ISSN 2808-2605 EISSN 2808-4497 *Forum Analisis Statistik*
Desember 2022, 2 (2): 92-103

Open Unemployment Rate Modeling in The Province of Jawa Timur in 2019-2021 using Data Panel Regression Method

Clarissa Putri Candraningtyas[1*], Agus Budhi Santosa[2]

[1,2]Departemen Statistika, Fakultas Sains dan Analitika Data, Institut Teknologi Sepuluh November,
Jl. Teknik Mesin No.175, Keputih, Kec. Sukolilo, Kota Surabaya, Jawa Timur;
*Penulis korespondensi. *e-mail*: clarissaputri793@gmail.com
(Diterima: 27 Desember 2022; Disetujui: 16 Januari 2023)

ABSTRACT

Unemployment is an essential indicator in the employment sector, where the unemployment rate could measure to what extent the existing field jobs can absorb the labor force. Based on BPS's data, the total population in Jawa Timur in 2019 was 39,70 million people, the second most populated province in Indonesia after West Java. This number increases every year until 2021 to 40,88 million people. This increase resulted in a greater labor force than demand, so unemployment emerged. Therefore, further analysis of factors affecting Jawa Timur's open unemployment rate in 2019-2021 is needed. This study will use a data panel regression method which combines time series data and cross-section. Besides getting a panel regression model of the open unemployment rate, the characteristics of the open unemployment rate and factors that are suspected to affect it will be analyzed. The estimation results from this panel regression give the best model, the Fixed Effect Model (FEM). The significant variable is the Economic Growth Rate, Minimum Wages of District/City, and Total Population, with a coefficient of determination of 92.90 percent.

Keywords: fixed effect model, jawa timur, panel regression, open unemployment rate

INTRODUCTION

Economic growth is one of the indicators used to assess the performance of an economy, especially parameters that reflect the results of a region's economic development. Economic growth in an area can show improvement, indicating that the regional economy is doing well (Amir, 2007).

Jawa Timur is one of the developing regions where socioeconomic problems are still complex problem to solve. Among the many problems that exist are related to the rapid growth of the population, namely the imbalance between job growth and the number of workers that increases yearly. Based on BPS data, Java Timur had a total population of 39.7 million in 2019, making it the second largest in Indonesia after West Java. This number will continue to increase yearly until it reaches 40.88 million in 2021 (BPS, 2021). This increase will lead to labour oversupply and unemployment (Pramusinto & Daerobi, 2019).

In the field of employment, unemployment can be a reference in employment. Unemployment is caused by many factors, including economic conditions, low levels of education, inequality between employment and the number of job seekers, government policies that do not match the needs of society, and the gap between the skills of job seekers and the labor market. (Atriyan, 2013).

To overcome the problem of unemployment, the Jawa Timur government has made many efforts, including preparing quality human resources (HR), preparing appropriate and accurate job vacancies for the community, and attracting investors to develop the economy in Indonesia. These efforts have

been able to lower the unemployment rate and increase the absorption of the labor force in the company over the past few years. However, in 2020, it was noted that the unemployment rate showed an increase of 5.84 percent compared to 2019 of 3.82 percent (BPS, 2021). Therefore, to reduce the unemployment rate, which had increased due to Covid-19, the government began to develop strategies for labor productivity because it saw potential labor opportunities in Jawa Timur increase yearly.

According to previous research, several factors affect the open unemployment rate, namely the labor force and education level. (Prayogo, 2020). Another study using the Classification and Regression Tree (CART) method found that the factors that affect the open unemployment rate include the average length of schooling, poverty, and school participation rates. (APS) (Yulistiani, 2020). In addition, a study was conducted using regression panel data obtained factors affecting the open unemployment rate, including economic growth, minimum wage, and an average length of education (Cahyo, 2016).

This study will analyze the factors influencing the Open Unemployment Rate of Jawa Timur Province 2019-2021. Judging from the increasing unemployment rate, it is suspected that there is a time effect in the calculation. Therefore, the panel regression method was used in this study, a combination of cross-section and time series data. In addition to obtaining a regression model of the Open Unemployment Rate panel, TPT characteristics and factors suspected to have an influence will also be analyzed.

