International Journal of Academic and Applied Research (IJAAR)
ISSN: 2643-9603
Vol. 7 Issue 6, June - 2023, Pages: 67-71

Modeling of Provincial Food Security Index in Indonesia based on Probit Ordinal Panel Regression Approach with Random Effects

Anisa Laila Azhar[1], Suliyanto[2], Nur Chamidah[3], Dita Amelia[4], Elly Ana[5]

[1,2,3,4,5] Statistics Study Program, Department of Mathematics, Faculty of Science and Technology
Airlangga University
Surabaya, Indonesia
Corresponding author: suliyanto@fst.unair.ac.id

***Abstract**: Indonesia has been renowned as an agricultural nation, giving paramount importance to agriculture in meeting its food requirements. Due to its large population, there is a significant demand for food, and insufficient food supply can jeopardize the country's food security. This research aims to develop a model for the food security index and determine the factors influencing provincial food security in Indonesia using the probit ordinal regression method with panel data and random effects. The analysis reveals that variables such as the percentage of households with access to electricity, the percentage of stunted toddlers, and rice production have a substantial impact on the provincial food security index in Indonesia. Moreover, the panel probit regression model with random effects achieves a classification accuracy of 44.83% and an AIC value of 236.79, surpassing the standard ordinal probit regression model. This conclusion is further supported by the model suitability test, which yields a p-value of 0.000, indicating that the panel ordinal probit regression model with random effects is more suitable than the standard probit regression model.*

Keywords—agriculture; food security; probit ordinal regression; efek acak; panel data

1. INTRODUCTION

Indonesia has been known as an agricultural country and makes agriculture an important field in maintaining food needs. The large population makes food needs needed in large quantities so that when the amount of food available is less than the national food needs this can damage the condition of national food security [5]. According to [13] if a country has strong food security and is able to meet national food needs, then self-sufficiency the country will have strong national security. The Ministry of Agriculture realizes efforts in maintaining national food security through sustainable development programs in the Sustainable Development Goals (SGD's) program to achieve No Poverty and Zero Hunger. There are three aspects of food security according to the World Health Organization (WHO) which include food availability, food affordability, and food utilization.

The value of food security in Indonesia has fluctuated from 2019 to 2022 and in 2022 has improved when compared to 2021, which is at a score of 60.2. The trend of fluctuations in the food security index during the period 2019 to 2022 has resulted in changes in food security status in each province and based on Global Food Security Index (GFSI) data, Indonesia's food security score in 2022 is still below the global average and Asia Pacific average of 62.2 and 63.4 respectively [1]. In addition, the food supply in Indonesia is considered inadequate with a score of 50.9. Nutrients contained in food also only get a score of 56.2, while sustainability and adaptation get a score of 46.3. In these three indicators, Indonesia's resilience is considered worse than the average of countries in the Asia Pacific region. This proves that significant results have not been seen after the government made various efforts to improve food security in Indonesia.

Research on food security has been carried out by several researchers. [14] Modeling food security in Indonesia using the binary logistic regression method. The results stated that the variables of the rice field area, corn production, soybean production, chicken meat production, beef production, population density, the consumer price index for housing, electricity, and gas; consumer price index for health, the consumer price index for transportation and financial services; and the Food Insecurity Multidimensional Index has a significant effect on Indonesia's food security. [12] Modeling household food security using the ordinal logistic regression method. The results showed that rice price variables, other consumer prices, and other household income had a positive effect on household food security. The availability of various nuts, the price of finished food and beverages, and the number of household members negatively affect household food security in Central Lampung Regency. [11] Modeling food security in Indonesia based on bivariate binary probit regression approach. The results showed that the variables of gross regional domestic product, percentage of poor people, and human development index had a significant effect on food security indicators.

Based on previous research, the authors suspect that changes in food security status are influenced by the percentage of *stunted* toddlers, the percentage of households with access to electricity, rice production, and the rate of economic growth. In addition, there are allegations that there are other factors that randomly affect the food security index but cannot be observed directly. In this context, researchers are

interested in modeling the food security index in Indonesia using a probit ordinal regression approach of panel data with random effects. The novelty of this study is that there has been no previous study that modeled the food security index in Indonesia using probit ordinal regression panel data with random effects. The expected outcome of this study is to identify factors influencing food security status in Indonesia, estimate the probability of food security status in each province in Indonesia, and provide guidance in government policy-making related to food security in the country

