International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

SEM (Structural Equation Modelling) Analysis Using PLS (Partial Least Square) Method on Poverty Data in Indonesia

Muh. Qardawi Hamzah[1], Georgina Maria Tinungki[2], Erna Tri Herdiani[3]
[1, 2, 3]Department of Statistics, Faculty of Mathematics and Natural Sciences, Hasanuddin University, Makassar 90245, Indonesia
Corresponding Author's: qardawihamzah@gmail.com

Abstract— *SEM (Structural Equation Modelling) is a multivariate analysis technique used to explain the relationship between latent variables and indicators to their latent variables. There are two types of SEM, namely covariant-based SEM or known as CB-SEM and variant-based SEM. PLS (Partial Least Square) is one of the methods that can be used to determine the relationship that occurs in variables with a variant approach. This study used four latent variables, namely poverty, economy, health and human resources (Human Resources). This study used secondary data obtained from the publication of BPS (Central Statistics Agency) Indonesia. Based on the results of SEM research with PLS, a significant relationship was obtained between the latent variables of health and HR (Human Resources) to poverty, and the relationship that was not significant between the latent variables of the economy to poverty. The R-square value on the endogenous latent variable of poverty is* 0.61 *, it can be expressed that the variable l exogenous aten affects the endogenous latent variable by* 61% *.*

Keywords— *Masukkan kata atau frasa kunci dalam urutan abjad, dipisahkan oleh titik dua.*

I. INTRODUCTION

Poverty is a problem faced by all elements of society, both at the regional and national levels, so special and sustainable attention is needed to deal with this problem. Many programs have been or are temporarily implemented to reduce the poverty rate of an area. However, the current level of poverty is still at the numbers and is9.8% still a problem that has not been able to be solved. Improving the welfare of society is the main goal of economic development. Welfare is closely related to the status of ownership of goods, so that people who are in the category of poor in the future are defined as people who do not have sufficient income or consumption to fall into the category of prosperous. Poverty can also simply be interpreted as a society that has no place to live, lacks food or poor welfare (Maipita: 2014). Research on poverty involves many factors or variables.

SEM (Structural Equation Modelling) is one of the multivariate analysis techniques used to analyze data involving many variables. SEM (Structural Equation Modelling) is a statistical analysis technique that can be used to analyze the relationship between variables. In the SEM (Structural Equation Modelling) method, there are two types of variables, namely latent variables and observed variables. Latent variables are variables that cannot be observed or measured directly, there are two types of latent variables, namely endogenous latent variables and exogenous latent variables. Observed variables are variables that can be observed or measured and are used to define latent variables, observed variables also known as indicators (Schumaker & Lomax, 2004). There are two important reasons underlying the use of SEM. First, SEM has the ability to estimate the relationship between variables that are multiple relationships which in this relationship are formed in the structural model described by the relationship between endogenous and exogenous latent variables. Second, SEM has the ability to describe models of relationships between latent variables and each of their indicators (Otok, 2013). One of the steps in metoode SEM (Structural Equation Modelling) is parameter estimation.

Parameter estimation is theestimation of the model variant-covariant matrix with the sample variant-covariant matrix using the estimation function. The purpose of parameter estimation is to obtain a convergent model variant-covariance matrix on a sample or population variant-covariant matrix with a smaller residual matrix (Hair et al., 2010). Many estimation methods have been developed, initially, SEM was developed using covariance or better known as CB-SEM (covariance based SEM) whose use is based on the theory that has been developed by experts. So the purpose of CB-SEM is to confirm or test a model built on the basis of existing theories. Before using CB-SEM there is an assumption that must be met, namely that the data must be normally distributed multivariately. If the data is not distributed normally, the estimated results obtained are not good and have not been able to approach the sample or population (Haryono & Wardoyo, 2015). To overcome the normal distribution in CB-SEM (Covariance Based SEM), a variant-based SEM (Structural Equation Modelling) was developed or more popularly referred to as PLS (Partial Least Square) which was first developed in 1975. SEM with PLS can overcome deficiencies in CB-SEM, namely that it can be used to analyze data that does not meet the normal assumptions of multivariates and a strong theoretical basis (Chin, 2002) . PLS-SEM is a statistical technique that combines factor analysis and regression analysis with separate estimation procedures between latent variables and indicators (Ghozali, 2011).

