RESEARCH ARTICLE OPEN ACCESS

Information and Communication Technology Development Index Modeling in Indonesia Using Panel Data Ordinal Probit Regression with Random Effects

Ainaya Zakiyah Nabila[1], Suliyanto[2],Sediono[3], Dita Amelia[4]
[1,2,3,4]Statistic Study program, Departement of Mathematics, Airlangga University, Surabaya-Indonesia
[1]ainaya.zakiyah.nabila-2020@fst.unair.ac.id
[2]Corresponding author :suliyanto@fst.unair.ac.id
[3]ahmad.s@fst.unair.ac.id
[4]dita.amelia@fst.unair.ac.id

--************************----------------------------------

Abstract:

Technology, Information, and Communication have become influential elements in various aspects of life. The existence of the Information and Communication Technology Development Index (ICT-DI) provides a precise parameter for systematically measuring ICT progress. This research aims to identify factors that can influence the ICT-DI by applying the ordinal probit regression method on panel data with random effects. The response variable used is the ICT-DI categorized as high, medium, and low. Meanwhile, predictor variables involve the average percentage of household telecommunication consumption, per capita Gross Domestic Product (GDP), the percentage of the population below the poverty line, and the percentage of completion of high school or equivalent education. This study utilizes data from the years 2017 to 2022. The research results indicate that the ordinal probit regression method has a classification accuracy rate of 75.49%, and the variables such as the average percentage of household telecommunication consumption, per capita GDP, the percentage of the population below the poverty line, and the percentage of completion of high school or equivalent education have been proven to significantly impact the ICT-DI variable.

Keywords **—Technology, ICT-DI, Ordinal Probit Panel Data Random Effects Regression.**

--************************----------------------------------

I. INTRODUCTION

In the context of the ongoing globalization and digital transformation, Technology, Information, and Communication (ICT) play a crucial role as a dominant force influencing various aspects of human life [1]. A profound understanding of the developments in the ICT sector is vital, especially for countries like Indonesia that are striving to excel in the global competition [2]. In this perspective, the progress of the ICT sector in Indonesia has the potential to enhance productivity, create job opportunities, and improve national economic competitiveness [3]. Innovations in the field of ICT continue to drive transformation across various sectors, including Health, Education, Business, and Government [4]. A deep understanding of the ICT sector's development can contribute positively to formulating policies that support innovation and sustainable development across sectors. .

Assessing the progress of the ICT sector in a country requires a systematically measurable indicator, namely the Information and Communication Technology Development Index (ICT-DI). The Central Statistics Agency (BPS) states that the composite index reflecting the level of ICT progress is measured through the ICT-DI developed by the International Telecommunication Union (ITU), known as the Information and Communication Technology Development Index (ICT-DI) [5]. The ICT-DI is derived from the calculation of 11 indicators divided into three sub-indices: access and infrastructure, usage, and skills [6]. By using appropriate indicators, the government and stakeholders can evaluate the achievement of ICT development targets and identify necessary actions. This approach aligns with the Sustainable Development Goals (SDGs), specifically

Goal 9, emphasizing resilient infrastructure development, fostering inclusive and sustainable industry, promoting innovation, with one of its indicators focusing on ICT progress [7]. Using the ICT-DI as the core indicator allows Indonesia to compare the progress of the ICT sector with other countries.

The ICT-DI in Indonesia shows an increasing trend from 2017 to 2022. Despite the improvement in the ICT-DI value, Indonesia ranks last in the Asia-Pacific region with a score of 1.2, as revealed by the Digital Skills Gap Index (DSGI) in the category of global science and technology skills [8]. Meanwhile, according to the World Digital Competitiveness, Indonesia is ranked 51 out of 63 countries globally and 12 out of 14 in the Asia-Pacific region [9]. This comparison indicates that the ICT condition in Indonesia lags significantly behind other countries. Although the Indonesian government has endeavored to enhance the development of the ICT sector, the efforts made have not yielded significant results. This statement is reinforced by the Minister of Communication and Information, Rudiantara, as quoted on the Kominfo website, stating that since 2017, the main issue faced is the inequality between regions, referring to the four categories set by BPS: high, medium, low, and very low [10]. This categorization aims to map regions based on their ICT development status, where a higher index value indicates rapid ICT development progress in an area, and vice versa. In line with the government's efforts to improve the ICT-DI, a more realistic approach can be taken through regression modeling to identify the factors influencing the ICT-DI. Thus, recommendations obtained from such analysis can guide the government in efforts to enhance the ICT-DI in Indonesia.

