VARIANSI: Journal of Statistics and Its Application on Teaching and Research **ISSN 2684-7590 (Online)**
Vol. 08 No. 02 (2026), 157-166

DOI: *10.35580/variansiunm561*

The Effects of Poverty, Average Years of Schooling, and Expected Years of Schooling on the Net Enrolment Rate for Senior High School and Equivalent Levels in the Sulawesi and Maluku Regions, 2018–2023

Rahul Fikri Alamsyah[1]*

Statistics, UIN Sjech M Djamil Djambek Bukittinggi, Indonesia

Keywords: net Enrolment rate, poverty, years of schooling, panel regression.

Abstract:

The Net Enrollment Rate (NER) for senior high school and equivalent levels in Indonesia still shows disparities, particularly in Maluku and Sulawesi, where rates fall below the national average, posing a challenge to national development. Poverty rates, average years of schooling, and expected years of schooling are estimated to influence the Net Enrollment Rate (NER) for senior high school and equivalent levels. The method used is panel data regression analysis because the indicators under study exhibit numerical instability and changes take time to occur. The estimation approaches used in the panel data analysis are the Common Effects Model (CEM), the Fixed Effects Model (FEM), and the Random Effects Model (REM). The best model was selected using the Chow test, the Hausman test, and the Lagrange Multiplier test. The test results indicate that the Random Effects Model is the best fit. Based on the estimation results, the percentage of the poor population has a negative and significant effect on NER, while the percentage of the poor population has a positive and significant effect on NER. The variable expected years of schooling does not have a significant partial effect at the 5% significance level. Simultaneously, the independent variables have a significant effect on NER with an adjusted R-value of 0.4105, which means that the panel regression model indicates that the independent variables account for 41.05% of the variation in the dependent variable, while the remaining 58.95% is influenced by factors other than these three variables.

1. Introduction

Data from the Central Statistics Agency show that the Net Enrolment Rate (NER) for senior high school and equivalent levels in Indonesia still exhibits regional disparities, particularly in eastern Indonesia. Several provinces in the Sulawesi and Maluku regions have NERs for senior high school and equivalent levels that fall below the national average. Provinces such as West Sulawesi, Maluku, and North Maluku still face various educational challenges characterized by high poverty rates, limited access to education, and low-quality human resources. These conditions indicate that ensuring equitable access to upper secondary education in eastern Indonesia remains a challenge for national development (BPS, 2023/2018; Indrasetianingsih & Wasik, 2020).

The low rate of upper secondary education Enrolment is believed to be influenced by various socioeconomic factors. The poverty rate is one of the factors affecting households' ability to finance their children's education through high school or its equivalent. In addition, educational indicators such as Average Years of Schooling and Expected Years of Schooling are also believed to be correlated with educational participation rates. Regions with higher levels of public education tend to have greater awareness of and better access to education compared to regions with lower-quality education. Thus, poverty, Average Years of Schooling, and Expected Years of Schooling are estimated to be

* Corresponding author.
E-mail address: putrarahmat366@gmail.com

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

important factors influencing the participation rate in senior high school or equivalent (BPS, 2018; Indrasetianingsih & Wasik, 2020; Wardhana & Kharisma, 2020).

Several previous studies have shown that socioeconomic variables have a significant impact on indicators of human development and community well-being, as determined using a panel data regression approach. In addition, poverty levels are also known to have a negative correlation with educational participation in Indonesia. These findings indicate that economic factors and the quality of education in a community are important aspects in determining a region's level of educational participation (Azizah, 2021; Indrasetianingsih & Wasik, 2020; Rochmatin, 2019).

Nevertheless, research on the determinants of the net Enrolment rate for senior high school and equivalent levels in the Sulawesi and Maluku regions using a panel data regression approach remains relatively limited. Previous studies have generally focused on only one specific region or have not simultaneously incorporated the variables of poverty, average years of schooling, and expected years of schooling into a panel data model. The variables under study are macroeconomic indicators whose changes take time to materialize, using panel data allows researchers to examine how fluctuations in poverty rates or improvements in the quality of education in a province directly influence changes in the net enrollment rate (Fadilla, 2021; Rahminawati, 2023)

The limitations of this study highlight the need for an analysis that identifies the factors associated with the Net Enrollment Rate for senior high school and equivalent levels in the Sulawesi and Maluku regions, taking into account changes in conditions in each regency and city during the observation period. Therefore, this study uses panel data regression to analyze the effects of poverty, average years of schooling, and expected years of schooling on the Net Enrollment Rate for senior high school and equivalent levels in regencies and cities in Maluku Province for the 2018–2023 period.

