



Spatial Spillover of Human Development Index in West Java: An Application of Enclave Panel Threshold Spatial Autoregressive (E-PT-SAR) Model

Nidia Ayu^{1*}, Henny Pramodyo¹, Suci Atutik¹, and Friansyah Gani¹

¹*Department Statistics, Faculty of Mathematics and Natural Science, Brawijaya University, Malang, Indonesia*

Abstract

This study examines the spatial spillover dynamics of the Human Development Index (HDI) in West Java Province, Indonesia, using the Enclave Panel Threshold Spatial Autoregressive Model (E-PT-SAR) and panel data from 27 districts during 2021–2025. The analysis extends the conventional PT-SAR framework by incorporating enclave identification to distinguish between core prosperous regions, prosperous enclaves, and poor regions. The estimated poverty threshold separates West Java into prosperous and poor regimes, while the enclavity index identifies two prosperous enclave districts: Kabupaten Bandung and Ciamis. The results indicate that spatial spillovers differ between spatial regimes. Significant spillover effects were detected in the core prosperous and poor regions, while no statistically significant spillover effect was identified in the prosperous enclave regions. These findings are consistent with the hypothesis of the enclave effect, suggesting that spatially isolated prosperity may not diffuse to surrounding poorer areas. Among formally comparable specifications, the PT-SAR model provides the best statistical fit, while the E-PT-SAR framework offers additional insight into spatial heterogeneity through enclave identification. The findings imply that regional development policies should be adapted to different spatial regimes.

Keywords: Spatial Econometrics; E-PT-SAR; Enclave; Spatial Spillover; HDI; West Java.

Copyright © 2025 by Authors, Published by CAUCHY Group. This is an open access article under the CC BY-SA License (<https://creativecommons.org/licenses/by-sa/4.0>)

1. Introduction

West Java Province, Indonesia's most populous province, exhibits striking spatial disparities in human development. The northern region, comprising Bekasi, Karawang, and Purwakarta, has achieved high scores on the Human Development Index (HDI), driven by industrialization and urban agglomeration. In contrast, the southern region, including Cianjur, Sukabumi, Garut, and Tasikmalaya, consistently shows lower HDI, characterized by agricultural dependence and limited infrastructure. This spatial cluster pattern is not unique to West Java; it reflects a broader phenomenon in regional data where development indicators often exhibit spatial dependence, which means that regions with similar characteristics tend to cluster together [1]. This clustering raises a critical question: does prosperity in high-HDI regions spill over to neighboring areas, or do these regions develop independently?

*Corresponding author. E-mail: nidiaayu@srudent.ub.ac.id

A particularly interesting case in West Java is Kabupaten Bandung. Although classified as a relatively prosperous district with low poverty levels, it is geographically surrounded by neighboring districts with considerably higher poverty rates, including Bandung Barat, Garut, Cianjur, and Tasikmalaya. This configuration suggests the presence of an enclave phenomenon, namely a prosperous region spatially surrounded by poorer regions. The concept of an enclave was originally introduced in a mathematical context as a region isolated from others of its kind [2], and has subsequently been applied in regional development studies to describe economically advantaged areas that remain weakly connected to their surrounding local economies [3]. Similar enclave phenomena have been observed elsewhere. In Poland, wealthy regions surrounded by poor regions acted as spatial outliers, failing to generate spillovers to their poorer neighbors [4]. Farole found that Special Economic Zones (SEZs) sometimes develop as enclaves with limited links to local economies, potentially worsening conditions in surrounding areas by attracting investment and skilled labor away from them. These findings suggest that enclaves may exhibit fundamentally different spillover behavior compared to prosperous regions that form contiguous clusters.

The phenomenon of enclaves is important because it challenges the conventional assumption that prosperous regions naturally drive growth in neighboring areas. Spatial econometric models, such as the Spatial Autoregressive (SAR) model, generally assume that the relationship between neighboring regions is homogeneous across space, implying that spillover mechanisms operate similarly in all regions [5]. However, regional development processes are often heterogeneous. Prosperous industrial clusters in northern West Java can generate strong interregional spillovers due to integrated labor markets, transportation networks, and industrial links, whereas enclave regions can remain relatively isolated despite their economic advantages. If such heterogeneity is ignored, important regime-specific spillover behavior may be obscured.

To address the limitation of parameter homogeneity, Hansen developed the panel threshold model, which allows parameters to vary between regimes depending on whether an indicator crosses a certain threshold [6]. Hansen developed the panel threshold model, which allows parameters to vary between regimes depending on whether an indicator crosses a certain threshold. This framework was later integrated into spatial panel econometrics by Wei et al. through the Panel Threshold Spatial Durbin (PT-SDM) model. The PT-SDM framework improves upon homogeneous spatial models by allowing spillover effects to differ between regimes, such as prosperous and poor regions. Nevertheless, an important limitation remains: regions within the same regime are still assumed to behave homogeneously. In practice, prosperous regions can consist of both contiguous development clusters and spatially isolated enclaves. Consequently, the standard threshold spatial framework may not adequately distinguish the unique spillover behavior of enclave regions.

This study applies the Enclave Panel Threshold Spatial Autoregressive (E-PT-SAR) model to HDI data from 27 districts in West Java for the period 2021-2025, yielding 135 observations (27 districts $\times$ 5 years). HDI is used as the dependent variable, while explanatory variables include the open unemployment rate (TPT), the logarithm of GRDP per capita (log PDRBK), and the poverty rate [8], [9], [10], [11]. The study has three specific objectives. First, identify the optimal poverty threshold that distinguishes prosperous from poor regions in West Java. Second, identify prosperous enclaves using the enclavity index. Third, estimate and compare spillover effects across the three regimes: the prosperous core, prosperous enclave, and poor regions. This study contributes to the spatial econometrics literature in two ways. First, it extends the spatial threshold framework by incorporating enclave identification into a spatial autoregressive panel model. Second, it provides empirical evidence that prosperous enclave regions may exhibit different spillover behaviors from both the main prosperous and poor regions, particularly in the context of regional human development disparities in West Java.

2. Methodology

2.1. Data and Variables

This study uses panel data from 27 districts and cities in West Java Province, Indonesia, for the period 2021-2025, producing 135 observations (27 regions $\times$ over 5 years). All data were obtained from Statistics Indonesia (BPS). The dependent variable is the Human Development Index (HDI), while the explanatory variables include the Open Unemployment Rate (TPT), the logarithm of GRDP per capita (log PDRBK), and the poverty rate. Poverty is also used as the threshold variable in the PT-SAR and E-PT-SAR models. Table 1 summarizes the variables used in this study, including their definitions and units of measurement.

Table 1: Variable Definitions and Descriptions

Variable	Symbol	Description	Unit
Human Development Index (HDI)	y	Composite measure of health, education, and living standards	Index (0–100)
Open Unemployment Rate (TPT)	x_1	Percentage of labor force unemployed and seeking work	Percent (%)
GRDP per capita (PDRBK)	x_2	Total economic output per person	Thousand IDR
Poverty Rate	x_3	Percentage of population below the poverty line	Percent (%)

Note: all variables are measured on a ratio scale and obtained from BPS.

2.2. Panel Data Fixed Effects Model

Prior to spatial analysis, three non-spatial panel models are estimated: Pooled OLS, Fixed Effects (FE), and Random Effects (RE). The general specification is as follows:

$$y_{it} = \underline{x}_{it}'\beta + \mu_i + \varepsilon_{it} \quad (1)$$

where y_{it} is the dependent variable for region i at time t ; $\underline{x}_{it}$ is the vector of independent variables ($K \times 1$); β is the vector of coefficients ($K \times 1$); μ_i represents individual fixed effects that capture time-invariant regional characteristics; ε_{it} is the idiosyncratic error term assumed $\varepsilon_{it} \sim N(0, \sigma^2)$.

