Media Statistika

p-ISSN 1979 – 3693 e-ISSN 2477 – 0647
MEDIA STATISTIKA 17(1) 2024: 81-92
http://ejournal.undip.ac.id/index.php/media_statistika

ANALYSIS OF MULTILEVEL STRUCTURAL EQUATION MODELING WITH GENERALIZED STRUCTURED COMPONENT ANALYSIS METHOD

Fitri Amanah[1], Abdurakhman[2]
[1] Master of Mathematics Study Programme, Universitas Gadjah Mada, Yogyakarta, Indonesia
[2] Department of Mathematics, Universitas Gadjah Mada, Yogyakarta, Indonesia

e-mail: *fitriamanah@mail.ugm.ac.id*

DOI: 10.14710/medstat.17.1.81-92

Article Info:
Received: 7 November 2022
Accepted: 20 December 2024
Available Online: 26 December 2024

Keywords:
SEM; *GSCA*; *Multilevel GSCA*; *ALS. poverty*.
.

Abstract: Generalized Structured Component Analysis (GSCA) is a component-based SEM. One of the developments of GSCA is the GSCA method for multilevel data known as multilevel GSCA. Multilevel data is data that has a nested, grouped, or nested structure. This study aims to apply multilevel GSCA to the data on factors that affect poverty. The data used is on Indonesia's health, education and poverty in 2023.. The result is that all indicators are significant to the latent variables. The structural model shows that the quality of health has a negative and significant effect on poverty, education has a negative and significant effect on poverty, and the quality of health has a positive and significant effect on education. The results of between group show that health quality has a positive and significant effect on education in all regions, health quality has a negative and significant effect on poverty in Bali & Nusa Tenggara, Sulawesi, as well as Maluku and Papua, education has a negative and significant effect on poverty in Sumatra, Java, and Maluku & Papua. The overall goodness of fit value (FIT) is 0.622, meaning the model can explain 62.2% of data variation.

1. INTRODUCTION

Structural Equation Modeling (SEM) is a statistical analysis that combines factor, regression and path analysis. SEM aims to test the relationship between latent variables and the relationship between latent variables and each indicator simultaneously. SEM is widely used in various research, for example the analyzes of relationship between relational coordination and the quality of online education systems using SEM (Gallego Sánchez et al., 2021) and the analyzes of structural relationship between health anxiety and social health among health workers exposed to COVID-19 in Iran using Structural Equation Modeling (SEM) (Javadi et al, 2022).

There are two types of SEM, namely SEM covariance and SEM components. The covariance SEM has limitations because it is greatly influenced by parametric assumptions, namely normally distributed data, large sample sizes and reflective indicators (Tenenhaus, 2008). To overcome the limitations of SEM covariance, SEM components are developed namely Partial Least Square (PLS) and Generalized Structured Component Analysis (GSCA). The advantage of GSCA compared to PLS is that it provides goodness of fit mechanism that is used to assess the overall feasibility of the model (Hwang & Takane, 2014). Research on GSCA includes its application in analyzing the relationship between remuneration, work motivation and employee performance at UIN Sunan Kalijaga Yogyakarta (Supandi, 2020) and in examining factors influencing student achievement based on campus environmental characteristics (Suhriani & Abdurakhman, 2019).

One of the developments of SEM is analysis for multilevel data. Multilevel data is data that has a nested, grouped, or nested structure. According to Rabe-Hesketh et al., (2007), multilevel SEM is a combination of multilevel and SEM. Hierarchical or multilevel structured data often causes the assumption of independence to be violated. That is because individuals in the same group tend to be more homogeneous than different groups. Research on multilevel SEM includes Burić & Kim (2020) examining the relationship between teacher self-efficacy (TSE), teaching quality and student performance, with student and teacher levels, Long et al. (2023) utilized energy-saving behavior at individual and city levels, and also Hwang & Takane (2014) introduced multilevel GSCA for customer satisfaction across companies.

The aim of this research is to apply multilevel GSCA in analyzing factors affecting poverty. Past studies on poverty include Amanah & Rahmawati (2023) applying RGSCA in East Java, Zebua & Harefa (2022) using SEM in Sumatra, and Anggita et al (2019) applying Finite Mixture PLS in Indonesia. Poverty is a complex issue with widespread societal impacts, particularly in Indonesia, comprising diverse provinces and districts/cities with varying poverty characteristics. Multilevel analysis can be an alternative to comprehend poverty factors nationwide, examining districts/cities as the lowest level and region as the second level. So the Multilevel GSCA method was chosen as a suitable analysis for this research because of its ability to handle hierarchical data structures. Apart from that, the multilevel GSCA method also does not require assumptions like covariance-based SEM, and still provides goodness of fit criteria.