Parameter estimation carried out in PLS-SEM goes through three stages, namely weight estimate, path estimate to obtain inner model and outer model, finally means and location parameters (Hoyle, 1999). However, in PLS-SEM, the distribution of data is unknown so that no significance value is obtained from the model. To overcome this, a resampling technique is used. A commonly used method is Bootstrap. Bootstrapping methods have been developed as tools to help

Muh. Qardawi Hamzah, Georgina Maria Tinungki, and Erna Tri Herdiani, "SEM (Structural Equation Modelling) Analysis Using PLS (Partial Least Square) Method on Poverty Data in Indonesia," *International Research Journal of Advanced Engineering and Science*, Volume 7, Issue 4, pp. 131-135, 2022.

International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

overcome the assumption of normality, and assess the degree of significance or probability of the results of the analysis (Efron & Tibshirani, 1993).

Based on the explanation above, a study was carried out with SEM (Structural Equation Modelling) with the PLS (Partial Least Square) parameter estimation method to obtain a structural equation model in the case of poverty in Indonesia. The latent variables in this study are poverty, human resources, health and economy. The purpose of this study is to obtain a poverty model with the PLS-SEM method. In the end, a model of poverty will be obtained in Indonesia.

II. Literatur Review

SEM (Structural Equation Modelling) is an analytical method used to find out causal relationships between variables or even involve moderating or intervening variables (Haryono & Wardoyo, 2015). SEM is built by two models namely (Wijaya. T, 2011)

1. Measurement model or better known as measurement model, which explains the relationship that occurs between indicators and latent variables. The endogenous variable measurement model can be seen from the following equation

$$Y = \Lambda_y \eta + \epsilon \quad (1)$$

In the form of a matrix is expressed as follows

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{bmatrix} = \begin{bmatrix} \lambda_{y_{11}} & \lambda_{y_{12}} & \cdots & \lambda_{y_{1m}} \\ \lambda_{y_{21}} & \lambda_{y_{22}} & \cdots & \lambda_{y_{2m}} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{y_{p1}} & \lambda_{y_{p2}} & \cdots & \lambda_{y_{pm}} \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_m \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_p \end{bmatrix}$$

Description :

Y : endogenous manifest variable vector
Λ_y : matrix of measurement coefficients (loading factor)
ϵ : Measurement error
p : number of indicators of endogenous variables
m : number of endogenous variables

The measurement model for exogenous variables can be seen from the following equation

$$X = \Lambda_x \xi + \delta \quad (2)$$

In the form of a matrix is expressed as follows

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix} = \begin{bmatrix} \lambda_{x_{11}} & \lambda_{x_{12}} & \cdots & \lambda_{x_{1n}} \\ \lambda_{x_{21}} & \lambda_{x_{22}} & \cdots & \lambda_{x_{2n}} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{x_{q1}} & \lambda_{x_{q2}} & \cdots & \lambda_{x_{qn}} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_q \end{bmatrix}$$

X : exogenous manifest variable vector
Λ_x : matrix of measurement coefficients (loading factor)
δ : Measurement error
q : Number of indicators of endogenous variables
n : number of exogenous variables

2. Structural model or known as structural model is to explain the relationship that occurs between latent variables through simultaneous equations. The structural model can be seen from the following equation.

$$\eta = B\eta + \Gamma\xi + \zeta \quad (3)$$

In the form of a matrix we state the following

$$\begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_m \end{bmatrix} = \begin{bmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1n} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{m1} & \beta_{12} & \cdots & \beta_{mn} \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_m \end{bmatrix} + \begin{bmatrix} \Gamma_{11} & \Gamma_{12} & \cdots & \Gamma_{1n} \\ \Gamma_{21} & \Gamma_{22} & \cdots & \Gamma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \Gamma_{m1} & \Gamma_{12} & \cdots & \Gamma_{mn} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_m \end{bmatrix} + \begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \vdots \\ \zeta_m \end{bmatrix}$$

β : matrix coefficient for endogenous latent variables
Γ : matrix coefficient for endogenous latent variables
η : endogenous latent variables
ξ : exogenous latent variables
ζ : error

There are three stages that will be passed in estimating SEM (Structural Equation Modelling) parameters with the PLS (Partial Least Square) method, namely weight estimation (weight estimate) which will produce variable values latent, path estimate that will produce relationships between variables and variables and their respective indicators, means and location of parameters on latent variables and indicators (Otok, 2013) .