METHODOLOGY

Descriptive Statistics

Descriptive statistics will provide an overview of data concentration, variability in data, and the general nature of data distribution. Besides being able to be displayed in numeric, descriptive statistics can also be visualized with graphs (W. R. E., 2007). To obtain a more accurate description of the data, it is necessary to obtain numerical values for the location or data center and the amount of variability that exists (Johson, 2010).

Panel Data Regression

Panel data regression model is a regression model obtained from panel data, namely data consisting of a combination of cross section and time series data (Gujarati, 2004). Cross section data is collected in a shared time for many objects; objects here can be individuals or groups. At the same time, time series data is data that has been collected within a certain period of time on an object. The advantage of panel data is that it allows much flexibility in a specific time span, so the panel data regression equation is written in the form of the following equation 1 (Greene, 2003).

$$Y_{it} = \alpha_{it} + \beta' X_{it} + \varepsilon_{it} \quad (1)$$

Description:

$$i = 1,2,\ldots,n;\ t = 1,2,\ldots,T$$

Y_{it}: vector of the dependent variable of the i-th individual unit in the t-th time period

α_{it}: vector of the intercept coefficient for each i-th individual in the t-th time period

β': $(\beta_1, \beta_2, \ldots, \beta_k)$ The slope coefficient vector measures 1 x k, where k is the number of independent variables in the observation

X_{it}: observation matrix of independent variables of the i-th individual and the t-th time period

ε_{it}: residual vector of the i-th individual in the t-th time period by satisfying the assumptions ε_{it} $IIDN\ (0, \sigma^2)$

The advantages of using panel data include the following (Gujarati, 2004).

1. The data panel can provide more data to provide complete information and obtain a greater degree of freedom (df) for better estimation results.

2. Combining information from time series and cross section data can solve problems with omitted variables.
3. Panel data helps reduce the collinearity between variables.
4. Panel data is better at detecting and measuring effects that purely and pure cross section data cannot achieve.
5. Can test and build more complex behavioral models such as economies of scale and technological change.
6. Panel data can minimize the bias generated by individual aggregates as more data is observed

Panel Data Regression Model Estimation

In panel data analysis, there are three-panel data models, namely the Common Effect Model (CEM), Fixed Effect Model (FEM), and Random Effect Model (REM).

The Common Effect Model (CEM) is a model with an estimation of multiple linear regression models in general. This model approach is carried out by regressing all the combined data without considering the effects of time and individuals. The general equation for the regression model of this approach is shown in equation 2 (Gujarati, 2004).

$$Y_{it} = \alpha_{it} + \beta' X_{it} + \varepsilon_{it} \quad (2)$$

In the CEM estimation model, the parameter approach used is Ordinary Least Square (OLS) to estimate the panel model.

The FEM model assumes that the intercept values of cross section or time series units are different but with a fixed slope between individuals and time. The FEM model equation is written as follows (Greene, 2004).

$$Y_{it} = \alpha + \alpha_{it} + \beta' X_{it} + \varepsilon_{it} \quad (3)$$

The REM model involves a correlation between error terms due to changes in time and observation units. The REM model uses estimation with the Generalized Least Square (GLS) technique. The general equation of the REM model can be written in equation 4 (Gujarati, 2004).

$$Y_{it} = \alpha_i + \beta' X_{it} + w_{it} \quad (4)$$

With the w_{it} value in equation 5.

$$w_{it} = \varepsilon_{it} + \mu_i \quad (5)$$

Testing and Selection of Panel Data Regression Models

Using hypothesis testing, a best model determination is used to determine a good estimation model. Some hypothesis tests used to determine the best model are the Chow Test and the Haussman Test. The Chow test was conducted to select the best estimate between CEM and FEM; the following is the hypothesis.

