JURNAL STATISTIKA DAN APLIKASINYA

e-ISSN 2620 – 8369
VOLUME 8 (ISSUE 1) JUNE 2024: 63 - 74
https://doi.org/10.21009/JSA.08106

PANEL DATA REGRESSION MODELING FACTORS AFFECTING PARTICIPATION IN INDONESIA'S NATIONAL HEALTH INSURANCE CONTRIBUTION ASSISTANCE PROGRAM

Vera Maya Santi[1*], **Muhammad Rafli** [2]

[1,2]Department of Statistics, Faculty of Mathematics and Natural Sciences, Jakarta State University Rawamangun Muka, East Jakarta 13320, Indonesia

Corresponding author's e-mail: [*] **vmsanti@unj.ac.id**

ABSTRACT

Article History:

Received: 18 May 2024
Revised: 05 June 2024
Accepted: 27 June 2024
Published: 30 June 2024
Available online.

Keywords:

Panel Data Regression, National Health Insurance, Random Effects Model

National Health Insurance (JKN) is a health insurance system that provides health protection to individuals who have paid premiums (non-PBI) or whose premiums are paid by the government (PBI). JKN participation in Indonesia from 2018 to 2021 has consistently increased, particularly for JKN PBI, yet the program has not achieved the target of 100% membership coverage globally or nationally. A deeper understanding of the factors influencing JKN is essential to increase the program's effectiveness and formulate targeted policies. This study aims to analyze the factors affecting the participation of JKN PBI annually in Indonesia. The data used in this study cover the period from 2018 to 2021 and involve 34 provinces in Indonesia, thus forming a panel data structure. A panel data regression model is employed to identify the factors influencing JKN PBI participation. The predictor variables analyzed include income levels, access to health facilities, employed population, and education levels. The results indicate that the most suitable model for determining the factors influencing participation in the National Health Insurance PBI with a significance level of $\alpha = 5\%$ is the Random Effect Model (REM). The simultaneous and partial tests show that the variables of the employed population and education levels significantly impact JKN PBI participation. Therefore, the REM is considered appropriate and effective in explaining the factors influencing participation in the JKN PBI program. This study provides insights that can help in formulating more targeted policies to enhance the effectiveness of the JKN program in Indonesia.

How to cite this article:
V.M. Santi, M. Rafli, "PANEL DATA REGRESSION MODELING FACTORS AFFECTING PARTICIPATION IN INDONESIA'S NATIONAL HEALTH INSURANCE CONTRIBUTION ASSISTANCE PROGRAM", Jurnal Statistika dan Aplikasinya, vol. 8, iss. 1, pp. 63 – 74, June 2024

Journal homepage: *https://journal.unj.ac.id/unj/index.php/statistika*
Journal e-mail: *jsa@unj.ac.id*
Research Article · Open Access

1. INTRODUCTION

Health is one of the most important aspects of human life, supporting various activities and enhancing the well-being of society. According to Law Number 36 of 2009, health is a state of being physically, mentally, spiritually, and socially well, enabling each person to live productively both socially and economically [1]. Health is also a goal of the Sustainable Development Goals (SDGs), which aim to ensure healthy lives and promote well-being for all at all ages [2]. Because fundamentally, everyone has the right to live prosperously, have adequate housing, and enjoy a good and healthy environment, as well as the right to receive healthcare services. Therefore, affordable, and quality healthcare services need to be provided. To ensure that health is not only a right but also equally accessible to all citizens, many countries have introduced the concept of national health insurance [3].

In Indonesia, this insurance is known as the National Health Insurance (Jaminan Kesehatan Nasional or JKN) system [4]. According to Minister of Health Regulation No. 6 of 2020, National Health Insurance is a guarantee in the form of health protection to ensure that participants receive benefits for health maintenance and protection to meet basic health needs, provided to every individual who has paid premiums or whose premiums are paid by the government [5]. In the National Social Security System (SJSN) Law, it is mandated that all residents must become participants in the national health insurance [6].

Based on data from the National Social Security Council (DJSN), JKN participation has increased annually from 2018 to 2021. By 2020, the number of JKN participants reached 248,771,083 individuals. However, the number of inactive participants continued to rise from 2019 to 2021, reaching its peak of 48,723,718 individuals in 2021, up from 24,591,275 individuals the previous year [7]. However, this program has not yet achieved the target of 100% membership coverage globally or nationally [8].

