional units. The test uses both the W-statistic and the $\bar{Z}$-bar standardized statistic, along with p-values to determine significance. All variables were used in their first-differenced log form (DL), indicating short-run relationships.

Table 8: Pairwise Dumitrescu Hurlin Panel Causality Tests Result

Null Hypothesis	W-Stat.	Zbar-Stat.	p-value
DLDR does not homogeneously cause DLHDI	11.6109	9.55829	0.0000*
DLHDI does not homogeneously cause DLDR	1.06956	-1.97932	0.4132
DLFLP does not homogeneously cause DLHDI	8.60728	5.42136	0.0000*
DLHDI does not homogeneously cause DLFLP	4.53926	2.67727	0.0135*
DLFR does not homogeneously cause DLHDI	12.2073	10.1635	0.0000*
DLHDI does not homogeneously cause DLFR	1.42521	-0.77820	0.4364
DLLEB does not homogeneously cause DLHDI	10.5926	7.94661	0.0000*
DLHDI does not homogeneously cause DLLEB	7.40776	6.16125	0.0000*
DLMMR does not homogeneously cause DLHDI	4.56993	2.41307	0.0158*
DLHDI does not homogeneously cause DLMMR	1.20161	-1.00512	0.3148
DLPD does not homogeneously cause DLHDI	2.37465	0.18529	0.8530
DLHDI does not homogeneously cause DLPD	1.39237	-0.81153	0.4171
DLPG does not homogeneously cause DLHDI	11.8517	9.92032	0.0000*
DLHDI does not homogeneously cause DLPG	1.14591	-1.06164	0.2884
DLUR does not homogeneously cause DLHDI	11.0704	9.00982	0.0000*
DLHDI does not homogeneously cause DLUR	1.69511	-0.50431	0.6140
DLYP does not homogeneously cause DLHDI	5.08097	2.93168	0.0034*
DLHDI does not homogeneously cause DLYP	7.74283	3.45589	0.0005*

The Dumitrescu-Hurlin panel causality test reveals multiple significant short-run causal relationships between demographic indicators and human development (HDI). Notably, there is strong unidirectional causality from the dependency ratio (DLDR) to HDI, with no reverse causality is detected. This suggests that variations in the dependency burden significantly influence changes in HDI, while HDI does not significantly predict shifts in dependency ratio. A similar pattern emerges for fertility rate (DLFR), population growth (DLPG), and maternal mortality rate (DLMMR)—each Granger-causes HDI without reciprocal influence, indicating that these demographic pressures are key drivers of short-term fluctuations in development outcomes.

Conversely, life expectancy at birth (DLLEB) exhibits bidirectional causality with HDI, suggesting a feedback relationship: not only does improved life expectancy foster human development, but rising HDI also contributes to gains in life expectancy, possibly through better access to healthcare and education. Youth population (DLYP) and female labour force participation (DLFLP) also demonstrate mutual causality with HDI. These results indicate that investing in youth and promoting gender inclusion in the workforce are not only beneficial to human development but are themselves reinforced by improvements in development indicators.

Meanwhile, urbanization rate (DLUR) Granger-causes HDI, but HDI does not Granger-cause urbanization, underscoring the importance of urban migration and infrastructural expansion in shaping development outcomes. On the other hand, population density (DLPD) and HDI show no significant causal relationship in either directions, suggesting that population concentration alone does not have a consistent short-term effect on development across regions in Nigeria.

Overall, the causality results emphasize the predominant role of demographic characteristics particularly fertility, dependency, maternal mortality, and urbanization in influencing human development in the short run. They also highlight the reinforcing effects of development on life expectancy, gender inclusion, and youth empowerment. These findings reinforce the need for comprehensive socio-demographic-sensitive policies aimed at sustaining human development gains.

CONCLUSION

This study examined the socio-demographic determinants of human development in Nigeria using panel data from the six geopolitical zones covering the period 1994-2024. Specifically, the study sought to determine the long-run effects of selected socio-demographic variables on the Human Development Index (HDI), investigate the short-run adjustment dynamics among the variables, and establish the direction of causality between human development and key demographic indicators. Using Dynamic Fully Modified Ordinary Least Squares (FMOLS), Panel Vector Error Correction Model (PVECM), and Dumitrescu-Hurlin panel causality techniques, the study provided comprehensive evidence on the demographic drivers of human development in Nigeria.

The findings revealed that life expectancy, urbanization, female labour force participation, and youth population positively contribute to human development, while population growth, fertility rate, maternal mortality, dependency ratio, and population density exert significant negative effects on HDI. The PVECM results further confirmed the existence of a stable long-run equilibrium relationship among the variables, while the causality analysis identified important directional linkages between socio-demographic factors and human development outcomes.

A major contribution of this study to the panel human development literature is the integration of FMOLS, PVECM, and panel causality techniques within a single analytical framework to simultaneously examine long-run relationships, short-run dynamics, and causal interactions among socio-demographic variables and human development. Furthermore, by utilizing a balanced panel dataset across Nigeria's six geopolitical zones over three decades, the study provides more robust and regionally representative evidence than many previous studies that relied solely on national time-series data or single-method approaches.

Despite these contributions, the study has some limitations. First, the analysis was conducted at the geopolitical-zone level, which may mask important variations among individual states within each region. Second, the study was constrained by the availability of consistent historical data, limiting the inclusion of some potentially relevant variables such as poverty incidence, income inequality, governance quality, healthcare expenditure, educational attainment, and environmental sustainability indicators. Third, although the selected variables explain important dimensions of human development, other social, institutional, and economic factors may also influence HDI and were not explicitly incorporated into the model.

