Jurnal Varian
Vol. 8, No. 1, October 2024, pp 55–66
e-ISSN: 2581-2017

Analysis of Underdeveloped Regency in Indonesia using Logistic Threshold Regression Model

Annisa Nur Salsabila, Siskarossa Ika Oktora
Politeknik Statistika STIS, Jakarta Timur, Indonesia

Article Info

Article history:

Received : 11-16-2023
Revised : 07-22-2024
Accepted : 08-22-2024

Keywords:

KHDI;
Threshold Regression;
Underdeveloped Regency.

ABSTRACT

Regional development inequality causes some regions to lag behind other regions. An underdeveloped regency is a regency where territories and people are less developed than other regions nationally. The government has set a Human Development Index (HDI) target of 62.2 to 62.7 to accelerate the development of underdeveloped regency and prevent the regions from lagging. This study aims to evaluate the HDI target and obtain the HDI value that reduces the risk of underdeveloped regency and acquires variables that affect underdeveloped regency's status. The logistic threshold regression model is used in this study with HDI as the threshold variable, 22 indicators for determining underdeveloped regency as explanatory variables, and the underdeveloped regency's status as the response variable. Threshold regression can handle non-linear relationships between response and explanatory variables, including various types of threshold models such as step, segmented, hinge, stegmented, and upper hinge. By applying a hinge threshold regression model using the R package 'chngpt,' this study addresses non-linear relationships and categorical responses. The results showed a threshold effect with a threshold value of 62.9, indicating that the HDI target can reduce the region's risk of being left behind.

Accredited by Kemenristekdikti, Decree No: 200/M/KPT/2020
DOI: https://doi.org/10.30812/vvarian.v8i1.3570

Corresponding Author:

Name of Corresponding Author,
Department of Engineering and Design, Universitas Bumigora
Email: corresponding-author@universitasbumigora.ac.ad

How to Cite:
Salsabila, A. N., & Oktora, S. I. (2024). Analysis of Underdeveloped Regency in Indonesia using Logistic Threshold Regression Model. *Jurnal Varian*, *8*(1), 55-66.

A. INTRODUCTION

Development is a multidimensional process that includes major changes in social structures, national institutions, economic growth, inequality, and the elimination of poverty (Todaro & Smith, 2020). Development in Indonesia is carried out to improve the welfare of society at the national and regional levels. The central government delegates its authority to regional governments to implement political decentralization policies. Fiscal decentralization involves the transfer of financial responsibilities from the central government to subnational jurisdictions. Administrative decentralization focuses on the transfer of administrative functions and decision-making authority to subnational jurisdictions (Shoesmith et al., 2020). In this way, regions can achieve fiscal independence to increase regional income (PAD), which spurs accelerated and equitable development. However, implementing decentralization is considered not to be able to reduce development inequality between regions.

Decentralization is likely to widen regional disparities due to inefficient resource allocation in countries with poor quality of government. Decentralization will reduce disparities between regions in richer countries because they will have better quality government (Kyriacou et al., 2015). Therefore, the gaps between regions are getting bigger with the implementation of decentralization. This inequality can be seen from the Gini ratio, which, since 2002-2021, has not experienced a significant decline, as shown in Figure 1. In addition, there is a fairly large inequality in the distribution of Indonesia's Gross Domestic Product (GDP) between islands. The distribution of GDP on Java Island is 57.89 percent of Indonesia's total GDP in 2021. Meanwhile, the distribution of GDP in the eastern region of Indonesia is only 12 percent of Indonesia's total GDP in 2021 (BPS, 2023). Regional disparities lead to some regions being less developed or known as underdeveloped regency.

In Presidential Regulation No. 63 of 2020, an underdeveloped regency is defined as a regency where territories and people are less developed than other regions nationally. The determination of an underdeveloped regency's status is based on several criteria, namely community economy, human resources, infrastructure, financial capacity, accessibility, and regional characteristics. Since 2019, it has been determined that 62 regencies have been eradicated from underdeveloped status, so in 2020, the government determined only 62 underdeveloped regencies, most of which are in est. In the National Strategy for Accelerating Development of Underdeveloped Regency for 2020-2024, the government plans several development targets, including reducing the poverty rate from 25.85 percent to 24 percent, increasing the HDI (Human Development Index) to 62.2 to 62.7, reducing the number of underdeveloped regencies, and developing the underdeveloped regency which has been eradicated in 2019. The value of poverty rate and HDI that are used as the basis for determining underdeveloped regencies and have been determined by the government have never been tested statistically. So it is not yet known whether the cut-offs prepared by the government can be validated through quantitative research.

