ISSN 2808-2605 EISSN 2808-4497 *Forum Analisis Statistik*
Juni 2023, 3 (1): 34 - 41

Analysis Of Factors Affecting The Percentage Of Poverty In West Kalimantan With Panel Data Regression

Dinda Kinanti Ayuningtyas[1]*, Khairina Auliannisa[2], Shantika Martha[3]

[1,2,3] Universitas Tanjungpura,
Jl. Prof Dr. Hadari Nawawi, Kota Pontianak
* Correspondence Writer. *e-mail*: h1091201036@student.untan.ac.id
(Received: Juni 26th 2023; Approved: July 30th 2023)

ABSTRACT

Indonesia has a high poverty rate, which is evidenced by the annual increase in the percentage of people living in poverty. This happens because the components of high life expectancy, as well as the higher level of education, Reducing the poverty rate is closely related to improving the welfare of the community. This has a significant impact on the overall well-being of the community. Therefore, this study aims to determine the best panel data regression model of Life Expectancy and Expected Years of Schooling on the percentage of poor people. The three model approaches tested are Common Effect Model (CEM), Fixed Effect Model (FEM), and Random Effect Model (REM). Based on the estimation of panel data regression parameters, it was found that the best model approach to analyze the poverty rate in West Kalimantan in 2018-2022 is the FEM. The F test showed that the result is 0.00 < 0.05. Which means that the Life Expectancy and Expected Years of Schooling variables have a significant influence on poverty rate. the Life Expectancy and Expected Years of Schooling were found to be the variables that influence poverty rate in West Kalimantan, with an adjusted R^2 value of 98,89%.

Keywords*: CEM, FEM, REM, Panel Data Regression, Poverty*

INTRODUCTION

Poverty hinders a country from improving its economy (Nafi'ah, 2021). Currently, poverty is still a problem in several countries around the world, and one of them is Indonesia; this is because poverty is multidimensional the needs of each individual are pretty diverse (Zuhdiyaty & Kaluge, 2018). Not only is it challenging to fulfill secondary conditions, but even primary needs are also tricky to fulfill because of the difficulty of finding income.

Reducing poverty in Indonesia can be overcome by improving education and health, which is seen in life expectancy. The higher the level of education, the lower the poverty rate. This is because the education factor is one of the determinants of reducing the poverty rate. In addition, an individual's health greatly affects the overall welfare of the community and is closely connected to poverty (Hasanah et al., 2021) and has a significant effect on poverty in West Kalimantan (Gunawan, 2020). However, research by Bancin & Usman (2020) states that life expectancy has no impact on the poverty rate.

This study uses panel regression to analyze what factors affect the poverty rate (Y) percentage in West Kalimantan. The factors are the Life Expectancy (X_1), and Expected Years of Schooling (X_2). The approaches used are Common Effect Model (CEM), Fixed Effect Model (FEM), and Random Effect Model (REM). Testing methods are also used: the Chow test, Hausman test, Lagrange Multiplier test, and residual classical assumption test. The data used for this study was obtained from the West Kalimantan BPS website and covers the observation period from 2018 to 2022.

METHODOLOGY

This study analyzed secondary data from the Central Bureau of Statistics of West Kalimantan Province, namely 14 districts/cities located in West Kalimantan province from 2018 to 2022. The study focuses on dependent and independent variables, namely the variable Percentage of Poor Population as the dependent variable (Y), *the Life Expectancy and Expected Years of Schooling* as independent variables (X_1, X_2).

Panel Data Regression Model

Panel data regression is a statistical method that combines time series and cross-sectional data, also known as panel data, to determine the impact of independent variables on dependent variables. The panel data regression model equation is typically written as follows:

$$Y_{it} = \beta_{0it} + \sum_{k=1}^{n} \beta_k X_{kit} + e_{it} e_{it}; \quad (1)$$

In equation (1), it is known that Y_{it} is the observation of the i unit, β_0 is the intercept of the i unit, β_k is the slope of the i unit, $X_{it} = x_{1it}, x_{2it}, \dots, x_{kit}$ are independent variables in the i observation, while μ_i is the error component for the i individual unit and is entirely observed in the t period.