III. Research Methods

The data used in this study is secondary data. The data was taken from the publication of BPS (Central Statistics Agency) RI. The latent variables used in this study are poverty, health, human resources (Human Resources) and the economy. Some of the indicators used on each latent variable will be presented in table 1.

The method and analysis used to achieve the objectives of this study is to analyze SEM-PLS modeling through the following steps

1. Develop theory-based conceptual models
2. Create a path chart that explains the pattern of relationships between latent variables and their indicators. In PLS-SEM, it is divided into 3 steps, namely

a. Inner model (structural model), which is a pattern that describes the relationship of latent variables, can be seen in the following equation

$$\eta = B\eta + \Gamma\xi + \zeta$$

b. Outer model (measurement model) which is a pattern that describes the relationship between latent variables and their indicators, can be seen in the following equation

$$Y = \Lambda_y \eta + \epsilon$$
$$X = \Lambda_x \xi + \delta$$

c. Weight relation is defined because the inner model and outer model provide specifications that are followed in the estimation of the PLS algorithm. here is the case value for each PLS estimated latent variable

$$\xi_b = \Sigma_{kb} w_{kb} x_{kb} \quad (4)$$
$$\eta_i = \Sigma_{ki} w_{ki} y_{ki} \quad (5)$$

w_{kb} and w_{ki} is k the weight used to form the estimation of latent changers ξ_b and η_i

3. Perform parameter estimation
4. Model evaluation

a. Outer model, evaluation of the measurement model is carried out with two types of testing, namely validity and

Muh. Qardawi Hamzah, Georgina Maria Tinungki, and Erna Tri Herdiani, "SEM (Structural Equation Modelling) Analysis Using PLS (Partial Least Square) Method on Poverty Data in Indonesia," *International Research Journal of Advanced Engineering and Science*, Volume 7, Issue 4, pp. 131-135, 2022.

International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

reliability. The validity test consists of two stages, namely convergent validity and discriminant. The reliability test is seen by paying attention to the composite reliability value.

b. Inner model, the evaluation of which is carried out by looking at the R-square criteria.

TABLE 1. Latent variables and indicators

No.	Latent Variables	Indicators
1	Poverty	Percentage of poor population (Y1)
		Poverty depth index (Y2)
		Poverty severity index (Y3)
2	Economics	Percentage of households with non-owned homeownership status (X1)
		Percentage of informal labor other than the agricultural sector (X2)
		Percentage of households with rental house ownership (X3)
		Percentage of per capita expenditure for non-food per month (X4)
3	HR (Human Research)	Human development index (X5)
		Average length of schooling (X6)
		Literacy rate (X7)
		School participation rate of children aged 7-12 years (X8)
4	Health	Life expectancy figures (X9)
		Percentage of households whose lighting uses electricity (X10)
		Percentage of households using proper sanitation (X11)
		Percentage of households with per capita floor area> $8m^2$ (X12)
		Percentage of households with a non-ground house floor area (X13)
		Percentage of households whose last birth process was helped by health workers (X14)

IV. PLS-SEM Analysis Results

1. Outer model (measurement model)

Evaluation carried out on the outer model (measurement model) will assess the level of validity and ralibility of the model. The model is first evaluated by taking into account the validity consisting of two, namely first the convergent validity consisting of the loading factor and the Average Validity Extracted (AVE), the second validity of discriminants by paying attention to cross loading values. Level of reliability by paying attention to composite reliability and Cronbach's alpha.