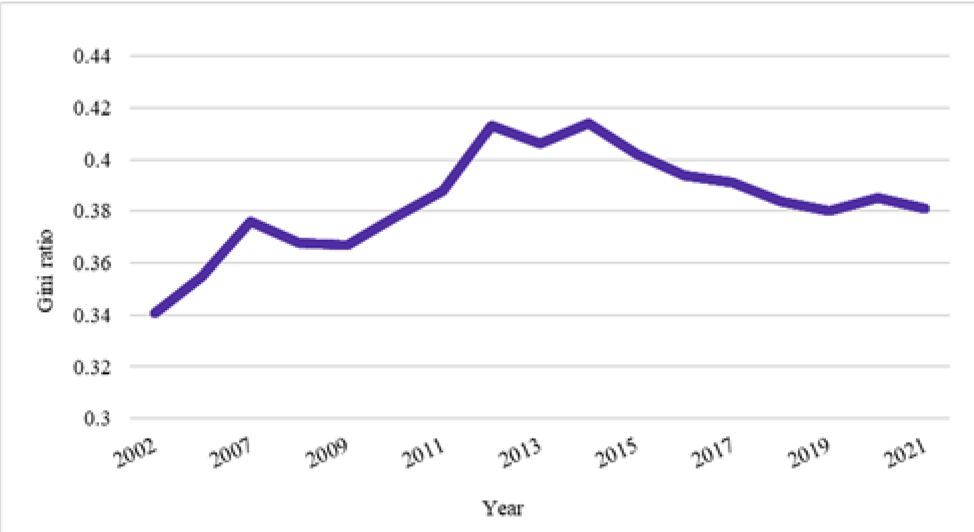

Figure 1. Gini ratio in Indonesia 2002-2021

HDI and poverty indicators have a mutually influential relationship, so these two indicators are important in developing underdeveloped regencies. Sofilda et al. (2014) shows that HDI and poverty have a negative two-way relationship, with HDI having a greater influence on poverty than vice versa. Higher HDI achievements will tend to reduce poverty through good education, health, and economic aspects to improve human quality and reduce poverty (Pudjianto & Syawie, 2015). Based on the results of previous research, HDI has a negative influence on poverty so increasing HDI through increasing life expectancy, expected years of education, average years of education, and per capita expenditure can reduce poverty.

At the end of the 2015-2019 period for determining underdeveloped regency, regions that had been developed had an average HDI of 66.32, while other regions that were still underdeveloped had an average HDI of 58.91. If underdeveloped regency lagging can increase their HDI, these regions have a greater opportunity to escape being left behind. The HDI target of 62.2 to 62.7 in 2020-2024 is still lower than the national average HDI in 2019 was 71.92. Apart from that, in 2010, the average HDI in underdeveloped regency was only 52.45, and it will increase to 59.33 in 2021 (BPS, 2022).

Most of the research on underdeveloped regency aims to identify variables that influence regional underdevelopment. According to Maulidina & Oktora (2020), using the Geographically Weighted Logistic Regression (GWLR) model shows that increasing the percentage of the poverty rate and the percentage of workers in the agricultural sector will increase the tendency for a region to be classified as an underdeveloped regency. However, increasing life expectancy, expected years of schooling, the percentage of households using clean water, and the percentage of villages with internet access will reduce the tendency for a region to be classified as an underdeveloped regency. This is also in line with research by Otok et al. (2018), who found that life expectancy, access to health facilities, per capita expenditure, and average years of schooling significantly affect regional underdevelopment. In addition, research by Fitrah et al. (2021) conducted using a logistic regression model also shows that the poverty rate, life expectancy, and the number of community health centers significantly affect regional underdevelopment.

Recent studies show that decentralization policies in Indonesia often increase regional disparities, particularly in regions with weak governance. Traditional linear regression models have been widely used but may not fully capture the complex relationships between underdeveloped regency indicators and underdeveloped regency's status. As a result, advanced methods such as threshold logistic regression models are being adopted to better understand the influence of threshold variables, like the Human Development Index (HDI), on regional development. Furthermore, the effectiveness of using HDI as a key macro indicator for assessing and developing underdeveloped regencies has yet to be statistically validated. This requires additional research to evaluate the effectiveness of HDI as a macro indicator in determining underdeveloped regency's status while still considering the impact of other regional development indicators. Therefore, this research aims to evaluate the HDI target as an indicator of development targets for underdeveloped regency to obtain HDI values that influence the tendency of a region to be classified as underdeveloped and to find out other variables that influence regional underdevelopment. The HDI value can be obtained using the logistic threshold regression model. The threshold regression model or threshold regression is a regression model that divides the sample into several subsamples based on a certain variable known as the threshold variable. Logistic threshold regression is used if the response variable used is categorical (Fong et al., 2017). With this model, the HDI value, which influences the tendency of a region to be classified as underdeveloped, can be obtained and can be used as evaluation material for setting targets for accelerating regional development.

Although there have been several studies exploring disparities and decentralization, there remain gaps in research regarding the validation of development indicators, the effectiveness of threshold regression models, and cross-temporal analysis of underdeveloped regions. Further research is needed to understand how decentralization can be optimized to more effectively reduce regional disparities in Indonesia.

B. RESEARCH METHOD

This study encompasses all regions in Indonesia, except the Seribu Islands Regency, totaling 415 regencies. The research relies on secondary data from various sources. Descriptive analysis was presented through box plots and thematic maps, while inferential analysis was performed using logistic threshold regression. The research employs a binary response variable based on the underdeveloped status of each region, as defined by the Presidential Regulation of the Republic of Indonesia No. 63 of 2020 for the period 2020-2024. The HDI, as reported in 2021 BPS (Statistics Indonesia) publications, serves as the threshold variable.