2. LITERATURE REVIEW

2.1. Multilevel Generalized Structured Component Analysis

Generalized Structured Component Analysis (GSCA) is a component-based SEM approach, in which latent variables are defined as components or weighted composites of indicator variables. One of the developments of SEM is SEM analysis for multilevel data. Hwang & Takane (2014) introduced multilevel GSCA as a component-based SEM approach for multilevel sample data. Multilevel data is a data structure that consists of several units of analysis, i.e. one unit is nested within another unit. For example, a study of voting focuses on voters in different regions. In this case study, there are two units of analysis, namely voters and regions, where voters nest within regions. In general, data is said to have a nested structure if some units of analysis are considered as subsets of other units (Steenbergen & Jones, 2002).

2.2. Multilevel Generalized Structured Component Analysis Model Specification

The multilevel GSCA model used is a two-level GSCA analysis, with level-1 units as individuals and level-2 units as groups, and using reflective indicator variables. The multilevel GSCA model consists of 3 submodels, namely measurement, structural and weighting models (Hwang & Takane, 2014).

Let N_g and G represent the number of individuals in each group and the number of groups $\left(i = 1, \dots, N_g; g = 1, \dots, G\right)$. The measurement model is as follows:

$$Z_{ig} = C'_g\gamma_{ig} + \varepsilon_{ig} \tag{1}$$

The two-level measurement model states the C_g matrix in the equation (1) as the sum of the average loading factors and the loading factor standard deviations. So that the measurement model in equation (1) can be rewritten as follows:

$$Z_{ig} = \left[L' + L'_g\right]\gamma_{ig} + \varepsilon_{ig} \tag{2}$$

where Z_{ig} is a vector of indicator variables, γ_{ig} is a vector of latent variable, L is matrix of the average loading factor, L_g is the standard deviation matrix of the loading factor, and ε_{ig} is the residual vector for Z_{ig}.

The structural model is as follows:

$$\gamma_{ig} = B'_g\gamma_{ig} + \zeta_{ig} \tag{3}$$

The two-level structural model states the B_g matrix in the equation (3) as the sum of the average path coefficient and the standard deviations of path coefficient. So that the structural model in equation (3) can be rewritten as follows:

$$\gamma_{ig} = [Q' + Q'_g]\gamma_{ig} + \zeta_{ig} \tag{4}$$

where Q is the average path coefficient matrix, Q_g is the standard deviation matrix of the path coefficient, and ζ_{ig} is the residual vector for γ_{ig}.

The weighting model is as follows:

$$\gamma_{ig} = W'Z_{ig} \tag{5}$$

where W is the weight components matrix from the indicator variable to the latent variable. The two-level GSCA weighting model is the same as the one-level model because the latent variable scores already depend on the group.

Equations (2) and (4) can be written as follows:

$$\left[z_{ig}, \gamma_{ig}\right] = \left[C'_g, B'_g\right]\gamma_{ig} + [\varepsilon_{ig}, \zeta_{ig}]$$

$$\left[z_{ig}, \gamma_{ig}\right] = \left[L' + L'_g, Q + Q'_g\right]\gamma_{ig} + [\varepsilon_{ig}, \zeta_{ig}] \tag{6}$$

Then by substituting equation (5), equation (6) becomes:

$$\left[z_{ig}, W'z_{ig}\right] = \left[L' + L'_g, Q + Q'_g\right]W'z_{ig} + [\varepsilon_{ig}, \zeta_{ig}]$$

$$[I, W']z_{ig} = \left[L' + L'_g, Q + Q'_g\right]W'z_{ig} + [\varepsilon_{ig}, \zeta_{ig}$$

$$V'z_{ig} = A'_gW'z_{ig} + e_{ig} \tag{7}$$

where I is the identity matrix, $V' = [I, W']$, $A'_g = \left[L' + L'_g, Q + Q'_g\right]$, and $e_{ig} = [\varepsilon_{ig}, \zeta_{ig}]$. Equation (7) is the multilevel GSCA model (Hwang & Takane, 2014).