Parameter Estimation

a. Common Effect Model (CEM)

In the context of CEM, it is known that the intercept and slope for every variable remain consistent across all cross-sections and time series units. The CEM equation can be written in the following manner:

$$Y_{it} = \beta_0 + \sum_{k=1}^{n} \beta_k X_{kit} + e_{it} \quad (2)$$

In equation (2), Y_{it} is the observation of the i unit. β_0 is the regression model intercept of the i observation unit. β_k is the slope coefficient of the i unit. X_{it} is the independent variable in the i observation of the t period. e_{it} is the error in the i observation.. i is the cross-section unit (1, 2, 3, …, N). t is the number of predictor variables (1, 2, 3, …, n).

b. Fixed Effect Model (FEM)

In FEM, it is known that there are different effects between individuals, one of which is the difference in the intercept. Therefore, each individual in FEM is an unknown parameter and will be estimated called a dummy variable. The FEM equation can be written as follows:

$$Y_{it} = \beta_{0it} + \sum_{k=1}^{n} \beta_k X_{kit} + e_{it} \quad (3)$$

c. Random Effect Model (REM)

REM researchers consider using panel data to address potential correlations between disturbance variables and time/individuals. In addition, REM also aims to fix various problems caused by the previous model, FEM. The REM equation is written as follows:

$$Y_{it} = \beta_{0it} + \sum_{k=1}^{n} \beta_k X_{kit} + \mu_i + e_{it} \quad (4)$$

RESULTS AND DISCUSSION

Variable

This research utilizes data from the official website of BPS West Kalimantan as a secondary source. The study examines the percentage of poor individuals in the population, with less per capita income than their total expenditure. The dependent variable is the percentage of poor people, while the independent variables include the Life Expectancy (X_1), and Expected Years of Schooling (X_2). Life expectancy estimates the average additional age a person is expected to live, with units being percent (%). Expected Years of Schooling is the length of schooling (in years) a child is likely to experience at a certain age, with the unit being percent (%).

Descriptive Statistic

The purpose of descriptive statistical analysis is to determine the mean, median, standard deviation, minimum, and maximum values of data. You can find the results of this analysis in Table 1.

Table 1. Descriptive Statistic

No	Variable	N	Mean	Median	Standar Deviation	Maximum	Minimum
1	Y (percent)	70	7,57	7,06	2,62	12,83	4,34
2	X_1 (percent)	70	71,54	71,68	1,51	74	67,71
3	X_2 (percent)	70	12,39	12,08	0,94	15,02	11,13

Source: Author

According to Table 1, the percentage of impoverished individuals in West Kalimantan's districts/cities has an average (mean) value of 7.57% between 2018 and 2022. Kubu Raya district had the lowest percentage of impoverished individuals in 2022, with a figure of 4.34%, while Melawi district had the highest percentage of impoverished individuals in 2018, at 12.83%.

It is known that the district/city with the lowest the Life Expectancy is North Kayong district in 2018 at 67,71%, while the highest the Life Expectancy is Bengkayang district in 2022 at 74%. It is known that the district/city with the lowest Expected Years of Schooling was Melawi Regency in 2018 at 11,13%, while the highest Expected Years of Schooling was Pontianak City in 2022 at 15,02%.

Model Parameter Estimation

a. *Common Effect Model* (CEM)

Table 2. Results of CEM

No	Variable	Coefficient
1	*Intercept*	21,840
2	X_1 (the Life Expectancy)	0,066
3	X_3 (Expected Years of Schooling)	-1,534

Based on Table 2, the CEM panel data regression model obtained the following results:

$$\widehat{Y_{it}} = 21{,}840 + 0{,}066X_{1it} - 1{,}534X_{2it} \quad (5)$$

In the regression model, it is known that a constant has a value of 21,840 which states that without the influence of the Life Expectancy and Expected Years of Schooling the percentage of poor people is 21,840. Every increase of 1 unit of the Life Expectancy can higher the percentage of people living in poverty by 6,6%. Every increase of 1 unit of Expected Years of Schooling can lower the percentage of people living in poverty by 153%.