$H_0: \alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$ (The corresponding model is CEM)

H_1: There is at least one $\alpha_i \neq 0;\ i = 1, 2, \ldots, n$ (The corresponding model is FEM)

The statistics of the Chow Test are as follows.

$$F_{count} = \frac{(R^2_{LSDV} - R^2_{Pooled})/(n-1)}{(1-R^2_{LSDV})/(nT - n - K)} \quad (6)$$

Description:

R^2_{LSDV} : *R-square for Fixed Effect Model (FEM)*

R^2_{Pooled}: *R-square* for *Common Effect Model (CEM)*

n : Number of Cross Section Units

T : Number of Time Series Units

K : number of independent variables

Areas of rejection H_0, if $F_{count} > F_{\alpha;(n-1,nT-n-K)}$, then in estimating the regression equation, can use the FEM model. Then a Haussman Test will be carried out to select the FEM or REM model as a model that is suitable for use. Here is the hypothesis of the Haussman Test.

$H_0: corr(X_{it}, \varepsilon_i) = 0$ (The corresponding model is CEM)

$H_1: corr(X_{it}, \varepsilon_i) \neq 0$ (The corresponding model is FEM)

Haussman Test statistics are written in the following equation.

$$W = [(\hat{\beta}_{FEM} - \hat{\beta}_{REM})]'[var(\hat{\beta}_{FEM}) - var(\hat{\beta}_{REM})]^{-1} - [(\hat{\beta}_{FEM} - \hat{\beta}_{REM})] \quad (7)$$

Description:

$\hat{\beta}_{FEM}$: slope estimation vector FEM

$\hat{\beta}_{REM}$: v slope estimation vector REM

Reject H_0 if the value is $W > X^2_{(k,\alpha)}$ then the corresponding model is FEM (Fixed Effect Model); otherwise, if it fails to reject, the suitable model is REM.

Meanwhile, when conducting the Chow Test and Haussman Test, the decision was to fail to reject H_0 the Lagrange Multiplier test was carried out to compare the CEM and REM models. The hypotheses in the LM test are as follows.

$H_0: \sigma^2_1 = \sigma^2_2 = \ldots = \sigma^2_n = \sigma^2$

$H_1: \sigma^2_i \neq \sigma^2;\ i = 1, 2, \ldots, n$

The test statistics on the Lagrange Multiplier Test are as follows.

$$LM = \frac{nT}{2(T-1)}\left[\frac{\sum_{i=1}^{n} (T\underline{e}_i)^2}{\sum_{i=1}^{n} \sum_{t=1}^{T} e^2_{it}} - 1\right]^2 \quad (8)$$

Description:

n : Number of cross section units

T : Number of time series units

e_{it}: *Cross section and time series errors*

n : Number of cross section units

If the value $LM > X^2_{(n-1,\alpha)}$ then H_0 rejected, which also means that the best model is the Random Effect Model (REM). It also means that FEM has indications of heteroskedasticity that the estimation of the white method can overcome.

Regression Parameter Testing

Regression parameter tests determine the relationship between independent and dependent variables. Regression parameter testing is performed using simultaneous tests and partial tests.

Simultaneous tests were conducted to determine the influence of all independent variables on the dependent variables with the following hypothesis (Drapper & Smith, 1992).

$H_0: \beta_1 = \beta_2 = \ldots = \beta_k$

H_1: there is at least 1 $\beta_k \neq 0$

Test Statistics:

$$F_{count} = \frac{MS_{regression}}{MS_{residual}} \quad (9)$$

Reject H_0 if $F_{hitung} > F_{tabel}(F_{\alpha;(k-1,n-k)})$, where n is the number of observations and k is the number of parameters.

Partial tests determine the effect of independent variables on individual dependent variables. The hypotheses used for partial tests are as follows.

$H_0: \beta_k$

$H_1: \beta_k \neq 0$

Test Statistics:

$$t_{count} = \frac{\widehat{\beta}k}{SE(\widehat{\beta}k)} \quad (10)$$

Reject H_0 if $|t_{hitung}| > t_{tabel}\left(\frac{\alpha}{2}, nt-k-1\right)$ or p-value of < α where n is the number of observations and k is the number of parameters.