Furthermore, the research includes a set of explanatory variables in the threshold logistic regression modeling encompassing all 22 indicators specified for identifying underdeveloped regencies, as outlined in Minister of Villages, Regional Development, and Transmigration Regulation No. 11 of 2020. Several indicators are derived from the 2021 Village Potential (Podes) data, including the percentage of villages with shops (X1), the percentage of villages with health facilities (X2), the percentage of villages with doctors (X3), the percentage of villages with elementary schools (X4), the percentage of villages with junior high schools (X5), the percentage of villages with asphalt/concrete main roads (X10), the percentage of villages with easy access to health facilities (X11), the percentage of villages with easy access to junior high schools (X12), the percentage of villages without disaster incidents (X13), and the percentage of villages did not experience social conflicts (X14). Additionally, several indicators are sourced from the 2021 National Socio-Economic Survey (Susenas), such as the percentage of households with access to electricity (X6), the percentage of households with access to telephones/cell phones (X7), the percentage of internet usage among the population (X8), the percentage of clean water supply in households (X9), the percentage of non-food household expenditure (X16), the percentage of the population employed in non-agricultural sectors (X17), the percentage of women aged 15-49 who have given birth in the last two years with medical attendants (X18), the percentage of fully immunized toddlers (X19), junior high school enrollment rates (X20), and senior high school enrollment rates (X21). Moreover, the research incorporates the Gross Regional Domestic Product (GRDP) indicator per capita based on constant prices, obtained from the 2021 BPS publication, and the regional income (PAD) per capita from the 2020 Ministry of Finance publication.

Inferential analysis was conducted to identify the threshold HDI value and determine the influence of explanatory variables on an underdeveloped regency's status. Threshold regression can handle non-linear relationships between response and explanatory variables, which include various types of threshold models, such as step, segmented, hinge, segmented, and upper hinge (Elder & Fong, 2019; Fong et al., 2017). One threshold regression model suitable for handling categorical response variables is the hinge model. For hypothesis testing, a significance level of 10% ($\alpha = 0.1$) was utilized. The inferential analysis was executed using Rstudio software with the 'chngpt' package. The following steps outline the process of inferential analysis.

1) Multicollinearity
Checking for multicollinearity in the dataset involved examining the correlation values between explanatory variables. High multicollinearity is considered present if the correlation value between explanatory variables is at least 0.8 (Gujarati & Porter, 2012).
2) Threshold effect testing
Threshold effect testing is conducted before estimating model parameters. The null hypothesis used is as follows:
   - $H_0 : \beta 1 = 0$ (indicating there is no threshold effect on the model)
   - $H_0 : \beta 1 \neq 0$ (indicating at least there is a threshold effect on the model)

   A significance level of 10% is applied, leading to the rejection of H0 if the p-value is less than 0.1.
3) Parameter estimation
Parameter estimation is conducted to estimate threshold values, threshold effects, and estimate coefficient values of other explanatory variables. The hinge model can be expressed in logit form as follows:

$$logit\left[P\left(Y_i = 1\right)\right] = logit\left[\pi\right] = ln\left[\frac{\pi}{1-\pi}\right] = \alpha_1 + \alpha_2' x + \beta_1 I(HDI - e)_+ \tag{1}$$

   The function $P\left(Y_i = 1\right)$ or π is the probability when $Y_i = 1$; α_1 is the intercept; α_2' is avector of explanatory variable coefficients, $i = 1, 2, \ldots, p$; where p is the number of explanatory variables; e is threshold parameter; x is vector of explanatory variables; HDI is human development index as threshold variable; β_1 is the threshold effect coefficient, and $I(HDI - e)_+$ will have the value of $(HDI - e)$ if $HDI > e$ and 0 if others.
   Estimation is conducted using the smooth method in the 'chngpt' package which estimates parameters through iterative optimization. According to Fong et al. (2017), for all initial estimates of all parameters, the following stages are follows:
   a. Update the threshold parameter e and each coefficient $\hat{\beta}$ associated with the threshold variable, conditional on $\hat{\alpha}$.
   b. Updates all coefficients $\hat{\alpha}$ and $\hat{\beta}$. Conditional on the estimated threshold $\hat{e}$. The algorithm is stopped when the relative change falls below a predetermined tolerance level. Additionally, confidence interval estimation was conducted using a robust model, which allows the model to converge when model misspecification occurs.
4) Model performance evaluation
Several important measurements can be obtained through the classification table, including correct classification rate, sensitivity, and specificity. Furthermore, the area under the Receiver Operating Characteristic (ROC) curve or Area Under Curve (AUC) is a more complete measure of model classification accuracy than sensitivity and specificity. The ROC curve is constructed by plotting sensitivity against 1-specificity for all possible cutoff points. AUC is a measure of the model's ability to determine the classification of observation units with values in the range 0 to 1. In general, the model will be better at performing model classification if the AUC is closer to 1 (Hosmer et al., 2013). Futhermore, steps in analyse Underdeveloped Regency in Indonesia using Logistic Threshold Regression Model are shown in Figure 2.