Pooled OLS assumes $\mu_i = 0 \forall i$. Fixed Effects (FE) allows μ_i to be correlated with x_{it} and eliminates μ_i via within transformation. Random Effects (RE) assumes μ_i is uncorrelated with x_{it} and treats μ_i as part of the error term [12].

2.2.1. Chow Test (Pooled OLS vs Fixed Effects)

The Chow test examines whether individual effects exist [13]. The null hypothesis is $H_0 : \mu_1 = \mu_2 = \dots = \mu_{N-1} = 0$ (Pooled OLS is appropriate), while the alternative hypothesis is H_1 : at least one $\mu_i \neq 0$ (FE is appropriate). The F-statistic is calculated as:

$$F = \frac{(RSS_{Pooled} - RSS_{FE})/(N - 1)}{RSS_{FE}/(NT - N - K)} \quad (2)$$

where RSS_{Pooled} is the residual sum of squares from the Pooled OLS Model; RSS_{FE} is the residual sum of squares from the Fixed Effect model; N is the number of regions; T is the number of time periods; K is the number of independent variables.

Under the null hypothesis, the F-statistic follows an F-distribution with $(N - 1)$ and $(NT - N - K)$ degrees of freedom. If the F-statistic exceeds the critical value (or p -value < 0.05), the null hypothesis is rejected, and the Fixed Effects model is preferred [12].

2.2.2. Hausman Test (Fixed Effects vs Random Effects)

The Hausman test compares the RE and FE estimators [14]. The null hypothesis is $H_0 : \text{Cov}(\mu_i, x_{it}) = 0$ (RE is consistent and efficient), while the alternative hypothesis is H_1 :

$\text{Cov}(\mu_i, x_{it}) \neq 0$ (FE is consistent, RE is inconsistent). The test statistic is calculated as:

$$H = (\hat{\beta}_{RE} - \hat{\beta}_{FE})' [\text{Var}(\hat{\beta}_{FE}) - \text{Var}(\hat{\beta}_{RE})]^{-1} (\hat{\beta}_{RE} - \hat{\beta}_{FE}) \quad (3)$$

where $\hat{\beta}_{FE}$ is the coefficient vector from the Fixed Effects model; $\hat{\beta}_{RE}$ is the coefficient vector from the Random Effects model; $\text{Var}(\hat{\beta}_{FE})$ and $\text{Var}(\hat{\beta}_{RE})$ are the respective variance-covariance matrices.

Under the null hypothesis, the H-statistic follows a chi-square distribution with K degrees of freedom. If the H-statistic exceeds the critical value (or p -value < 0.05), the null hypothesis is rejected, and the Fixed Effects model is preferred [12].

2.3. Spatial Weight Matrix (Queen Contiguity)

The spatial weight matrix defines the neighborhood structure among regions. This study adopts the Queen contiguity criterion, where two regions are considered neighbors if they share a common border or a common vertex, analogous to the movement of a queen in chess [1], [15], [16], [17]. The elements of the spatial weight matrix are defined as:

$$w_{ij} = \begin{cases} 1, & \text{if regions } i \text{ and } j \text{ share a border or a vertex} \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

The diagonal elements are set to zero ($w_{ii} = 0$) to prevent a region from being considered its own neighbor.

To ensure that the spatial autoregressive coefficient (ρ) can be interpreted as the average influence of neighboring regions, the matrix is standardized in rows so that each row sums up to one [1]. The row-standardized weight is computed as:

$$w_{ij}^* = \frac{w_{ij}}{\sum_{j=1}^N w_{ij}} \quad (5)$$

where N is the total number of regions. This transformation ensures that $\sum_{j=1}^N w_{ij}^* = 1$ for each row i . The resulting spatial weight matrix is used in all subsequent spatial tests and models.

2.4. Spatial Autocorrelation: Moran's I

To detect the presence of spatial dependence, Moran's I test is applied [15], [1], [16]. The test is applied to two objects: (i) the dependent variable y , and (ii) the residuals of the Fixed Effects model.

Moran's I on Dependent Variable (Initial Exploration). This test aims to determine whether the dependent variable (y) exhibits a spatial clustering pattern descriptively. The formula for Moran's I for panel data on the dependent variable is defined as follows [15]:

$$I_y = \frac{\sum_{i=1}^N \sum_{j=1}^N w_{ij} (y_{it} - \bar{y})(y_{jt} - \bar{y})}{\sum_{t=1}^T \sum_{i=1}^N (y_{it} - \bar{y})^2} \times \frac{N}{\sum_{i=1}^N \sum_{j=1}^N w_{ij}} \quad (6)$$

where N is the number of spatial observations (regions), T is the number of time periods, y_{it} is the value of the dependent variable in region i at time t , $\bar{y}$ is the mean of y across all regions and time periods, and w_{ij} is the element of the spatial weight matrix W . The hypotheses tested are:

1. $H_0 : I_y = E(I)$ (no spatial dependence in the dependent variable)
2. $H_1 : I_y \neq E(I)$ (spatial dependence exists in the dependent variable)

The expected value of Moran's I under the null hypothesis of no spatial dependence is [15]:

$$E(I) = -\frac{1}{N-1} \quad (7)$$

Statistical significance is calculated using a permutation (bootstrap) approach with 999 permutations. If the p -value < 0.05 , then H_0 is rejected and the dependent variable exhibits significant spatial autocorrelation.

Moran's I on FE Residuals. This test is a more formal procedure for deciding whether a spatial model is needed in regression analysis. The steps are as follows [1]:

1. Estimate the non-spatial panel model (Fixed Effects or Random Effects) as in equation (1).
2. Extract the residuals from the model, denoted $\hat{e}_{it}$.
3. Compute Moran's I on the residuals $\hat{e}_{it}$ using the following formula:

$$I_e = \frac{\sum_{t=1}^T \sum_{i=1}^N \sum_{j=1}^N w_{ij} \hat{e}_{it} \hat{e}_{jt}}{\sum_{t=1}^T \sum_{i=1}^N \hat{e}_{it}^2} \times \frac{N}{\sum_{i=1}^N \sum_{j=1}^N w_{ij}} \quad (8)$$

The hypotheses tested are:

1. $H_0 : I_e = 0$ (no spatial autocorrelation in the residuals)
2. $H_1 : I_e \neq 0$ (spatial autocorrelation exists in the residuals)

Statistical significance is calculated using a permutation (bootstrap) approach with 999 permutations [1]. If the p -value < 0.05 , then H_0 is rejected and the residuals still contain spatial dependence, indicating that a spatial econometric model is required.

2.5. Lagrange Multiplier (LM) Tests

After estimating the non-spatial panel model, the next step is to determine the most appropriate spatial specification. For this purpose, Lagrange Multiplier (LM) tests developed for panel data by [1] are employed [16], [17]. There are two main tests, each with a standard and a robust version.

Standard LM Tests. These tests detect the presence of spatial dependence in general, either in the form of a spatial lag or a spatial error.

1. LM_{error} detects spatial dependence in the error term ($H_0 : \lambda = 0$):

$$LM_{error} = \frac{[\underline{e}'(I_T \otimes W)\underline{e}/\hat{\sigma}^2]^2}{T \cdot \text{tr}(W^2 + W'W)} \sim \chi^2(1) \quad (9)$$

2. LM_{lag} detects spatial dependence in the dependent variable ($H_0 : \rho = 0$):

$$LM_{lag} = \frac{[\underline{e}'(I_T \otimes W)\underline{y}/\hat{\sigma}^2]^2}{D} \sim \chi^2(1) \quad (10)$$

with $D = T \cdot \text{tr}(W^2 + W'W) + \frac{T^2(WX\underline{\beta})'M(WX\underline{\beta})}{\hat{\sigma}^2}$, $M = I_N - X(X'X)^{-1}X'$, and $\underline{e}$ is the vector of residuals, X is the matrix of independent variables, $\underline{y}$ is the vector of the dependent variable.