Based on the outer model outer model produced in the figure below, there are 17 indicators that reflect four latent variables. The first step is to evaluate the standardized loading factor with the criteria that the value of the loading factor ≥ 0.7 can be said to be valid to be able to measure latent variables. The value of the loading factor > 0.5 is still acceptable but if the value is< 0.5 then the indicator must be removed or removed from the model. In the latent variable of health there are 6 indicators with values X9 of 0.7, X10 amount 0.929, X11 amount 0.770, X12 amount of, 0.737, X13 amount of 0.757, X14 amount of 0.847. On the latent variable of health, all its indicators have values ≥ 0.7 and can be said to be valid. In economic variables there are 4 indicators with values X1 of 0.916, X2 amount 0.823, X3 amount 0.922 and X4 amount 0.657

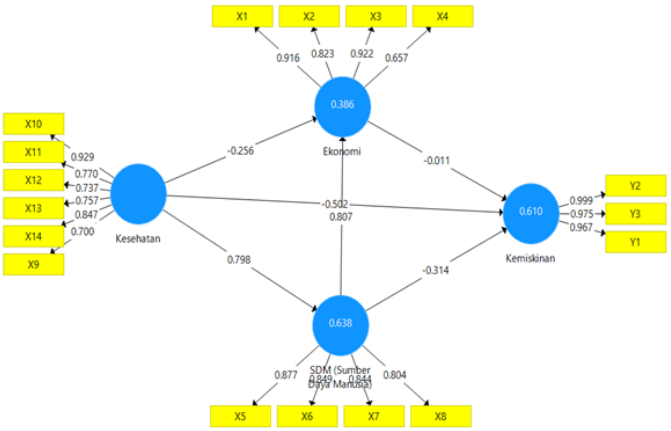

Fig. 1. Outer model

All indicators on economic variables can be said to be valid. In the latent variable HR (Human Resources) there are 4 indicators with X5 a value 0.877 of X6, 0.849, x7 as large as 0.844, X8 0.804. all indicators in the latent variable HR (Human Resources) can be said to be valid. The latent variable of poverty there are 3 indicators with a Y1 value0.967, Y2 of 0.999, Y3 0.975.

The second stage of evaluation is by paying attention to the value of the AVE (Average Variance Extraxted). The criterion that must be met for an AVE is the value of> 0.5 . The following will be shown the results of the analysis of AVE data for each latent variable.

TABLE 2. AVE Value

Variable	Average Variance Extracted (AVE)
Economics	0.699
Poverty	0.961
Health	0.63
HR (Human Resources)	0.712

Based on the table above, the AVE value on the economic latent variable is0.699 , the latent variable of poverty is equal, the 0.961latent variable of health is as much, 0.63and the latent variable of HR (Human Resources) is as much 0.712as. All latent variables on the model have an AVE value> 5. So it can be stated that the PLS model can meet the requirements of a good convergent.

The next evaluation is the validity of the discriminant which is carried out by paying attention to the value of the cross loading. The condition that must be met in order for the model to meet the validity of the discriminant is that the value obtained must be ≥ 0.7. Below the crossloading value of each indicator.

From the table below, it can be seen that each indicator that measures its latent variable has a greater cross loading value compared to indicators against other latent variables. So it can be concluded that the model meets the requirements of discriminant validity.

Evaluate the reliability to corner the outer model. It aims to find out the accuracy, consistency of the indicators in measuring their latent variables. On reliability will be measured composite reliability and Cronbach's alpha. An acceptable value is ≥ 0.7 .

Muh. Qardawi Hamzah, Georgina Maria Tinungki, and Erna Tri Herdiani, "SEM (Structural Equation Modelling) Analysis Using PLS (Partial Least Square) Method on Poverty Data in Indonesia," *International Research Journal of Advanced Engineering and Science*, Volume 7, Issue 4, pp. 131-135, 2022.