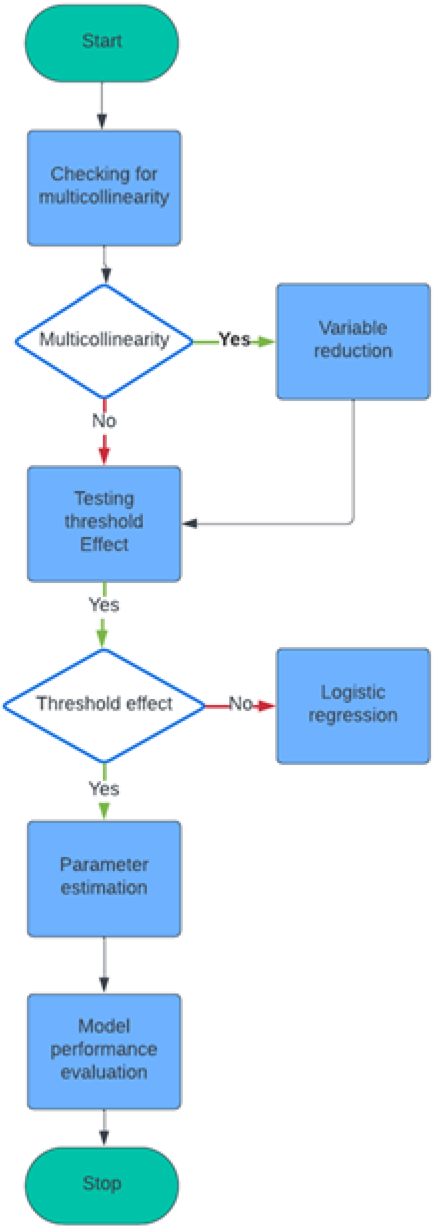

Figure 2. Analysis of Underdeveloped Regency in Indonesia using Logistic Threshold Regression Model

C. RESULT AND DISCUSSION

1. Descriptive Analysis

a. Underdeveloped regency

Disparities in regional development in Indonesia have led to certain regions being more underdeveloped than others. In 2020, the government identified 62 regencies across Indonesia as underdeveloped regencies. Referring to Figure 3, the majority of these underdeveloped regencies, indicated in purple, are concentrated in the eastern region of Indonesia, while the remainder are in the western region. Specifically, there are 55 underdeveloped regencies in eastern Indonesia, spanning 14 regencies on Nusa Tenggara Island, three regencies on Sulawesi Island, eight regencies on Maluku Island, and 30 regencies on Papua Island. On the other hand, the underdeveloped regency in the western region only consists of 7 regencies, all of which are located on Sumatra Island.

Figure 3. Underdeveloped regency in Indonesia in 2021

b. Human Development Index (HDI

The HDI is a significant macro indicator in measuring the achievement of targets for underdeveloped regency. An increase in HDI implies an improvement in the quality of human resources within a particular region. The regions with high-quality human resources are more capable of effectively advancing development. Thereby, the risk of being underdeveloped will decrease. In 2021, the average HDI for all regencies and cities in Indonesia stood at 72. The highest regency HDI achievement is in Sleman Regency at 84, indicating excellent human resource quality in that region. Conversely, Nduga Regency, classified as an underdeveloped regency, had the lowest HDI achievement of only 33. The difference in the highest and lowest HDI shows significant development inequality between regions.

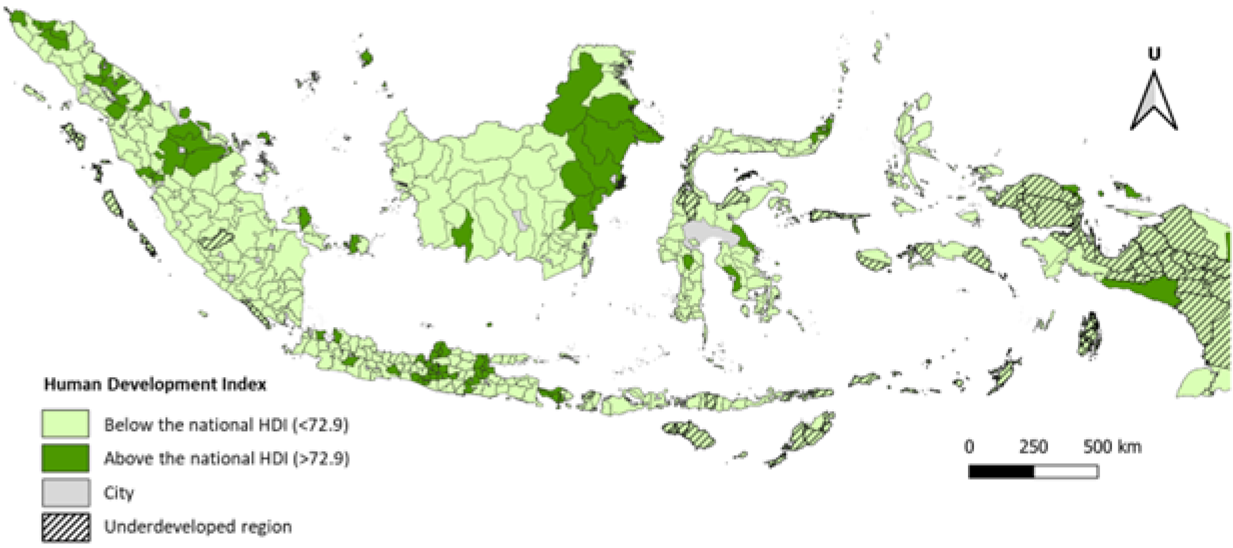

Figure 4. Underdeveloped regency in Indonesia in 2021

As illustrated in Figure 4, most underdeveloped regencies have HDI values below the national average HDI. Furthermore, most non-underdeveloped regencies also have an HDI below the national HDI, with only a limited number of regions surpassing the national HDI. Non-underdeveloped regencies with HDI above the national HDI are mostly spread across KBI and several areas in the eastern region of Indonesia, such as Sulawesi Island and Papua Island. This shows that human development in the eastern region has begun to increase despite inequality between regions because most regions in Papua are still lagging.