Robust LM Tests (RLM). Standard LM tests may be biased when both types of dependence (lag and error) are present simultaneously. To address this issue, robust LM tests are used [1], which can more accurately distinguish the source of dependence.

1. Robust LM Lag tests for spatial lag dependence ($H_0 : \rho = 0$) while controlling for the presence of spatial error:

$$RLM_{lag} = \frac{[\underline{e}'(I_T \otimes W)\underline{y}/\hat{\sigma}^2 - \underline{e}'(I_T \otimes W)\underline{e}/\hat{\sigma}^2]^2}{T \cdot \text{tr}(W^2 + W'W) - [\underline{e}'(I_T \otimes W)\underline{e}]^2 / (\hat{\sigma}^2 T \cdot \text{tr}(W^2 + W'W))} \sim \chi^2(1) \quad (11)$$

2. Robust LM Error tests for spatial error dependence ($H_0 : \lambda = 0$) while controlling for the presence of spatial lag:

$$RLM_{error} = \frac{\left[\underline{e}'(I_T \otimes W)\underline{e}/\hat{\sigma}^2 - \underline{e}'(I_T \otimes W)\underline{y}/\hat{\sigma}^2 \right]^2}{T \cdot \text{tr}(W^2 + W'W) - \left[\underline{e}'(I_T \otimes W)\underline{y} \right]^2 / (\hat{\sigma}^2 T \cdot \text{tr}(W^2 + W'W))} \sim \chi^2(1) \quad (12)$$

The LM test for spatial lag dependence examines whether the spatially lagged dependent variable (Wy) should be included in the model. The null hypothesis is $H_0 : \rho = 0$ (no spatial lag dependence), while the alternative hypothesis is $H_1 : \rho \neq 0$ (spatial lag dependence exists). Similarly, the LM test for spatial error dependence examines whether the error terms are spatially correlated, with $H_0 : \lambda = 0$ and $H_1 : \lambda \neq 0$. The procedure for selecting the appropriate spatial model is as follows [18]:

1. Estimate the non-spatial panel model (Fixed Effects) and obtain the residuals $\hat{e}_{it}$.
2. Compute the four LM test statistics: LM_{lag} , LM_{error} , RLM_{lag} , RLM_{error} .
3. Compare these values with the critical chi-square value at 1 degree of freedom ($\chi_{0.05,1}^2 = 3.841$).

The decision rule is:

1. If RLM_{lag} is significant and RLM_{error} is not, the SAR Model is chosen.
2. If RLM_{error} is significant and RLM_{lag} is not, the SEM Model is chosen.
3. If both are significant, the Spatial Durbin Model (SDM) may be considered.
4. If neither is significant, the non-spatial Fixed Effects model is sufficient.

2.6. Spatial Autoregressive (SAR) Model

The Panel Spatial Autoregressive (SAR) model with fixed effects captures spatial spillovers by including the spatially lagged dependent variable as an additional regressor [5]. The model is specified as:

$$y_{it} = \rho \sum_{j=1}^N w_{ij} y_{jt} + \underline{x}_{it}' \underline{\beta} + \mu_i + \varepsilon_{it} \quad (13)$$

In this specification, the term $\sum_{j=1}^N w_{ij} y_{jt}$ captures the spatial spillover effect: the dependent variable in region i is influenced by the weighted average of the dependent variable in its neighboring regions. The coefficient ρ measures the strength and direction of this spillover:

1. If $\rho > 0$ and significant, positive spatial spillover (neighbors' values positively influence region i),
2. If $\rho < 0$ and significant, negative spatial spillover (neighbors' values negatively influence region i),
3. If $\rho = 0$ or not significant, no spatial spillover effect exists.

The spatial autoregressive coefficient ρ must satisfy the condition $|\rho| < 1$ for the spatial process to be stable [1]. This condition ensures that the spatial multiplier matrix $(I_N - \rho W)^{-1}$ exists and can be expanded as an infinite series.

The spatial lag term $\sum_{j=1}^N w_{ij} y_{jt}$ is endogenous because it is correlated with the error term [15]. Therefore, the model cannot be estimated consistently by OLS. Instead, Maximum Likelihood (ML) or Spatial Two-Stage Least Squares (S2SLS) is employed [19]. The fixed effects μ_i are eliminated using the within transformation prior to estimation [1].

2.7. Panel Threshold SAR (PT-SAR) Model

The standard SAR model in equation (13) assumes parameter homogeneity across all regions. However, in regional development contexts, the relationship between variables may change when

an indicator crosses a certain threshold [6]. The Panel Threshold SAR (PT-SAR) model allows parameters to vary across two regimes determined by a threshold variable. Following Hansen and Wei et al., the PT-SAR model with fixed effects is specified as [6], [7]:

$$y_{it} = \begin{cases} \rho_1 \sum_{j=1}^N w_{ij} y_{jt} + \underline{x}_{it}' \underline{\beta}_1 + \mu_i + \varepsilon_{it}, & q_{it} \leq \gamma \\ \rho_2 \sum_{j=1}^N w_{ij} y_{jt} + \underline{x}_{it}' \underline{\beta}_2 + \mu_i + \varepsilon_{it}, & q_{it} > \gamma \end{cases} \quad (14)$$

where $\bar{q}_i = \frac{1}{T} \sum_{t=1}^T q_{it}$ is the time-averaged threshold variable; γ is the unknown threshold parameter to be estimated; ρ_1, ρ_2 are the spatial autoregressive coefficients for low and high regimes; $\underline{\beta}_1, \underline{\beta}_2$ are the coefficient vectors for low and high regimes.

The indicator function implicitly defines the two regimes: Regime 1 (Low/Prosperous) $q_{it} \leq \gamma$, and Regime 2 (High/Poor) $q_{it} > \gamma$.

The threshold parameter γ splits the sample into two distinct regimes. Regions with average poverty below or equal to γ are classified as prosperous (Regime 1), while regions with average poverty above γ are classified as poor (Regime 2). The threshold is estimated endogenously from the data using a grid search procedure [6]. The threshold parameter γ is estimated by minimizing the sum of squared residuals from the panel threshold model:

$$\hat{\gamma} = \arg \min_{\gamma \in \Gamma} SSR(\gamma) \quad (15)$$

where Γ is the set of candidate threshold values (typically the trimmed quantiles of the threshold variable distribution), and $SSR(\gamma)$ is the sum of squared residuals when the sample is split using γ .

The PT-SAR model allows the spillover effect to differ across regimes:

1. If ρ_1 is significant but ρ_2 is not, spillover exists only in prosperous regions,
2. If ρ_2 is significant but ρ_1 is not, spillover exists only in poor regions,
3. If both are significant but different in magnitude, spillover strength differs across regimes,
4. If both are not significant, no spatial spillover in either regime.

2.8. Enclave Identification

While the PT-SAR model distinguishes between prosperous and poor regions based on the poverty threshold γ , it treats all prosperous regions as homogeneous. However, a prosperous region may either be part of a contiguous cluster of prosperous regions or be spatially isolated in the midst of poor regions. The latter is referred to as a prosperous enclave [3].

To identify enclaves, the enclavity index (E) is calculated for each prosperous district. E measures the proportion of neighbors belonging to the poor regime [20]:

$$E_i = \frac{\sum_{j=1}^N w_{ij} \cdot I(r_j \neq r_i)}{\sum_{j=1}^N w_{ij}} \quad (16)$$

where w_{ij} are the elements of the spatial weight matrix (1 if regions i and j are neighbors, 0 otherwise); r_i denotes the regime classification of region i (1 for prosperous, 2 for poor); $I(\cdot)$ is the indicator function equal to 1 if the condition holds and 0 otherwise.