Multicollinearity Assumption Test

Multicollinearity is a perfectly linear relationship between some or all predictor variables. This relationship's closeness indicates a linear correlation between related variables (Gujarati, 2014). To detect multicollinearity, we use VIF (Variance Inflation Factor) with the formula in equation 11.

$$VIF_j = \frac{1}{1-R^2{}_j} \quad (11)$$

Where j in a VIF denotes the j-th variable tested for multicollinearity. With $R^2{}_j$ the value of the auxiliary regression coefficient of determination between the j-th predictor variable and the remaining predictor variable (k-1) (Drapper & Smith, 2014). If the value of VIF≤10, then there is no multicollinearity.

Residual Assumption Testing

After obtaining the model, a residual assumption test is carried out, namely a homogeneity test and a normal distribution test on the residual. The testing of identical residual assumptions aims to determine whether the residuals meet identical assumptions, homoskedasticity, or have the same variance. Data is said to be identical if its residual plots spread randomly and do not form a specific pattern.

In this study, the Glejser Test was used to test homogeneity, where this test was carried out by regressing predictor variables and absolute residual values as response variables. Here's the hypothesis used for the Glejser Test.

H_0: $\beta_k = 0$ (Identical residuals or homoskedasticity)

H_1: $\beta_k \neq 0$ (Residuals are not identical or heteroskedasticity)

With $k = 1,2,\ldots,K$

The test statistics used in this test use equation 12. Test Statistics:

$$t_{count} = \frac{\widehat{\beta}k}{SE(\widehat{\beta}k)} \quad (12)$$

Reject H_0 $|t_{count}| > t_{tabel}\left(\frac{\alpha}{2}, nt-k-1\right)$ or p-value < α.

Normality tests are also required for residual assumptions of regression models. Graphically, the distribution of residual data is obtained from the regression model. Here is the hypothesis used for the Kolmogorov-Smirnov Test.

H_0: Normally distributed residuals

H_1: Residuals are not normally distributed

The test statistics used can be written on equation 13.

$$D = Sup_x|F(x) - F_0(x)| \quad (13)$$

Where $F(x)$represents the cumulative distribution function, while $F_0(x)$ is the observed chance function. The decision of the region of refusal for the Kolmogorov-Smirnov test was to reject H_0 if $|D| = q_{1-\alpha}$ with $q_{1-\alpha}$ obtained from the Kolmogorov-Smirnov table or with p-value < α.

Open Unemployment Rate

Open unemployment includes people who are unemployed and looking for work, people who are unemployed and preparing to start a business, people who are unemployed and not looking for work and people who already have a job but have not yet started working. In comparison, the Open Unemployment Rate (TPT) is the percentage of the number of unemployed to the number of the labor force. The formula for calculating the TPT is:

$$TPT = \frac{Unemployment\ Number}{Labor\ Force\ Number} \times 100\% \quad (14)$$

The Open Unemployment Rate (TPT) is used to indicate the percentage of the labor force included in the unemployed. A high tax return may indicate the presence of an unabsorbed labor force in the job market (BPS,2021).

Data Source

The data used in this study is secondary data provided by the Statistics of Jawa Timur (BPS) of Jawa Timur Province. The data obtained are data on the Open Unemployment Rate (TPT) and factors suspected to be influential with district/city research units in Jawa Timur Province from 2019 to 2021.

Variable

The variables in this study consist of response and predictor variables because they follow the analysis to be used, namely regression analysis. The response variable to be used is the Open Unemployment Rate. Variables that allegedly affect the Open Unemployment Rate are referred to as predictor variables. The following are the predictor variables to be used.

Table 1. Variables

Variables	Description
Y	Open Unemployment
X_1	Economic Growth Rate
X_2	District/City Minimum Wage (UMK)
X_3	Average Length of Schooling
X_4	Population

Analysis Steps

In conducting this research, several analytical steps must be done, namely as follows.