A panel data approach was used to identify relationships among variables, taking into account the differing characteristics of each regency and city as well as changes in social and educational conditions from year to year. The best model was selected through a series of statistical tests to identify the model best suited to explaining the factors associated with upper secondary education participation rates in Maluku Province.

2. Research Methods

2.1 Research Types and Data Sources

This study employs a quantitative research approach using panel data regression analysis. The study uses secondary data from official publications of the Central Statistics Agency (BPS) of the Republic of Indonesia via the official website www.bps.go.id. The observation objects consist of cross-sectional data from eight provinces in the Sulawesi and Maluku regions, namely North Sulawesi, Central Sulawesi, South Sulawesi, Southeast Sulawesi, Gorontalo, West Sulawesi, Maluku, and North Maluku. The time-series data used covers the period from 2018 to 2023, resulting in a total of 48 observations (8 provinces × 6 years) for this study, based on the data available in the system for the 2018–2023 period.

The variables used in this study consist of one dependent variable and three independent variables. The dependent variable is the Net Enrolment Rate (NER) for senior high school and equivalent levels, expressed as a percentage, while the independent variables are the percentage of the poor population (POOR), the average years of schooling (AYS), and the expected years of schooling (EYS). All research variables are presented in Table 1.

Tabel 1. Research Variables

Variable	Description	Unit
NER (Y)	Net Enrolment Rate	Percent (%)
POOR (X1)	Percentage of the Poor	Percent (%)
AYS (X2)	Average Years of Schooling	Years
EYS (X3)	Expected Years of Schooling	Years

2.2 Data Analysis

In this study, the author uses the panel data regression analysis method. Panel data is a combination of cross-sectional data and time-series data. Cross-sectional data consists of many individual units observed at a specific point in time, while time-series data consists of data from a single observation unit tracked over several consecutive time periods. The combination of the two results in panel data that has both an individual dimension (i) and a time dimension (t) simultaneously (Amaliah et al., 2020).

The general form of the panel data regression equation is as follows:

$$Y_{\mathrm{i}}^{\mathrm{t}} = \beta_0 + \Sigma^{\mathrm{k}}{}_{\mathrm{ki}}{}^{\mathrm{t}} X_{\mathrm{ki}}{}^{\mathrm{t}} + \varepsilon_{\mathrm{i}}{}^{\mathrm{t}} (k = 1,2,\ldots,K)$$

where: $Y_{\mathrm{i}}{}^{\mathrm{t}}$ is the value of the i-th individual's dependent variable for period t, $i = 1,2,\ldots,N$ and $t = 1,2,\ldots,T$; $X_{\mathrm{ki}}{}^{\mathrm{t}}$ is the value of the k-th independent variable for the i-th individual in year t; β is the parameter to be estimated; $\varepsilon_{\mathrm{i}}{}^{\mathrm{t}}$ is the error term for the i-th individual in period t; and K is the number of regression parameters to be estimated.

The equations for this research model can be written specifically as follows:

$$NER_{\mathrm{i}}{}^{\mathrm{t}} = \beta_0 + \beta_1 POOR_{\mathrm{i}}{}^{\mathrm{t}} + \beta_2 AYS_{\mathrm{i}}{}^{\mathrm{t}} + \beta_3 EYS_{\mathrm{i}}{}^{\mathrm{t}} + e_{\mathrm{i}}{}^{\mathrm{t}}$$

Panel data regression models can be estimated using three approaches: the Common Effect Model (CEM), the Fixed Effect Model (FEM), and the Random Effect Model (REM). The Common Effect Model (CEM) and Fixed Effect Model (FEM) estimation techniques use Ordinary Least Squares (OLS), while the Random Effect Model (REM) uses the Generalized Least Squares (GLS) estimation technique (Mona et al., 2022).