The enclavity index E_i ranges from 0 to 1. A value of 0 indicates that all neighbors belong to the same regime (spatial cluster), while a value of 1 indicates that all neighbors belong to the opposite regime (fully isolated enclave). This index is an extension of the mathematical definition of enclaves proposed by Suzuki and Inohara [2]. Based on this classification, prosperous regions are divided into two distinct categories:

1. **Core Prosperous Regions:** $\bar{q}_i \leq \gamma$ and $E_i < \tau$
2. **Prosperous Enclaves:** $\bar{q}_i \leq \gamma$ and $E_i \geq \tau$

Together with the poor regions ($\bar{q}_i > \gamma$), this yields three mutually exclusive regimes that form the basis for the E-PT-SAR model.

2.9. Enclave PT-SAR Model

After identifying the three regimes (core prosperous, prosperous enclave, and poor), the Enclave Panel Threshold Spatial Autoregressive (E-PT-SAR) model is specified. This model extends the PT-SAR framework by incorporating a three-regime structure based on both poverty levels and spatial configuration [20]. The E-PT-SAR model with fixed effects is specified as:

$$y_{it} = \begin{cases} \rho_1 \sum_{j=1}^N w_{ij} y_{jt} + \underline{x}_{it}' \underline{\beta}_1 + \mu_i + \varepsilon_{it}, & \bar{q}_i \leq \gamma \text{ and } E_i < \tau \\ \rho_2 \sum_{j=1}^N w_{ij} y_{jt} + \underline{x}_{it}' \underline{\beta}_2 + \mu_i + \varepsilon_{it}, & \bar{q}_i \leq \gamma \text{ and } E_i \geq \tau \\ \rho_3 \sum_{j=1}^N w_{ij} y_{jt} + \underline{x}_{it}' \underline{\beta}_3 + \mu_i + \varepsilon_{it}, & \bar{q}_i > \gamma \end{cases} \quad (17)$$

where y_{it} is the dependent variable (HDI); $\underline{x}_{it}$ is the vector of independent variables (TPT, Log PDRBK, Poverty Rate); $\bar{q}_i$ is the time-averaged threshold variable (poverty rate); γ is the poverty threshold parameter; E_i is the enclavity index; τ is the enclave threshold parameter; ρ_1, ρ_2, ρ_3 are the spatial autoregressive coefficients for each regime; $\underline{\beta}_1, \underline{\beta}_2, \underline{\beta}_3$ are the slope coefficients for each regime; μ_i is the individual fixed effects; ε_{it} is the idiosyncratic error term.

The three regimes are mutually exclusive and collectively exhaustive:

1. **Core Prosperous Regime:** $\bar{q}_i \leq \gamma$ and $E_i < \tau$ (prosperous districts that form a contiguous cluster with other prosperous districts)
2. **Prosperous Enclave Regime:** $\bar{q}_i \leq \gamma$ and $E_i \geq \tau$ (prosperous districts that are spatially isolated, surrounded by poor neighbors)
3. **Poor Regime:** $\bar{q}_i > \gamma$ (districts with high poverty rates, regardless of their neighbors' status)

This specification allows each regime to have its own spatial spillover coefficient (ρ_1, ρ_2, ρ_3) and its own slope coefficients ($\underline{\beta}_1, \underline{\beta}_2, \underline{\beta}_3$). This enables the model to test whether prosperous enclaves exhibit different spillover behavior compared to core prosperous regions.

The fixed effects μ_i are eliminated using the within transformation prior to estimation. The model is estimated using Spatial Two-Stage Least Squares (S2SLS) with instruments $H = [\tilde{X}, W\tilde{X}, W^2\tilde{X}]$ [19]. For complete derivation, estimation procedures, and asymptotic properties, we refer readers to our companion paper [20].

3. Results and Discussion

This section presents the empirical results of the E-PT-SAR model applied to HDI data in West Java Province for the period 2021–2025. The analysis proceeds in a sequential manner, beginning with descriptive statistics and spatial dependence tests, followed by panel model selection, threshold identification, PT-SAR estimation, enclave identification, E-PT-SAR estimation, decay effect analysis, and model comparison. The section concludes with a discussion of the findings and their policy implications.

3.1. Descriptive Statistics

This subsection provides an overview of the variables used in this study over the period 2021–2025. Table 2 presents the descriptive statistics for all variables.

Table 2: Descriptive Statistics (2021–2025)

Variable	N	Mean	Median	Std. Dev	Min	Max
HDI	135	74.10	73.30	4.40	67.20	84.70
TPT (%)	135	7.49	7.65	2.29	1.52	13.07
log PDRBK	135	10.55	10.46	0.98	8.12	12.59
Poverty Rate (%)	135	8.45	8.73	2.75	2.34	13.13

Source: BPS Jawa Barat, authors' calculation

The descriptive statistics reveal substantial variation across districts. HDI ranges from 67.20 (Cianjur) to 84.70 (Kota Bandung), indicating significant development disparities between the most and least developed regions. The mean HDI of 74.10 suggests moderate human development on average, though the standard deviation of 4.40 indicates notable heterogeneity among the 27 districts.

Poverty rates range from 2.34% (Kota Depok) to 13.13% (Kota Tasikmalaya), with a mean of 8.45%. TPT shows considerable variation, ranging from 1.52% (Pangandaran) to 13.07% (Kota Cimahi), reflecting diverse labor market conditions. The log PDRBK has a mean of 10.55 with a standard deviation of 0.98, indicating moderate variation in economic output per capita.

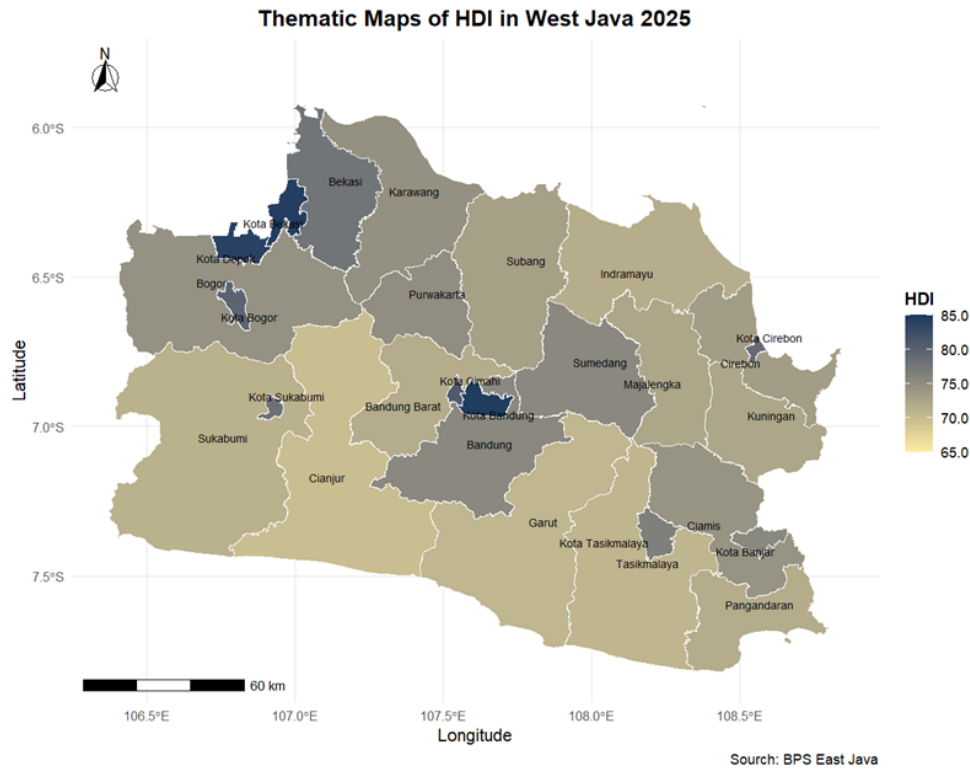


Figure 1: Thematic Maps of HDI in West Java 2025

Figure 1 presents the thematic map of HDI distribution across 27 districts in West Java for 2025. Darker shades indicate higher HDI values. The map reveals a clear north-south divide: northern districts (Bekasi, Karawang, Purwakarta) and metropolitan areas (Kota Bekasi, Kota Bogor, Kota Depok, Kota Bandung) exhibit higher HDI scores (darker shades), while southern districts (Cianjur, Sukabumi, Garut, Tasikmalaya) show lower HDI values (lighter shades). Notably, Kota Bandung appears as a high-HDI district (84.70) surrounded by lower-HDI neighbors (Bandung Barat with 71.65 and Cianjur with 69.84), suggesting a potential enclave phenomenon. This spatial pattern remained relatively stable throughout the 2021–2025 period, with only gradual HDI improvements across all districts.