The National Health Insurance is divided into two categories: Recipients of Premium Assistance (PBI) and Non-Recipients of Premium Assistance (Non-PBI). PBI participants have their monthly premiums subsidized by the government, hence they do not need to pay them themselves, whereas non-PBI participants pay their monthly premiums personally [9]. Data from the National Social Security Council shows that the number of users of JKN PBI is higher compared to non-PBI. In 2019, JKN PBI users reached 135,359,642 individuals, while non-PBI amounted to 88,789,877 individuals. In 2020, the difference between JKN PBI and non-PBI users was 43,072,416 individuals, and by 2021, this difference increased to 54,817,369 individuals [7]. To maintain and increase participation in the JKN program, especially JKN PBI, research is needed to identify factors influencing JKN PBI membership. This study will serve as a basis for formulating effective policies in the future.

In the study by Darmayanti and Raharjo (2020), it was found that education levels, knowledge level, occupation, income level, information about JKN, and support are factors influencing National Health Insurance. This research attempts to identify factors related to community participation in voluntary JKN [9]. Another study by Yosalli and As Shidieq (2020) also conducted research related to participation in national health insurance in Sleman Regency. This study found that membership in the National Health Insurance in Sleman Regency is influenced by variables such as age, gender, education, occupation, marital status, population status, location of residence, health status, and health coverage status [8].

The data used in this study covers the period from 2018 to 2021 and involves 34 provinces in Indonesia, thus forming a panel data structure. Therefore, one approach that can be used to analyze it is through panel data regression models. In a study conducted by Sitorus and Yuliana (2018) on "Penerapan Regresi Data Panel pada Analisis Pengaruh Infrastruktur terhadap Produktivitas Ekonomi Provinsi-Provinsi di Luar Pulau Jawa Tahun 2010-2014," they successfully identified factors influencing the economic productivity of provinces outside Java, namely road infrastructure, healthcare, and regional budget (APBD) [10]. In addition, this study found that the appropriate model for such analysis is the fixed effect model [10]. The study by Alamsyah et al. (2022) on "Analisis Regresi Data Panel untuk Mengetahui Faktor yang Memengaruhi Jumlah Penduduk Miskin di Kalimantan Timur" indicates that the appropriate model for the research is the random effects model. The factors influencing the number of poor populations in East Kalimantan identified in the study are the Human Development Index and the GDP growth rate [11].

The panel data regression study conducted by Lestari and Setyawan (2017) on "Analisis Regresi Data Panel untuk Mengetahui Faktor yang Mempengaruhi Belanja Daerah di Provinsi Jawa Tengah" indicates that the factors influencing regional expenditure in Central Java Province include Gross Regional Domestic Product, Population Size, Local Own-source Revenue, General Allocation Fund, and Economic Growth [12]. Moreover, the most appropriate model for this analysis is the random effects model [12]. Furthermore, research by Venosial et al. (2022) on "Pemodelan Persentase Kepesertaan BPJS Non Penerima Bantuan Iuran dengan Pendekatan Regresi Data Panel" indicates that the most suitable model is the fixed effects model, and significant factors influencing the enrollment of Non-Recipients of Premium Assistance BPJS are the population living in poverty and the unemployment rate [13]. Quantitative analysis studies on National Health Insurance for Premium Assistance Recipients using panel regression are still rare, so this research is expected to provide new insights into developing more effective policies related to JKN PBI.

2. METHODS

Material and Data

This study uses secondary data obtained from the National Social Security Council (DJSN) and the Central Statistics Agency (BPS). The data covers 34 provinces in Indonesia over the period 2018-2021. The research variables consist of 5 variables as shown in Table 1.

Table 1. Research Variables

Notation	Variables	Scale
Y	Participation in National Health Insurance for Premium Assistance Recipients (JKN PBI)	Ratio
X_1	Income Level	Ratio
X_2	Number of Health Facilities	Ratio
X_3	Number of Employed Population	Ratio
X_4	Education levels	Ratio

Research Method

The analytical approach used in this research is panel data regression. This approach is a regression model used to examine the influence of predictor variables on a response variable using panel data structure [11]. Panel data is a combination of cross-sectional data and time series data [11]. The procedural steps of analysis conducted in this research are as follows:

1. Collecting panel data consisting of cross-sectional data (data from various provinces) and time series data (data from 2018-2021).
2. Presenting descriptive statistics used to understand the characteristics of factors and data on JKN PBI participation, facilitating further data analysis.
3. Applying the panel data regression method