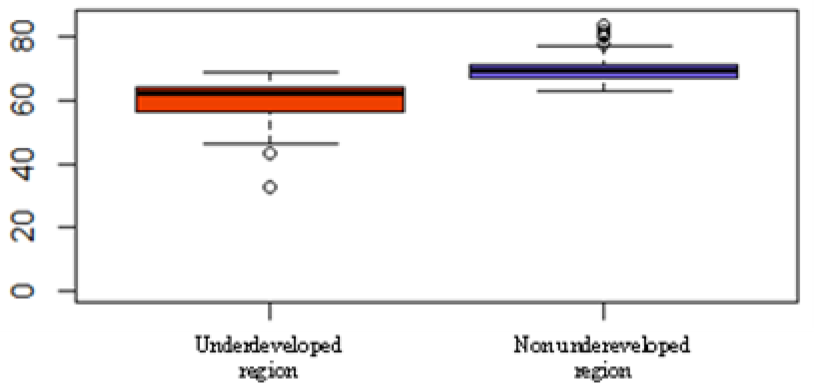

Figure 5. Box plot of HDI in Indonesia in 2021

Based on Figure 5, the median HDI achievement in an underdeveloped regency is 62.075, which remains below the HDI in a non-underdeveloped regency. This suggests that underdeveloped regencies in Indonesia have relatively lower HDI characteristics than non-underdeveloped regencies. Moreover, two regencies are underdeveloped regencies with shallow HDI values, namely Nduga Regency and Puncak Regency.

c. Indicators for Determining Underdeveloped regency

The characteristics of all indicators for determining underdeveloped areas in Indonesia can be seen in Table 1. Based on Table 1, the characteristics of underdeveloped regencies for all indicators tend to be lower than for non-underdeveloped regencies except for the variable percentage of villages that did not experience a disaster.

Table 1. Mean and median of indicators for determining underdeveloped regency

No.	Indicators	Mean		Median	
		Underdeveloped	Non-underdeveloped	Underdeveloped	Non-underdeveloped
1	The percentage of villages with shops	12.96	28.44	9.42	26.00
2	The percentage of villages with health facilities	39.72	55.50	36.96	54.29
3	The percentage of villages with doctors	10.09	23.42	9.08	20.00
4	The percentage of villages with elementary schools	69.06	91.09	81.77	97.26
5	The percentage of villages with junior high schools	31.99	48.68	31.08	48.32
6	The percentage of households with access to electricity	83.17	98.88	89.21	99.67
7	The percentage of households with access to telephones/cell phones	72.28	92.20	83.83	92.86
8	The percentage of internet usage among the population	26.03	50.46	30.57	50.53
9	The percentage of clean water supply in households	41.11	69.87	42.19	72.18
10	The percentage of villages with asphalt/concrete main roads	43.22	82.19	48.99	90.84
11	The percentage of villages that can easily reach health facilities	80.36	96.94	83.10	99.03
12	The percentage of villages that can easily reach junior high school	72.22	96.02	78.51	98.86
13	The percentage of villages without disaster incidents	66.08	60.72	72.18	63.81
14	And the percentage of villages did not experience social conflicts	96.37	98.51	97.97	99.19
15	Gross Regional Domestic Product (GRDP) per capita	18.26	29.22	14.52	24.03

No.	Indicators	Mean		Median	
		Underdeveloped	Non-underdeveloped	Underdeveloped	Non-underdeveloped
16	The percentage of non-food household expenditure	41.42	46.58	43.27	46.41
17	The percentage of the population employed in non-agricultural sectors	32.39	54.16	35.56	52.89
18	The percentage of women aged 15-49 who have given birth in the last two years with medical attendants	70.52	95.31	74.01	97.39
19	The percentage of fully immunized toddlers	43.96	56.99	46.71	60.74
20	Junior high school enrollment rates	88.06	95.71	93.85	96.38
21	Senior high school enrollment rates	68.66	73.97	71.86	73.86
22	Regional income per capita	328524.28	408402.63	280875.28	347952.92

2. Inferential Analysis

a. Threshold Effect Testing

Threshold effect testing is conducted to determine whether a threshold exists in the model. Before performing this test, multicollinearity is assessed by examining the correlation between the explanatory variables in the dataset. Based on the results obtained in Appendix 1, there is a high correlation between the variable the percentage of internet usage among the population (X8) and HDI as well as the variable percentage of villages that can easily reach health facilities (X11) and the percentage of villages that can easily reach junior high school (X12). Thus, variables X8 and X12 were removed from the model to avoid multicollinearity in the model.

Subsequently, a threshold effect was tested using the hinge model hypothesis test. The hinge model is a form of threshold regression that assumes that the threshold variable has no effect on the model before a certain threshold value and has an effect after that threshold value (Fong et al., 2017). The results of testing the existence of a threshold effect with the hinge model produce a p-value of less than 0.1, so there is sufficient evidence to reject the null hypothesis. This shows a threshold effect in the hinge model with the HDI value as the threshold variable.

3. Parameter Estimation

When performing parameter estimation with the hinge model, it is assumed that the threshold variable, in this case, the HDI, has no impact on the model before a specific value is reached. Furthermore, the explanatory variables are considered to significantly affect the model if the associated p-value is less than 0.1.