International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

TABLE 3. Cross loading value

Indicators	Economics	Poverty	Health	HR (Human Resources)
X1	0.916	-0.244	0.263	0.529
X2	0.823	-0.496	0.369	0.522
X3	0.922	-0.279	0.276	0.497
X4	0.657	-0.262	0.396	0.456
X5	0.685	-0.675	0.8	0.877
X6	0.712	-0.496	0.504	0.849
X7	0.397	-0.575	0.602	0.844
X8	0.19	-0.671	0.755	0.804
X9	0.433	-0.564	0.7	0.579
X10	0.279	-0.733	0.929	0.788
X11	0.402	-0.476	0.77	0.632
X12	-0.18	-0.587	0.737	0.482
X13	0.426	-0.644	0.757	0.671
X14	0.354	-0.57	0.847	0.583
Y1	-0.481	0.967	-0.717	-0.67
Y2	-0.383	0.999	-0.754	-0.719
Y3	-0.302	0.975	-0.754	-0.73

TABLE 4. Reliability test

Variable	Cronbach's Alpha	Composite Reliability
Economics	0.85	0.901
Poverty	0.98	0.987
Health	0.88	0.91
HR (Human Resources)	0.866	0.908

Based on the table of data analysis results above. The value of cronbach's alpha for the economic latent variable is as large 0.85 , the latent variable of poverty is by 0.98, the latent variable of health is as large 0.88, and the latent variable hr (Human Resources) is as large 0.866 as. composite reliability value for economic latent variables of 0.901 latent variables of poverty by 0.987, latent variables of health by 0.91, and latent variables of HR (Human Resources) of 0.908. The values of cronbach's alpha and composite reliability are obtained from each latent variable $>$ 0.7. So, it can be stated that it meets the requirements of good reliability.

2. Inner model (structural model)

At the inner stage of the model, the relationship between the latent variables will be predicted. The initial step is to pay attention to the R-square value generated by the measurement for each latent changer.

TABLE 5. R-square values

Variable	R Square
Economics	0.386
Poverty	0.61
HR (Human Resources)	0.638

Based on the table above.

$$Q^2_{PLS} = 1 - (1 - R^2_1)(1 - R^2_2)(1 - R^2_3)$$
$$= 1 - (1 - 0.61)(1 - 0.386)(1 - 0.638)$$
$$= 1 - (0.39)(0.614)(0.362)$$
$$= 0.913$$

Partial Least Square poverty can be explained by variables of economic, HR (Human Resources), and health as much and the rest is explained by variables outside the model. 91.3%

Next is to measure the inner model with bootstrapping techniques on the outer model by paying attention to the T-statistical value. It is carried out as a basis for knowing the significance of the model. The value to be met is the t-statistic$\geq$ 1.96 for $\alpha = 0.05$

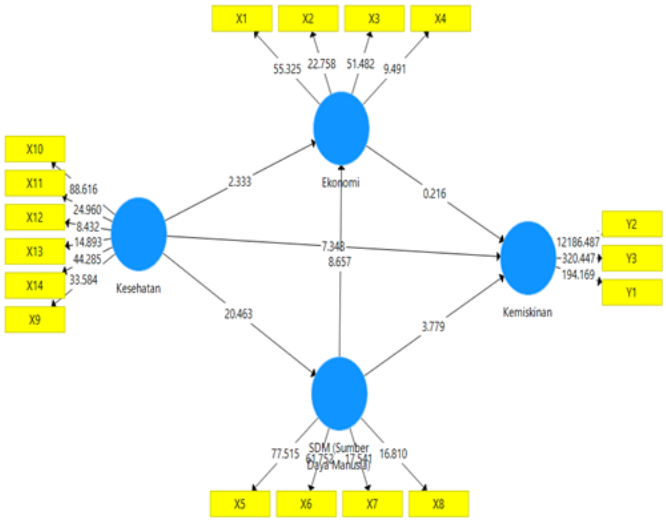

Fig. 2. Inner model

Based on the figure above, the bootstrapping results show that the latent variable of health to poverty has a t-count of7.348 . The latent variable of HR (Human Resources) to poverty is. 3.779 The latent variables of the economy to poverty are as large as0.216. Varibel latent exogenous health and HR (Human Resources) has a t-count$>$ 1.96 with $\alpha = 0.05$ so that it can besignificantly calculated. The exogenous latent variable of the economy has a t-count of$<$ 1.96 with $\alpha = 0.05$ so it can be concluded that it is insignificant.