   3.1 Estimating the panel data regression model

   Panel data regression estimation consists of three types of estimation methods:

   1. Common Effect Model (CEM)

      Common effect model is a model that combines all data, both cross-sectional and time series, without considering the time and place of the study [11]. This method assumes that the intercept values for each variable are the same, as are the slope coefficients for all cross-sectional and time series units [14]. In general, the equation for the common effect model is written as follows:

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

Where Y_{it} is the response variable for observation unit i at time t, β_0 is the intercept of the regression model, β_k is the slope coefficient, X_{it} is the predictor variable for observation unit i at time t, e_{it} is the error or residual component for observation unit i at time t, i represents the cross-sectional unit $(1,2,3,\dots,N)$, t represents the time series unit $(1,2,3,\dots,T)$, and k is the number of predictor variables $(1,2,3,\dots,n)$.

2. Fixed Effect Model (FEM)
   The fixed effect model is a regression method that estimates panel data by adding dummy variables [15]. This model assumes that the slope is constant, but the intercept is constant over time and varies across individuals. The equation for the fixed effect model can be expressed as follows:

$$Y_{it} = \beta_{0_{it}} + \sum_{k=1}^{n} \beta_k X_{kit} + e_{it}$$

Where Y_{it} is the response variable for observation unit i at time t, β_0 is the intercept of the regression model, β_k is the slope coefficient, X_{it} is the predictor variable for observation unit i at time t, e_{it} is the error component for observation unit i at time t, i represents the cross-sectional unit $(1,2,3,\dots,N)$, t represents the time series unit $(1,2,3,\dots,T)$, and k is the number of predictor variables $(1,2,3,\dots,n)$.

3. Random Effect Model (REM)
   The random effect model is a method for estimating panel data where the error term may be correlated over time or between individuals [11]. The equation for the model is as follows:

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

Where Y_{it} is the response variable for observation unit i at time t, β_0 is the intercept of the regression model, β_k is the slope coefficient, X_{it} is the predictor variable for observation unit i at time t, μ_i is the error term for observation unit i, e_{it} is the error component for observation unit i at time t, i represents the cross-sectional unit $(1,2,3,\dots,N)$, t represents the time series unit $(1,2,3,\dots,T)$, and k is the number of predictor variables $(1,2,3,\dots,n)$.

3.2 Selection of Panel Data Regression Estimation Model

1. Chow Test
   The Chow test is used to determine the best model between two models: the common effect model and the fixed effect model. The Chow test statistic is as follows:

$$Chow = \frac{\dfrac{[RRS - URRS]}{(n-1)}}{\dfrac{URRS}{nT - n - K)}}$$

2. Hausman Test
   The Hausman test is used to determine the best model between the fixed effect model and the random effect model. The Hausman test statistic is as follows:

$$\chi^2(K) = (b-\beta)'[var(b-\beta)]^{-1}(b-\beta)$$

3.3 Testing Parameter Significance

The significance test of parameters is conducted after obtaining the regression model estimates. This test aims to identify the independent variables that have a significant impact on the model [16].

1 Simultaneous Test (F-Test)
   The F-test is used to evaluate the overall influence of predictor variables on the response variable. The F-Test statistic is as follows:

$$F_{statistic} = \frac{\dfrac{R^2}{(N+K-1)}}{\dfrac{(1-R^2)}{(NT-N-K)}}$$

2 Partial Test (t-Test)

The partial test (t-test) aims to evaluate the significance of each predictor variable individually on the response variable [17]. The t-test statistic is as follows:

$$t = \frac{b_k}{s.e(b_k)}$$

3.4 Classic Assumption Testing

1 Normality Test

The normality test is used to evaluate whether the residual values follow a normal distribution [18], [19]. In this normality test, the Jarque-Bera (JB) method is employed. The statistical test from the Jarque-Bera method is as follows:

$$JB = n\left[\frac{s^2}{6} + \frac{(K-3)^2}{24}\right]$$

2 Multicollinearity Test

This test aims to determine whether there is a relationship among predictor variables. Multicollinearity can be addressed by using Variance Inflation Factors (VIF) [20]. Multicollinearity can be identified by examining the correlation matrix of the data and the values of Variance Inflation Factors (VIF). If the VIF values are less than 10, it indicates that there is no multicollinearity issue in the model [18], [21].

3 Heteroskedasticity Test

The heteroskedasticity test is used to determine whether there is inconsistency in the residual variation in the formed regression model [12].