According to Hsiao (2003), panel data is a collection that combines elements of cross-sectional and time-series data. The advantage of panel data lies in its ability to analyze dynamics that cannot be accomplished with cross-sectional or time-series data alone [11]. If the response variable has more than two categories, applicable analysis methods include nominal logistic regression or ordinal logistic regression, depending on the measurement scale of the response variable [12]. The ordinal logistic regression model for panel data reflects the relationship between an ordinal-scale response variable and one or more predictor variables, whether continuous or categorical [13]. Based on the fact that the ICT-DI is categorized into four levels [6] and the existence of unobserved factors that have a random impact on the ICT-DI, the researcher is interested in applying the ordinal logistic regression method for panel data with random effects to model the ICT-DI. This approach is expected to provide a deep understanding of the factors influencing ICT development in Indonesia and formulate more effective policy recommendations to support the growth of this sector.

II. METHODOLOGY

The data applied in this study is secondary data related to ICT and factors suspected to have an impact on the ICT Development Index (ICT-DI) in each province, covering the period from 2017 to 2022. The study encompasses 34 provinces across Indonesia as research units. The variables used are categorized into two groups: the response variable (Y) and the predictor variable (X), with specific details outlined in Table 1.

TABLE I
RESEARCH VARIABLES

Variables	Variable Description	Data Scale	Data Type
Y	index value of ICT development in each province in Indonesia.	ratio	Categories 1 High 2 Medium 3 Low
X1	percentage of average household telecommunication consumption	ratio	continuous.
X2	GDP per capita (idr)	ratio	continuous.
X3	percentage of the population below the poverty line (%)	ratio	continuous.
X4	percentage of completion of high school education (%).	ratio	continuous.

In this study, the method employed is ordinal probit panel data regression with random effects, which can be written in the form of a linear latent response variable and an ordinal response variable y_{it} derived from the continuous latent response variable and expressed as follows:

$$y_{it}^* = \boldsymbol{X}_{it}\boldsymbol{\beta} + v_i + \varepsilon_{it} \#(1)$$

with

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

Next, simultan and individual tests are conducted using the Wald test statistics with a target P-value < 5%. Subsequently, the adequacy of the ordinal probit panel data regression model is tested to examine whether the obtained model is suitable for the observed data using the likelihood ratio test, as follows [14]:

$$\Lambda = -2\ln\left(\frac{L_{H0}}{L_{H1}}\right) \#(3)$$

Afterward, the Apparent Error Rate (APPER) is calculated, representing the value used to determine the probability of error in object classification. The APPER value is calculated using the following formula [15]:

$$APPER = \frac{\sum_{i\neq j=1}^{q} n_{ij}}{\sum_{i,j=1}^{q} n_{ij}} \times 100\%;\ i = 1,2,\ldots \#(4)$$

Where n_ij is the number of occurrences of Y in category i from observations classified as occurrences of Y in category j from predictions. The classification accuracy of the ordinal probit regression model is 1−APPER.

III. RESULT

A. Modeling the Information and Communication Technology Development Index in Indonesia Based on the Ordinal Logistic Probit Panel Data Approach.

