JIGE 6 (2) (2025) 186-194
JURNAL ILMIAH GLOBAL EDUCATION
ejournal.nusantaraglobal.ac.id/index.php/jige
DOI: https://doi.org/10.55681/jige.v6i2.3675

Application of Zero Inflated Ordered Logit (ZIOL) (Case Study: The Employment Status of the Working-Age Population in Banten Province)

Jessie Reyna Marshiela[1], Vita Ratnasari[1*], Santi Puteri Rahayu[1]
[1] Department of Statistics, Institut Teknologi Sepuluh Nopember, Surabaya 61111, Indonesia
**Corresponding author email: vita_ratna@its.ac.id*

Article Info

Article history:

Received April 10, 2025
Approved May 03, 2025

Keywords:

Banten Province, Employment Status, Zero Inflated Ordered Logit (ZIOL)

ABSTRACT

Unemployment remains a major economic issue in Indonesia, particularly in Banten Province, which has the highest open unemployment rate. Traditional models struggle to capture the zero inflation characteristics in labor force data, where most individuals are employed. This study applies the Zero-Inflated Ordered Logit (ZIOL) model to better analyze labor force status in Banten by distinguishing between genuinely unemployed individuals and those appearing unemployed due to external factors.Using data from the National Labor Force Survey (SAKERNAS) 2023, this study examines the impact of gender, education, residence, job training access, and work experience on employment. The results show that women, individuals with lower education, and those lacking work experience are more likely to be unemployed or underemployed. ZIOL outperforms traditional ordinal logit models in capturing these dynamics.The findings provide insights for policymakers to design more effective employment strategies, particularly in regions facing high unemployment.

How to cite: Marshiela, J. R., Ratnasari, V., & Rahayu, S. P. (2025). Application of Zero Inflated Ordered Logit (ZIOL) (Case Study: The Employment Status Of The Working-Age Population In Banten Province) . *Jurnal Ilmiah Global Education*, *6*(2), 186–194. https://doi.org/10.55681/jige.v6i2.3675

INTRODUCTION

Employment is a fundamental issue in a country's economic development, especially for developing countries such as Indonesia, which has a large population. Every year, the number of labor force continues to increase along with population growth, but the availability of employment opportunities is not always able to absorb the existing workforce. This gap between the number of workers and employment opportunities causes an increase in the unemployment rate, which has an impact on various social and economic aspects. Unemployment not only hampers economic growth but also reduces community welfare, increases poverty, and widens social inequality. According to Yanthiani, the imbalance between job seekers and job availability, as well as the mismatch between job seekers' competencies and market needs, are among the main causes of unemployment in Indonesia(Yanthiani, 2023).

Furthermore, structural transformation plays an important role in the creation of decent jobs. Research by Sander and Yoong highlights that labor productivity growth in Indonesia is more influenced by productivity improvements within certain sectors rather than by labor shifts to more productive sectors (Yoong & Gil Sander, 2020). Additionally, policies aimed at improving labor productivity and workforce skills are essential for addressing unemployment issues. The Asian Development Bank (ADB)emphasizes the importance of enhancing productivity through quality jobs by ensuring that skilled workers are available in line with market demands while guiding low-skilled workers into other productive activities (Allen et al., 2018).

Indonesia continues to grapple with significant employment challenges, with the open unemployment rate reaching 5.2% in 2024, marking the highest rate among ASEAN countries. This concerning trend highlights the persistent structural issues within Indonesia's labor market, which is characterized by an imbalance between the rapid expansion of the labor force and the limited growth of available job opportunities. The country's reliance on capital-intensive industries further exacerbates this issue, as these sectors require advanced machinery and automation, thereby reducing the demand for labor. Additionally, technological advancements and digitalization, while beneficial for economic efficiency, have led to the displacement of low-skilled workers, contributing to the rising unemployment rate. One of the regions most affected by this trend is Banten Province, which has recorded the highest open unemployment rate in Indonesia at 7.97% (Pardosi & Septriani, 2023). Despite its strategic proximity to the capital city, Jakarta, and its role as an industrial hub, Banten faces substantial employment challenges due to the predominance of industries that are less labor-intensive. The growing labor force in the region surpasses the number of new jobs created each year, creating a surplus of job seekers who struggle to find employment that matches their skills and qualifications. A study conducted by Ayu (2024) using panel data regression analysis revealed that several factors contribute to the high unemployment rate in Banten, including economic fluctuations, population growth, and a skills mismatch between workers and industry demands. These findings indicate that without adequate intervention, the labor market in Banten will continue to experience difficulties in absorbing its expanding workforce (Annisa Qurrota Ayu et al., 2024).