4 Autocorrelation Test

The autocorrelation test aims to detect correlations between values in specific time periods with values in previous time periods [18]. One of the methods used for this purpose is the Durbin-Watson test.

3. RESULTS

Descriptive Statistics

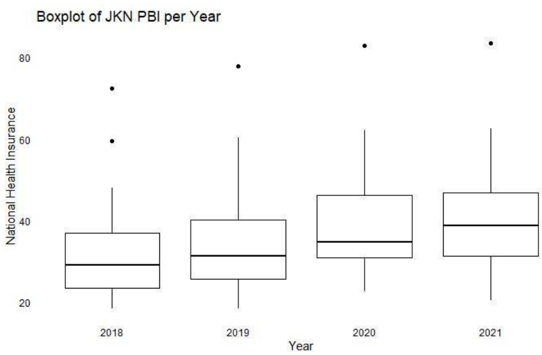

Figure 1. Boxplot of JKN PBI

The distribution of National Health Insurance for Contribution Assistance Recipients (JKN PBI) per year indicates that in 2018, the JKN PBI values exhibited relatively low variability with a narrow interquartile range (IQR) and a median of approximately 29, indicating values more concentrated around the median. In 2019, the variability increased with a wider IQR, while the median rose slightly to 31, indicating a more dispersed value distribution. In 2020, variability further increased, with the IQR widening and the median rising to 35, signifying greater value variation. In 2021, the variability remained high with a wide IQR and the median increasing to 39, showing that the distribution of JKN PBI values continued to vary significantly from year to year. Overall, the variability of JKN PBI values increased from 2018 to 2021, indicating greater variation in the value distribution each year.

Descriptive statistics aim to provide a clear picture of the characteristics of data related to an event [22]. The results of descriptive statistical analysis for the research variables are presented in Table 2.

Table 2. Descriptive Statistics

No.	Variables	Mean	Std. Deviation
1	Participation in National Health Insurance for Premium Assistance Recipients (JKN PBI)	36.53	12.8
2	Income Level	4295	3208.8
3	Number of Health Facilities	761.8	825.9
4	Number of Employed Population	41.02	10.4
5	Education levels	62.86	21.9

Estimation of Panel Data Regression Model

1) Common Effect Model (CEM)

In the estimation of CEM, it is assumed that the model does not consider effects across time and individuals, thus the intercept and slope remain constant across all time periods and individuals [11]. The estimation results of the common effect model can be seen in the following table:

Table 3. Common Effect Model Result

No.	Variables	Estimated Coefficient	p-value
1	*Intercept*	39.673	3.131×10^{-10}
2	Income Level	1.979×10^{-5}	0.669
3	Number of Health Facilities	6.832×10^{-4}	0.598
4	Number of Employed Population	-0.295	0.004
5	Education levels	0.137	5×10^{-4}

Based on the results from the table, the panel data regression model using the common effect estimation is as follows:

$$Y = 39.673 + (1.979 \times 10^{-5})X_{it} - (6.832 \times 10^{-4})X_{2it} - (0.295)X_{3it} + (0.137)X_{4it} + e_{it}$$

From the analysis of the common effect model with a significance level of $\alpha = 5\%$, it is found that the variables significantly influencing participation in the National Health Insurance for the Poor (JKN PBI) are the number of employed population and education levels. The constant term of 39.673 indicates that, without considering the number of employed population and education levels, the participation in JKN PBI is 39.673 individuals. The regression coefficient for the number of employed populations, -0.295, indicates that a one-unit increase in the number of employed populations decreases participation in JKN PBI by 0.295 individuals. Meanwhile, the regression coefficient for education levels, 0.137, shows that a one-unit increase in education levels increases participation in JKN PBI by 0.137 individuals.