The ordinal probit panel data regression model with random effects can be presented in Equation 1, where x_1it is the average percentage of household telecommunication consumption, x_2it is the Gross Domestic Product (GDP) per capita, x_3it is the percentage of the population below the poverty line, and x_4it is the percentage of completion of high school or equivalent education in province i in year t as follows:

$$\hat{y}_{it}^{*} = -3.603227x_{1it} - 0.0001596x_{2it} + 0.221182x_{3it} - 0.1079405x_{4it}\#$$

with

$$\hat{y}_{it} = \begin{cases} 1\ if & \hat{y}_{it}^{*} \le -62.1405 \\ 2\ if & -62.1405 < \hat{y}_{it}^{*} \le -17.38997\ \# \\ 3\ if & -17.38997 < \hat{y}_{it}^{*} \end{cases}$$

Where $\hat{y}_{it}$ represents the ICT-DI, where $\hat{y}_{it} = 1$ indicates a suspected high ICT-DI, $\hat{y}_{it} = 2$ indicates a suspected medium ICT-DI, and $\hat{y}_{it} = 3$ indicates a suspected low ICT-DI.

B. Simultan Test.

Simultan testing of parameter significance aims to determine whether predictor variables collectively influence the response variable. The Wald test statistic is used with a significance level α of 5%. The results of simultaneous parameter significance testing are presented in Table 2.

TABLE II
SIMULTAN TEST

Measurement	Values
W	12.43
$X^2_{(4;0.05)}$	9.488
$p-value$	0.0144

Based on Table 2, the W value is 12.43, and the $X^2_{(4;0.05)}$ value is 9.488 with a $p-value$ of 0.0144. It is decided to reject H_0 because $W > X^2_{(4;0.05)}$ or the $p-value < \alpha$, concluding that at least one predictor variable significantly influences the ICT-DI in Indonesia.

C. Individual Test

Individual parameter significance testing is conducted to determine the individual influence of predictor variables on the response variable. The Wald test statistic is used with a significance level α 5% resulting in a $Z_{0,025}$ value of 1,96. The testing results for each predictor variable are presented in Table 3.

TABLE III
INDIVIDUAL TEST

Variable	Z_{hitung}#	$p-val$	Decisions
percentage of average household telecommunication consumption.	-3.00	0.003	reject H_0
gdp per capita	-2.38	0.017	reject H_0
percentage of the population below the poverty line.	2.07	0.038	reject H_0
percentage of completion of high school education.	-2.83	0.005	reject H_0

Based on Table 3, all predictor variables have a significant impact on the ICT-DI. This is evidenced by the $Z_{calculated}$ value for the variable average percentage of household telecommunication consumption, which is -3.00 with a p-value of 0.003, where the $|Z_{hitung}| > Z_{(0,025)}$ or the p-value $<\alpha$, leading to the decision to reject H_0. Next, the computed Z_{hitung} for the variable Gross Domestic Product (GDP) per capita is -2.38 with a p-value of 0.017. Considering that the absolute value of $|Z_{hitung}|$ score is greater than the critical value $Z_{(0,025)}$ or p-value $<\alpha$, the decision is to reject H_0. Furthermore, the Z_{hitung} score for the variable percentage of the population below the poverty line is 2.07 with a p-value of 0.038 Again, as the absolute value of $|Z_{hitung}|$score is greater than $Z_{(0,025)}$ or p-value $<\alpha$, the decision is to reject H_0. Lastly, the Z_{hitung} score for the variable percentage of completion of high school education is -2.83, with a p-value of 0.005. With the absolute value of$|Z_{hitung}|$score greater than $Z_{(0,025)}$or p-value $<\alpha$, the decision is to reject H_0. Therefore, the conclusion is drawn that the variables, namely the average household telecommunication consumption, GDP per capita, percentage of the population below the poverty line, and percentage of completion of high school education, individually influence the Information and Communication Technology Development Index (ICT-DI) in Indonesia.

D. Model Fit Test

The fitness of the ordinal probit panel data regression model with random effects is examined through the Likelihood Ratio Test with a significance level α of 5%. The results of the model fit test are presented in Table 4.

TABLE IV
UJI KESESUAIAN MODEL

Measurement	Value
Λ	13.39
$\chi^2_{(4;0.05)}$	3.842
$p-value$	0.0001

Based on the results from Table 4, the value of Λ is 13.39 and the value of $\chi^2_{(4;0.05)}$is 3.842 with the p-value of 0.0001. The decision is made to reject H_0 because $\Lambda > \chi^2_{(4;0.05)}$ or p-value < α, concluding that there is enough variability among provinces to justify the ordinal probit panel data regression model with random effects.