2. LITERATURE REVIEW

2.1 Probit regression Ordinal Random Effects Panel Data

Panel data probit regression model is a form of development of the probit regression model. In this model, response variables are divided into two categories and the data used for analysis is panel data. The probit regression model is expressed as follows:

$$y_{it}^{*} = \mathbf{x}_{it}\boldsymbol{\beta} + u_i + v_{it} \quad (1)$$
$$\varepsilon_{it} = u_i + v_{it}$$

With

$$y_{it} = \begin{cases} 1 \text{ if} & y_{it}^{*} \le \kappa_1 \\ 2 \text{ if} & \kappa_1 < y_{it}^{*} \le \kappa_2 \\ & \vdots \\ q \text{ if} & \kappa_{q-1} < y_{it}^{*} \end{cases} \quad (2)$$

When y_{it} is an observation for the i-th data at the t-th time, $\mathbf{x}_{it}$ is a vector that is 1 x p for the predictor variable, $\boldsymbol{\beta}$ is the vector p x 1 for the parameter coefficient. u_i is an unobserved individual effect and v_{it} is a random error. The probit error model also assumes errors are independent and standard normally distributed [7]. In addition, random effect models also assume variation at the unit level [4].

2.2 Parameter Estimation

Testing this parameter aims to determine the significance of the influence of predictor variables on response variables in the model. In probit regression, there are two types of tests performed, namely overall testing (simultaneous testing) and individual testing (partial testing). Here is a further explanation of both types of testing.

2.2.1 Simultaneous Testing

Concurrent tests are used to test the effect of overall cefficient β in the model. The test hypothesis is $H_0 : \boldsymbol{\beta} = \mathbf{0}$ dan $H_1 : \boldsymbol{\beta} \neq \mathbf{0}$. The statistical test used is the Wald Test, with the following requirements [6].

$$W = \left(\widehat{\boldsymbol{\beta}} - \mathbf{0}\right)^T I\left(\widehat{\boldsymbol{\beta}}\right)\left(\widehat{\boldsymbol{\beta}} - \mathbf{0}\right) = \left(\widehat{\boldsymbol{\beta}} - \mathbf{0}\right)^T \mathbf{V^{-1}}\left(\widehat{\boldsymbol{\beta}} - \mathbf{0}\right)$$
$$= \widehat{\boldsymbol{\beta}}^T \mathbf{V}^{-1}\widehat{\boldsymbol{\beta}} \sim \chi_p^2 \quad (3)$$
$$= \widehat{\boldsymbol{\beta}}^T \mathbf{V}^{-1}\widehat{\boldsymbol{\beta}} \sim \chi_p^2$$

This Wald statistical test follows the Chi-Square distribution with a degree of freedom equal to the number of parameters in the model. Test decisions are made by comparing values of W with values of χ^2. The test criterion is that H_0 will be rejected if the value $W > \chi^2_{(p;\,\alpha)}$, where p is the number of predictor variables in the model, or H_0 will be rejected if the p-value is < α.

2.2.2 Individual Testing

Individual tests are used to test the significance of each parameter, to find out whether the predictor variables are significantly influential in the model or not. This partial test is also known as the Wald test (W) or Wald test. The Wald test is obtained by comparing the β estimate with the standard estimate of error, with the test hypothesis $H_0 : \beta_k = 0$ and $H_1 : \beta_k \neq 0, k = 1, 2, \ldots, p$ and test statistics as follows [9].

$$Z_k = \frac{\widehat{\beta}}{SE(\widehat{\beta}_k)}; k = 1,2, \ldots, p \quad (4)$$

Wald's statistics follow a normal distribution, so test decisions are made by comparing values Z_k with value of $Z_{\alpha/2}$. The test criterion is H_0 will be rejected if the value $|Z_k| > Z_{\alpha/2}$ or value of *p-value* < *α*.

2.3 Model Compatibility

Model conformity testing is used to evaluate whether the model that has been created is suitable or not yet matches the observed data. The hypotheses used are *H₀* : The ordinal probit regression model of panel data is not appropriate whereas *H₁* : The ordinal probit regression model of panel data is appropriate. The test statistics used are the Likelihood Ratio Test formulated as follows [8]:

$$\Lambda = -2\ln\left(\frac{L_{H_0}}{L_{H_1}}\right) \quad (5)$$

The statistics of the Likelihood Ratio Test test follow the Chi-Square distribution with a degree of freedom of 1. The test decision is obtained by comparing the value of the Likelihood Ratio Test with the value of χ^2 of the table. The H_0 decision is rejected if the value $\Lambda > \chi^2_{(1;\,\alpha)}$, or H_0 will be rejected if the p-value < α.