a) *Common Effect Model (CEM)*

The Common Effects model combines cross-sectional and time-series data into a single model without accounting for differences in time or individuals. This model is the simplest of the three panel data regression models. The estimation technique used is Ordinary Least Squares (OLS). The general equation for the Common Effects model is as follows:

$$Y_{\mathrm{i}}{}^{\mathrm{t}} = \beta_0 + \beta_1 X_{1\mathrm{i}}{}^{\mathrm{t}} + \beta_2 X_{2\mathrm{i}}{}^{\mathrm{t}} + \ldots + \beta^{\mathrm{k}} X^{\mathrm{k}}{}_{\mathrm{i}}{}^{\mathrm{t}} + \varepsilon_{\mathrm{i}}{}^{\mathrm{t}}$$

b) *Fixed Effect Model (FEM)*

The fixed-effects model estimates panel data regression using the Least Squares Dummy Variable (LSDV) method. This model assumes that each individual has a different intercept but the same slope across all time periods. The general equation for the fixed-effects model is as follows:

$$Y_{\mathrm{i}}{}^{\mathrm{t}} = \alpha_{\mathrm{i}} + \beta_1 X_{1\mathrm{i}}{}^{\mathrm{t}} + \beta_2 X_{2\mathrm{i}}{}^{\mathrm{t}} + \ldots + \beta^{\mathrm{k}} X^{\mathrm{k}}{}_{\mathrm{i}}{}^{\mathrm{t}} + \varepsilon_{\mathrm{i}}{}^{\mathrm{t}}$$

where α_{i} is an intercept that varies for each individual i.

c) *Random Effect Model (REM)*

In the Random Effects model, differences in individual characteristics and time are accounted for through the error component of the model. There are two components that contribute to the error term: the individual error component (u_{i}) and the pooled error component $(\varepsilon_{\mathrm{i}}{}^{\mathrm{t}})$, therefore, the random error in the Random Effects model must be decomposed into these two parts. The estimation technique used is Generalized Least Squares (GLS). The general equation for the Random Effects model is as follows:

$$Y_{\mathrm{i}}{}^{\mathrm{t}} = \beta_0 + \beta_1 X_{1\mathrm{i}}{}^{\mathrm{t}} + \beta_2 X_{2\mathrm{i}}{}^{\mathrm{t}} + \ldots + \beta^{\mathrm{k}} X^{\mathrm{k}}{}_{\mathrm{i}}{}^{\mathrm{t}} + u_{\mathrm{i}} + \varepsilon_{\mathrm{i}}{}^{\mathrm{t}}$$

where u_{i} is the error component that reflects differences in characteristics among individuals, and $\varepsilon_{\mathrm{i}}{}^{\mathrm{t}}$ is the combined error component (Amaliah et al., 2020; Firdaus, 2021).

2.3 Selection of the Best Model

In panel data regression, to determine the best model among the three available models, three types of statistical tests are used: the Chow test, the Hausman test, and the Lagrange multiplier test.

1) Chow Test

The Chow test is used to determine the best model between the Fixed Effects Model and the Common Effects Model. The hypotheses in the Chow test are:

H_0 : The Common Effects Model provides better estimates than the Fixed Effects Model

H_1 : The Fixed Effects Model estimates are better than those of the Common Effects Model

If p-value is less than the significance level $\alpha = 0{,}05$ then H_0 rejected.

2) Hausman Test

The Hausman test is used to determine the best model between the Fixed Effects Model and the Random Effects Model. The hypotheses in the Hausman test are:

H_0 : Estimates from the Random Effects Model are better than those from the Fixed Effects Model

H_1 : Estimates from the fixed-effects model are better than those from the random-effects model

If p-value is less than the significance level $\alpha = 0{,}05$ then H_0 rejected.

3) Lagrange Multiplier Test

The Lagrange Multiplier test is used to determine the best model between the Common Effects Model and the Random Effects Model. The hypotheses in the Lagrange Multiplier test are:

H_0 : Estimates from the Common Effects Model are better than those from the Random Effects Model

H_1 : Estimates from the Random Effects Model are better than those from the Common Effects Model

If p-value is less than the significance level $\alpha = 0{,}05$ then H_0 rejected.

