Response-Based Units Segmentation Partial Least Square (REBUS PLS) Analysis on Human Development Index Data in Indonesia

Esra Rombeallo[1], Georgina Maria Tinungki[2], Anna Islamiyati[3]

[1, 2, 3]Department of Statistics, Faculty of Mathematics and Natural Sciences, Hasanuddin University, Makassar 90245, Indonesia

Corresponding Author's: esrarombeallo01@gmail.com

Abstract— Structural Equation Modeling is a multivariate analysis technique that can be used in analyzing path equation models. There are two types of Structural Equation Modeling, namely Covariance-Based Structural Equation Modeling and Partial Least Squares Structural Equation Modeling. Partial Least Square is an analytical method that can be used for various scales of data measurement, can be used for small sample sizes, and demands on the residual distribution. Partial Least Square Structural Equation Modeling assumes that the data is homogeneous. This assumption is irrational because data can come from several different segments. This problem is called heterogeneity which greatly influences the analysis results and conclusions. Response-Based Units Segmentation Partial Least Square is a method that can segment observation units and detect heterogeneity in structural equations and measurement equations. In this study, the variables used consisted of four latent variables with 14 manifest variables. The values of the closeness measure are used as a basis for the results of this study, it can be seen that when 2 classes/segments are formed, the composition of class/segment 1 consists of 9 provinces, while class/segment 2 consists of 25 provinces. The conclusion is that REBUS PLS can detect heterogeneity in the Partial Least Square Structural Equation Modeling model with the Prediction Relevance (Q^2) value in each formed segment being greater than the Q^2 value in the global model. Meanwhile, the global model's Goodness of Fit (GoF) value is 0.598, local model 1 is 0.856, and local model 2 is 0.434. A high GoF value also indicates better structure and measurement.

Keywords— Structural Equations Modelling; Partial Least Square Structural Equations Modelling; Response Based Units Segmentation Partial Least Square; Closeness Measure; Human Development Index.

I. INTRODUCTION

According to the United Nations Development Program (UNDP), human development is a development model aimed at expanding options that can be grown through efforts to empower the population. Human development uses a measurement introduced by UNDP in 1990, namely the Human Development Index (HDI) (Suntono and A'laa, 2021). In 2017, Indonesia's human development index (HDI) experienced an increase compared to the previous year, which reached 70.81 or grew by 0.90 percent (Statistics, 2020). Based on the results of the HDI calculation using the new method, Indonesia's HDI figure for 2020 is 71.94, while the HDI value for 2019 is 71.92. Research on the human development index involves many factors and variables.

Structural Equation Modeling (SEM) is a statistical analysis that can be used to solve complicated problems involving many variables, namely being able to estimate the relationship between variables that are multiple relationships with the output being a measurement model of some indicators and a structural model composed of some variables latent (Anggita et al., 2019). The two types of SEM are Covariance-Based Structural Equation Modeling (CB-SEM) and Partial Least Squares Structural Equation Modeling (PLS-SEM) (Hair et al, 2012). CB-SEM is required for a strong theoretical base, fulfills various parametric assumptions, and meets the model is goodness of fit test. Partial Least Square (PLS) is a powerful and flexible analytical method because it can be used for various scales of data measurement, can be used for small sample sizes, and demands on residual distribution (does not require multivariate normal distribution) (Patih and Halistin, 2020).

In PLS-SEM, it is generally assumed that the data is homogeneous or that the respondents come from the same characteristics (Reagan, 2016). This assumption is irrational because data can come from several segments or units of the respondents collected. Because of this, without paying attention to the homogeneity, it will produce a biased analysis. This problem is called heterogeneity which greatly influences the analysis results and conclusions. Ignoring unobserved heterogeneity will cause estimates of parameter values to be biased, leading to type I or II inference errors and generating invalid conclusions (Jedidi et al., 1997).

Response Based Units Segmentation Partial Least Square (REBUS PLS) was introduced by Trinchera (2007) and Vinzi et al., (2008). REBUS PLS can segment the observation units and estimate the parameters of each local model in the formed segments (Trinchera, 2007). REBUS PLS has advantages; for example, REBUS PLS can detect heterogeneity in structural equations and measurement equations. Another advantage is that REBUS PLS does not require distribution assumptions on either manifest or latent variables (Trinchera, 2007).