b. *Fixed Effect Model* (FEM)

Table 3. Results of FEM

No	Variable	Coefficient
1	*Intercept*	86,499
2	X_1 (the Life Expectancy)	-1,215
3	X_2 (Expected Years of Schooling)	0,647

Source : Author

Based on Table 3, the FEM panel data regression model obtained the following results:

$$\widehat{Y_{it}} = 86{,}499 - 1{,}215X_{1it} + 0{,}647X_{2it} \quad (6)$$

In the regression model, it is known that a constant has a value of 86,499 which states that without the influence of the Life Expectancy and Expected Years of Schooling, the percentage of poor people is 86,499. Every increase of 1 unit of the Life Expectancy can lower the percentage of people living in poverty by 121,5%. Every increase of 1 unit of Expected Years of Schooling can higher the percentage of people living in poverty by 64,7%.

c. *Random Effect Model* (REM)

Table 4. Results of REM

No	Variable	Coefficient
1	*Intercept*	77,781
2	X_1 (The Life Expectancy)	-0,928
3	X_2 (Expected Years of Schooling)	-0,302

Based on Table 4, the REM panel data regression model obtained the following results:

$$\widehat{Y_{it}} = 77{,}781 - 0{,}928X_{1it} - 0{,}302X_{2it} \quad (7)$$

In the regression model, it is known that a constant has a value of 77,781 which states that without the influence of the Life Expectancy and Expected Years of Schooling, the percentage of poor people is 77,781. Every increase of 1 unit of the Life Expectancy can lower the percentage of people living in poverty by 92,8%. Every increase of 1 unit of Expected Years of Schooling can lower the percentage of people living in poverty by 30,2%.

Best Model Selection

Chow Test

This test determines which panel data regression model is better and whether FEM is better than CEM. The test hypothesis is as follows.

H_0: $\alpha_1 = \alpha_2 = \cdots = \alpha_{10} = 0$ (CEM is better than FEM).

H_1: One intercept (α_i) is not equal (FEM is better than CEM)

If the probability value in the Chow test is higher than the significant level of 0.05, then H_0, should be accepted. This means that the CEM model is superior to the FEM model.

Table 5. Results of Chow Test

No	Effects Test	Probability
1	*Cross-section F*	0,0000
2	*Cross section Chi-Square*	0,0000

The Chow test results indicate that the FEM model is a better estimation model than the CEM model. This is because the Cross Section Chi-Square probability value is 0.00 < 0.05, This means that the hypothesis is rejected. To determine which model is better between FEM and REM, the Hausman Test will be conducted. This is because in the Chow test, FEM was found to be the best estimation model.

Haussman Test

This test determines which panel data regression technique is better and whether FEM is better than REM. The hypothesis is as follows.

H_0: $corr(x_{it}, u_{it}) = 0$ (*REM* is better than *FEM*)

H_1: $corr(x_{it}, u_{it}) \neq 0$ (FEM is better than REM)

When conducting the Hausman test, if the probability value is higher than the significant level value of 0.05, it indicates that the REM model is superior to the FEM model and should be accepted.

Table 6. Results of Haussman Test

No	Effects Test	Probability
1	*Cross-section random*	0,0026

According to Table 6, the probability value for the cross-section random effect is 0.0026 < 0,05. The results indicate that H_0 has been rejected, implying that the FEM model is a valid estimation method for comparing FEM and REM. Hence, we will proceed with the panel data regression model using Fixed Effect Model (FEM) estimation. And the Lagrange multiplier test will not be carried out because the best regression model is the FEM.

From the results obtained, it can be concluded that an excellent model to use in analyzing the percentage of poor people is the FEM. Next, the parameter significance test will be conducted.