Economic instability and the low level of entrepreneurship have further aggravated the unemployment situation in Indonesia. Research conducted by Haya, Lestari, and Crisanty (2023) found that the high unemployment rate in Banten is influenced by several macroeconomic factors, including inflation, gross regional domestic product (GRDP), and labor force participation rate (Aqilla Haya et al., 2023).The Indonesian government has set a target of reducing the open unemployment rate in the 2020-2024 National Medium-Term Development Plan (RPJMN), which is in the range of 3.6-4.3%. However, achieving this target still faces many challenges, especially in areas with high unemployment rates such as Banten (Presiden Republik Indonesia, 2020). Based on BPS data, the number of labor force in Banten in August 2024 reached 6.21 million people, an increase compared to the previous year which amounted to 5.97 million people. Meanwhile, the Labor Force Participation Rate (TPAK) also increased by 1.73 percentage points. This increase indicates an increase in the number of job seekers who are not necessarily absorbed in the world of work.

In the context of the labor force, data often exhibits zero inflation characteristics, where most individuals are employed, leading to an excess of zero values in unemployment data. Traditional models struggle to capture this pattern, making the Zero-Inflated Ordered Logit

(ZIOL) method a more suitable alternative for analyzing unemployment in Indonesia, particularly in Banten Province. This approach distinguishes between individuals who are truly unemployed and those appearing unemployed due to external factors, ensuring more accurate estimations. Previous studies have successfully applied similar models in various fields. The Zero-Inflated Ordered Probit (ZIOP) model to analyze sports participation in Spain, showing that economic, social, and demographic factors significantly influence participation (Downward et al., 2011). Applied ZIOP to examine the impact of curbs on single-vehicle crash injuries, finding that while curbs increase crash likelihood, they reduce the probability of severe or fatal injuries (Jiang et al., 2013).Further developed a zero-inflated model for ordinal data to improve the accuracy of analyzing excess zero responses, particularly in public health studies (Kelley & Anderson, 2008).By adopting a similar approach, this study aims to identify key factors influencing labor force status and provide insights for more effective employment policies in Indonesia. The ZIOL model enables a better understanding of unemployment patterns by addressing excess zero values, allowing for more precise workforce strategies and policy recommendations (Stata, n.d.).

METHODS

The In this study, Zero Inflated Ordered Logit (ZIOL) is applied to labor force data in Banten Province. The data source used in this study is secondary data. The data was obtained from the results of the Survei Angkatan Kerja Nasional (SAKERNAS) in August 2023 conducted by the Badan Pusat Statistik (BPS) of Banten Province.

Table 1. Data

Variabel	Keterangan	Skala Data
Y	Labor Force Status	Ordinal
X_1	Gender	Nominal
X_2	Highest Education	Ordinal
X_3	Place of Residence	Nominal
X_4	Access to Job Training	Nominal
X_5	Work Experience	Nominal

The analysis steps that will be used are as follows.

1. Prepare individual data on the labor force from the August SAKERNAS data collection in 2023 Banten Province.
2. Identify response variables and predictor variables.
3. Identify the response variable Labor Force Status which is Ordinal scale.
4. Perform data processing and presentation in descriptive statistics.
5. Checking the assumption of free multicollinearity in the predictor variables.
6. Modeling labor force status using Zero Inflated Ordered Logit (ZIOL) regression and Ordinal Logit.
7. Comparison test between Zero Inflated Ordered Logit (ZIOL) and Ordinal Logit models using Vuong Test.
8. Interpreting the best model.