3.2. Panel Data Model Selection

Before introducing spatial effects, the appropriate non-spatial panel model must be selected. Three specifications were estimated: Pooled OLS, Fixed Effects (FE), and Random Effects (RE). Table 3 presents the results of the Chow and Hausman specification tests.

Table 3: Panel Model Selection Tests

Test	Hypothesis	Statistic	<i>p</i> -value	Decision
Chow Test	H_0 : Pooled OLS, H_1 : FE	$F = 1070.67$	< 0.001	Reject H_0 , Choose FE
Hausman Test	H_0 : RE, H_1 : FE	$\chi^2 = 1318.04$	< 0.001	Reject H_0 , Choose FE

The Chow test examines whether individual effects exist. The null hypothesis assumes that Pooled OLS is appropriate (all $\mu_i = 0$), while the alternative hypothesis assumes that Fixed Effects are appropriate (at least one $\mu_i \neq 0$). The F-statistic of 1070.67 with p -value < 0.001 leads to rejection of the null hypothesis. Thus, the Fixed Effects model is preferred over Pooled OLS, indicating the presence of significant unobserved individual heterogeneity across districts [12].

The Hausman test compares the Fixed Effects and Random Effects estimators. The null hypothesis assumes that Random Effects is consistent and efficient (no correlation between μ_i and x_{it}), while the alternative hypothesis assumes that Fixed Effects is consistent (correlation exists). The chi-square statistic of 1318.04 with p -value < 0.001 leads to rejection of the null hypothesis. This indicates that individual effects are correlated with the regressors, favoring the Fixed Effects model over Random Effects [12].

Based on both specification tests, the Fixed Effects (FE) model is selected as the baseline non-spatial panel model for subsequent analysis. This model controls for time-invariant unobserved heterogeneity across districts, such as geographic characteristics, historical factors, and institutional differences that may affect human development outcomes.

3.3. Spatial Dependence Test

Before estimating spatial models, the presence of spatial dependence must be formally tested. Three sets of tests are employed: (i) Moran's I on the dependent variable (HDI), (ii) Moran's I on Fixed Effects residuals, and (iii) Lagrange Multiplier (LM) tests to determine the appropriate spatial specification.

Moran's I on HDI. Table 4 presents Moran's I test results for HDI from 2021 to 2025.

Table 4: Moran's I on HDI (2021–2025)

Year	Moran's I	p -value	Conclusion
2021	0.283	0.0107	Significant
2022	0.281	0.0112	Significant
2023	0.277	0.0121	Significant
2024	0.275	0.0127	Significant
2025	0.284	0.0105	Significant

The positive and significant Moran's I values for all years ($p < 0.05$) indicate the presence of positive spatial autocorrelation. Districts with high HDI tend to be surrounded by other high-HDI districts, while districts with low HDI cluster with low-HDI districts [15]. This confirms that spatial dependence exists in the dependent variable, justifying the use of spatial econometric models.

Moran's I on FE Residuals. To test whether the Fixed Effects model adequately captures spatial dependence, Moran's I was computed on the FE residuals.

Table 5: Moran's I on FE Residuals

Statistic	Value	p -value
Moran's I	0.167	0.072

The Moran's I on FE residuals is not significant at the 5% level ($p = 0.072$). However, as argued by Anselin and Florax (1995), Moran's I has lower power than LM tests in detecting spatial dependence, particularly in finite samples with limited time periods ($T = 5$ in this study). Therefore, the non-significance of Moran's I on residuals does not preclude the existence of spatial spillovers. The LM tests presented in the next subsection provide more reliable evidence.

Lagrange Multiplier (LM) Tests. To determine whether the Spatial Lag Model (SAR) or Spatial Error Model (SEM) is more appropriate, LM tests were conducted on the FE residuals. Table 6 presents the results.

Table 6: Lagrange Multiplier Tests

Test	Statistic	<i>p</i> -value	Conclusion
LM Lag	23.79	< 0.001	Significant
LM Error	2.75	0.097	Not significant
Robust LM Lag	41.57	< 0.001	Highly significant
Robust LM Error	20.54	< 0.001	Significant

The robust LM lag test (41.57) is larger than the robust LM error test (20.54), and both are significant. This indicates that the Spatial Lag Model (SAR) is more appropriate than the Spatial Error Model (SEM) [18]. The strong significance of the robust LM lag test ($p < 0.001$) provides clear evidence for including a spatially lagged dependent variable ($\rho W y$) in the model. Based on the LM test results, the Spatial Lag Model (SAR) is selected as the appropriate spatial specification. The analysis proceeds with the estimation of the baseline SAR model in Section 3.4.

3.4. Baseline SAR Model

Based on the LM test results, the Spatial Lag Model (SAR) is selected as the appropriate spatial specification. The Panel SAR model with fixed effects captures spatial spillovers by including the spatially lagged dependent variable as an additional regressor [5]. Table 7 presents the estimation results.

Table 7: Baseline SAR Model Estimation (2021–2025)

Parameter	Coefficient	Std. Error	<i>p</i> -value
λ (rho)	0.5306	0.0712	< 0.001
TPT	-0.0062	0.0186	0.739
log PDRBK	6.0367	1.0108	< 0.001
Poverty	0.0305	0.0564	0.589

The spatial autoregressive coefficient λ (rho) is positive and highly significant (0.5306, $p < 0.001$), confirming the presence of positive spatial spillover effects among districts. A 1% increase in neighboring HDI is associated with a 0.53% increase in a district's own HDI. This indicates that human development outcomes are not independent across regions; improvements in one district tend to benefit its neighbors.

Among the independent variables, only log PDRBK is significant ($p < 0.001$), with a coefficient of 6.04. This indicates a strong positive relationship between economic output and human development. A 1% increase in GRDP per capita is associated with a 6.04 unit increase in HDI. TPT and poverty are not significant in this specification, suggesting that their effects may be captured by other variables or may vary across regimes. The AIC value of -139.98 serves as a baseline for comparison with the threshold models in subsequent sections.

3.5. Poverty Threshold Identification

The baseline SAR model assumes parameter homogeneity across all districts. However, the relationship between poverty and HDI may change when poverty levels cross a certain threshold. Following Hansen, this study identifies an optimal poverty threshold that separates districts into two regimes: prosperous (low poverty) and poor (high poverty) [6]. The threshold variable is the time-averaged poverty rate for each district, calculated as $\bar{q}_i = \frac{1}{T} \sum_{t=1}^T q_{it}$, where q_{it} is the poverty rate of district i at time t . Following the grid search procedure described in Section 2.7, the optimal poverty threshold was identified as the median of the time-averaged poverty rates across districts, yielding $\gamma = 8.45\%$. This threshold separates 14 prosperous districts from 13 poor districts.

Table 8: Poverty Threshold Identification

Aspect	Value
Threshold variable	Poverty rate
Optimal threshold (γ)	8.45%
Prosperous regions (poverty $\leq$ 8.45%)	14 districts
Poor regions (poverty $>$ 8.45%)	13 districts

Prosperous regions (14 districts): Bekasi, Bogor, Karawang, Kota Bandung, Kota Banjar, Kota Bekasi, Kota Bogor, Kota Cimahi, Kota Depok, Kota Sukabumi, Purwakarta, Sukabumi, Kabupaten Bandung, Ciamis. Poor regions (13 districts): Bandung Barat, Cianjur, Cirebon, Garut, Indramayu, Kota Cirebon, Kota Tasikmalaya, Kuningan, Majalengka, Pangandaran, Subang, Sumedang, Tasikmalaya.