1. Conduct literature studies and data collection
2. Analyze the characteristics of the Poor Percentage variable and the factors suspected to be influential.
3. Detecting cases of multicollinearity with VIF test criteria
4. Perform testing to discover the best panel regression model with the following:
   a. Chow test to select CEM or FEM models. If the decision is to reject H_0 then the selected model is FEM and proceeds to the Haussman Test. However, if the decision is Failed to reject H_0 then, the model is selected as CEM and tested for LM.
   b. Haussman test to determine the most appropriate model between FEM and REM. If the decision is to reject H_0 then the selected model is FEM then It is not necessary to perform the Lagrange multiplier test. However, if the decision is Failed to reject H_0 then the selected model is REM and proceeds to the LM Test.
   c. Test the Lagrange Multiplier to determine the most appropriate model for REM or CEM. If the decision is to reject H_0 then the model used is REM. However, if the decision fails to reject then the model used is CEM.

5. Perform panel regression model parameter significance tests simultaneously and partially.
6. Test the residual assumptions identical and normally distributed.
7. Obtain estimates of panel regression models and interpret obtained models.
8. Conclude.

RESULTS AND DISCUSSION

Descriptive Statistics

In 2020, the unemployment rate in Jawa Timur increased, accompanied by a decrease in the economic growth rate. All regions also follow this in Indonesia. Table 2 describes the characteristics of each variable over three years. At the minimum and maximum variables, Kabupaten Sampang and Kota Madiun became the cities with the lowest and highest average school lengths for three consecutive years. Meanwhile, Kota Mojokerto and Kota Surabaya have become the cities with the smallest and largest population for the past three years. The highest economic growth rate in 2019 was in Batu city, but in 2020 Batu City became the city with the lowest economic growth rate reaching -6.46. This figure is considered to have decreased quite drastically due to the Covid-19 pandemic. However, this decline was felt not only by Kota Batu but all regions in Jawa Timur. However, the Covid-19 pandemic has had no impact on the Regency/City Minimum Wage because it is proven that every year it always increases.

Table 2. Characteristics of Research Data

Variable	Mean	Minimum	Maximum
Y	4.93	0.91	10.97
X_1	1.88	-6.46	6.51
X_2	2395190	1763268	4300479
X_3	7.93	4.55	11.37
X_4	10.88	3.81	23.76

Multicollinearity Detection

Multicollinearity is a linear relationship between several or all predictor variables. The strong relationship between these variables suggests a linear correlation between independent variables. Therefore, it is necessary to detect whether independent variables do not affect each other since it will affect the created model.

The following table shows the VIF values of each predictor variable. If the VIF value is more than 10, there will be a case of multicollinearity.

Table 3. VIF Values of Research Variables

Variable	VIF
Economic Growth Rate (X_1)	1.01
District/City Minimum Wage (X_2)	1.93
Average length of schooling (X_3)	1.53
Population (X_4)	1.64

Based on Table 3, the VIF value on all independent Variables is less than 10, so it can be concluded that all independent Variables do not occur in multicollinearity. Thus, all Variables can proceed to the next step.

Regression Model Estimation

Panel regression model estimates are divided into three general models: CEM, FEM, and REM. CEM is the most superficial regression model estimation because it uses all the data regardless of individual effects or time. The CEM estimate can be written in the following equation.

$$Y_{it} = -1.990603 - 0,193368X_{1it} + 0,000000773X_{2it} + 0,632826X_{3it} + 0,000000389X_{4it}$$

Based on the estimates above results, the model's goodness was obtained by 64.43%. This means that in the CEM equation, the predictor variable can explain the percentage figure of the poor population of 64.43%, while other Variables explain the rest outside the model.

FEM is an estimation of a panel regression model that uses dummy variables in its calculations. The FEM estimate can be written in the following equation.

$$Y_{it} = -8.728912 + \alpha_i - 0,090963X_{1it} + 0,00000682X_{2it} + 0,427538X_{3it} + 0,00000555X_{4it}$$

FEM has a model goodness of 91.93%. Another variable outside the model describes the TPT with an estimated FEM of 8.07%.