2.4 Selection of The Best Model

1. Akaike's Information Criterion (AIC)

Akaike's Information Criterion (AIC) is an evaluation method used to select the best model from several existing models. AIC can be used as a model selector that is able to evaluate the relative quality of each model using maximum likelihood estimation as a precise calculation [2]. The selection of the best model is based on the smallest AIC criteria, this is because the AIC value is directly proportional to the value of model devians. The smaller the devians value, the error rate value is also smaller, so the resulting model will be more precise. The following is the formula for calculating the AIC value [3].

$$AIC(p) = -2\frac{\ln L(P)}{n} + 2\frac{p}{n} \quad (6)$$

Where $\ln L(P)$ is the maximum likelihood with *k* predictor variables, p is a lot of β parameters with p = 1.2,... k, and *n* represents the sample size.

2. Apparent Error Rate (APPER)

Apparent Error Rate (APPER) is a metric used to assess the degree of accuracy of model classification in classifying

objects. APPER describes the probability of error in the classification process. The formula used to calculate the APPER value is as follows:

$$APPER = \frac{\sum_{i\neq j=1}^{q} n_{ij}}{\sum_{i,j=1}^{q} n_{ij}} \times 100\% \quad (7)$$

In ordinal probit regression, the value of the accuracy of the model classification is measured using equation 1 - APPER, where n_{ij} is the number of Y events in category *i* of observations classified in category *j* of the prediction results [10].

3. Result and Discussion

3.1 Descriptive Statistics

The descriptive statistical analysis outcomes of the variables related to the food security index over a span of four years are presented in **Table** 1, as shown below.:

Table 1: The Food Security Index Descriptive Statistics

	Year	Minimum	Maximum	Average
Food Security Index	2019	25.13	85.15	65.95
	2020	34.79	84.54	72.11
	2021	35.48	83.82	72.43
	2022	37.80	85.19	71.68

Referring to Table 1, it can be observed that Indonesia's food security index values from 2019 to 2022 averaged at 65.95, 72.11, 72.43, and 71.68, respectively. The highest values recorded were 85.15, 84.54, 83.82, and 85.19, with Bali Province achieving the top ranks. Conversely, the lowest index values were 25.13, 34.79, 35.48, and 37.80, respectively, and were reported by Papua Province..

Table 2: Percentage of Regions by Food Security Category

Category / Year	High Food Vulnerability	Medium Food Vulnerable	Low Food Vulnerability	Low Food Security	Medium Food Security	High Food Security
2019	5,88	2,94	11,76	14,71	32,35	32,35
2020	2,94	0	2,94	11,76	32,35	50
2021	2,94	2,94	0	8,82	32,35	52,94
2022	0	5,88	0	11,76	38,23	44,12
Total	2,94	2,94	3,68	44,44	33,82	44,85

Based on **Table 2,** it can be seen that provinces with the food security index values are in the high food security category of 44.85%; medium food security category of 38.82%; low food security category of 44.44%; low food vulnerable category of 3.68%; medium food vulnerable category of 2.94%; and high food vulnerable category of 2.94%.

3.2 Food Security Index Modeling Provincial in Indonesia Based on Probit Ordinal Regression Approach Panel Data

The Probit ordinal regression model for panel data with random effects can be expressed as a linear representation of latent response variables. These latent response variables are transformed into ordinal response variables, denoted as y_{it}, which are obtained from continuous latent response variables.

$$\hat{y}_{it}^{*} = 0{,}1414026x_{1it} - 0{,}0825694x_{2it} + \\ +0{,}000000374x_{3it} - 0{,}0247079x_{4it} \quad (8)$$

as

$$\hat{y}_{it} = \begin{cases} 1 \text{ if } & y_{it}^{*} \leq 5{,}523174 \\ 2 \text{ if } & 5{,}523174 < y_{it}^{*} \leq 7{,}333561 \\ 3 \text{ if } & 7{,}333561 < y_{it}^{*} \leq 8{,}602362 \\ 4 \text{ if } & 8{,}602362 < y_{it}^{*} \leq 10{,}28933 \\ 5 \text{ if } & 10{,}28933 < y_{it}^{*} \leq 13{,}17075 \\ 6 \text{ if } & 13{,}17075 < y_{it}^{*} \end{cases} \quad (9)$$

In this case $\hat{y}_{it}$ it refers to the estimated food security index. If $\hat{y}_{it}$= 1 it means that the alleged food security index is in the category of high food vulnerability. If $\hat{y}_{it}$= 2 it means that the alleged food security index is in the category of medium food vulnerability. If $\hat{y}_{it}$= 3 it means that the alleged food security index is in the low food vulnerable category. If $\hat{y}_{it}$= 4 it means that the alleged food security index falls into the category of food security lace. If $\hat{y}_{it}$= 5 it means the alleged food security index is in the medium food security category and $\hat{y}_{it}$= 6 it means that the alleged food security index is in the high food security category.