TABLE 6. Path coefficient value , P-Value,

Latent Variables	path coeficient	T Statistics	P-Value	information
Economy -> Poverty	-0.011	0.216	0.829	Insignificant
Kesehatan_ -economic >	-0.256	2.333	0.02	Significant
Kesehatan_ -> Poverty	-0.502	7.348	0	Significant
Kesehatan_ -> HR (Human Resources)	0.798	20.463	0	Significant
HR (Human Resources) -Economic >	0.807	8.657	0	Significant
Human Resources (HR) -> Poverty	-0.314	3.779	0	Significant

Based on the table above, it can be stated that the latent variables of health and HR (Human Resources) have a negative and significant influence on the latent variables of poverty. The latent variable health has a coefficient path -0.502 of magnitude-0.314 and $(-)$. coefficient has a sign, it indicates

Muh. Qardawi Hamzah, Georgina Maria Tinungki, and Erna Tri Herdiani, "SEM (Structural Equation Modelling) Analysis Using PLS (Partial Least Square) Method on Poverty Data in Indonesia," *International Research Journal of Advanced Engineering and Science*, Volume 7, Issue 4, pp. 131-135, 2022.

that the better the quality of health and HR (Human Resources) it will be lowering the poverty rate. Economic latent variables have a path coefficient -0.011 so coefficient a negative $(-)$ sign this indicates that the economy of better can lower the poverty but not significantly.

V. Conclusion

Based on the results of the tests that have been carried out, the total influence score of exogenous latent variables on endogenous is obtained. 0.913 It can be interpreted that the latent variables of health, HR (Human Resources), and Economics are able to explain the latent variables of poverty by 91.3% and the rest 8.7% are explained by other latent variables and are not explained in this study. It can be stated that the model from SEM-PLS is powerful. So it can be stated that exogenous latent variables have a significant effect on endogenous latent variables. The resulting model is

$$\text{kemiskinan} = -0.011\ \text{ekonomi} - 0.502\ \text{kesehatan} - 0.314\ \text{SDM}$$

The advice that can be given in later studies is to add indicators for each latent variable and include other latent variables.

References

1. Chin, W. W. (2002). Partial Least Squares For Researchers : An overview and presentation of recent advances using the PLS approach. January 2000.
2. Efron, B., & Tibshirani, R. J. (1993). An Introduction to the Bootstrap. In D.R. Cox, D.V. Hinkley, N. Reid, D. B. Rubin, & B. W. Silverman (Eds.), An Introduction to the Bootstrap (1 st). Springer-Science+Business Media, B.V. https://doi.org/10.1007/978-1-4899-4541-9
3. Ghozali I. 2011. Concepts, Techniques and Applications Using SmartPLS 2.0M3 Program. Semarang (ID): UNDIP Issuing Agency
4. Hair, J. F. J., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). Multivariate Data Analysis (7th Edition) by Joseph F. Hair, William C. Black, Barry J. Babin, Rolph E. Anderson (z-lib.org).pdf (p. 761).
5. Haryono, S., & Wardoyo, P. (2015). Structural Equation. In S. Hamid Mintardja (Ed.), Dictionary of Statistics & Methodology (1st ed.). Pt. Major Personnel Intermedia. https://doi.org/10.4135/9781412983907.n1909
6. Hoyle, R. H. (1999). Structural Equation Modeling Analysis with Small Samples using Partial Lesst Squares. Statistical Strategies for Small Sample Research, January 1999, 34.
7. Maipita, Indra. 2014. Measuring Poverty and Income Distribution, Yogyakarta. UPP STIM YKPN.
8. Otok, B. W. (2013). Modeling Poverty In East Java With. Statistics, 1(2).
9. Schumaker, R. E., & Lomax, R. G. (2004). A Beginner's Guide to Structural Equation Modeling (2nd ed.). Lawrance Erlbaum Associates.
10. Wijaya. T. (2011). Structural Equation Modeling Analysis Using Amos. Yogyakarta. Atma Jaya University Yogyakarta, 1315201008.

Muh. Qardawi Hamzah, Georgina Maria Tinungki, and Erna Tri Herdiani, "SEM (Structural Equation Modelling) Analysis Using PLS (Partial Least Square) Method on Poverty Data in Indonesia," *International Research Journal of Advanced Engineering and Science*, Volume 7, Issue 4, pp. 131-135, 2022.