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 human development index in Indonesia. The latent variables in this study are poverty, human resources, health and economy. The purpose of this study is to obtain a human development index model with the REBUS PLS method. In the end, a model of human development index will be obtained in Indonesia.

II. LITERATUR REVIEW

The Partial Least Square Structural Equation Modeling (PLS-SEM) method is a method that maximizes the variance expressed in dependent variables explained by independent variables. PLS-SEM has two types of variables (variables), namely latent variables and manifest variables. Latent variables are usually called variables that cannot be measured. The two types of latent variables are exogenous latent variables (ξ) and endogenous latent variables (η). Exogenous latent variables can be called independent variables, while endogenous latent variables are dependent variables. Manifest variables are variables that can be observed or can be measured empirically.

1. A structural model is a model that describes the relationship between latent variables.

$$\eta = \beta\eta + \Gamma\xi + \zeta \quad (1)$$

In matrix form, as follows

$$\begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_m \end{bmatrix} = \begin{bmatrix} \beta_{11} & \beta_{12} & \dots & \beta_{1n} \\ \beta_{21} & \beta_{22} & \dots & \beta_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{m1} & \beta_{12} & \dots & \beta_{mn} \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{bmatrix}$$
$$+ \begin{bmatrix} \gamma_{11} & \gamma_{12} & \dots & \gamma_{1n} \\ \gamma_{21} & \gamma_{22} & \dots & \gamma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{m1} & \gamma_{12} & \dots & \gamma_{mn} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix} + \begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \vdots \\ \zeta_m \end{bmatrix}$$

Where, η is the vector of endogenous latent variables, β is the path coefficient matrix of endogenous latent variables, Γ is the path coefficient matrix connecting endogenous and exogenous latent variables, ξ is the exogenous latent variable vector, and ζ is the residual vector.

2. The relationship between the latent variable and the indicator is specified with the measurement model. The relationship formed is a reflexive relationship with a causality direction from latent variables to indicators. In the case of exogenous latent variables, a measurement model is as follows.

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

In matrix form, as follows

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_q \end{bmatrix} = \begin{bmatrix} \lambda_{x11} & \lambda_{x12} & \dots & \lambda_{x1n} \\ \lambda_{x21} & \lambda_{x22} & \dots & \lambda_{x2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{xq1} & \lambda_{xq2} & \dots & \lambda_{xqn} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_q \end{bmatrix}$$

Where, X is the indicator variable vector for exogenous latent variables, Λ_X is the coefficient matrix of X to ξ (loading factor), ξ is the exogenous latent variable vector, and δ is the residual vector.

The measurement model for endogenous latent variables is as follows.

$$Y = \Lambda_y \eta + \varepsilon_y$$

In matrix form, as follows

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{bmatrix} = \begin{bmatrix} \lambda_{y11} & \lambda_{y12} & \dots & \lambda_{y1m} \\ \lambda_{y21} & \lambda_{y22} & \dots & \lambda_{y2m} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{yp1} & \lambda_{y12} & \dots & \lambda_{ypm} \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_m \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_p \end{bmatrix}$$

Where, Y is the indicator variable vector for endogenous latent variables, Λ_Y is the coefficient matrix of Y to η (loading factor),

η is the vector of endogenous latent variables, and ε is the residual vector.

Parameter estimation in Partial Least Square Structural Equation Modeling (PLS-SEM) uses the Ordinary Least Square method. The calculation process is carried out by iteration and will stop when it has found a convergent point (Muliati and Handoyo, 2019). Wold proposed $\left|\tilde{w}_{jk}^{S-1} - \tilde{w}_{jk}^S\right| < 10^{-5}$ as a convergence criterion (Wold, 1982). Parameter estimation of Partial Least Square Structural Equation Modeling (PLS-SEM) includes a weight estimate to obtain a latent variable score, a path estimate that links between latent variables, namely the beta coefficient (β), gamma coefficient (γ), and lambda coefficient (λ), as well as the average estimate-average and location parameters (Henseler et al., 2009).