The summary of the parameter estimation results for the hinge model, as displayed in Table 2, indicates that the threshold effect (HDI-e)+ significantly affects the model with a p-value of less than 0.1. Additionally, five variables have a significant influence on the model, so the hinge model formed is as follows:

$$\begin{aligned} ln\left(\frac{\hat{\pi}}{1-\hat{\pi}}\right) = {} & 37,925 + 0,080X1^{*} + 0,019X2 - 0,120X3^{*} - 0,010X4 - 0,003X5 - 0,055X6 - 0,053X7 - 0,018X9 - \\ & 0,009X10 - 0,049X11 - 0,015X13 - 0,159X14^{*} - 0,035X15 - 0,085X16 - 0,007X17 - 0,009X18 - \\ & 0,011X19 - 0,071X20 + 0,104X21^{*} + 0,042X22^{*} - 1,036\ I(HDI - e)_{+}{}^{*} \end{aligned}$$

According to this model, it is evident that HDI influences the underdeveloped regency's status after surpassing the threshold value of 62.9. The coefficient value of -1.036 indicates that when HDI exceeds the threshold value, each increase in HDI reduces the likelihood of an area becoming underdeveloped by a factor of 2.82. The estimation results for each explanatory variable show that only five explanatory variables have a significant effect on the model, namely:

1) The percentage of villages with shops (X1);
2) The percentage of villages with doctors (X3);
3) The percentage of villages that do not experience social conflicts (X14)
4) Senior high school enrollment rates (X21); and
5) Regional income per capita (X22).

Table 2. Parameter estimation

Variable	Coefficient	Standard Error	Odds Ratio	p-value
(1)	(2)	(3)	(4)	(5)
Intercept	37,925	14,060		0,007*
X1	0,080	0,034	1,019	0,018*
X2	0,019	0,029	0,887	0,511
X3	-0,120	0,042	0,990	0,004*
X4	-0,010	0,029	0,997	0,739
X5	-0,003	0,026	0,946	0,910
X6	-0,055	0,076	0,948	0,473
X7	-0,053	0,073	0,982	0,467
X9	-0,018	0,022	0,991	0,396
X10	-0,009	0,020	0,952	0,627
X11	-0,049	0,054	0,985	0,358
X13	-0,015	0,013	0,853	0,268
X14	-0,159	0,081	0,966	0,051*
X15	-0,035	0,037	0,919	0,351
X16	-0,085	0,085	0,993	0,314
X17	-0,007	0,023	0,991	0,745
X18	-0,009	0,038	0,989	0,822
X19	-0,011	0,020	0,931	0,583
X20	-0,071	0,118	1,110	0,545
X21	0,104	0,043	1,043	0,016*
X22	0,042	0,022	0,355	0,056*
(HDI-e)+	-1,036	0,310	1,083	0,001*

a. Performance Evaluation of Logistic Threshold Regression Model

The performance evaluation of the hinge model is conducted using the classification table and the measurement of the AUC value derived from the ROC curve. The results of the model classification for all regencies are presented in Table 3.

Table 3. Classification table

Classification	Observasi	
	Underdeveloped	Non-underdeveloped
(1)	(2)	(3)
Underdeveloped	51	5
Non-underdeveloped	11	348
Classification Accuracy Percentage	82,26%	98,58%
Correct Classification Rate (CCR)	96%	

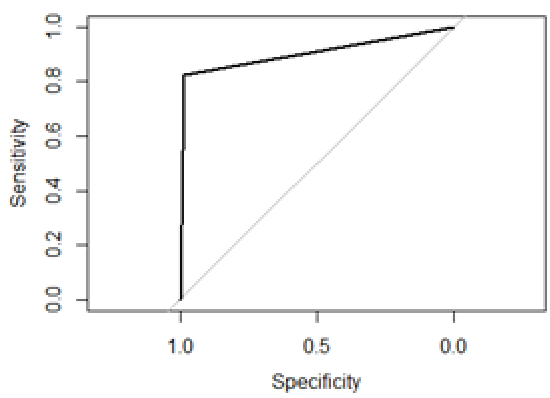

Figure 6. Effects of selecting different switching under dynamic conditions

According to Table 3, it can be seen that 82.26% of regency were accurately classified as underdeveloped regency, while 98.58% of regencies were correctly classified as non-underdeveloped regencies. Thus, the overall accuracy rate of the hinge model in classifying an underdeveloped regency's status is 96%. However, 11 regencies are underdeveloped but have been classified as non-underdeveloped, and five regencies are non-underdeveloped but have been classified as underdeveloped, so there are 16 regencies that are misclassified. Furthermore, as shown in Figure 6, the ROC curve produces an AUC value of 0.90, indicating that the model is highly effective in performing classifications.

The parameter estimation results in Table 2 show a significant threshold effect, and not all explanatory variables significantly affect the hinge model. However, these variables are considered crucial for understanding and determining the underdeveloped regency's status so that the variables are retained in the model. The analysis of each variable within the model is conducted using the Odds Ratio (OR), derived from the exponential coefficient associated with each variable in the hinge model.

The threshold effect with HDI as the threshold variable significantly affects the model with a p-value = 0.001 and a coefficient of -1.036 (1/OR=2.82). This means that a one-unit increase in HDI beyond the threshold value of 62.9 reduces the tendency of a region to be classified as underdeveloped by a factor of 2.82 compared to a non-underdeveloped regency. Increasing human development in a region will reduce poverty and the risk of regional underdevelopment (Hasan, 2022; Maulidina & Oktora, 2020). This result is in line with the research results of Purwandari & Hidayat (2017), which stated that increasing HDI can reduce the tendency for a region to be classified as an underdeveloped regency. Furthermore, the results indicating an estimated HDI threshold value of 62.9 confirm that the government's HDI target, ranging from 62.2 to 62.7, has been statistically proven to reduce the tendency of regions to become underdeveloped regencies.