3.2.1 Parameter Estimation

Test results for each predictor variable are presented in the following.

Table 3: Simultaneous Test Results and Individual Tests

Variables	Simultaneous Test	Individual Test	
		Z_{count}	*P-Value*
Percentage of Households with Access to Electricity (X_1)	W = 31,98 $\chi^2_{(4;0,05)}$ = 9,488 *p-value* = 0,000	4,33	0,000
Percentage of Stunting Toddlers (X_2)		-3,81	0,000

Percentage of Economic Growth Rate (X_3)		-1,12	0,263
Rice Production (X_4)		2,14	0,033

Based on simultaneous tests, it is known that the W value is 31,98 and value of $\chi^2_{(4;0,05)}$ is 9,488 with a p-value of 0,000. It was decided to reject H_0 because $W > \chi^2_{(4;0,05)}$ or $p - value < \alpha$ so that the conclusion obtained is that there is at least one predictor variable that has a significant effect on the food security index in Indonesia. In addition, individual tests show that there are three variables that have a significant effect on the food security index in Indonesia, namely the percentage of stunted toddlers, the percentage of households with access to electricity, and rice production. This is evidenced by the value of Z_{count} In the variable percentage of stunting toddlers is -3,81 and p-value of 0,000 which is the value $|Z_{count}| > Z_{(0,025)}$ or $p\text{-}value < \alpha$ thus resulting in a decision to reject H_0. On the other hand, the value of Z_{count} In the variable percentage of households with access to electricity is 4,33 and the p-value is 0,000 where the value $|Z_{count}| > Z_{(0,025)}$ or p-value < α thus resulting in a decision to reject H_0. As for the value Z_{count} in the variable rice production of 2,14 and p-value of 0,033 which is the value $|Z_{count}| > Z_{(0,025)}$ or $p\text{-}value < \alpha$ thus resulting in a decision to reject H_0. Therefore, it was concluded that the variables of the percentage of stunted toddlers, the percentage of households with access to electricity, and individual rice production affect the food security index in Indonesia.

3.2.2 Likelihood Ratio Test

The results of model conformity testing are presented in **Table 5** as follows.

Table 5: Model Conformity Test Results

Measurement	Value
Λ	59,99
$\chi^2_{(1;0,05)}$	3,842
p-value	0,0000

Based on **Table 5** it is known that the value Λ is 59.99 and the value $\chi^2_{(4;0,05)}$ is 3,842 with $p - value$ of 0,000. The decision was rejected H_0 because $\Lambda > \chi^2_{(4;0,05)}$ or $p - value < \alpha$, so it can be concluded that there is enough variability between provinces to satisfy the ordinal probit regression model of random effects panel data compared to the standard probit regression model or in other words the ordinal probit regression model of random effects panel data is appropriate.

3.2.3 Best Model Determination

After the modeling and model estimation process is complete, the next step is to calculate the accuracy value of the model classification based on 1 – *APPER*. This study compares the goodness of the ordinal probit regression model of random effects panel data (xtologit) with the standard ordinal probit regression model (ologite) based on classification accuracy values and AIC values.

Table 4: Results of Model Classification Accuracy Calculation

Classification Accuracy	
Model xtoprobit	Model oprobit
0,4483	0,0441

The determination of the model goodness criteria is also based on the AIC value by comparing the AIC value of the xtologit model with the ologite model.

Table 5: AIC Value Calculation Results

Models	AIC Value
Model xtoprobit	236,79
Model oprobit	294,78

According to the calculations, the xtoprobit model has a classification accuracy value of 0.4483 or 44.83%. This indicates that the xtoprobit model has been correctly classified by 44.83%, while the remaining 55.17% is classified incorrectly or differently. On the other hand, the oprobit model has a classification accuracy value of 0.0441 or 4.41%. This means that the oprobit model has only been correctly classified by 4.41%, while the remaining 95.59% are classified incorrectly or differently. In addition, based on the AIC value, it is known that the AIC value of the xtoprobit model is smaller than the AIC value of the oprobit model. Based on the calculation of classification accuracy and AIC value, it can be concluded that the ordinal probit regression model of random effects panel data (xtoprobit*)* is better because it is able to produce a more precise classification model and the AIC value is smaller when compared to the standard ordinal probit regression model (oprobit).