2) Fixed Effect Model (FEM)

This model assumes that the slope remains constant, but the intercept varies across time and differs among individuals. The estimated results of the fixed effect model can be seen in the following table:

Table 4. Fixed Effect Model Result

No.	Variables	Estimated Coefficient	p-value
1	Income Level	6.690×10^{-5}	0.592
2	Number of Health Facilities	0.048	0.008
3	Number of Employed Population	-0.667	3.92×10^{-7}
4	Education levels	0.088	6.48×10^{-10}
5	Province of Aceh	64.35	3.04×10^{-6}
..	…	…	…
39	Province of North Sumatra	-19.51	0.444

Based on the results from the table, the panel data regression model using fixed effect estimation is as follows:

$$Y = \beta_{0_{it}} + (6.690 \times 10^{-5})X_{it} + (0.048)X_{2_{it}} - (0.667)X_{3_{it}} + (0.088)X_{4_{it}} + 64.35X_{D_1} + \cdots - (19.51X)_{D_{34}}$$

Based on the fixed effect model with a significance level of $\alpha = 5\%$, the variables influencing participation in the National Health Insurance for the Poor (JKN PBI) are the number of health facilities, the number of employed populations, and education levels. The regression coefficients indicate that each additional health facility increases JKN PBI participation by 0.048 individuals, each additional employed person decreases participation by 0.667 individuals, and each unit increase in education levels increases participation by 0.088 individuals. Dummy variables D_1 to D_{34} capture differences in intercepts across provinces regarding JKN PBI participation.

3) Random Effect Model (REM)

In the random effects model (REM) estimation is used to estimate panel data where there are differences in intercepts and slopes due to errors. The estimated results of this random effects model can be seen in the following table:

Table 5. Random Effect Model Result

No.	Variables	Estimated Coefficient	p-value
1	*Intercept*	50.459	2.2×10^{-16}
2	Income Level	9.164×10^{-5}	0.154
3	Number of Health Facilities	5.337×10^{-4}	0.839
4	Number of Employed Population	-0.593	2.739×10^{-7}
5	Education levels	0.096	3.956×10^{-14}

Based on the results from the table, the panel data regression model using random effects estimation is as follows:

$$Y = 50.459 + (9.164 \times 10^{-5})X_{it} + (5.337 \times 10^{-4})X_{2it} - (0.593)X_{3_{it}} + (0.096)X_{4_{it}} + e_{it}$$

From the equation of the random effects model with a significance level of $\alpha = 5\%$, it is known that the variables influencing participation in the National Health Insurance for the Poor (JKN PBI) are the number of employed population and education levels. The constant term of 50.459 indicates that without considering the effects of the number of employed population and education levels, participation in JKN PBI is 50.459 individuals. The regression coefficient of -0.593 for the employed population suggests that each additional unit increase in the number of employed populations decreases participation in JKN PBI by 0.593 individuals. Meanwhile, the regression coefficient of 0.0963 for education levels indicates that each additional unit increase in education levels increases participation in JKN PBI by 0.0963 individuals.

Regression Model Selection

1. Chow Test

The Chow test is used to determine which estimation method is better to use between the Common Effect Model (CEM) or Fixed Effect Model (FEM). Here are the results of the Chow test:

Table 6. Results of the Chow Test

Chow Test	
p-value	2.2×10^{-16}

Hypothesis:
H_0 : Common Effect Model
H_1 : Fixed Effect Model

With a significance level of $\alpha = 0.05$, the obtained p-value is 2.2×10^{-16}. Based on the Chow test, the decision is to reject H_0 because the p-value of $2.2 \times 10^{-16} < \alpha(0.05)$. From these results, it can be concluded that the fixed effect model is better than the common effect model.

2. Hausman Test

The Hausman test is used to determine which estimation method is better to use between the Fixed Effect Model (FEM) or Random Effects Model (REM). Here are the results of the Hausman test:

Table 7. Results of the Hausman Test

Hausman Test	
p-value	0.052

Hypothesis:
H_0 : Random Effect Model
H_1 : Fixed Effect Model

Using a significance level of $\alpha = 0.05$, the obtained p-value is 0.052. Based on the Hausman test, the decision is to accept H_0 because the $p - value\ (0.052) > \alpha = 0.05$. Therefore, it is concluded that the Random Effects Model (REM) is better than the Fixed Effect Model (FEM). From these results, it can be inferred that the most suitable model for analyzing participation in the National Health Insurance for the Poor (JKN PBI) data is the random effects model.

Significance Test of Parameters

1. Simultaneous Test

The F-test is used to evaluate the overall influence of predictor variables on the response variable. Here are the results of the simultaneous test:

Table 8. Results of the Simultaneous Test

Simultaneous Test	
p-value	2.2×10^{-16}

Hypothesis:
$H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4$
$H_1: \exists\ \beta_k \neq \beta_{k*}\ with\ k = 1, 2, 3, 4$

Based on the results of the simultaneous test, the obtained p-value is 2.2×10^{-16}. Because the p-value of $2.2 \times 10^{-16} < \alpha(0.05)$, H_0 is rejected. From these results, it is concluded that the predictor variables collectively have a significant influence on participation in the National Health Insurance for the Poor (JKN PBI), and the model is deemed appropriate.