2.4 Test of Classical Assumptions

The multicollinearity test is one of the classical assumptions that must be satisfied in panel data regression. The multicollinearity test is conducted to determine whether there is a high degree of correlation among the independent variables in the regression model. A model is considered free of multicollinearity if the correlation coefficient between independent variables does not exceed 0.85 or if the Variance Inflation Factor (VIF) is less than 10. The multicollinearity test must be performed on any model selected as the best fit.

The heteroscedasticity test is required only for models that use the OLS estimation technique, such as the Common Effects Model and the Fixed Effects Model. If the best model selected is the Random Effects Model, which uses GLS estimation, then the heteroscedasticity test is not required because GLS already accounts for the heterogeneity of variance across individual units (Amaliah et al., 2020).

2.5 Test for the Best Model Parameters

Hypothesis testing was performed on the best model that had been selected. The parameter tests used included the coefficient of determination, simultaneous significance tests, and partial significance tests.

The coefficient of determination (Adjusted R^2) is used to determine the percentage of the dependent variable's variation that is explained by the independent variables collectively. An adjusted R^2 value close to one indicates that the independent variables account for nearly all the information needed to predict the dependent variable.

The simultaneous significance test (F-test) is used to determine whether the independent variables collectively have a significant effect on the dependent variable. If the probability value of the F-statistic is less than $\alpha = 0{,}05$ then H_0 is rejected, meaning that all independent variables collectively have a significant effect on the dependent variable.

The partial significance test (t-test) is used to determine whether each independent variable has a significant effect on the dependent variable individually. If the probability value of the t-statistic for each variable is less than $\alpha = 0.05$, then H_0 is rejected, meaning that the independent variable has a significant partial effect on the dependent variable (Septianingsih, 2022).

3. Results and Discussion

3.1 Descriptive Statistics

Data exploration in this study used descriptive statistics to obtain a general overview of the characteristics of each research variable. The results of the descriptive analysis are presented in Table 2.

Table 2. Descriptive statistics

Variable	Description	Max.	Min.	Average
NER (Y)	Net Enrolment Rate	96,75	52,18	73,29
POOR (X1)	Percentage of the Poor	27,74	4,61	12,93
AYS (X2)	Average Years of Schooling	11,27	7,36	8,77
EYS (X3)	Expected Years of Schooling	16,71	12,37	14,08

Based on Table 2, the average NER for senior high school and equivalent levels in the eight provinces of Sulawesi and Maluku during the 2018–2023 period was 73.29%. The maximum NER value of 96.75% indicates that there are provinces with very high upper secondary education participation rates, while the minimum value of 52.18% shows that there are still provinces with relatively low participation rates. The average percentage of the poor population in the study area was 12.93%, with the highest value reaching 27.74%, reflecting that poverty levels remain quite high in several provinces. The average years of schooling (AYS) stood at 8.77 years, while the average expected years of schooling (EYS) reached 14.08 years, indicating that people in the Sulawesi and Maluku regions are expected to pursue education through upper secondary school or college.

3.2 Estimation of a Panel Data Regression Model

Following the descriptive analysis, parameter estimation and the formulation of the equations for the three panel data regression models were conducted.

a) Common Effect Model (CEM)

The results of the Common Effect Model estimation indicate that there are two independent variables that have a significant effect on the NER for high school and equivalent levels: variable X2 (average years of schooling) with a probability of 0.0000, and variable X3 (expected years of schooling) with a probability of 0.0317. Variable X1 (percentage of the poor population) does not have a significant effect, with a probability of 0.2910. The CEM estimation results are presented in Table 3

Table 3. Estimation Results of the Common Effect Model (CEM)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	22,5178	6,8619	3,2816	0,0017
X1 (POOR)	-0,0731	0,0686	-1,0653	0,2910
X2 (AYS)	2,4519	0,3640	6,7358	0,0000
X3 (EYS)	1,4525	0,6603	2,1998	0,0317
R-squared	0,5889	Adj. R-squared	0,5683	
F-statistic	28,6497	Prob (F-statistic)	0,0000	