REM is a panel regression model estimation using the Generalized Least Square (GLS) method with the following estimation result equations.

$$Y_{it} = -2.525211 - 0.167879X_{1it} + 0,00000135X_{2it} + 0,567391X_{3it} + 0,0000000343X_{4it}$$

This REM model has a model goodness value of 55.18%, which can be described by all four Variables, while other Variables explain the rest outside the model. In the three models, there is an equation of negative influence on the Economic Growth Rate Variable (X1) and a positive influence on the Average Length of Schooling Variable (X3) and the Number Population (X4).

Regression Model Selection

A model selection test will be carried out of the three-panel regression models formed by comparing models. The first test performed was the Chow Test. The Chow test chooses between CEM and FEM with the following hypothesis.

$H_0: \alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$ (The corresponding model is CEM)

H_1: There is at least one $\alpha_n \neq 0; i = 1,2,\ldots,n$ (The corresponding model is FEM)

From the results of the Chow Test, a calculated F value of 7.45 and a p-value of 0.000 were obtained. Using 0.05, the decision to reject H_0. Thus, the model selected in the Chow Test is FEM.

Then the Haussman Test was carried out to determine the best model between FEM and REM with the following hypothesis.

$H_0: corr(X_{it}, \varepsilon_i) = 0$ (The corresponding model is CEM)

$H_1: corr(X_{it}, \varepsilon_i) \neq 0$ (The corresponding model is FEM)

The statistical value of the Haussman Test is in the form of a calculated Chi-Square value of 70.02 with a p-value of 0.000. Using 0.05, then the decision Rejects H_0. Thus, the model chosen in the Haussman Test is FEM. The two tests above state that FEM is the best model, so there is no need to proceed to the Lagrange Multiplier (LM) test.

Parameter Significance Test

The parameter significance test is carried out using the Simultaneous Test, which aims to determine the linear relationship between the response and predictor variables. The following is a hypothesis for the Simultaneous test.

$H_0: \beta_1 = \beta_2 = \ldots = \beta_k$

H_1: there is a minimum of 1 $\beta_k \neq 0$ for $k = 1,2,3,\ldots,k$

Based on the parameter estimation results, a calculated F value of 61.38 and a p-value of 0.000 were obtained. By using 0.05, the p-value will be less $\propto$ than or reject H_0. This indicates that simultaneously, the predictor Variable has a significant effect on the response Variable.

Then a partial test was carried out to determine the effect of predictor variables on individual response variables. Variables are said to be insignificant when the p-value$<\propto$. The Average Variable

of School Length is an insignificant Variable to the Open Unemployment Rate. Then elimination will be carried out on these insignificant Variables.

Table 4. FEM Partial Tests

Variable	t-count	*p-value*
X_1	-3.95	0.0002
X_2	6.58	0.0000
X_3	0.52	0.6018
X_4	2.17	0.0332

The results are obtained in the table below when removing the Variable Average Length of Schooling.

Table 5. Second Partial Test of FEM

Variable	t-count	*p-value*
X_1	-4.06	0.0001
X_2	10.41	0.0000
X_4	2.12	0.0377

The FEM model explains that three variables significantly affect the model so that an exact residual assumption test will be carried out.

Residual Assumption Test

The FEM model obtained has an estimated value for the response variable; from the estimation results, it is different from the actual value of the response variable, called the residual. The assumptions that need to be met for residuals are identical and normally distributed. Furthermore, an exact test was carried out using the Glejser Test, which was carried out by regressing the predictor Variable and absolute residual value as the response Variable.

Hypothesis:

H_0: $\beta_k = 0$ (Identical residuals or homoskedasticity)

H_1: $\beta_k \neq 0$ (Residuals are not identical or heteroskedasticity)

Reject H_0 if $|t_{hitung}| > t_{tabel}\left(\frac{\alpha}{2}, nT - k - 1\right)|$ or $p - value < \alpha$. Based on the table below, all Variables fail to reject, so the assumptions for identical residuals have been met.

Table 6. Glejser Test

Variable	t-count	*p-value*
X_1	-4.06	0.0781
X_2	10.41	0.0515
X_4	-2.12	0.0521

The following residual assumption test is the normality test. One of the normality tests is the Kolmogorov Smirnov test, with the following hypothesis.