1. Multiconierity

Multicollinearity assumption check is used to determine whether there is a linear relationship or correlation between significant predictor variables in the regression model. In ordinal logistic regression analysis, multicollinearity is not allowed to determine whether there is a case of multicollinearity, it is necessary to check using VIF (Variance Inflation Factor). If the VIF value is greater than 10, it indicates that there is a case of multicollinearity. The equation of VIF can be written as follows(Hosmer and & Lemeshow, 2000).

2. Binary Logistics

The binary logit model is a data analysis method used to find the relationship between response variables (y) that are binary or dichotomous with predictor variables (x) that are polycotomous.The outcome of the response variable y consists of 2 categories namely "success" and "failure". where, Y=1 for success and Y=0 for failure(Hosmer and & Lemeshow, 2000). The logistic regression function can be written in equation (1).

$$\pi(x) = \frac{e^{-\mathbf{x}^T\boldsymbol{\beta}}}{1 + e^{-\mathbf{x}^T\boldsymbol{\beta}}}, \qquad (1)$$

3. Ordinal Logistics

Logistic regression is an analytical method used to determine the relationship between response variables and predictor variables, where the response variable is polycotomous and each category has its level(Hosmer and & Lemeshow, 2000). The ordinal logistic model in the form of a logit model is as follows in equation (2).

$$\text{Logit}\,[P(Y \le j|\mathbf{x})] = \ln\left(\frac{P(Y\le j|\mathbf{x})}{1-P(Y\le j|\mathbf{x})}\right) = \mu_j + \mathbf{x}^T\boldsymbol{\delta} = \frac{\exp(\mu_j+\mathbf{x}^T\boldsymbol{\delta})}{1+\exp(\mu_j+\mathbf{x}^T\boldsymbol{\delta})} \qquad (2)$$

Significance Testing of Ordinal Logit Parameters

1. Simultaneous Significance Test

The hypothesis used is as follows.

$H_0: \delta_1 = \delta_2 = \cdots = \delta_p = 0$

H_1 : There is at least one $\delta_t \neq 0; t = 1,2,\ldots,p$

2. Partial Significance Test

The hypothesis used is as follows.

$H_0: \delta_k = 0$

$H_1: \delta_k \neq 0; k = 1,2,\ldots,p$

3. Zero Inflated Ordered Logit (ZIOL)

The Zero Inflated Ordered Logit (ZIOL) model is a development model of the ordinal logit regression model. This model has an inflated variable, which is a binary variable that has the possibility of the observation experiencing zero-inflation. ZIOL is expected to overcome zero-inflated conditions and provide more accuracy than the ordinal logistic regression model alone by modeling inflated variables as binary variables, this model is a combination of binary logit and ordinal logit and aims to analyze data with an excessive number of zero categories and has a level or order on the response variable(Stata, n.d.).

4. Vuong Test

In this study, the Vuong test will be used to compare the two ZIOL models with Ordinal Logit. The Vuong test is used to compare the two models. The hypothesis used in the comparison of the two models(Rejeki et al., 2024).

H_0: ZIOL model is as good as Ordinal Logit

H_1: ZIOL model is different from Ordinal Logit model

The Vuong test is recommended in testing the comparison between the zero inflated model and the conventional model. The Vuong Statistic equation, as illustrated in (3).

$$v = \frac{\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} m_i\right)}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(m_i - \bar{m})^2}}, \quad (3)$$

where n is the sample size, Vuong's statistic, v is the asymptotic standard normal distribution (Harris & Zhao, 2007).

5. Interpretation

Model interpretation is a form of definition of response change units caused by predictor variables and determining the functional relationship between response variables and predictor variables. To make it easier to interpret the model, the odds ratio value is used. Odds Ratio is a set of odds divided by other odds(Agresti, 2002).