The variable percentage of villages with shops (X1) significantly affects the model, as indicated by a p-value of 0.018 and a coefficient of 0.080 (OR=1.083). This implies that for every 1 percent increase in the number of villages with shops, the likelihood of a region being classified as underdeveloped is 1.083 times higher than becoming a non-underdeveloped regency, assuming other variables remain constant. Interestingly, this result contradicts the findings of Hochard & Barbier (2017) who suggest that the presence of markets or access to markets significantly influences economic growth. However, inequality of market accessibility can constrain growth potential. Generally, regions with robust economic growth or those serving as economic hubs tend to have a higher concentration of markets than other regions. Table 1 also suggests that the percentage of villages with shops is relatively lower in underdeveloped regency than in non-underdeveloped regency. However, the parameter estimation results reveal a different trend. This inconsistency could be attributed to outliers or interactions between variables, which can cause a change in the regression coefficient's sign when other variables are included in the model (Kennedy, 2005).

Furthermore, the percentage of villages with doctors (X3) has a significant effect on the model with a p-value of 0.004 and a coefficient of -0.120 (1/OR=1.127), which means that with an increase of 1 percent of villages that have doctors, the tendency for a region to become an underdeveloped regency is greater 1.127 times compared to non-underdeveloped regency. The presence of health professionals and equipment in healthcare facilities positively influences the decision to seek care. Increasing the availibility health professionals and equipment would benefit the individuals living closer to health facilities (Anselmi et al., 2015). Good health services will improve the quality of population health. However, Hermawan (2019) shows that in 2013, health professionals were still only concentrated in several regencies/cities. Apart from that, the target ratio of doctors per 100,000 population has not been achieved to date. This inequality can lead to an increasing trend in a region with a shortage of health professionals.

The variable percentage of villages that did not experience social conflict (X14) also significantly affects the model with a p-value of 0.051 and a coefficient of -0.159 (1/OR=1.172). This implies that for every 1 percent increase in villages without social conflict, the likelihood of an area being classified as an underdeveloped regency is 1.172 times lower than for a non-underdeveloped regency, assuming other variables remain constant. Social conflict that occurs in an area can hamper physical, social, and economic development so that an area will have a higher risk of being left behind (Kemendes PDTT, 2019). The government even noted that in 2016, 20.49% of 122 underdeveloped regencies had low conflict resilience. This means that some underdeveloped areas have the potential to experience conflict, and it will be difficult for them to escape from underdevelopment.

Furthermore, the high school enrollment rate (X21) influences the model with a p-value = 0.016 and a coefficient of 0.104 (OR=1.11). This means that an increase of 1 high school enrollment rate will increase the tendency of an area to become an underdeveloped regency by 1.11 times compared to being a non-underdeveloped regency. Better education should reduce the propensity of a region to become an underdeveloped regency because it can improve the quality of human resources and reduce poverty (Pratama, 2014).

However, the results of this study show that the higher the high school enrollment rate, the tendency to become an underdeveloped regency will also increase. Additionally, Maulidina & Oktora (2020) suggest that the education variable only has a significant and positive effect in the areas around Yogyakarta, indicating that variable X21 may not have a significant effect globally.

Additionally, the variable regional income per capita (X22) significantly affects the model with p-value = 0.056 and a coefficient of 0.042 (OR=1.043). This means that an increase of 1 unit of regional income per capita will increase the tendency of a region to become an underdeveloped regency by 1.043 times compared to being a non-underdeveloped regency. Regional financial capacity shows the fiscal independence of a region, which tends to be a characteristic of a non-underdeveloped regency. However, the positive sign in the parameter estimation for regional income per capita indicates the opposite. The OR value of 1.043 shows that the tendency for a regency with higher regional income per capita to become an underdeveloped regency is not significantly different from that of a non-underdeveloped regency. This aligns with the data presented in Table 1, which illustrates that the distribution of regional income per capita between underdeveloped and non-underdeveloped regencies is relatively similar.

D. CONCLUSION AND SUGGESTION

Based on the results, it can be concluded that in 2021, disadvantaged regions in Indonesia are concentrated in the eastern region with HDI characteristics that tend to be lower than the national HDI and the average HDI for regions that are not left behind. Additionally, underdeveloped areas have relatively lower values across all explanatory variables compared to non-disadvantaged areas, except for the variable indicating the percentage of villages that did not experience a disaster. Furthermore, the Human Development Index (HDI) as a threshold variable significantly affects the model with a threshold value of 62.9. This means HDI will impact the region's underdeveloped status beyond 62.9. Consequently, the government's HDI target for underdeveloped areas, ranging from 62.2 to 62.7, is considered capable of reducing the risk of being left behind and that regions can come out of being left behind. Additionally, the variables influencing regional underdevelopment include the percentage of villages that have shops, the percentage of villages that have doctors, the percentage of villages that did not experience social conflict, the high school enrollment rate, and local income per capita.

The influence of HDI is crucial in accelerating the development of underdeveloped areas. Therefore, the government is expected to focus on development strategies prioritizing human resources in health, education, and the economy. Thus, the HDI target can be achieved so underdeveloped areas can develop. The government can achieve this by augmenting the quantity and quality of health facilities and healthcare professionals across all regions. Additionally, facilitating proper access for students to schools can increase school participation.