H_0: Normally distributed residuals

H_1: Residuals are not normally distributed

From the calculation of the value of Kolmogorov Smirnov of 6.46 with a p-value of 0.039. If the α value is 0.05, the decision will be Reject H_0, with a α value of 0.01, the residual can be said to have been distributed normally.

Best Model Interpretation

The best model in the study is FEM. Here is the equation of the obtained model.

$$Y_{it} = -6.609689 + \alpha_i - 0.086608\, X_{it} + 0.00000722\, X_{2it} - 0.00000522\, X_4$$

The model obtained different values for each Regency/City in Jawa Timur, as shown in Table 7. As is known, the Fixed Effect Model (FEM) concept in the model estimation process uses the Least Square Dummy Variable (LSDV) method. LSDV is a method that estimates linear regression parameters and uses OLS in models involving dummy Variables as its predictor Variables for different intercepts in each individual.

The model also explains that the number of inhabitants affects the TPT. When the population increases by 1%, the TPT will increase by 0.00000522. When the economic growth rate increases by 1%, the TPT figure will decrease by 0.086. This is because when the market's demand for goods and services increases, the output will also increase. Increased production can increase employment, which then impacts reducing the unemployment rate.

Theoretically, the minimum wage positively correlates to the TPT because an increase in the minimum wage will raise the TPT. An increase in the minimum wage reduces the demand for labor, while on the labor side, wages are wages received as compensation for labor time. The average variable length of schooling has a positive and insignificant influence on TPT. This is because the longer time to find work for a highly educated workforce is longer than for the poorly educated. In this case, it is the educated labor force that is more selective in choosing jobs; the expected wage rate is higher than the labor force whose education level is lower.

Table 7. District/City Intercepts

No	Jawa Timur Regency/City	α_i	No	Jawa Timur Regency/City	α_i
1	Kabupaten Pacitan	-2.002533	17	Kabupaten Jombang	0.911708
2	Kabupaten Ponorogo	2.286002	18	Kabupaten Nganjuk	3.113433
3	Kabupaten Trenggalek	0.751578	19	Kabupaten Madiun	1.530347
4	Kabupaten Tulungagung	2.800389	20	Kabupaten Magetan	0.280151
5	Kabupaten Blitar	2.753681	21	Kabupaten Ngawi	2.158900
6	Kabupaten Kediri	5.734827	22	Kabupaten Bojonegoro	3.826834
7	Kabupaten Malang	4.137468	23	Kabupaten Tuban	-1.104427
8	Kabupaten Lumajang	1.775365	24	Kabupaten Lamongan	1.057730
9	Kabupaten Jember	7.900838	25	Kabupaten Gresik	-8.885661
10	Kabupaten Banyuwangi	4.211840	26	Kabupaten Bangkalan	5.625539
11	Kabupaten Bondowoso	0.988540	27	Kabupaten Sampang	1.444485
12	Kabupaten Situbondo	0.329963	28	Kabupaten Pamekasan	0.746362
13	Kabupaten Probolinggo	-0.465424	29	Kabupaten Sumenep	1.177836
14	Kabupaten Pasuruan	-8.566941	30	Kota Kediri	-0.780663
15	Kabupaten Sidoarjo	-2.898460	31	Kota Blitar	-0.282774
16	Kabupaten Mojokerto	-11.91998	32	Kota Malang	-0.841856

No	Jawa Timur Regency/City	α_i	No	Jawa Timur Regency/City	α_i
33	Kota Probolinggo	-2.561252	36	Kota Madiun	0.903137
34	Kota Pasuruan	-5.406042	37	Kota Surabaya	0.492192
35	Kota Mojokerto	-4.349042	38	Kota Batu	-6.874191

CONCLUSION AND RECOMMENDATION

The following conclusions can be drawn based on the existing analysis and discussion.