2.3. Multilevel Generalized Structured Component Analysis Parameter Estimation

The Multilevel Generalized Structured Component Analysis parameters were estimated using the Alternating Least Square (ALS), i.e. by minimizing the sum of square (SS) of all the residual. Let z_{ig} be a vector of indicator variables denote $J \times 1$ measured at the i-th observations from N observations in the g-group $\left(i = 1, \ldots, N_g\right)$. The unknown parameters $\left(W, L, L_g, Q, Q_g\right)$ are estimated by minimizing the sum of square of e_{ig}. Based on equation (7) is obtained:

$$e_{ig} = V'z_{ig} - A'_gW'z_{ig} \tag{8}$$

So, the sum of square (SS) criteria is:

$$\emptyset = \sum_{g=1}^{G} \sum_{i=1}^{N_g} \boldsymbol{e_{ig}}'\boldsymbol{e_{ig}} = \sum_{g=1}^{G} \sum_{i=1}^{N_g} \boldsymbol{SS}(\boldsymbol{e_{ig}}) \tag{9}$$

Let $\boldsymbol{Z}_g$ be the indicator matrix for the g-group denote $N_g \times J$ with $\boldsymbol{Z}_g = [\boldsymbol{z}_1, \dots, \boldsymbol{z}_{Ng}]'$ for $i = 1, \dots, N_g$ and $g = 1, \dots, G$. So, the equation (9) can be written more concisely without writing the sum of N_g observations.

$$\begin{aligned}
\boldsymbol{\phi} &= \sum_{g=1}^{G} \sum_{i=1}^{N_g} \boldsymbol{SS}(\boldsymbol{e_{ig}}) \\
&= \sum_{g=1}^{G} \boldsymbol{SS}(\boldsymbol{e_g}) \\
&= \sum_{g=1}^{G} \boldsymbol{SS}(\boldsymbol{Z_g V} - \boldsymbol{Z_g W A_g}) \\
&= \sum_{g=1}^{G} \boldsymbol{SS}(\boldsymbol{Z_g V} - \boldsymbol{Z_g W}[\boldsymbol{L} + \boldsymbol{L_g}, \boldsymbol{Q} + \boldsymbol{Q_g}]) \\
&= \sum_{g=1}^{G} \boldsymbol{SS}\left(\boldsymbol{Z_g V} - \boldsymbol{Z_g W}[\boldsymbol{L}, \boldsymbol{I}, \boldsymbol{Q}, \boldsymbol{I}] \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{L_g} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I} \\ \boldsymbol{0} & \boldsymbol{Q_g} \end{bmatrix} \right) \\
&= \sum_{g=1}^{G} \boldsymbol{SS}(\boldsymbol{Z_g V} - \boldsymbol{Z_g W M H_g})
\end{aligned} \tag{10}$$

where $\boldsymbol{M} = [\boldsymbol{L}, \boldsymbol{I}, \boldsymbol{Q}, \boldsymbol{I}]$, and $\boldsymbol{H}_g = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{L}_g & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I} \\ \boldsymbol{0} & \boldsymbol{Q}_g \end{bmatrix}$.

The Alternating Least Square (ALS) method is a parameter technique that involves dividing parameters into subsets. It calculates the least squares for one subset while assuming the other parameters remain constant. Hwang & Takane (2014) proposed three main steps to estimate the parameters.

Step 1. Matrix $\boldsymbol{W}$ is updated with matrix $\boldsymbol{L}, \boldsymbol{L_g}, \boldsymbol{Q}$ and $\boldsymbol{Q_g}$ are fixed. This is equivalent to minimizing

$$\begin{aligned}
\boldsymbol{\phi} &= \sum_{g=1}^{G} \boldsymbol{SS}(\boldsymbol{Z_g V} - \boldsymbol{Z_g W M H_g}) \\
&= \sum_{g=1}^{G} \boldsymbol{SS}(\boldsymbol{Z_g V} - \boldsymbol{Z_g W A_g})
\end{aligned} \tag{11}$$

with respect to $\boldsymbol{W}$ and $\boldsymbol{A}_g = \boldsymbol{M H}_g$. This criterion is essentially equivalent for GSCA (Hwang and Takane, 2014).