Based on Table 3, the CEM model equation obtained is as follows:

$$\widehat{NER}_i^t = 22{,}5178 - 0{,}0731\, X1_i^t + 2{,}4519\, X2_i^t + 1{,}4525\, X3_i^t$$

b) Fixed Effect Model (FEM)

The results of the Fixed Effects Model estimation indicate that none of the independent variables has a significant partial effect on the NER for senior high schools and equivalent schools at the 5% significance level, although the

model is significant when considered simultaneously. This is likely due to stronger multicollinearity when the inter-provincial dummy variables are included in the model. The results of the FEM estimation are presented in Table 4.

Table 4. Estimation Results of the Fixed Effect Model (FEM)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	14,9247	47,6732	0,3131	0,7555
X1 (POOR)	-0,5554	0,4097	-1,3557	0,1809
X2 (AYS)	2,8025	1,7452	1,6059	0,1142
X3 (EYS)	2,1907	4,4340	0,4941	0,6233
R-squared	0,7925	Adj. R-squared	0,7533	
F-statistic	20,2376	Prob (F-statistic)	0,0000	

Based on Table 4, the FEM model equation obtained is as follows:

$$\widehat{NER}_{i}^{t} = 14{,}9247 - 0{,}5554\, X1_{i}^{t} + 2{,}8025\, X2_{i}^{t} + 2{,}1907\, X3_{i}^{t}$$

c) Random Effect Model (REM)

The results of the Random Effects Model (REM) estimation indicate that there are two independent variables that have a significant partial effect on the NER for senior high schools and equivalent schools: variable X1 (percentage of the poor population) with a probability of 0.0049 and variable X2 (average years of schooling) with a probability of 0.0112. The REM estimation results are presented in Table 5.

Table 5. Estimation Results of the Random Effect Model (REM)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	15,9729	10,8670	1,4699	0,1851
X1 (POOR)	-0,1911	0,0473	-4,0429	0,0049
X2 (AYS)	3,0547	0,8945	3,4148	0,0112
X3 (EYS)	1,6388	0,9909	1,6539	0,1421
Adj. R-squared	0,4105	F-statistic	0,0000	

Based on Table 5, the FEM model equation obtained is as follows:

$$\widehat{NER}_{i}^{t} = 15{,}9729 - 0{,}1911\, X1_{i}^{t} + 3{,}0547\, X2_{i}^{t} + 1{,}6388\, X3_{i}^{t}$$

3.3 Best Model Selection Test

a) Chow Test

A Chow test was conducted to determine the best model between the Fixed Effects Model and the Common Effects Model. The test results are presented in Table 6.

Table 6. Chow Test Results

Effects Test	Statistic	Prob.
Cross-section F	7,4265	0,0000
Cross-section Chi-square	43,7461	0,0000

Based on Table 6, the results of the Chow test show that the probability value for the cross-sectional chi-square test (p-value) is 0.0000, which is less than the significance level of $\alpha = 0.05$. This leads to the rejection of H_0, so it can be concluded that the fixed-effects model is superior to the common-effects model. Thus, the analysis proceeds to the next stage, namely the Hausman test.

b) Hausman Test

The Hausman test was conducted to determine the best model between the fixed-effects model and the random-effects model. The test results are presented in Table 7.

Table 7. Hausman Test Results

Test Summary	Chi-Sq. Statistic	Chi-Sq. df	Prob.
Cross-section random	2,9384	3	0,4012

Based on Table 7, the results of the Hausman test show that the probability value for the cross-section random hypothesis (p-value) is 0.4012, which is greater than the significance level of $\alpha = 0.05$. This means that H_0 is not rejected, so it can be concluded that the Random Effects model is better than the Fixed Effects model. Thus, the analysis proceeds to the Lagrange Multiplier test to determine whether the Random Effects model is better than the Common Effects model.

c) Lagrange Multiplier Test

The Lagrange Multiplier test was conducted to determine the best model between the Common Effects Model and the Random Effects Model. The results of the Breusch-Pagan test are presented in Table 8.