The threshold of 8.45% naturally separates West Java into two distinct groups. Prosperous regions are predominantly located in the northern industrial belt and major metropolitan areas, while poor regions are concentrated in the southern agricultural areas. This spatial pattern aligns with the visual clustering observed in Figure ??.

3.6. PT-SAR Model (Two Regimes)

After identifying the optimal poverty threshold ($\gamma = 8.45\%$), the PT-SAR model is estimated. Unlike the baseline SAR model which assumes all districts have the same spillover behavior, the PT-SAR model allows parameters to differ between prosperous and poor regimes. Table 9 presents the estimation results.

Table 9: PT-SAR Model Estimation (Two Regimes)

Parameter	Prosperous Regime	Poor Regime
ρ (spatial lag)	0.3311	0.2558
TPT	0.0412	-0.0410
log PDRBK	8.1639	10.5397
Poverty	-0.1523	0.1836
N (observations)	70	65

Note: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

The PT-SAR model yields several findings that distinguish it from the baseline SAR model. First, spatial spillovers are significant in both regimes, but they are stronger in prosperous regions ($\rho = 0.3311$) than in poor regions ($\rho = 0.2558$). This means that a 1% increase in neighboring HDI increases a district's own HDI by 0.33% if it is in a prosperous area, compared to 0.26% if it is in a poor area. The baseline SAR model assumed a single spillover coefficient of 0.5306 for all districts, masking this important difference.

Second, logarithmic PDRBK has a positive and highly significant effect in both regimes ($p < 0.001$), but the magnitude differs. In prosperous regions, a 1% increase in GRDP per capita increases HDI by 8.16 units, while in poor regions the same increase raises HDI by 10.54 units. This suggests that economic growth has a stronger marginal impact on human development in poorer areas, consistent with diminishing returns.

Third, poverty significantly reduces HDI only in prosperous regions (-0.1523 , $p = 0.015$). In poor regions, poverty is not statistically significant ($p = 0.109$), possibly due to limited variation in poverty within this group.

Fourth, TPT is significant only in prosperous regions (0.0412 , $p = 0.014$). This counterintuitive positive relationship may reflect the dynamics of the labor market where higher wages attract more job seekers, leading to higher measured unemployment alongside higher HDI.

In summary, the PT-SAR model demonstrates that spillover behavior and covariate effects differ systematically between prosperous and poor districts, a nuance that the homogeneous SAR model fails to capture.

3.7. Enclave Identification

While the PT-SAR model distinguishes between prosperous (14 districts) and poor (13 districts) regions, it treats all prosperous regions as homogeneous. However, a prosperous region may either be part of a contiguous cluster of prosperous regions or be spatially isolated in the midst of poor regions. The latter is referred to as a prosperous enclave [3].

To identify enclaves, the enclavity index (E) is calculated for each prosperous district, as defined in equation (16) of Section 2.8. In this study, a prosperous district is classified as an enclave if its enclavity index meets or exceeds a threshold of $\tau = 0.7$. This value was selected as the optimal parameter during a grid search procedure that jointly determines the poverty threshold (γ) and the enclave threshold (τ) by minimizing the sum of squared residuals [6]. The detailed procedure is described in our companion paper [20].

Table 10: Enclavity Index for Prosperous Districts (14 districts)

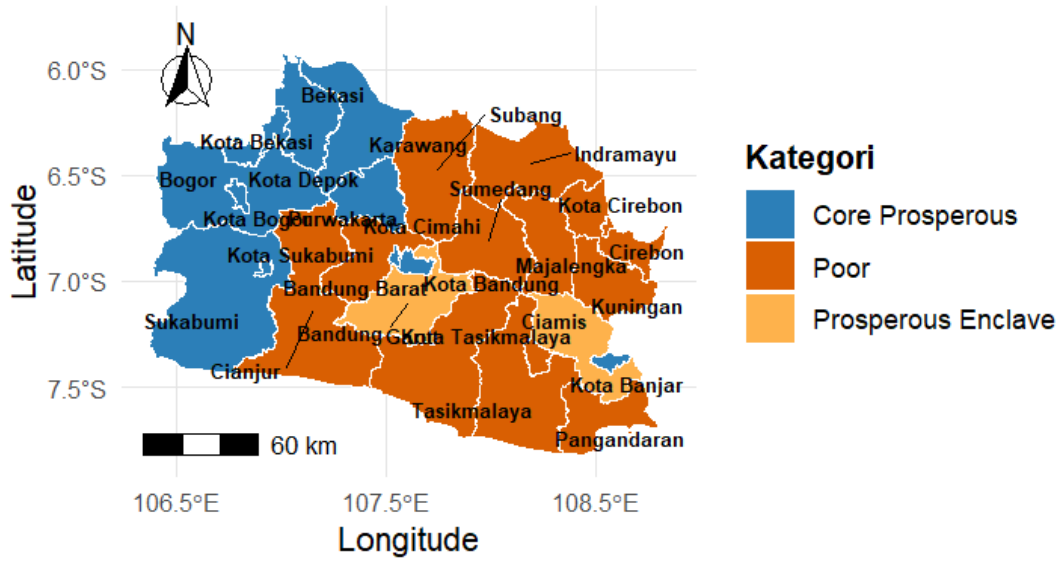
District	Total Neighbors	Poor Neighbors	Enclavity (E)	Classification
Ciamis	6	5	0.833	Prosperous Enclave
Bandung	7	5	0.714	Prosperous Enclave
Purwakarta	5	3	0.600	Core Prosperous
Kota Bandung	6	2	0.333	Core Prosperous
Kota Cimahi	3	1	0.333	Core Prosperous
Sukabumi	3	1	0.333	Core Prosperous
Karawang	4	1	0.250	Core Prosperous
Bogor	8	1	0.125	Core Prosperous
Bekasi	6	0	0.000	Core Prosperous
Kota Banjar	3	0	0.000	Core Prosperous
Kota Bekasi	6	0	0.000	Core Prosperous
Kota Bogor	4	0	0.000	Core Prosperous
Kota Depok	5	0	0.000	Core Prosperous
Kota Sukabumi	3	0	0.000	Core Prosperous

Based on the enclavity index, two districts are classified as prosperous enclaves: Kabupaten Bandung ($E = 0.714$) and Ciamis ($E = 0.833$). The remaining 12 prosperous districts are classified as core prosperous ($E < 0.7$). Notably, Kota Bandung (the city) is not an enclave because its enclavity index is only 0.333; it has a significant number of prosperous neighbors, including Kabupaten Bandung itself (which is an enclave but still prosperous). The 13 poor districts (poverty $> 8.45\%$) are not included in the table because enclavity is only defined for prosperous regions.

Figure 2 presents the spatial classification of all 27 districts into three categories: core prosperous (12 districts, blue), prosperous enclave (2 districts, yellow), and poor (13 districts, orange).

Classification Map of West Java Districts

Poverty threshold = 8.45% | Enclave threshold = 0.7



Source: BPS West Java

Figure 2: Classification Map of West Java Districts (Core Prosperous, Enclave, Poor)

Figure 2 shows the spatial distribution of the three regimes. Core prosperous districts (blue) are concentrated in northern West Java, forming a contiguous cluster. The two yellow enclaves are Kabupaten Bandung (surrounding the city of Bandung) and Ciamis (in the south). Notice that Kota Bandung (the city) is colored blue (core prosperous) because its enclavity is only 0.333; it has many prosperous neighbors, despite being inside an enclave. This distinction is important: an enclave is defined by the proportion of poor neighbors, not by the prosperity of the district itself. Kabupaten Bandung has 5 poor neighbors out of 7 (71%), while Kota Bandung has only 2 poor neighbors out of 6 (33%). Thus, the enclave (yellow) surrounds the core prosperous city (blue) in the Bandung region. The map clearly shows that enclaves are isolated pockets of prosperity surrounded by poor districts (orange).