Step 2. Matrix $\boldsymbol{L}$ and $\boldsymbol{Q}$ (or equivalently $\boldsymbol{M}$) are updated with matrix $\boldsymbol{W}, \boldsymbol{L}_g$ and $\boldsymbol{Q}_g$ are fixed. Equation (10) can be re-written as:

$$\boldsymbol{\phi} = \sum_{g=1}^{G} \boldsymbol{SS}\left((\boldsymbol{Z_g V}) - (\boldsymbol{H}'_g \otimes \boldsymbol{Z_g W}) \boldsymbol{vec}(\boldsymbol{M}) \right) \tag{12}$$

Let $\boldsymbol{m}$ is a vector of free parameters in $vec(\boldsymbol{M})$, $\boldsymbol{\Omega}_g$ is a matrix formed by eliminating columns of $(\boldsymbol{H}'_g \otimes \boldsymbol{Z}_g \boldsymbol{W})$ which corresponds to the zero elements in $vec(\boldsymbol{M})$, $\widetilde{\boldsymbol{\Omega}}_g$ is a matrix formed by eliminating columns of $(\boldsymbol{H}'_g \otimes \boldsymbol{Z}_g \boldsymbol{W})$ which corresponds to the ones in $vec(\boldsymbol{M})$, and $\widetilde{\boldsymbol{m}}$ is a vector of ones in $vec(\boldsymbol{M})$. Then, the least square estimate of $\boldsymbol{m}$ is obtained by:

$$\widehat{\boldsymbol{m}} = \left(\sum_{g=1}^{G} \boldsymbol{\Omega}'_g \boldsymbol{\Omega}_g \right)^{-1} \left(\sum_{g=1}^{G} \boldsymbol{\Omega}'_g \left(\boldsymbol{vec}(\boldsymbol{Z_g V}) - \widetilde{\boldsymbol{\Omega}}_g \widetilde{\boldsymbol{m}} \right) \right) \tag{13}$$

Step 3. Matrix $\boldsymbol{L}_g$ and $\boldsymbol{Q}_g$ (or equivalently $\boldsymbol{H}_g$) are updated with matrix $\boldsymbol{W}, \boldsymbol{L}$ and $\boldsymbol{Q}$ are fixed. Equation (10) can be re-written as:

$$\boldsymbol{\phi} = \sum_{g=1}^{G} SS\left(\left(\boldsymbol{Z}_g\boldsymbol{V}\right) - \left(\boldsymbol{I}\otimes\boldsymbol{Z}_g\boldsymbol{W}\boldsymbol{M}\right)\boldsymbol{vec}\left(\boldsymbol{H}_g\right)\right) \quad (14)$$

Let $\boldsymbol{h}_g$ is a vector of free parameters in $vec\left(\boldsymbol{H}_g\right)$, Ξ_g is a matrix formed by eliminating columns of $\left(\boldsymbol{I}\otimes\boldsymbol{Z}_g\boldsymbol{W}\boldsymbol{M}\right)$ which corresponds to the zero elements in $vec\left(\boldsymbol{H}_g\right)$, $\widetilde{\Xi}_g$ is a matrix formed by eliminating columns of $\left(\boldsymbol{I}\otimes\boldsymbol{Z}_g\boldsymbol{W}\boldsymbol{M}\right)$ which corresponds to the ones in $vec\left(\boldsymbol{H}_g\right)$, and $\widetilde{\boldsymbol{h}}_g$ is a vector of ones in $vec\left(\boldsymbol{H}_g\right)$. Then, the least square estimate of $\boldsymbol{h}_g$ is obtained by:

$$\widehat{\boldsymbol{h}_g} = \left(\sum_{g=1}^{G} \Xi_g' \Xi_g\right)^{-1} \Xi'_g\left(\boldsymbol{vec}\left(\boldsymbol{Z}_g\boldsymbol{V}\right) - \widetilde{\Xi}_g\widetilde{\boldsymbol{h}}_g\right) \quad (15)$$

2.4. Model Evaluation

The measurement evaluation model assesses model validity using convergent and discriminant validity, as well as composite reliability. Convergent validity is indicated by loading factors above 0,50 for each indicator. Discriminant validity is measured by Average Variance Extracted (AVE), with a value ≥ 0,50 considered acceptable (Chin, 1998). The AVE formula is defined as equation (16) where λ_i is the component of loading factor.

$$\text{AVE} = \frac{\sum(\boldsymbol{\lambda}_i^2)}{\sum(\boldsymbol{\lambda}_i^2)+\sum(\boldsymbol{1}-\boldsymbol{\lambda}_i^2)} \quad (16)$$

The third evaluation is composite reliability, which measures a latent construct. Composite reliability can be calculated using GSCA output with formula:

$$\text{Composite Reliability} = \frac{(\boldsymbol{\Sigma\lambda}_i)^2}{(\boldsymbol{\Sigma\lambda}_i)^2+\sum(\boldsymbol{1}-\boldsymbol{\lambda}_i^2)} \quad (17)$$

The recommended composite reliability value is ≥ 0,70 (Ghozali, 2008).