RESULTS AND DISCUSSION

Variance Inflation Factor (VIF) is used to detect the presence of multicollinearity in regression. Multicollinearity occurs if there is a strong relationship between the independent variables, which can cause the regression coefficient estimates to be unstable. Table 1 presents the VIF value of each variable which has a value <10 which indicates that the data used has fulfilled the multiconierity assumption.

Table 2. Multicollinearity Detection

Variable	VIF
X1	1.03
X2	1.22
X3	1.09
X4	1.14
X5	1.03

Table 1 shows that all variables in this study have VIF values in the range of 1.03 - 1.22, this indicates that there is no multicollinearity problem in the model. This means that the independent variables in the regression model are relatively free from each other and do not significantly affect each other. Multicollinearity testing is one of the classical assumption tests used to determine whether there is a strong correlation among the independent variables in a regression model. In this study, multicollinearity was assessed using tolerance and Variance Inflation Factor (VIF) indicators. A VIF value below 10 suggests that multicollinearity is not an issue. Based on the results, all predictor variables had VIF values under 10, indicating that they are statistically reliable and appropriate to include in the model. In this study, the status of the labor force is divided into 3 parts, Employed, Underemployed, and Open Unemployment. The division of this variable is based on the number of hours worked by individuals as follows:

- Employed (t = ≥35 hours/week)
- Underemployed (0 < t < 35 hours/week)
- Open Unemployment (t = 0 hours/week)

Where t represents time.

Table 2. Results of ZIOL Parameter Estimation

Variable predictor		**Coefficient**	**Std. err.**	**P-Value**
Y				
X1	Female	0.756	0.077	0
X2				
	elementary school & junior high school	0.319	0.090	0
	≥ Senior High School	0.374	0.100	0
X3	City	0.590	0.072	0
X4	Never	0.302	0.087	0.001
X5	Never	0.819	0.072	0
inflate				
X1	Female	1.352	0.038	0
X2				
	elementary school & junior high school	-0.249	0.114	0.029
	≥ Senior High School	-0.731	0.143	0
X3	City	-0.155	0.112	0.168
X4	Never	0.357	0.055	0
X5	Never	0.637	0.064	0
Constant		-0.422	0.346	0.223
Constant (Y=0)		-1.800	1.728	
Constant (Y=1)		0.868	0.281	

Table 2 above is the result of parameter estimation based on the output of the Stata computer program, it is found that the inflated variables and predictor variables have an effect and have no effect. The Zero-Inflated Ordered Logit (ZIOL) model is a statistical approach used to analyze data with a high proportion of zeros and an outcome variable that is ordinal in nature. In this study, the ZIOL model is applied to examine the status of the labor force, which is divided into three categories: Employed (≥35 hours/week), Underemployed (1–34 hours/week), and Open Unemployment (0 hours/week).

The model consists of two main components. The first is a binary logit model, which is used to estimate the probability that an individual falls into the zero-inflated group—in this case, those who are completely unemployed and work zero hours per week. The second component is an ordered logit model, which is used to analyze the employment status among individuals who are active in the labor force, distinguishing between those who are fully employed and those who are underemployed. These two components are estimated simultaneously within a single ZIOL framework, allowing the model to capture the specific characteristics of the data, especially the overrepresentation of individuals with zero working hours. Therefore, the ZIOL model provides

a more comprehensive understanding of employment dynamics than a standard ordered logit model, which does not account for the excess zeros in the data.

Based on the results of the Zero-Inflated Ordered Logit (ZIOL) model estimation shown in Table 2, several factors significantly influence labor force status, which in this study is categorized into three ordered groups: Employed Underemployed, and Open Unemployment. The ZIOL model is particularly suitable for this analysis due to the high proportion of zero outcomes (representing open unemployment), as it accounts for excess zeros and the ordinal nature of the dependent variable.

The analysis indicates that gender significantly affects labor force status. Female individuals tend to fall into categories with lower labor force participation, as shown by the positive and significant coefficients in both the ordinal and inflation components of the ZIOL model.

The education level of the individuals also plays a crucial role. Those with higher educational attainment (senior high school or above) are more likely to be employed and less likely to fall into open unemployment. This is reflected in the negative and statistically significant coefficients in the inflation part of the model, suggesting that better-educated individuals are less likely to be part of the excess-zero groups (not in the labor force at all).