This classification provides the basis for the three-regime E-PT-SAR model estimated in the next section, where we allow spillover coefficients to differ between core prosperous, enclave, and poor regions.

3.8. E-PT-SAR Model (Three Categories)

Based on the enclave identification in Section 3.7, the Enclave Panel Threshold Spatial Autoregressive (E-PT-SAR) model was estimated with three regimes: core prosperous (12 districts), prosperous enclave (2 districts: Kabupaten Bandung and Ciamis), and poor (13 districts). The model extends the PT-SAR framework by allowing spatial spillover effects and regression coefficients to differ across these three categories. Table 11 summarizes the results.

Table 11: E-PT-SAR Model Estimation (Three Categories)

Parameter	Core Prosperous	Prosperous Enclave	Poor
ρ (spatial lag)	0.2111***	0	0.2558*
TPT	0.0794***	-0.1794	-0.0410
log PDRBK	10.1446***	0.0923	10.5397***
Poverty	-0.2296***	-1.4201	0.1836
N (observations)	60	8	65

Note: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

The results indicate that spatial spillovers are statistically significant in both core prosperous regions ($\rho = 0.2111$, $p < 0.001$) and poor regions ($\rho = 0.2558$, $p < 0.05$). In contrast, no statistically significant spatial spillover effect was detected in prosperous enclave regions. This finding suggests that prosperous enclaves may exhibit different spatial interaction patterns compared to both core prosperous and poor regions, consistent with the enclave hypothesis proposed by Farole.

The variable log PDRBK has a positive and statistically significant effect on HDI in both core prosperous and poor regions, indicating that regional economic output remains an important driver of human development in these areas. However, the coefficient is not statistically significant in enclave regions, suggesting that economic growth alone may not translate into broader gains in human development when prosperous regions are spatially isolated from surrounding areas.

Poverty has a significant negative effect on HDI only in core prosperous regions, while the unemployment rate (TPT) is positively associated with HDI in the same regime. One possible explanation is that economically advanced districts tend to attract job seekers from surrounding areas, increasing measured unemployment despite relatively high levels of human development.

Overall, the e PT-SAR model reveals additional spatial heterogeneity that cannot be fully captured by the standard two-regime PT-SAR framework. In particular, the results suggest that prosperous enclaves may behave differently from core prosperous regions in terms of spatial spillover dynamics. However, these findings should be interpreted cautiously due to the limited number of enclave observations, which restricted the use of fixed effects estimation within the enclave regime.

3.9. Enclave Effect Test

To examine whether distinguishing prosperous regions into core prosperous and prosperous enclave categories provides additional explanatory value, a likelihood ratio (LR) test was conducted by comparing the PT-SAR model (two regimes) and the E-PT-SAR model (three regimes). The null hypothesis assumes that prosperous regions behave homogeneously, implying that the enclave classification is unnecessary.

Table 12: Likelihood Ratio Test for Enclave Effect

Statistic	Value	df	Critical Value ($\alpha = 0.05$)	p-value	Conclusion
LR	10.9	2	5.99	0.0043	Reject H_0

The LR statistic exceeds the critical value at the 5% significance level, suggesting that separating prosperous regions into core and enclave categories may capture additional spatial heterogeneity not reflected in the standard PT-SAR framework. Substantively, this result indicates that prosperous enclave regions may exhibit different spillover behavior compared to core prosperous regions.

These findings are consistent with the theoretical expectation that spatially isolated prosperous regions do not necessarily interact with neighboring regions in the same manner as prosperous regions forming contiguous clusters [3]. In particular, enclave regions may experience weaker economic linkages, limited regional integration, or reduced diffusion of development benefits to surrounding districts.

However, the LR results should be interpreted cautiously. The enclave regime consists of only two districts, which requires a different estimation treatment from the other regimes. Consequently, the LR test should be viewed as supporting rather than definitive evidence of enclave-specific spatial behavior. Nevertheless, the results provide empirical motivation to consider enclave regions as a distinct spatial category within the broader threshold spatial framework.

3.10. Decay Effect

The decay effect analysis examines how far spatial spillovers propagate through the neighborhood structure. Following the spatial multiplier expansion in spatial autoregressive models, the influence transmitted through the m -th order neighbors is given by ρ^m [1]. The decay effect was calculated for regimes with statistically significant spatial spillovers, namely the core prosperous and poor regimes. Since the spatial lag coefficient in the enclave regime was not statistically significant, no significant decay pattern was identified for enclave regions.

Table 13: Decay Effect for Core Prosperous and Poor Regions

Order (m)	Core Prosperous ($\rho = 0.2111$)	Poor ($\rho = 0.2558$)
1	0.2111	0.2558
2	0.0446	0.0654
3	0.0094	0.0167
4	0.0020	0.0043
5	0.0004	0.0011

Using a 1% threshold as negligible limit [21], the spatial reach is the smallest m such that $\rho^m < 0.01$. For a core prosperous $\rho^2 = 0.0446$, $\rho^3 = 0.0094 < 0.01$, the spatial reach is 2 orders. For poor $\rho^2 = 0.2558$, $\rho^3 = 0.0167$, $\rho^4 = 0.0043 < 0.01$, the spatial reach is 3 orders.

These findings suggest that development shocks in poor regions propagate across a wider spatial network than in core prosperous regions. One possible explanation is that poorer districts in the southern part of Java tend to exhibit stronger interdependence due to similarities in economic structure, infrastructure limitations, and regional connectivity patterns. In contrast, spillovers in core prosperous regions appear to be more localized within industrial and metropolitan clusters.

For prosperous enclave regions, no substantial decay effect was observed because the estimated spatial spillover coefficient was not statistically significant. This result is consistent with the interpretation that enclave regions may be relatively disconnected from surrounding spatial development dynamics.

3.11. Model Comparison

Model comparison using information criteria is appropriate only when models are estimated under the same likelihood framework. In this study, both the SAR and PT-SAR models were estimated using maximum likelihood with fixed effects, allowing formal comparison through the Akaike Information Criterion (AIC). In contrast, the E-PT-SAR model required a different estimation treatment for the enclave regime due to the limited number of enclave observations, making direct likelihood-based comparison statistically inappropriate.

Table 14: Model Comparison (SAR vs PT-SAR)

Model	AIC	LogLik	Parameters
SAR	-139.98	73.99	4
PT-SAR	-153.45	84.73	8

The PT-SAR model produces a substantially lower AIC value than the baseline SAR model, indicating that allowing spatial spillover behavior to vary across prosperous and poor regimes improves the fit of the model, even after accounting for additional parameters. Among the formally comparable specifications, the PT-SAR model therefore provides the best overall statistical fit.

A formal AIC comparison involving the E-PT-SAR model was not conducted because the enclave regime contained only two districts, which required a pooled estimation rather than the fixed effects likelihood framework used in the SAR and PT-SAR models. Consequently, the resulting likelihood values are not directly comparable between models.

Nevertheless, the E-PT-SAR model remains substantively important because it captures additional spatial heterogeneity that cannot be identified within the standard two-regime PT-SAR framework. In particular, the model distinguishes between contiguous prosperous regions and spatially isolated prosperous enclaves, revealing differences in spillover behavior across these categories. Therefore the E-PT-SAR model should be interpreted primarily as a theoretical and exploratory extension of the PT-SAR framework, rather than as a formally superior statistical specification.

3.12. Discussion

The findings demonstrate that the dynamics of human development in West Java are spatially heterogeneous. The estimated poverty threshold separates the province into prosperous northern districts and poorer southern districts, indicating that spatial spillover mechanisms differ between development regimes.

The PT-SAR results suggest that prosperous regions benefit from stronger localized spillovers, likely due to industrial concentration, infrastructure connectivity, and agglomeration effects. In contrast, spillovers in poorer regions spread over a wider spatial range, reflecting stronger interdependence among less developed districts.