Structural model evaluation is measured by the Critical Ratio. The Critical Ratio formula is defined as:

$$\text{Critical Ratio} = \frac{\hat{\boldsymbol{b}}_i}{\boldsymbol{SE}(\hat{\boldsymbol{b}}_i)} \quad (18)$$

where $\hat{b}_i$ the estimated path coefficient and $SE\left(\hat{b}_i\right)$ is the standard error of the path coefficient. If Critical Ratio ≥ $t_{(\alpha,n-k)}$, the relationship between latent variables is significant (Abell et al., 2009).

Overall goodness of fit (FIT) value reflects the total variance explained by the model for all endogenous latent variables. FIT formula is (Hwang & Takane, 2014):

$$\text{FIT} = 1 - \frac{\sum_{g=1}^{G} \boldsymbol{SS}(\boldsymbol{Z}_g\boldsymbol{V}-\boldsymbol{Z}_g\boldsymbol{W}\boldsymbol{A}_g)}{\sum_{g=1}^{G} \boldsymbol{SS}(\boldsymbol{Z}_g\boldsymbol{V})} \quad (19)$$

3. RESEARCH METHOD

3.1. Data Source

The research utilizes secondary data from the online publication of the Central Statistics Agency (BPS), accessible at https://www.bps.go.id/id. The data comprises health, education and poverty indicators for all 514 districts/cities in Indonesia. This research employs a two-level GSCA, with districts/cities as level-1 units and regions as level-2 units The regions is determined based on nearby islands and data adequacy.

Table 1. The Number of Districts/Cities

Region	Number of Districts/Cities
Sumatera	154
Jawa	119
Bali and Nusa Tenggara	41
Kalimantan	56
Sulawesi	81
Maluku and Papua	63
Total	514

An illustration of the data used in the case study is presented in Figure 1.

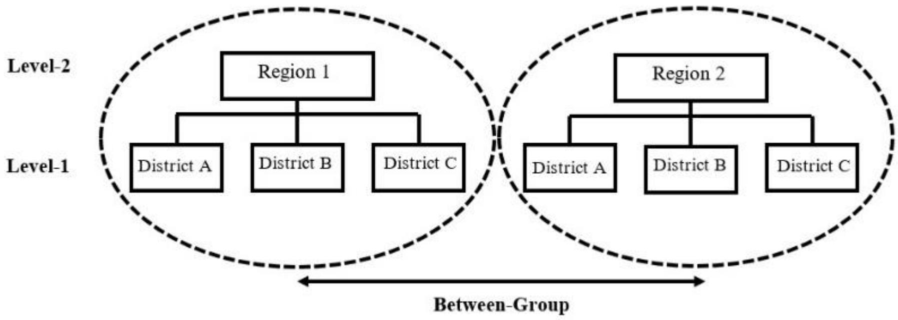

Figure 1. Illustration of Model Two Level's GSCA

3.2. Research Variables

This research uses 3 latent variables, namely health and education as exogenous latent variables, also poverty as the endogenous latent variable. According to Badan Pusat Statistik (2023), poverty is seen as an economic inability to meet basic food and non-food needs as measured from the expenditure side. The level of education affects the poverty rate. People who are more educated usually have a lower chance of being poor. Likewise, the quality of health also affects the poverty rate.

Table 2. The Indicators of Latent Variables

Variables	Symbols	Indicators	Symbols
Health (exogenous latent variables)	HE (γ_1)	Life expectancy	z_1
		Percentage of households using proper drinking water sources	z_2
		Percentage of poor households that use private/shared latrines	z_3
Education (exogenous latent variables)	ED (γ_2)	Completed education (SMA+)	z_4
		Expected length of schooling	z_5
		Average length of schooling	z_6
Poverty (endogenous latent variable)	POV (γ_3)	Percentage of poor people	z_7
		Poverty severity index	z_8
		Number of Poor Population (Thousand People)	z_9

3.3. Data Analysis Technique

This study employs Multilevel Generalized Structured Component Analysis with GSCA Pro software, which can be accessed for free at https://www.gscapro.com/. In general, the steps for GSCA analysis are as follows (Suhriani & Abdurakhman, 2019):

1. Determine data and models according to concepts and theories.
2. Constructing a path diagram that describes the relationship between latent variables and indicators.
3. Convert path diagrams to equations namely measurement, structural and weighting model.
4. Estimate model parameters using Alternating Least Square algorithm.
5. Evaluate the model, Interpret the results and make conclusions

4. RESULTS AND DISCUSSION

Figure 2 is a path diagram construction of the factors that influence poverty. The relationship between quality of health, education and poverty were modeled using the Multilevel Generalized Structured Component Analysis (MGSCA).