2. Partial Test

The partial test (t-test) aims to evaluate the significance of each predictor variable on the response variable individually. Here are the results of the partial test:

Table 9. Results of the Partial Test

No.	Variables	P-value
1	Income Level	0.155
2	Number of Health Facilities	0.839
3	Number of Employed Population	2.739×10^{-7}
4	Education levels	3.956×10^{-14}

Hypothesis:

$H_0: \beta_k = 0$

$H_1: \beta_k \neq 0 \; with \; k = 1, 2, 3,4$

Based on the results of the partial test above, with significance level of $\alpha = 0.05,$ the parameters that significantly influence the model are the number of Employed Population and education levels because their $p - value < \alpha$ lead to rejection of H_0. However, for variables such as income level and health facilities, H_0 is not rejected because $p - values > \alpha$, indicating that these variables do not significantly influence the model.

Classical Assumption Tests

1. Normality Test

For testing the assumption of normality, the Jarque-Bera test was employed. Here are the results of the normality test:

Table 10. Results of the Normality Test

	Normality Test
P-value	0.099

Hypothesis:

H_0: $\varepsilon_{it} \sim N(0, \sigma^2)$ (residuals are normally distributed)

H_1: residuals are not normally distributed

Based on the results of the Jarque-Bera test for normality, the obtained p-value is 0.099. Since $p - value$ (0.099) $> \alpha(0.05)$, we have failed to reject H_0. Therefore, it can be concluded that the residuals produced by the model follow a normal distribution.

2. Multicollinearity Test

For testing multicollinearity, the Variance Inflation Factors (VIF) method was used, where values should ideally be less than 10. Here are the results of the multicollinearity test:

Table 11. VIF Values Result

No.	Variables	VIF
1	Income Level	1.168
2	Number of Health Facilities	1.004
3	Number of Employed Population	1.204
4	Education levels	1.104

Based on the results from the table, the VIF values for each variable used are less than 10. Therefore, it can be concluded that there is no multicollinearity in this regression model, and the assumption of multicollinearity is satisfied.

3. Heteroskedasticity Test

The heteroskedasticity test is used to determine whether there is inconsistency in the residual variation within the formed regression model. This test is conducted using the Breusch-Pagan test. Here are the results:

Table 12. Heteroskedasticity Test Results

Heteroskedasticity Test	
p-value	0.691

Hypothesis:
H_0: There is no issue of heteroskedasticity
H_1: There is an issue of heteroskedasticity

Based on the results of the Breusch-Pagan test, the obtained p-value is 0.691. Since $p-value\ (0.691) > \alpha(0.05)$ there is not enough evidence to reject H_0. Therefore, it can be concluded that the residuals produced have constant variance across all levels of predictor variables, and the assumption of homoskedasticity is satisfied.

4. Autocorrelation Test

The autocorrelation test aims to detect correlations between values at specific time periods and values in previous time periods. This test uses the Durbin-Watson test. Here are the results

Table 13. Autocorrelation Test Results

Autocorrelation Test	
P-value	0.112

Hypothesis:
$H_0: \rho = 0$ (No autocorrelation)$H_1: \rho \neq 0$ (There is autocorrelation)

Based on the results obtained, the p-value is 0.112. Because the $p - value(0.112) > \alpha(0.05)$, there is not enough evidence to reject H_0 . Therefore, it can be concluded that there is no correlation between observations in the data, and the assumption of no autocorrelation is satisfied.

4. DISCUSSIONS

The number of Employed Population (X_3) significantly influences the enrollment in the National Health Insurance (JKN PBI) in Indonesia. The coefficient is negative, indicating that an increase in the Employed Population reduces the users of the National Health Insurance provided to premium assistance recipients. This finding aligns with previous studies [8] - [9], which identified employment as a factor affecting health insurance participation. Education levels (X_4) have a positive impact on JKN PBI enrollment in Indonesia. A positive coefficient for education levels implies that an increase in the educated population enhances JKN PBI participation. This is consistent with Darmayanti and Raharjo's findings that education levels correlate with community participation in JKN [9].