Table 8. Lagrange Multiplier (Breusch-Pagan) Test Results

	Cross-section	Time	Both
Breusch-Pagan	29,7478(0,0000)	0,0195 (0,8889)	29,7673 (0,0000)
Honda	5,4542(0,0000)	-0,1396 (0,5555)	3,7579 (0,0001)

Based on Table 8, the results of the Lagrange Multiplier test show that the probability value for the Cross-section Breusch-Pagan test is 0.0000, which is smaller than the significance level of $\alpha = 0.05$. This leads to the rejection of H_0, so it can be concluded that the Random Effects model is better than the Common Effects model. Based on the entire series of model selection tests, it can be concluded that the best model for estimating the relationship between the independent variables and the Net Enrolment Rate for senior high school and equivalent levels in the Sulawesi and Maluku regions is the Random Effects Model (REM).

3.4 Classical Assumptions Test

Since the best model selected was the Random Effects Model (REM), which uses the Generalized Least Squares (GLS) estimation technique, the only classical assumption test that must be conducted is the multicollinearity test. The multicollinearity test was conducted using the correlation matrix among the independent variables. The test results are presented in Table 9.

Table 9. Correlation Matrix Between Variables

	Y	X1	X2	X3
NER (Y)	1,000000	-0,079709	0,750369	0,558322
POOR (X1)	-0,079709	1,000000	-0,050994	0,196690
AYS (X2)	0,750369	-0,050994	1,000000	0,597482
EYS (X3)	0,558322	0,196690	0,597482	1,000000

Based on Table 9, the correlation coefficients between the independent variables are as follows: the correlation between X1 and X2 is -0.0510, the correlation between X1 and X3 is 0.1967, and the correlation between X2 and X3 is 0.5975. All correlation coefficients between the independent variables are below 0.85, indicating that there is no serious multicollinearity among the independent variables in the model. Thus, the classical assumption of multicollinearity is satisfied.

3.5 Test for the Best Model Parameters

Hypothesis testing was performed on the best-fitting model, namely the Random Effects Model. The parameter tests conducted included the coefficient of determination, simultaneous significance tests, and partial significance tests.

a) Coefficient of Determination Test

The coefficient of determination analysis is used to determine the percentage of the independent variables' combined contribution to the dependent variable. The results of the Random Effects Model estimation yielded an adjusted R-squared value of 0.4105. This means that the independent variables comprising the percentage of the poor population, average years of schooling, and expected years of schooling explain 41.05% of the variation in the dependent variable (NER for high school or equivalent), while the remaining 58.95% is explained by other variables outside the model.

b) Simultaneous Significance Test

A simultaneous significance test was conducted to determine whether the independent variables collectively have a significant effect on the dependent variable. Based on the results of the Random Effects Model estimation, the F-statistic probability value was 0.0000, which is far smaller than the significance level of $\alpha = 0.05$. This leads to the rejection of H_0, so it can be concluded that the variables percentage of the poor population, average years of schooling, and expected years of schooling collectively have a significant effect on the Net Enrolment Rate for senior high school and equivalent levels in the Sulawesi and Maluku regions.

c) Partial Significance Test

The partial significance test is used to determine whether each independent variable has a significant individual effect on the dependent variable. The results of the partial test from the Random Effects Model are presented in Table 10.

Table 10. Results of the Partial Significance Test for the Random Effects Model

Variable	Coefficient	Std. Error	t-Statistic	Prob.
NER (Y)	15,9729	10,8670	1,4699	0,1851
POOR (X1)	-0,1911	0,0473	-4,0429	0,0049**
AYS (X2)	3,0547	0,8945	3,4148	0,0112*
EYS (X3)	1,6388	0,9909	1,6539	0,1421

Note: * significant at $\alpha = 5\%$; ** significant at $\alpha = 1\%$

Based on Table 10, the results of the hypothesis testing for each independent variable are as follows:

1) Variable X1 (Percentage of the Poor Population): The p-value for variable X1 is 0.0049, which is less than $\alpha = 0.05$. This leads to the rejection of H_0, so it can be concluded that the percentage of the poor population has a significant partial effect on the dependent variable, NER for high school and equivalent levels.
2) Variable X2 (Average Years of Schooling): The p-value for variable X2 is 0.0112, which is less than $\alpha = 0.05$. This leads to the rejection of H_0, so it can be concluded that the average length of schooling has a significant partial effect on the dependent variable, NER for high school and equivalent levels.
3) Variable X3 (Expected Years of Schooling): The p-value for variable X3 is 0.1421, which is greater than $\alpha = 0.05$. Consequently, H_0 is not rejected, so it can be concluded that the expected schooling duration variable does not have a significant partial effect on the dependent variable, NER for high school and equivalent levels, at the 5% significance level

3.6 Interpretation of the Random-Effects Model

Based on the results of the above analysis, the following equation for the Random Effects Model was obtained:

$$\widehat{NER}_i{}^t = 15{,}9729 - 0{,}1911\ X1_i{}^t + 3{,}0547\ X2_i{}^t + 1{,}6388\ X3_i{}^t$$

Based on the estimation model above, the regression coefficients can be interpreted as follows. First, the coefficient for the percentage of the poor population (X1) is -0.1911. This finding confirms the initial hypothesis stated by (Indrasetianingsih & Wasik, 2020) that the poverty rate acts as the primary financial barrier that constrains households' purchasing power to fund their children's continued education. These results are consistent with previous studies by (Azizah, 2021; Rochmatin, 2019), which confirm that poverty is a factor influencing human well-being and development. In archipelagic regions such as Maluku, high poverty rates can increase the financial burden associated with education and reduce households' ability to support their children's continued education. Furthermore,

government policies aimed at increasing upper secondary education participation must be accompanied by targeted educational assistance for students from low-income households.

Second, the coefficient for the average years of schooling variable (X2) is positive at 3.0547, meaning that when the average years of schooling in a province increases by one year, the NER for senior high school and equivalent levels will increase by 3.0547%, assuming all other variables remain constant. This finding indicates that higher educational attainment within the population is associated with greater participation in upper secondary education. This result is theoretically consistent with the human capital perspective, in which higher educational attainment can strengthen awareness of the benefits of education and encourage educational continuation among younger generations. The finding also supports previous studies showing the relationship between educational attainment and educational participation (Azizah, 2021; Wardhana & Kharisma, 2020). Therefore, policies to prevent school dropout and improve educational continuity are important for increasing the NER in Maluku.

Third, the coefficient for the expected years of schooling variable (X3) is positive at 1.6388 but is not statistically significant at the 5% significance level, with a p-value of 0.1421. Although the coefficient indicates a positive relationship, the effect is not strong enough to be statistically proven at the significance level used. This finding may be related to the moderate correlation between expected years of schooling and average years of schooling (0.5975), which may cause overlapping information between the two educational indicators. From a policy perspective, increasing educational expectations alone may not be sufficient to increase actual school participation. Such expectations need to be supported by adequate economic capacity and access to education. Therefore, education policies should focus not only on increasing educational aspirations but also on ensuring that students are able to continue their education through the upper secondary level..

4. Conslusion

The best estimation model in this study is the Random Effects Model (REM), selected through the Chow Test, Hausman Test, and Lagrange Multiplier Test. Based on the REM estimation results, the percentage of the poor population (X1) has a significant negative effect on the Net Enrolment Rate (NER) for senior high school and equivalent levels, while the average years of schooling (X2) has a significant positive effect. Meanwhile, the expected years of schooling (X3) has a positive but statistically insignificant effect at the 5% significance level. These findings indicate that economic conditions and educational attainment are important factors associated with upper secondary education participation in the Sulawesi and Maluku regions.

The adjusted R-squared value of 0.4105 indicates that the three independent variables explain 41.05% of the variation in NER, while the remaining 58.95% is explained by other factors outside the model. The simultaneous test also confirms that the three independent variables jointly have a significant effect on NER. Therefore, government policies to increase upper secondary education participation should focus on reducing economic barriers for low-income households and improving educational attainment. Targeted educational assistance and school dropout prevention programs can be considered, particularly in regions with high poverty rates and low average years of schooling.