E. Accuracy Classification Test

After modeling and estimating the model, the next step is to calculate the accuracy classification value based on 1-APPER.

TABLE V
ICT-DI CATEGORY CLASSIFICATION FUSION MATRIX XTOPROBIT MODEL

Actual Category	Prediction Category 1	2	3	Actual Totals
1	0	2	0	2
2	0	141	3	144
3	0	45	13	58
prediction totals	0	188	16	204

Then, the calculation of the APPER value is performed.

$$APPER_{xtoprobit} = \frac{2+3+45}{204} = 0.245098$$

$$1 - APPER_{xtoprobit} = 1 - 0.245098 = 0.754902$$

Based on the above calculation, it is evident that the accuracy classification value for the xtoprobit model is 0.7549 or 75.49%. This signifies that the xtoprobitmodel has been correctly classified at a rate of 75.49%,while the remaining 24.50% is classified less accurately or incorrectly.

IV. CONCLUSIONS

Fluctuating trends are observed in the ICT Development Index (ICT-DI) values in each province in Indonesia from 2017 to 2022. The results from the ordinal probit panel data regression modeling with random effects show that all predictor variables, such as the percentage of average household telecommunication consumption, GDP per capita, percentage of the population below the poverty line, and percentage of completion of high school education, have a significant influence on the response variable, namely the IC-TDI value. The accuracy classification of the ordinal probit panel data regression model with random effects reaches 75.49%. Based on these analysis results, the researcher recommends to the Ministry of Information and Communication to design appropriate programs to equalize the factors supporting ICT development among provinces. By optimizing the implementation of these programs, it is hoped that they can contribute to the improvement of the ICT-DI value in Indonesia.

REFERENCES

[1] Nasution, R. D. Pengaruh Modernisasi dan Globalisasi terhadap Perubahan Sosial Budaya di Indonesia. *Jurnal Penelitian Komunikasi Dan Opini Publik*, *21*(1), pp. 30–42, 2017.

[2] Yoga, S. Perubahan Sosial Budaya Masyarakat Indonesia Dan Perkembangan Teknologi Komunikasi. *Jurnal Al-Bayan*, *24*(1), pp. 29–46, 2019. https://doi.org/10.22373/albayan.v24i1.3175

[3] Wardhana, A., Kharisma, B., & Lisdiyanti, T. Teknologi Informasi Komunikasi Dan Pertumbuhan Ekonomi Wilayah Barat Dan Timur Indonesia Periode 2014-2018. *E-Jurnal Ekonomi Dan Bisnis Universitas Udayana*, *11*, pp. 1103, 2020. https://doi.org/10.24843/eeb.2020.v09.i11.p05

[4] CHOLIK, C. A. Perkembangan Teknologi Informasi Komunikasi/Ict Dalam Berbagai Bidang. *Jurnal Fakultas Teknik UNISA Kuningan*, *10*(2), pp. 39-46, 2021.

[5] Badan Pusat Statistik,*Indeks Pembangunan Teknologi, Informasi dan Teknologi,*2022.

[6] ITU, *International Communiaction Union*, 2016.

[7] United Nations Development Programme, *Sustainable Development Goals*, 2022.

[8] DSGI,*The Digital Skills Gap Index*, 2021

[9] World Digital Competitiveness Ranking.*IMD World Competitiveness Center,* 2022.

[10] Kominfo,*Dukung Transformasi Digital Indonesia, Pemerintah Bangun Infrastuktur Maju Satu Dekade,* 2020.

[11] Hsiao, C. *Analysis of Panel Data*. Cambridge University Press:New York, 2003.

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

[13] Menard, S. *Logistic Regression: From Introductory to Advanced Concepts and Aplications.* California: Sage Publication, 2010.

[14] Deb, P., Norton, E. C., & Manning, W. G. Health econometrics using Stata, 3, College Station, TX: Stata press, 2017

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