In the random effect model obtained, the random intercept indicates that without the influence of independent variables, the estimated JKN PBI enrollment is 50.459 individuals, but this value can vary across individuals and over time. Each coefficient for the independent variables (income level, number of healthcare facilities, number of Employed Population, and education levels) is also random and can differ between individuals and over time. For example, the coefficient for the Employed Population (X_3) (-0.593) and education levels (X_4) (0.096) demonstrates how the effect of these variables can vary stochastically. With this random effect model, we understand that both intercept and regression coefficients can vary depending on the individuals studied and the time range chosen.

The Random Effect Model (REM) has provided crucial insights into the factors influencing JKN PBI participation. The results indicate that the number of Employed Population (X_3) and education levels (X_4) are significant factors affecting participation in the JKN PBI program. Policies aimed at improving education and focusing on unemployed population groups can help increase participation in this program, thereby expanding the coverage of national health insurance.

5. CONCLUSION

Based on the analysis in this study, it can be concluded that using the panel data regression equation, the most suitable model for identifying factors influencing participation in the National Health Insurance for the Poor (JKN PBI) is the Random Effects Model (REM). The analysis results also indicate the formation of the panel data regression model $Y = 50.459 + (9.164 \times 10^{-5})X_{it} + (5.337 \times 10^{-4})X_{2it} - (0.593)X_{3_{it}} + (0.096)X_{4_{it}} + e_{it}$. Factors influencing participation in the National Health Insurance for the Poor (JKN PBI) with a significance level $\alpha = 5\%$ are the number of Employed Population (X_3) and education levels (X_4) because $p - value < \alpha$. From the simultaneous test, the obtained $p - value\ (2.22 \times 10^{-16}) < \alpha(0.05)$ indicates that both variables, the number of employed population and education levels, collectively have a significant influence on participation in JKN PBI, and the model is considered appropriate. In the partial test, the significant parameters influencing the model are the number of employed population and education levels. Further research remains open to include other analytical factors that could affect participation in JKN PBI, such as poverty, unemployment, healthcare service quality, and the use of other estimation methods.

6. ACKNOWLEDGMENTS

The author expresses gratitude to Universitas Negeri Jakarta for providing the support and assistance needed to complete this research.

7. REFERENCES

[1] P. S. Mustafa, "Peran Pendidikan Jasmani untuk Mewujudkan Tujuan Pendidikan Nasional," *Jurnal Ilmiah Wahana Pendidikan*, vol. 8, no. 9, pp. 68–80, 2022, doi: 10.5281/zenodo.6629984.

[2] F. Irhamsyah, "Sustainable Development Goals (SDGs) dan Dampaknya Bagi Ketahanan Nasional," Jun. 2019.

[3] Y. Syofyan and D. Gusman, "Studi Perbandingan Sistem Jaminan Sosial Antara Indonesia Dan Malaysia Dalam Rangka Pemenuhan Hak Kesehatan di Indonesia," *UNES Journal of Swara Justisia*, vol. 7, no. 1, pp. 208–219, 2023, doi: 10.31933/ujsj. v7i1.

[4] R. Mardiansyah, "Dinamika Politik Hukum Dalam Pemenuhan Hak Atas Kesehatan Di Indonesia," *Veritas et Justitia*, vol. 4, no. 1, pp. 227–251, Jun. 2018, doi: 10.25123/vej.2918.

[5] A. Dewi and D. Salsabilla, "Sosialisasi Dalam Upaya Peningkatan Pengetahuan Masyarakat Tentang Jaminan Kesehatan Nasional (Jkn) Melalui Badan Penyelenggara Jaminan Sosial (Bpjs) Kesehatan Di Kelurahan Surau Gadang," *Jurnal Abdimas Saintika*, vol. 3, no. 1, pp. 166–170, 2021, [Online]. Available: https://jurnal.syedzasaintika.ac.id

[6] F. Subkhan, "Analisis Faktor-Faktor Yang Mempengaruhi Kepesertaan Program Jaminan Kesehatan Nasional Pada Pekerja Informal Perkotaan," 2018.

[7] Dewan Jaminan Sosial Nasional, "Kepesertaan Jaminan Kesehatan Nasional."

[8] Yosalli and F. Hafidz As Shidieq, "Faktor Yang Mempengaruhi Kepesertaan Jaminan Kesehatan Nasional Di Kabupaten," *Jurnal Manajemen Pelayanan Kesehatan*, vol. 23, pp. 133–139, Dec. 2020.