Meanwhile, residence in urban areas (city), though not statistically significant in the inflation part, is positively associated with employment status in the ordered part of the model, implying better job access in urban regions. Other variables such as flooring quality, roofing materials, and wall type often used as proxies for housing quality also display significance in explaining labor force status, aligning with the notion that better living conditions are associated with greater participation in the labor market.

To strengthen the justification for using the ZIOL model over conventional approaches, a Vuong test was conducted to statistically compare the ZIOL model with the traditional Ordinal Logit model. The hypotheses for this test are:

H_0: ZIOL model is as good as Ordinal Logit

H_1: ZIOL model is different from Ordinal Logit model

Where the test statistic for the Vuong test (v) is as follows.

$$v = \frac{\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} m_i\right)}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(m_i - \bar{m})^2}},$$

If the value of the test statistic exceeds a certain threshold, then we can conclude that one model is significantly better at explaining the data than the other model. Based on equation 3, the value of v from the results of stata and excel processing is 38.068. This value shows that the ZIOL model is different from the ordinal logit model in modeling data on the level of labor force status in this study.

The value of $|v|$> 1.96 with an alpha of five percent, it can be concluded that the value of ZIOL and Ordinal Logistics is different in analyzing the level of labor force status in Banten Province. When viewed from the v value of 38.068 where the number is greater than 1.96, it can be interpreted that the first model, namely ZIOL, is better than the Conventional Logistic Ordinal model.

The interpretation of the ZIOL model is carried out using the odds ratio (OR) values to understand how much each predictor variable influences the labor force status. The OR value illustrates the likelihood of an individual being classified as underemployed or openly unemployed compared to the reference group.Where women are 2.130 times more likely to be underemployed or openly unemployed than men, indicating a gender-based disparity in employment status. Education level also plays a significant role. Individuals with elementary and junior high school education are 1.376 times more likely to be in the underemployed or openly unemployed category compared to those with no education. Surprisingly, individuals with senior high school education or higher are even more likely (OR = 1.453) to fall into this category, suggesting that higher education does not always guarantee employment. This may reflect a mismatch between qualifications and job market demands. Urban residency is associated with a 1.805 times greater likelihood of being underemployed or unemployed compared to rural residency, possibly due to more intense job competition in cities. Furthermore, individuals who have never participated in job training are 1.352 times more likely to face underemployment or open unemployment, emphasizing the importance of skill development. The highest risk is found among individuals with no prior work experience, who are 2.269 times more likely to be in this vulnerable employment category. In the inflate section, which models the likelihood of remaining employed (zero-inflated component), women are 3.864 times more likely than men to remain in the employed category. This suggests that while women face a higher risk of unemployment, there is also a segment of women who maintain employment, potentially in informal sectors.

Regarding education, those with elementary and junior high education are slightly less likely (OR = 0.780) to remain employed than individuals with no education. More strikingly, those with senior high school education or above are even less likely to remain employed (OR = 0.481), which might indicate their tendency to seek better opportunities, increasing the risk of temporary unemployment. Urban residents are also less likely to remain employed compared to rural residents, again pointing to urban job market challenges. Interestingly, individuals who have never received job training are 1.429 times more likely to stay employed, possibly because they are more likely to accept and stay in available jobs, even if they are not ideal. Similarly, individuals without work experience are 1.892 times more likely to remain employed, possibly in informal jobs that do not require prior experience.

CONCLUSION

The Vuong test result (v = 38.068) indicates that the Zero-Inflated Ordered Logit (ZIOL) model is significantly better than the conventional Ordinal Logit model for analyzing labor force status data in Banten Province. ZIOL effectively captures both the excess zeros (employed individuals) and the ordinal nature of labor force categories. The model shows that gender, education level, place of residence, job training, and work experience significantly influence labor status. Women, individuals with higher education, urban residents, and those without training or experience are more likely to be underemployed or unemployed, although some also remain employed—often in informal jobs. This suggests that ZIOL provides deeper insights into labor vulnerability, and supports the need for targeted policies to improve access to training, job creation, and support for vulnerable groups.