4. CONCLUSION

1. The value of the food security index of each province in Indonesia shows a fluctuating trend from 2019 to 2022. The distribution of food security levels in Indonesia is as follows: 44.85% falls under the high food security category, 33.82% falls under the medium food security category, 44.44% falls under the low food security category, 3.68% falls under the low food vulnerability category, 2.94% falls under the medium food vulnerability category, and 2.94% falls under the high food vulnerability category.
2. Using panel data probit regression modeling with random effects, it has been revealed that three variables exert a notable impact on the food security index. These variables encompass the percentage of stunted toddlers, which negatively affects the food security index, the considerable scale of food production, which holds a position in relation to the food security index, and the

percentage of households with access to electricity, which positively influences the food security index.

3. The Likelihood Ratio Test method was conducted to assess the suitability of the model. The results indicate that there is sufficient variability in each province to support the use of the ordinal probit regression model with random effects in panel data. This finding suggests that the random effects panel data model is more suitable compared to the standard probit regression model. Furthermore, the ordinal probit regression model with random effects yielded a calibration accuracy value of 44.83% and an AIC value of 236.79.

5. REFERENCES

[1] Ahdiat, A., 2022. *Ketahanan Pangan Indonesia Menguat pada 2022.* [Online] Available at: https://databoks.katadata.co.id/datapublish/2022/12/02/ketahanan-pangan-indonesia-menguat-pada-2022 [Accessed 6 6 2023].

[2] Anderson, D. R., Burnham, K. P. & White, G. C., 1998. Comparison of Akaike Information Criterion and Consistent Akaike Information Criterion for Mode Selection and Statistical Inference from Capture-Recapture Studies. *Journal of Applied Statistics,* 25(2), pp. 263-282.

[3] Aswi & Sukarna, 2006. *Analisis Deret Waktu.* Makassar: Andira.

[4] Butler, J. S. & Moffit, R., 1982. A Computationally Efficient Quadrature Procedure For The One-Factor Multinomial Probit Model. *Econometrica,* 50(3), pp. 761-764.

[5] Christiyanto, M. & Mayulu, H., 2021. Pentingnya Pembangunan Pertanian dan Pemberdayaan Petani Wilayah Perbatasan dalam Upaya mendukung Ketahanan Pangan Nasioanl: Studi Kasus di Wilayah Perbatasaan Kalimantan. *Journal of Tropical Agrifood,* 3(1), pp. 1-14.

[6] Harrel, F. E., 2015. *Regression Modeling Stratregies With Application to Linier Models, Logistic Regression, and Survival Analysis.* USA: Springer.

[7] Harris, M. N., Macquarie, l. R. & Siouclis, A. J., 2000. *A Comparison of Alternatif Estimators for Binary Panel Probit Models.* Victoria: The University of Melbourne.

[8] Hosmer, D. W., Lemeshow, S. & Sturdivant, R. X., 2013. *Applied Logistic Regressios.* 3rd ed. USA: Jhon Wiley and Sons.

[9] Hosmer, D. W. & Stanley, L., 2000. *Applied Logistic Regression: Second Edition.* New York: John Wiley & Sons, Inc..

[10] Johnson, R. A. & Wichern, D. W., 2002. *Applied Multivariate Statistical Analysis.* 5th ed. New Jersey: Hall.

[11] Khuroidah, I. A., 2018. *Pemodelan Ketahanan Pangan di Indonesia berdasarkan Pendekatan Regresi Probit Biner Bivariat,* Surabaya: Universitas Airlangga.

[12] Nanda, L. P., Mulyo, J. H. & Waluyati, L. R., 2019. Analisis Ketahanan Pangan Rumah Tangga di Kabupaten Lampung Tengah. *Jurnal Ilmu Ekonomi dan Agribisnis,* 3(2), pp. 219-232.

[13] Rhofita, E. I., 2022. Optimalisasi Sumber Daya Pertanian Indonesia untuk Mendukung Program Ketahanan Pangan dan Energi Nasional. *Jurnal Ketahanan Nasional,* 28(1), pp. 81-99.

[14] Widada, W. A., M. & Mulyo, J. H., 2017. Determinan Factors of Food Security in Indonesia. *Agro Ekonomi,* 28(2), pp. 205-219.