The E-PT-SAR model further indicates that prosperous enclaves may behave differently from core prosperous regions. No statistically significant spillover effect was detected in enclave districts, suggesting that spatially isolated prosperity does not necessarily generate broader regional development benefits. This finding is consistent with the enclave hypothesis proposed by Farole, where economically advantaged regions can remain weakly integrated with surrounding areas.

Several limitations should be noted. The enclave regime consists of only two districts, limiting statistical robustness and preventing the use of fully comparable estimation procedures across all regimes. In addition, the relatively short study period may affect the stability of the estimated threshold and enclave classification over time.

From a policy perspective, the results imply that development strategies should be adapted to regional characteristics. Core prosperous regions may benefit from integrated regional planning and interregional cooperation, while enclave regions require stronger connectivity with surrounding districts. For poorer regions, infrastructure improvement and poverty reduction remain essential priorities.

4. Conclusions

This study applied the Enclave Panel Threshold Spatial Autoregressive (E-PT-SAR) model to analyze spatial spillover dynamics of human development in West Java. The results show that the spillover effects differ between spatial regimes. Significant spillovers were identified in the core prosperous and poor regions, while no statistically significant spillover effect was detected in the prosperous enclave regions. These findings suggest that spatially isolated prosperous districts may exhibit different interaction patterns from prosperous regions forming contiguous development clusters.

Among the formally comparable models, PT-SAR provided the best statistical fit, while E-PT-SAR provided additional insight by distinguishing enclave regions from other prosperous districts. However, the enclave findings should be interpreted with caution due to the limited number of enclave observations and the short study period. Future research may extend the analysis to broader spatial coverage or longer panel periods and incorporate additional regional characteristics to better understand enclave behavior in regional development processes.

CRediT Authorship Contribution Statement

Nidia Ayu: Conceptualization, Methodology, Formal Analysis, Software, Writing–Original Draft Preparation. **Henny Pramodyo:** Supervision, Validation, Writing–Review & Editing. **Suci Astutik:** Supervision, Validation, Writing–Review & Editing. **Friansyah Gani:** Writing–Review & Editing.

Declaration of Generative AI and AI-assisted Technologies

Generative AI technology (ChatGPT, OpenAI) was used during the preparation of this manuscript for language refinement, grammar checking, code exploration, and L^AT_EX formatting assistance. In addition, Friansyah Gani contributed to the review and editing of the manuscript. All authors reviewed and edited the final manuscript and assume full responsibility for its content.

Declaration of Competing Interest

The authors declare no competing financial or personal interests.

Funding and Acknowledgments

This research received no external funding.

Data and Code Availability

The simulated data and code supporting the findings of this study are available from the corresponding author upon reasonable request.

References

- [1] J. P. Elhorst, *Spatial Econometrics: From Cross-Sectional Data to Spatial Panels*. Springer Berlin Heidelberg, 2014. DOI: [10.1007/978-3-642-40340-8](https://doi.org/10.1007/978-3-642-40340-8).
- [2] S. Suzuki and T. Inohara, “Mathematical definitions of enclave and exclave, and applications,” *Applied Mathematics and Computation*, vol. 268, pp. 728–742, 2015. DOI: [10.1016/j.amc.2015.06.114](https://doi.org/10.1016/j.amc.2015.06.114).
- [3] T. Farole, *Special Economic Zones in Africa: Comparing Performance and Learning from Global Experiences*. World Bank Publications, 2011.
- [4] M. B. Pietrzak, J. Wilk, R. S. Bivand, and T. Kossowski, “The application of local indicators for categorical data (LICD) in the spatial analysis of economic development,” *Comparative Economic Research*, vol. 17, no. 4, pp. 203–220, 2014. DOI: [10.2478/cer-2014-0041](https://doi.org/10.2478/cer-2014-0041).
- [5] J. P. LeSage and R. K. Pace, *Introduction to Spatial Econometrics*. CRC Press, 2009. DOI: [10.1111/j.1467-985x.2010.00681_13.x](https://doi.org/10.1111/j.1467-985x.2010.00681_13.x).
- [6] B. E. Hansen, “Threshold effects in non-dynamic panels: Estimation, testing, and inference,” *Journal of Econometrics*, vol. 93, no. 2, pp. 345–368, 1999. DOI: [10.1016/S0304-4076\(99\)00025-1](https://doi.org/10.1016/S0304-4076(99)00025-1).
- [7] L. Wei, C. Zhang, J. J. Su, and L. Yang, “Panel threshold spatial Durbin models with individual fixed effects,” *Economics Letters*, vol. 201, 2021. DOI: [10.1016/j.econlet.2021.109778](https://doi.org/10.1016/j.econlet.2021.109778).
- [8] Badan Pusat Statistik, “Tingkat pengangguran terbuka menurut kabupaten/kota di provinsi jawa barat,” *Badan Pusat Statistik Provinsi Jawa Barat*, 2024. [Available online](#).
- [9] Badan Pusat Statistik, “Indeks pembangunan manusia (IPM) jawa barat 2024,” *Badan Pusat Statistik Provinsi Jawa Barat*, 2025. [Available online](#).

- [10] Badan Pusat Statistik, “Jumlah dan persentase penduduk miskin menurut kabupaten/kota di provinsi jawa barat, 2024,” *Badan Pusat Statistik Indonesia*, 2025. [Available online](#).
- [11] Badan Pusat Statistik, “Laju pertumbuhan produk domestik regional bruto (PDRB) atas dasar harga konstan berdasarkan kabupaten/kota di jawa barat,” *Badan Pusat Statistik Provinsi Jawa Barat*, 2025. [Available online](#).
- [12] B. H. Baltagi, *Econometric Analysis of Panel Data*, 6th ed. Springer, 2021. DOI: [10.1007/978-3-030-53953-5](#).
- [13] G. C. Chow, “Tests of equality between sets of coefficients in two linear regressions,” *Econometrica*, vol. 28, no. 3, pp. 591–605, 1960. DOI: [10.2307/1910133](#).
- [14] J. A. Hausman, “Specification tests in econometrics,” *Econometrica*, vol. 46, no. 6, pp. 1251–1271, 1978. DOI: [10.2307/1913827](#).
- [15] L. Anselin, *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers, 1988. DOI: [10.1007/978-94-015-7799-1](#).
- [16] F. Gani, H. Pramoedyo, and A. Efendi, “Spatial variation of hdi in east java: A tricube-based geographically weighted regression–flower pollination algorithm modeling approach,” *CAUCHY: Jurnal Matematika Murni dan Aplikasi*, vol. 11, no. 1, pp. 239–254, 2025.
- [17] D. D. Kurniawati, H. Pramoedyo, S. Astutik, and F. Gani, “Haversine-based geographically weighted panel regression of human development in gorontalo (2016–2025),” *CAUCHY: Jurnal Matematika Murni dan Aplikasi*, vol. 11, no. 1, pp. 870–886, 2025.
- [18] L. Anselin and S. J. Rey, *Modern Spatial Econometrics in Practice: A Guide to GeoDa, GeoDaSpace and PySAL*. GeoDa Press LLC, 2014.
- [19] H. H. Kelejian and I. R. Prucha, “A generalized spatial two-stage least squares procedure,” *Journal of Real Estate Finance and Economics*, vol. 17, no. 1, pp. 99–121, 1998. DOI: [10.1023/A:1007707430415](#).
- [20] N. Ayu, H. Pramoedyo, and S. Astutik, “Extending panel threshold spatial autoregressive models with enclave identification: The E-PT-SAR framework,” 2026, Companion paper.
- [21] J. P. Elhorst, I. Tziolas, C. Tan, and P. Millionis, “The distance decay effect and spatial reach of spillovers,” *Journal of Geographical Systems*, vol. 26, no. 2, pp. 265–289, 2024. DOI: [10.1007/s10109-024-00440-5](#).