REFERENCES

Agresti. (2002). *Categorical Data Analysis*.

Allen, E., Francisco, R., Ginting, E., Indrio, V. T., Manning, C., Marshan, J., Mercer-Blackman, V., Naval, D., Pratomo, D. S., Suryahadi, A., Tadjoeddin, M. Z., & Taniguchi, K. (2018). Indonesia Enhancing Productivity Through Quality Jobs. In *Asian Development Bank*.

Annisa Qurrota Ayu, Subanti, S., & Sugiyanto. (2024). FACTORS AFFECTING OPEN UNEMPLOYMENT RATE IN BANTEN PROVINCE USING PANEL DATA REGRESSION. *Journal of Mathematics and Mathematics Education*.

Aqilla Haya, Risma Dwi Lestari, & Tengku Mashitah Crisanty. (2023). Analysis of Factors Affecting the Open Unemployment Rate (UOR) 2022 : A Case of Banten in Indonesia. *Proceedings of The International Conference on Data Science and Official Statistics*, *2023*(1), 639–649. https://doi.org/10.34123/icdsos.v2023i1.386

Downward, P., Lera-Lopez, F., & Rasciute, S. (2011). The zero-inflated ordered probit approach to modelling sports participation. *Economic Modelling*, *28*(6), 2469–2477. https://doi.org/10.1016/j.econmod.2011.06.024

Harris, M. N., & Zhao, X. (2007). A zero-inflated ordered probit model, with an application to modelling tobacco consumption. *Journal of Econometrics*, *141*(2), 1073–1099. https://doi.org/10.1016/j.jeconom.2007.01.002

Hosmer and, & Lemeshow. (2000). *Epdf.Pub_Applied-Logistic-Regression-Wiley-Series-in-Probab.Pdf* (pp. 1–30).

Jiang, X., Huang, B., Zaretzki, R. L., Richards, S., Yan, X., & Zhang, H. (2013). Investigating the influence of curbs on single-vehicle crash injury severity utilizing zero-inflated ordered probit models. *Accident Analysis and Prevention*, *57*, 55–66. https://doi.org/10.1016/j.aap.2013.03.018

Kelley, M. E., & Anderson, S. J. (2008). Zero inflation in ordinal data: Incorporating susceptibility to response through the use of a mixture model. *Statistics in Medicine*, *27*(18), 3674–3688. https://doi.org/10.1002/sim.3267

Pardosi, A. Y., & Septriani, S. (2023). Determination Factors of Unemployment in Banten. *Jurnal Ekonomi*, *12*(04), 2341–2347. https://ejournal.seaninstitute.or.id/index.php/Ekonomi/article/view/3337%0Ahttps://ejournal.seaninstitute.or.id/index.php/Ekonomi/article/download/3337/2725

Presiden Republik Indonesia. (2020). Peraturan Presiden Republik Indonesia Nomor 18 Tahun 2020 Tentang Rencana Pembangunan Jangka Menengah Nasional 2020-2024. *Sekretariat Presiden Republik Indonesia*, 1–7.

Rejeki, N., Ratnasari, V., & Ahsan, M. (2024). Modelling of Poor Household in East Kalimantan Using Zero Inflated Ordered Probit (ZIOP) Approach. *Procedia Computer Science*, *234*(2023), 278–285. https://doi.org/10.1016/j.procs.2024.03.002

Stata, S. (n.d.). *Zero - inflated ordered logit model*.

Yanthiani, L. (2023). The Impact of Unemployment on the Economy in Indonesia. *Journal of Islamic Economics and Business*, *2*(2), 112–130. https://doi.org/10.15575/jieb.v2i2.21310

Yoong, P. S., & Gil Sander, F. (2020). Structural Transformation and Labor Productivity in Indonesia. *Structural Transformation and Labor Productivity in Indonesia*, 1–33. https://doi.org/10.1596/35951