1. Results of descriptive statistics on the Open Unemployment Rate in Jawa Timur in 2020, the unemployment rate had increased due to Covid-19. This has caused the pace of economic growth also to decline. Kota Batu, the city with the highest economic growth in 2020, decreased drastically to reach the number -6.46. In 2021, the unemployment rate in Jawa Timur began to show a decline and an increase in economic growth.
2. The estimation results from the panel regression provide the best model, namely the Fixed Effect Model (FEM), and obtained three significant variables, namely the Economic Growth Rate, District/City Minimum Wage, and Population, with a coefficient of determination of 92.90%.

Advice to the government, especially the Jawa Timur region, to stabilize unemployment in Jawa Timur. Although currently, unemployment is starting to show a decline, it must still be watched out for if at any time it increases. Moreover, looking at the condition of the population that continues to increase every year, this will potentially increase the TPT rate in Jawa Timur. However, this can be minimized by continuing to provide additional skills to make the workforce more competitive in the labor market competition. In addition, the government needs to provide adaptive jobs following the times and technology and follow the era of the industrial revolution 4.0.

REFERENCES

Amir, A. (2007). Pengaruh inflasi dan pertumbuhan ekonomi terhadap pengangguran di Indonesia. Jurnal Inflasi dan Pengangguran, 1(1), 4-9.

Atriyan, S. T. (2013). Analisis Pengaruh Inflasi, Pertumbuhan Ekonomi, dan Investasi terhadap Pengangguran di Provinsi Jawa Tengah.

BPS, "Badan Pusat Statistik," [Online]. Available: https://www.bps.go.id/indicator/12/1886/1/jumlah-penduduk-hasil-proyeksi-menurut-provinsi-dan-jenis-kelamin.html. [Accessed 20 August 2022].

BPS, "Badan Pusat Statistik," 2021. [Online]. Available: https://jatim.bps.go.id/indicator/6/54/1/tingkat-pengangguran-terbuka-tpt-provinsi-jawa-timur.html. [Accessed 2020 August 2022].

BPS, "Badan Pusat Statistik," [Online]. Available: https://www.bps.go.id/subject/6/tenaga-kerja.html. [Accessed 20 August 2022].

Cahyo, Rizky Dwi. "Analisis Pengaruh Pertumbuhan Ekonomi, Upah Minimum, dan Rata-Rata Lama Pendidikan Terhadap Tingkat Pengangguran Terbuka Kabupaten/Kota di Jawa Timur Tahun 2009-2014." Jurnal Ilmiah Mahasiswa FEB 4.2 (2016).

Drapper, N.R. & Smith, H. (1992). Applied Regression Analysis, New York: John Wiley and Sons,.

Drapper, N. R., & Smith, H. (2014). *Analisis Regresi Terapan Edisi Kedua.* Jakarta: Gramedia Pustaka Utama.

Gujarati, D. N., & Porter, D. C. (2004). Basic Econometrics. McGraw-Hill Companies. New York, NY, USA,.

Gujarati, D. N. (2014). *Basic Econometrics (4th ed).* New York: Gramedia Pustaka Utama.

Greene, W.H. (2003). *Econometric Analysis*, 5th ed, New Jersey: Pearson Education.

Johson, G. K. B. R.A. (2010). *Statistics Principles & Methods*, 6th ed, United State: John Willey Sons, Inc, 2010.

Pramusinto, N. D., & Daerobi, A. (2019). Permintaan dan Penawaran Tenaga Kerja serta Pengangguran di Indonesia. Seminar Bisnis Magister Manajemen (SAMBIS-2019).

Prayogo, S. A. (2020). "Analisis Faktor-Faktor yang Mempengaruhi Tingkat Pengangguran Terbuka Kabupaten/Kota di Provinsi Jawa Timur Tahun 2014-2018," Jurnal Ilmiah, pp. 1-2,.

W. R.E. (2007). *Probability & Statistics for Engineers & Scientists*, 8 ed, New Jersey: Prentice Hall.

Yulistiani, R., Putra, N. C., Said, Q., & Ernawati, I. (2020). Klasifikasi dan Prediksi Tingkat Pengangguran Terbuka di Indonesia Menggunakan Metode Classification and Regression Tree (CART). Senamika, 1(1), 123-130.