Future research is recommended to include other factors that may influence NER, such as the Human Development Index (HDI), government expenditure on education, school availability, and geographical accessibility. Expanding the study to all provinces in Indonesia may also provide a broader understanding of the determinants of upper secondary education participation at the national level.

References

Amaliah, E. N., Darnah, D., & Sifriyani, S. (2020). Regresi Data Panel dengan Pendekatan Common Effect Model (CEM), Fixed Effect model (FEM) dan Random Effect Model (REM) (Studi Kasus: Persentase Penduduk Miskin Menurut Kabupaten/Kota di Kalimantan Timur Tahun 2015-2018). *ESTIMASI: Journal of Statistics and Its Application*, *1*(2), 106. https://doi.org/10.20956/ejsa.v1i2.10574

Azizah, N. N. N. (2021). *Pemodelan Data Kemiskinan Di Provinsi Jawa Barat Menggunakan Regresi Spasial Data Panel* [Universitas Negeri Jakarta]. http://repository.unj.ac.id/id/eprint/18127

BPS. (2018). *Jumlah dan Persentase Penduduk Miskin Menurut Provinsi, 2025—Tabel Statistik* [Dataset]. https://www.bps.go.id/id/statistics-table/3/UkVkWGJVZFNWakl6VWxKVFQwWjVWeTlSZDNabVFUMDkjMw==/jumlah-dan-persentase-penduduk-miskin-menurut-provinsi--2018.html?year=2018

BPS. (2018). *Statistika Pendidikan* (1st ed., Vol. 12). (Original work published 2023)

Fadilla, A. (2021). MACRO ECONOMY AND POVERTY REDUCTION IN INDONESIA (Analysis of Economic Growth Capability and Macroeconomic Indicators in Poverty Reduction in Indonesia). *Ascarya: Journal of Islamic Science, Culture, and Social Studies*, *1*. https://doi.org/DOI:%20https://doi.org/10.53754/iscs

Firdaus, M. (2021). *Aplikasi Model Ekonometrika dengan RStudio Edisi Kedua*. https://ipbpress.com/product/633-Aplikasi Model Ekonometrika dengan RStudio Edisi Kedua

Indrasetianingsih, A., & Wasik, T. K. (2020). Model Regresi Data Panel Untuk Mengetahui Faktor Yang Mempengaruhi Tingkat Kemiskinan Di Pulau Madura. *Jurnal Gaussian*, *9*(3), 355–363. https://doi.org/10.14710/j.gauss.v9i3.28925

Mona, I., Martha, S., & Perdana, H. (2022). PEMODELAN DATA PERSENTASE PENDUDUK MISKIN DI INDONESIA MENGGUNAKAN SUR-SEM DATA PANEL. *Buletin Ilmiah Mat. Stat. dan Terapannya (Bimaster)*, *11*(3), 403–412. https://doi.org/https://doi.org/10.26418/bbimst.v11i3.54617

Rahminawati, N. (2023). STRATEGI PENINGKATAN HARAPAN LAMA SEKOLAH (HLS) DAN RATA-RATA LAMA SEKOLAH (RLS). *Ta'dib : Jurnal Pendidikan Islam*, (2), 367–382. https://doi.org/DOI:%2010.29313/tjpi.v12i2.12852

Rochmatin, H. (2019). *Analisis Regresi Data Panel Untuk Mengetahui Faktor Yang Mempengaruhi Angka Harapan Hidup Di Provinsi Jawa Timur* [Universitas Airlangga]. http://lib.unair.ac.id

Septianingsih, A. (2022). Pemodelan Data Panel Menggunakan Random Effect Model Untuk Mengetahui Faktor Yang Mempengaruhi Umur Harapan Hidup Di Indonesia. *Jurnal Lebesgue : Jurnal Ilmiah Pendidikan Matematika, Matematika dan Statistika*, *3*(3), 525–536. https://doi.org/10.46306/lb.v3i3.163

Wardhana, A., & Kharisma, B. (2020). Angka Harapan Hidup Di Era Desentralisasi (kabupaten/Kota Jawa Barat). *E-Jurnal Ekonomi dan Bisnis Universitas Udayana*, 405. https://doi.org/10.24843/EEB.2020.v09.i05.p02