REBUS PLS is an iterative algorithm that can group or segment observation units and estimate the parameters of each local model in the segments formed without requiring any kind of distribution assumptions either on manifest or latent variables (Trinchera, 2007). The sources of heterogeneity are obtained through the distance of a unit of observation in the model. In REBUS PLS, the distance of a unit in the model is defined by considering the shape of the model formed based on the residuals in the structural model and the measurement model for all latent variables. The measure for assessing the distance between units and models in REBUS PLS is calculated from the residual commonality model, namely using the Closeness Measure (CM). The units in a cluster are updated by considering the Closeness Measurement (CM) value. CM is defined as follows (Vinzi et al., 2008).

$$CM_{ig} = \sqrt{\frac{\sum_{j=1}^{J} \sum_{q=1}^{P_j} \left[e_{iqjg}^2 / Com(\hat{\xi}_{qj}, \; x_{jq})\right]}{\frac{\sum_{i=1}^{N} \sum_{j=1}^{J} \sum_{q=1}^{P_j} \left[e_{iqjg}^2 / Com(\hat{\xi}_{qj}, \; x_{jq})\right]}{(n_g - m_g - 1)}}}$$
$$\times \sqrt{\frac{\sum_{j^*=1}^{J^*} \left[f_{ij^*g}^2 / R^2\left(\hat{\xi}_{j^*}, \{\hat{\xi}_{j's \; menjelaskan} \; \hat{\xi}_{j^*}\}\right)\right]}{\frac{\sum_{i=1}^{N} \sum_{j^*=1}^{J^*} \left[f_{ij^*g}^2 / R^2\left(\hat{\xi}_{j^*}, \{\hat{\xi}_{j's \; menjelaskan} \; \hat{\xi}_{j^*}\}\right)\right]}{(n_g - m_g - 1)}}} \quad (3)$$

Where, $Com(\hat{\xi}_{qj}, \; x_{jq})$ is the communal index of the q_{th} indicator of the j_{th} block within the g_{th} latent class; e_{iqjg} is the model residual size for the i_{th} unit in the g_{th} latent class that refers to the q_{th} manifest variable in the j_{th} block; f_{ij^*g} is the structural model residual on the i_{th} unit in the g_{th} latent class which refers to the j^*_{th} endogenous block; n_g is the number of units of the g_{th} latent class; m_g is the number of dimensions (because all manifest variables are reflective their value is always equal to 1) (Vinzi et al., 2008).

III. RESEARCH METHODS

The data to be used in this study is secondary economic data on the human development index, summarized from the publication of the Indonesian Statistics Center with research units of 34 provinces in the data year 2020. The variables used consist of four latent variables, where 3 are endogenous latent variables and 1 latent variable is exogenous, along with 14 manifest variables (indicators).

TABLE 1. Latent Variables and Manifest Variables

No	Latent Variables	Manifest Variables
1	Economy	X_1 = Percentage of households with non-owned housing status (%)
		X_2 = Percentage of informal workers other than the agricultural sector (%)
		X_3 = Percentage of households with rental housing (%)
2	Human Resources	Y_1 = Human development index
		Y_2 = Average length of schooling
		Y_3 = Literacy rate
		Y_4 = School enrollment rate for children aged 13-15 years
3	Health	Y_5 = Life expectancy
		Y_6 = Percentage of households whose lighting uses electricity
		Y_7 = Percentage of households using proper sanitation
		Y_8 = Percentage of households whose last birth was assisted by a health worker
4	Poverty	Y_9 = Percentage of Poor People (P_0) (%)
		Y_{10} = Poverty Depth Index (P_1)(%)
		Y_{11} = Poverty Severity Index (P_2) (%)

The methods and analysis carried out on the Response Based Units Segmentation Partial Least Square (REBUS PLS) algorithm are as follows.
1. The specification of the path model in PLS-SEM consists of three sets of relationships including:
   a. Designing a structural model is a model that describes the relationship between latent variables, the inner model equation (structural model) can be written as follows.

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

   b. Measurement model is the relationship between a latent variable and an indicator, the outer model equation (measurement model) can stated as follows.

   $$X = \Lambda_x\xi + \delta_x$$
   $$Y = \Lambda_y\eta + \varepsilon_y$$

   c. Weight Relation is to estimate latent variable scores based on how the inner model and outer model are specified. The values for each case of latent variables are estimated in PLS-SEM as follows.

   $$\xi_i = \sum_{h=1}^{H} w_{ki}X_{jh}$$
   $$\eta_j = \sum_{h=1}^{H} w_{kj}Y_{jh}$$

   Where w_{ki} and w_{kj} are the weights used to estimate the latent variables ξ_i and η_j.
2. Parameter estimation
3. Model evaluation
   a. The measurement model is evaluated by testing the validity and reliability tests. The validity test consists of convergent validation and discriminant validity. The reliability test is done by looking at the composite reliability value.
   b. The structural model was evaluated using the R-squared (R^2) criterion.
4. Cluster analysis of residuals
5. Local model estimation