The Parameter Significance Test

F-Test

A simultaneous test is conducted to evaluate the effect of both the independent and dependent variables at the same time. The hypothesis for a simultaneous test is as follows.

H_0: $\beta_1 = \beta_2$

H_1: $\beta_k \neq 0$ with k=1,2

When conducting a simultaneous F test, if the F value is less than the significant level (0.05), it indicates that the independent and dependent variables have a significant influence, and we accept H_0. After conducting the test, it was found that the F value was 0.000, which is less than the significant level (0.05). Therefore, the null hypothesis (H_0) was rejected, and conclude that the variables of the Life Expectancy and Expected Years of Schooling have a significant influence on the Percentage of Poor Population variable simultaneously.

T-Test

The purpose of the Partial T-Test is to determine the extent of the significant effect between the dependent and independent variables. The hypothesis for the partial T-test is as follows.

$H_0: \beta_k = 0$

$H_1: \beta_k \neq 0$ with k = 1, 2, dan 3.

When conducting a partial T-test, if the p-value exceeds the significant level of 0.05, it is appropriate to accept H_0, or it can be concluded that the independent and dependent variables have a partially significant effect.

Table 8. Results of T-test

No	Parameter	*P-Value*
1	β_1 (the Life Expectancy)	0,000
3	β_3 (Expected Years of Schooling)	0,214

Based on Table 8, the t-statistic probability value for each parameter is obtained for the Life Expectancy and Expected Years of Schooling, which have a value of 0.000 and 0.214, respectively, where the value of the significant level (0.05), Based on the analysis, it can be inferred that the Life Expectancy variable has a partial significant impact on the Poor Population Percentage variable.

Residual Classical Assumption Test

Normality Test

We will conduct a normality test to check if the data meets the criteria for normal distribution or not. This test will help us in hypothesis testing. Here is the hypothesis statement for the normality test.

H_0: $\varepsilon_i \sim N(0, \sigma^2)$ or residuals are normally distributed.

H_1: Residuals are not normally distributed.

To determine if data is normally distributed, we use the Jarque-Bera normality test. If the probability value from this test is greater than 0.05, we accept the null hypothesis (H_0), meaning the data is normally distributed.

Jarque Bera probability value is 0.318. If the value is more significant than α = 0.05, accept H_0; if the value is smaller than α = 0.05, reject H_0. So, it can be concluded that accepting H_0 so that the resulting residuals are normally distributed.

Multicollinearity Test

The multicollinearity test checks if there is a significant correlation among independent variables in the regression model. In this multicollinearity test, there will be no multicollinearity if it has a correlation value smaller than 0.8.

Table 9. Results of Multicollinearity Test

No	Parameter	*Correlation*
1	The Life Expectancy	0,074
3	Expected Years of Schooling	0,746

According to Table 9, the correlation value is less than 0.8. This means that there is no multicollinearity, which meets the requirements. So, in this case, the existing multicollinearity can be fulfilled.

Heteroscedasticity Test

The purpose of the Heteroscedasticity test is to check if there is consistent residual variance in each independent variable. The heteroscedasticity test hypothesis is as follows.

H_0: The heteroscedasticity problem doesn't exist.

H_1: The heteroscedasticity problem does exist..

If the p-value in the heteroscedasticity test exceeds the significance level, it can be concluded that there is no heteroscedasticity problem and should be accepted H_0 .

Table 10. Results of Heteroscedasticity test

No	Parameter	*P-Value*
1	Intercept	0,100
2	X_1 (The Life Expectancy)	0,179
3	X_2 (Expected Years of Schooling)	0,544

The test results show that the probability value obtained is higher than α = 0.05. so accept H_0, and it can be concluded that the resulting residuals have the same variance in each independent variable used. So, in this case, there is no heteroscedasticity problem.

Autocorrelation Test

The purpose of the Autocorrelation test is to establish if there is a correlation between the variables used. The autocorrelation test hypothesis is as follows.

H_0: $\rho = 0$ (There is no autocorrelation)

H_1: $\rho \neq 0$ (There is autocorrelation)

In the autocorrelation test, if the Durbin-Watson value is more significant than dL, it will accept H_0, or it can be concluded that there is no autocorrelation.