[9] L. Dwi Darmayanti and B. Budi Raharjo, "Keikutsertaan Masyarakat dalam Jaminan Kesehatan Nasional Mandiri," *Higeia Journal of Public Health Research and Development,* pp. 824–834, 2020, doi: 10.15294/higeia.v4iSpecial%204/38011.

[10] Y. M. Sitorus and L. Yuliana, "Penerapan Regresi Data Panel Pada Analisis Pengaruh Infrastruktur Terhadap Produktivitas Ekonomi Provinsi-Provinsi di Luar Pulau Jawa Tahun 2010-2014," *MEDIA STATISTIKA*, vol. 11, no. 1, pp. 1–15, Sep. 2018, doi: 10.14710/medstat.11.1.1-15.

[11] I. Firman Alamsyah, R. Esra, S. Awalia, and D. Andi Nohe, "Analisis Regresi Data Panel Untuk Mengetahui Faktor yang Memengaruhi Jumlah Penduduk Miskin di Kalimantan Timur," in *Prosiding Seminar Nasional Matematika, Statistika, dan Aplikasinya*, 2022, pp. 254–266.

[12] A. Lestari and Y. Setyawan, "Analisis Regresi Data Panel Untuk Mengetahui Faktor yang Mempengaruhi Belanja Daerah di Provinsi Jawa Tengah," *Jurnal Statistika Industri dan Komputasi*, vol. 2, no. 1, pp. 1–11, 2017.

[13] D. Venosia, Suliyanto, Sediono, and N. Chamidah, “Pemodelan Persentase Kepesertaan Bpjs Non Penerima Bantuan Iuran Dengan Pendekatan Regresi Data Panel,” *J Statistika*, vol. 15, no. 1, pp. 116–126, 2022, [Online]. Available: www.unipasby.ac.id
[14] D. Apriliawan and H. Yasin, “Pemodelan Laju Inflasi di Provinsi Jawa Tengah Menggunakan Regresi Data Panel,” 2013. [Online]. Available: http://ejournal-s1.undip.ac.id/index.php/gaussian
[15] S. Pangestika, “Analisis Estimasi Model Regresi Data Panel Dengan Pendekatan Common Effect Model (CEM), Fixed Effect Model (FEM), dan Random Effect Model (REM),” 2015.
[16] G. Eka Kusumawardhani and V. Maya Santi, “Analisis Survival dengan Model Regresi pada Data Tersensor Berdistribusi Log-logistik,” *Jurnal Statistika dan Aplikasinya (JSA)*, vol. 2, no. 2, 2018.
[17] W. Rahayu and V. Maya Santi, “Penerapan Imputasi Ganda dengan Metode Predictive Mean Matching (PMM) untuk Pendugaan Data Hilang Pada Model Regresi Linear,” 2018.
[18] N. Astriawati, “Penerapan Analisis Regresi Linier Berganda Untuk Menentukan Pengaruh Pelayanan Pendidikan Terhdap Efektivitas Belajar Taruna di Akademi Maritim Yogyakarta,” *Jurnal Ilmu-Ilmu Kemaritiman, Manajemen, dan Transportasi*, pp. 22–37, 2016.
[19] V. M. Santi, A. N. Mutia, and Q. Meidianingsih, “Geographically Weighted Regression dalam Menganalisis Faktor-Faktor yang Mempengaruhi Kasus Tuberkulosis di Sumatera Utara,” *Sainmatika: Jurnal Ilmiah Matematika dan Ilmu Pengetahuan Alam*, pp. 107–116, Dec. 2022, doi: 10.31851/sainmatika. v19i2.9020.
[20] V. M. Santi, D. Ambarwati, and B. Sumargo, “Zero Inflated Poisson Regression Analysis in Maternal Death Cases on Java Island,” *Pattimura International Journal of Mathematics (PIJMath)*, vol. 1, no. 2, pp. 59–68, Nov. 2022, doi: 10.30598/pijmathvol1iss2pp59-68.
[21] V. M. Santi and Y. Rahayuningsih, “Negative Binomial Regression in Overcoming Overdispersion in Extreme Poverty Data in Indonesia,” *Pattimura International Journal of Mathematics (PIJMath)*, vol. 2, no. 2, pp. 43–52, Nov. 2023, doi: 10.30598/pijmathvol2iss2pp43-52.
[22] N. Ruwah Ibnatur Husnul *et al.*, *Statistik Deskriptif*. Tangerang Selatan: Unpam Press, 2020. [Online]. Available: www.unpam.ac.id