6. Calculating the Closeness Measure (CM) value (Equation 3)
7. Segmentation based on the Closeness Measure (CM) value
8. Evaluate and interpret the output

The following conceptual framework illustrates the influence of latent variables and the manifest variables of each latent variable in this study.
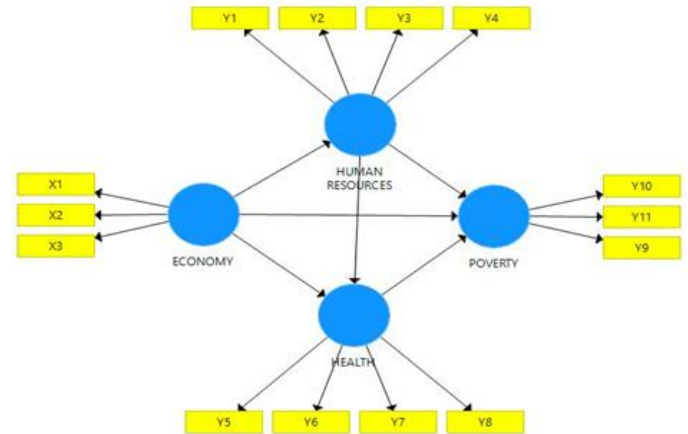

Fig. 1. Economic Pathway Model, Human Resources, Health, and Poverty on SmartPLS

IV. PLS-SEM AND REBUS-PLS ANALYSIS RESULTS

1. Evaluation of the Measurement Model
   a. Convergent Validity Test

TABLE 2. Convergent validity of latent variables with their manifest variables

No	Variable	λ
1	$X_1 \leftarrow \xi$	0.609
2	$X_2 \leftarrow \xi$	0.903
3	$X_3 \leftarrow \xi$	0.681
4	$Y_1 \leftarrow \eta_1$	0.901
5	$Y_2 \leftarrow \eta_1$	0.888
6	$Y_3 \leftarrow \eta_1$	0.782
7	$Y_4 \leftarrow \eta_1$	0.832
8	$Y_5 \leftarrow \eta_2$	0.655
9	$Y_6 \leftarrow \eta_2$	0.577
10	$Y_7 \leftarrow \eta_2$	0.898
11	$Y_8 \leftarrow \eta_2$	0.894
12	$Y_9 \leftarrow \eta_3$	0.974
13	$Y_{10} \leftarrow \eta_3$	0.999
14	$Y_{11} \leftarrow \eta_3$	0.982

Table 2 shows that in measuring the construct, the loading factor value for each indicator is ≥ 0.5, so all of these indicators are valid.
   b. Discriminant Validity Test

Discriminant validity testing is based on the Average Variance Extracted (AVE) value. Based on the calculation, the AVE value of the economic latent variable is 0.550, the Human Resources latent variable is 0.726, the health latent variable is 0.592, and the poverty latent variable is 0.970. These results indicate that the AVE value for each latent variable is more than 0.5. This shows that the latent variable is said to be valid.
2. Structural Model Evaluation

Esra Rombeallo, Georgina Maria Tinungki, and Anna Islamiyati, "Response-Based Units Segmentation Partial Least Square (REBUS PLS) Analysis on Human Development Index Data in Indonesia," *International Research Journal of Advanced Engineering and Science*, Volume 8, Issue 3, pp. 18-22, 2023.

Evaluation of the structural model can be measured by looking at the coefficient of determination R^2 value, which indicates the variability of the endogenous variables that exogenous variables can explain. The R^2 value for the Human Resources variable ($\eta 1$) is 0.455, which indicates that between variables describe a large relationship, the R^2 value for the Health variable ($\eta 2$) is 0.521, which indicates a large relationship, while the R^2 value for the Poverty variable ($\eta 3$) is 0.553 which shows a large relationship.

The next evaluation is the goodness of fit (GoF) to validate the model. Based on the processing results, a GoF value of 0.598 is obtained, which means it is included in the large GoF category so that it can be interpreted that the resulting model has a very good ability to explain empirical data.

3. Hypothesis Test

Based on the results of bootstrapping with a significance level of $\alpha = 5\%$ and test criteria: H_0 is rejected if $|T-statistics| > 1.96$ or $p-value < 0.05$, then the test for each path between latent variables can be seen in Table 5 below.