On the other hand, several variables have regression coefficient signs that are opposite to the results of descriptive analysis and previous research. This may occur because of outliers, variables' interactions, or spatial effects. These problems still cannot be resolved with the logistic threshold regression model, so further exploration of this model is needed. The threshold logistic model using the R package 'chngpt' currently does not support models that include outliers or interactions between variables

REFERENCES

Anselmi, L., Lagarde, M., & Hanson, K. (2015). Health service availability and health seeking behaviour in resource poor settings: Evidence from Mozambique. *Health Economics Review*, *5*(1), 26. https://doi.org/10.1186/s13561-015-0062-6

BPS. (2022). *Indeks Pembanguan Manusia 2021*.

BPS. (2023). *Produk Domestik Regional Bruto Provinsi-Provinsi di Indonesia menurut Pengeluaran 2018-2022*.

Elder, A., & Fong, Y. (2019). Estimation and inference for upper hinge regression models. *Environmental and Ecological Statistics*, *26*(4), 287–302. https://doi.org/10.1007/s10651-019-00428-1

Fitrah, R., Asdi, Y., & Maiyastri, M. (2021). Pemodelan Faktor-Faktor yang Mempengaruhi Status Daerah di Indonesia Menggunakan Metode Regresi Logistik Biner. *Jurnal Matematika UNAND*, *10*(1), 108. https://doi.org/10.25077/jmu.10.1.108-115.2021

Fong, Y., Huang, Y., Gilbert, P. B., & Permar, S. R. (2017). Chngpt: Threshold regression model estimation and inference. *BMC Bioinformatics*, *18*(1), 454. https://doi.org/10.1186/s12859-017-1863-x

Gujarati, D. N., & Porter, D. C. (2012). *Basic Econometrics*. McGraw-Hill/Irwin.

Hasan, Z. (2022). The Effect of Human Development Index and Net Participation Rate on the Percentage of Poor Population: A Case Study in Riau Province, Indonesia. *International Journal of Islamic Economics and Finance Studies*. https://doi.org/10.54427/ijisef.964861

Hermawan, A. (2019). Analisis Distribusi Tenaga Kesehatan (Dokter Perawat Dan Bidan) Di Indonesia Pada 2013 Dengan Menggunakan Gini Index. *Buletin Penelitian Sistem Kesehatan*, *22*(3). https://doi.org/10.22435/hsr.v22i3.1304

Hochard, J., & Barbier, E. (2017). Market Accessibility and Economic Growth: Insights from a New Dimension of Inequality. *World Development*, *97*, 279–297. https://doi.org/10.1016/j.worlddev.2017.04.018

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (Third edition). Wiley.

Kemendes PDTT. (2019). *Dukungan Infrastruktur untuk Daerah Tertinggal*.

Kennedy, P. E. (2005). Oh No! I Got the Wrong Sign! What Should I Do? *The Journal of Economic Education*, *36*(1), 77–92. https://doi.org/10.3200/JECE.36.1.77-92

Kyriacou, A. P., Muinelo-Gallo, L., & Roca-Sagalés, O. (2015). Fiscal decentralization and regional disparities: The importance of good governance. *Papers in Regional Science*, *94*(1), 89–108. https://doi.org/10.1111/pirs.12061

Maulidina, T. P., & Oktora, S. I. (2020). Analisis Spasial Ketertinggalan Daerah di Indonesia Tahun 2018 Menggunakan Geographically Weighted Logistic Regression. *Indonesian Journal of Statistics and Its Applications*, *4*(3), 528–544. https://doi.org/10.29244/ijsa.v4i3.690

Otok, B. W., Hidayat, R., Mahsyari, Z., Sa'diyah, S. H., & Fadhila, D. A. (2018). Classification of Underdeveloped Regions in Indonesia Using Decision Tree Method: *Proceedings of the 2nd International Conference Postgraduate School*, 879–883. https://doi.org/10.5220/0007553308790883

Pratama, Y. C. (2014). Analisis Faktor-Faktor yang Mempengaruhi Kemiskinan di Indonesia. *Esensi: Jurnal Bisnis dan Manajemen*, *4*(2). https://doi.org/10.15408/ess.v4i2.1966

Pudjianto, B., & Syawie, M. (2015). Kemiskinan dan Pembangunan Manusia. *Sosio Informa*, *1*(3). https://doi.org/10.33007/inf.v1i3.167

Purwandari, T., & Hidayat, Y. (2017). Pemodelan Ketertinggalan Daerah di Indonesia Menggunakan Analisis Diskriminan, 194–200. https://scholar.google.com/scholar?cluster=16071151654119604742&hl=en&oi=scholarr

Shoesmith, D., Franklin, N., & Hidayat, R. (2020). Decentralised Governance in Indonesia's Disadvantaged Regions: A Critique of the Underperforming Model of Local Governance in Eastern Indonesia. *Journal of Current Southeast Asian Affairs*, *39*(3), 359–380. https://doi.org/10.1177/1868103420963140

Sofilda, E., Hamzah, M., & Sholeh, A. (2014, January). Human Development and Poverty in Papua Province (An Analysis of Simultaneous Approach on Panel Data Regression). Retrieved October 29, 2024, from https://papers.ssrn.com/abstract=2382080

Todaro, M. P., & Smith, S. C. (2020). *Economic Development*. Pearson UK.