The findings have important implications for achieving the Sustainable Development Goals (SDGs), particularly SDG 1 (No Poverty), SDG 3 (Good Health and Well-being), SDG 4 (Quality Education), SDG 5 (Gender Equality), SDG 8 (Decent Work and Economic Growth), and SDG 10 (Reduced Inequalities). The results also align with Nigeria's national development priorities, including the National Development Plan (2021-2025) and related human capital development strategies aimed at improving health outcomes, enhancing women's economic participation, reducing maternal mortality, managing demographic pressures, and harnessing the demographic dividend for sustainable growth.

Future studies may extend this research by employing state-level data to capture subnational heterogeneity more effectively. Researchers may also explore spatial panel models to account for regional spillover effects in human development, as well as nonlinear and threshold panel approaches to examine whether the effects of socio-demographic variables vary across different levels of development. Additionally, incorporating institutional, environmental, and governance-related variables could provide a more comprehensive understanding of the factors shaping human development in Nigeria and other developing economies.

Based on the empirical findings of this study, the following recommendations are hereby presented:

i. The government should strengthen healthcare systems to improve life expectancy and reduce maternal mortality across all regions in Nigeria.
ii. Policies aimed at promoting female labour force participation should be enhanced to boost inclusive economic growth and human development.
iii. Effective population control measures, including family planning and reproductive health education, should be implemented to manage high fertility and population growth rates.
iv. Investments in youth development, education, and infrastructure should be increased to harness the potential benefits of a growing and dense population.

REFERENCES

Abamba, O., Aramide, V. O., Adesipo, A. E., & Akunna, L. C. (2023). Correlates of human development in Africa: Evidence across gender and income groups. *Research in Globalization, 6*, 100-115.

Anand, S., & Sen, A. (2000). Human development and economic sustainability. *World Development, 28*(12), 2029–2049.

Barro, R. J., & Lee, J.-W. (2015). *Education matters: Global schooling gains from the 19th to the 21st century*. Oxford University Press.

Bloom, D. E., Canning, D., & Sevilla, J. (2003). *The demographic dividend: A new perspective on the economic consequences of population change*. Santa Monica, CA: RAND Corporation.

Černák, J. (2017). Low fertility rate reversal and human development: A cross-country analysis. *MPRA Paper No. 76212*, University Library of Munich.

Chani, M. I., Shahid, M., & Hassan, M. (2012). Some socio-economic determinants of fertility in Pakistan: An empirical analysis. *Pakistan Journal of Social Sciences, 32*(1), 23–37.

Dumitrescu, E.-I., & Hurlin, C. (2012). Testing for granger non-causality in heterogeneous panels*. Economic Modelling*, *29*(4), 1450-1460

Gulcemal, T. (2020). Effect of human development index on GDP for developing countries: A panel data analysis. *Journal of Economics, Finance and Accounting, 7*(4), 401–410.

Johansen, S. (1988). Statistical analysis of cointegration vectors*. Journal of Economic Dynamics and Control*, *12*(2-3), 231-254.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models*. Econometrica*, *59*(6), 1551-1580.

Larsson, R., Lyhagen, J., & Lothgren, M. (2001). Likelihood-based cointegration tests in heterogeneous panels. *Econometric Journal*, *4*(1), 41-64.

Levin, A., Lin, C. F., & Chu, C. S. J., (2002). Unit root tests in panel data: asymptotic and finite-sample properties. *Journal of Econometrics*, *108*(1), 1-24.

Maddala, G. S., & Wu, S., (1999). A comparative study of unit root tests with panel data and a new simple test. *Oxford Bulletin of Economics and Statistics, 61, Supplement, 51*, 631-652.

Ogujiuba, K., Maponya, L., & Stiegler, N. (2024). Determinants of human development index in South Africa: A comparative analysis of different time periods. *World, 5*(3), 527–550. https://doi.org/10.3390/world5030027

Pedroni, P. (2000). Fully modified OLS for heterogeneous cointegrated panels. *Econometric Theory*, *16*(6), 1023-1051.

Pesaran, M. H. (2015). Testing weak cross-sectional dependence in large panels. *Econometric Reviews*, 34(6-10), 1089-1117.

Ranis, G., Stewart, F., & Samman, E. (2006). Human development: Beyond the human development index. *Journal of Human Development, 7*(3), 323–358.

Singh, K., Cheemalapati, S., Reddy, S., Kurian, G., Muzumdar, P., & Muley, A. (2025). Determinants of human development index (HDI): A regression analysis of economic and social indicators. *arXiv Preprint*, arXiv:2502.00006.

Suparman, S., & Wahyuningsih, W. (2022). Human development index and labour productivity: Evidence from Central Sulawesi Province. *Budapest International Research and Critics Institute (BIRCI-Journal): Humanities and Social Sciences, 5*(1), 4351–4361.

Todaro, M. P., & Smith, S. C. (2020). *Economic development* (13th ed.). Boston, MA: Pearson Education.

United Nations Development Programme. (2010). *Human development report 2010: The real wealth of nations—Pathways to human development*. New York, NY: UNDP.

United Nations Development Programme. (2022). *Human development report 2021/2022: Uncertain times, unsettled lives—Shaping our future in a transforming world*. UNDP.

United Nations Development Programme. (2023). *Human development report 2023/2024*. UNDP.

Wahab, A. (2024). Determinants of human capital development in Nigeria. *International Journal of Research and Innovation in Social Science, 8*(4), 1450–1463.