TABLE 3. T-statistics values

Relationship between Variable	Coefficient Parameter	Loading Factor ($\hat{\lambda}_h$)	$Se^*(\hat{\lambda}_h)$	$|T\text{-}statistics|$
$\xi \to \eta_1$	γ_{11}	0,715	0,102	7,01
$\xi \to \eta_2$	γ_{21}	0,422	0,206	2,05
$\xi \to \eta_3$	γ_{31}	-0,315	0,173	1,82
$\eta_1 \to \eta_2$	β_{21}	0,720	0,304	2,37
$\eta_1 \to \eta_3$	β_{31}	-0,685	0,396	1,73
$\eta_2 \to \eta_3$	β_{32}	-0,404	0,304	1,33

The decision, because $|T - statistics| > 1.96$, then H_0 is rejected for $\gamma_{11}, \gamma_{21}, \beta_{21}$, meaning there is a significant effect. And because $|T - statistics| < 1.96$, H_0 is accepted for $\gamma_{31}, \beta_{31}, \beta_{32}$, meaning that there is no significant effect.

4. Results of Response Based Units Segmentation Partial Least Square (REBUS PLS)

The grouping of observation units in REBUS-PLS is based on the hierarchical cluster analysis results using Ward's communal and structural residuals method. The results of the cluster analysis of communal residuals and structural residuals are as follows.
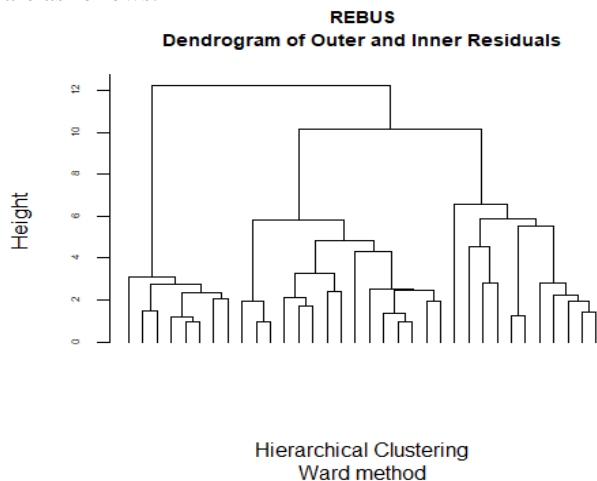

Fig. 2. Dendogram of Outer and Inner Residuals on RStudio

Based on the dendrogram that has been generated, it is only possible to group the observation units into two classes/segments. The grouping of provinces based on hierarchical cluster analysis is only used as an initial grouping stage. The results of the hierarchical cluster analysis are used to calculate the closeness measure (CM) value.

The essence of REBUS PLS is that each observation unit is included in a class that shows a smaller closeness measure (CM) value. The composition of the provinces from the 2 segments of the results of the analysis is based on the value of the closeness measure (CM), namely in class/segment 1, where there are 9 provinces consisting of Riau Islands, DKI Jakarta, DI Yogyakarta, Bali, East Nusa Tenggara, East Kalimantan, West Sulawesi, South Sulawesi and Papua. Whereas in class/segment 2 where there are 25 provinces consisting of Aceh, West Nusa Tenggara North Sumatra, West Kalimantan, West Sumatra, Central Kalimantan, Riau, South Kalimantan, Jambi, North Kalimantan, Bengkulu, North Sulawesi, South Sumatra, Gorontalo, Bangka Belitung Islands, Central Sulawesi, Lampung, Southeast Sulawesi, Banten, North Maluku, West Java, Maluku, Central Java, West Papua, East Java.

TABLE 4. Value of Q^2 and GoF global model and 2 segments

	Global Model	Segment 1	Segment 2
Q^2	0,883	0,996	0,797
GoF	0,601	0,856	0,432

Furthermore, if districts/cities have been grouped based on the segments formed, each of these segments then estimates the model for each class/segment (local model) using PLS-SEM.