Based on the Durbin-Watson test results, the Durbin-Watson value obtained is 1,565 with a dL value of 1.554 and a dU value of 1.671. The DW value obtained is higher than dL, so accept H_0 with the conclusion that there is non autocorrelation between the variables used in the study. So, in this case, the non-autocorrelation assumption can be fulfilled.

In the test results, the adjusted R^2 value obtained is 0.988. So from The Life Expectancy and Expected Years of Schooling influence 98,8% and the remaining 1,2% is influenced by other variables outside the research that has been done.

CONCLUSION AND RECOMMENDATIONS

Based on the data analysis and discussion above, it is concluded that the best model obtained in this study is FEM. The panel data regression model obtained with FEM is as follows:

$$\widehat{Y_{it}} = 86{,}499 - 1{,}215X_{1it} + 0{,}647X_{2it}$$

According to a study, The Life Expectancy has a significant impact on the percentage of poor people in West Kalimantan between 2018-2022. The Life Expectancy in each district or city in West Kalimantan can affect the poverty rate.

Total Life Expectancy has a positive influence. This means the government can improve public health facilities such as health centers, clinics, hospitals, etc. In addition, the government can strive for people experiencing poverty to get assistance facilities such as the Social Security Organizing Agency, Healthy Indonesia Card, and vitamins to maintain public health. Thus, this can help increase life

expectancy, and the community can work optimally to reduce the poverty rate in the district/city of West Kalimantan Province.

BIBLIOGRAPHY

Bancin, S. K., & Usman, U. (2020). Pengaruh AHH, AHLS Dan PDRB Terhadap Jumlah Penduduk Miskin Di Provinsi Aceh. *Jurnal Ekonomika Indonesia*, *9*(1), 17. https://doi.org/10.29103/ekonomika.v9i1.3169

BPS Kalimantan Barat, 2022. Indeks Pembangunan Manusia Provinsi Kalimantan Barat. Badan Pusat Statistik Provinsi Kalimantan Barat. Kalimantan Barat.

BPS Provinsi Kalimantan Barat. (2012, Juli). *Kemiskinan di Kalimantan BaratMaret 2012.* Retrieved from BPS Provinsi Kalimantan Barat: https://kalbar.bps.go.id/pressrelease/2012/07/02/341/kemiskinan-di- kalimantan-barat-maret-2012.html

BPS Kalimantan Barat. (2022). Berita Resmi Statistik. https://kalbar.bps.go.id/

Gunawan, M. W. (2020). Pengaruh Pertumbuhan Ekonomi, Pengangguran dan Angka Harapan Hidup Terhadap Kemiskinan di Provinsi Kalimantan Barat Tahun 2010-2018. *Jurnal Ekonomi Daerah (JEDA)*, *8*(3).

Hasanah, R., Syaparuddin, S., & Rosmeli, R. (2021). Pengaruh angka harapan hidup, rata-rata lama sekolah dan pengeluaran perkapita terhadap tingkat kemiskinan pada Kabupaten /Kota di Provinsi Jambi. *E-Jurnal Perspektif Ekonomi Dan Pembangunan Daerah*, *10*(3), 223–232. https://doi.org/10.22437/pdpd.v10i3.16253

Nafi'ah, B. (2021). Analisis Faktor-Faktor Yang Dapat Mempengaruhi Pengentasan Kemiskinan Di Indonesia (2016- 2019). *Jurnal Ilmiah Ekonomi Islam*, *7*(2), 953–960. https://doi.org/10.29040/jiei.v7i2.2206

Zuhdiyaty, N., & Kaluge, D. (2018). Analisis Faktor - Faktor Yang Mempengaruhi Kemiskinan Di Indonesia Selama Lima Tahun Terakhir. *Jurnal Ilmiah Bisnis Dan Ekonomi Asia*, *11*(2), 27–31. https://doi.org/10.32812/jibeka.v11i2.42