V. CONCLUSION

The conclusions obtained from the results and discussion of this study are that in the relationship model for latent economic variables, human resources, health, and poverty, there are 14 significant indicators in the structural model. Based on the dendrogram that has been produced, it is determined to group the observation units into two classes/segments. The following is the human development index equation model based on REBUS PLS analysis on two segments. Equation model of human development index in Indonesia segment 1 is

$$\eta_1 = 0,841\,\xi$$
$$\eta_2 = -0,367\,\xi + 1,229\,\eta_1$$
$$\eta_3 = -0,248\,\xi + 0,193\,\eta_1 - 0,939\,\eta_2$$

Model of human development index in Indonesia segment 2 is

$$\eta_1 = 0,364\,\xi$$
$$\eta_2 = -0,170\,\xi - 0,727\,\eta_1$$
$$\eta_3 = -0,286\,\xi + 0,263\,\eta_1 - 0,414\,\eta_2$$

Based on the value of the closeness measure (CM), it can be seen that when 2 classes/segments are formed, the composition of class/segment 1 consists of 18 provinces, while class/segment 2 consists of 16 provinces. REBUS-PLS can detect heterogeneity in the SEM-PLS model, with the Q^2 value in each formed segment (local model) being greater than the Q^2 value in the global model, which means the local model is more better than the global model. As for the good GoF values for the

Esra Rombeallo, Georgina Maria Tinungki, and Anna Islamiyati, "Response-Based Units Segmentation Partial Least Square (REBUS PLS) Analysis on Human Development Index Data in Indonesia," *International Research Journal of Advanced Engineering and Science*, Volume 8, Issue 3, pp. 18-22, 2023.

global model, local model 1 and local model 2 fall into the large GoF category. A high GoF value also indicates a better structural and measurement model.

For further research, we can compare REBUS PLS, FIMIX PLS (Finite Mixture), and PLS TPM (PLS Typological Path Model).

REFERENCES

1. Anggita, E. D., Hoyyi, A., and Rusgiyono, A. (2019). *Analisis Structural Equation Modelling Pendekatan Partial Least Square Dan Pengelompokan Dengan Finite Mixture PLS (FIMIX-PLS) (Studi Kasus: Kemiskinan Rumah Tangga di Indonesia 2017)*. *8*(1), 35–45. http://ejournal3.undip.ac.id/index.php/gaussian.
2. Henseler, J., Ringle, C. M., and Sinkovics, R. R. (2009). The use of partial least squares path modeling in international marketing. In *Advances in International Marketing* (Vol. 20, pp. 277–319). Emerald Group Publishing Ltd. https://doi.org/10.1108/S1474-7979(2009)0000020014.
3. Jedidi, K., Jagpal, H. S., and DeSarbo, W. S. (1997). Finite-Mixture Structural Equation Models for Response-Based Segmentation and Unobserved Heterogeneity. *Marketing Science*, *16*(1), 39–59. https://doi.org/10.1287/mksc.16.1.39
4. Muliati, W. D., and Handoyo, R. D. (2019). Public Budget Allocation and Poverty Allevation on the Island of Kalimantan. *Journal of Treasury, State Finance and Public Policy*, *4*. http://www.djpk.kemenkeu.go.id/?p=5412
5. Patih, T., and Halistin. (2020). *Structural Equation Modeling Dengan Partial Least Square (SEM-PLS) Serta Aplikasinya Pada Bidang Pendidikan* (T. Hidayati (ed.); 2020th ed.). CV. Pena Persada.
6. Reagan, H. A. (2016). *The Effect of Economic Growth on Human Development Dimensions Analysis Using Response Based Unit Segmentation in Partial Least Square (Rebus-PLS)*.
7. Statistics, C. B. (2020). *Human Development Index (HDI) 2020*. 97. https://www.bps.go.id/
8. Suntono, and A'laa, R. D. (2021). *Human Development Index for South Sulawesi Province 2021* (p. 57). Central Bureau of Statistics South Sulawesi. https://sulsel.bps.go.id/
9. Trinchera, L. (2007). *Unobserved Heterogeneity in SEM: a new approach to latent class detection in PLS Path Modeling*. Universit`a degli Studi di Napoli.
10. Vinzi, V. E., Trinchera, L., Squillacciotti, S., & Tenenhaus, M. (2008). REBUS-PLS: A response-based procedure for detecting unit segments in PLS path modelling. *Applied Stochastic Models in Business and Industry*, *24*(5 SPEC. ISS.), 439–458. https://doi.org/10.1002/asmb.728
11. Wold, H. (1982). Soft modeling: The basic design and some extensions. In K. G. Wold, *Systems under indirect observations: Part II* (pp. 1-54). Amsterdam: North Holland.