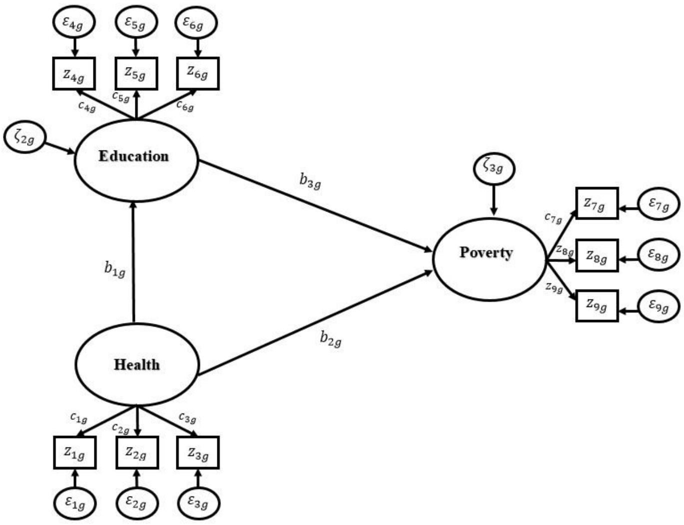

Figure 2. Graphs of Model Two Level's GSCA

The MGSCA model was solved using the Alternating Least Square (ALS) algorithms. The estimated MGSCA parameters from the GSCA Pro software are as follows:

Table 3. Estimate of Weights and Their 95% Confidence Intervals

Latent	Indicator	Estimate	95%CI
HE	z_1	0.426	0.37-0.47
	z_2	0.479	0.42-0.53
	z_3	0.375	0.33-0.42
ED	z_4	0.293	0.25-0.33
	z_5	0.345	0.30-0.37
	z_6	0.475	0.43-0.51
POV	z_7	0.400	0.35-0.45
	z_8	0.468	0.41-0.50
	z_9	0.548	0.49-0.60

Table 3 presents the weights and 95%-confidence intervals for all indicators. All weight estimates are statistically significant, indicating that each indicator contributes equally well to determining the latent variable.

Table 4. Estimate of Average Loading, the between-Group Standard Deviation and Their 95% Confidence Intervals

		Average Loading		Between-Group Standard Deviation	
Latent	Indicator	Estimate	95%CI	Estimate	95%CI
HE	z_1	0.766	0.65-0.84	0.248	0.18-0.35
	z_2	0.843	0.75-0.92	0.268	0.16-0.39
	z_3	0.722	0.63-0.78	0.212	0.13-0.31
AVE = 0.606; Composite Reliability = 0.821					
ED	z_4	0.891	0.81-0.96	0.200	0.15-0.28
	z_5	0.815	0.74-0.87	0.173	0.11-0.27
	z_6	0.963	0.93-0.99	0.006	0.03-0.10
AVE = 0.795; Composite Reliability = 0.921					
POV	z_7	0.799	0.78-0.84	0.282	0.22-0.33
	z_8	0.648	0.59-0.74	0.384	0.35-0.41
	z_9	0.690	0.61-0.76	0.515	0.47-0.56
AVE = 0.511; Composite Reliability = 0.757					

Table 4 summarizes the estimate of average loading, all statistically significant at 95% confidence intervals, indicating significant relationships between indicators and the latent variable. Table 4 also shows statistically significant standard deviation of the loading factor at 95% confidence interval, indicating substantial variations in loading factors between regions. The validity assessment using average loading factor shows that all values exceed 0.50, indicating significantly able to explain their latent variables. Additionally, AVE values for all latent variables greater than 0.50, and the composite reliability exceeds 0.70 for all latent variables. All latent variables are affectively explained by each indicator.

Table 5. Estimate of Average Path Coefficients, the between-Group Standard Deviation and Their 95% Confidence Intervals

	Average Path Coefficients		Between-Group Standard Deviation	
	Estimate	95%CI	Estimate	95%CI
HE→ED	0.497	0.42 - 0.56	0.127	0.09-0.22
HE→POV	-0.158	-0.29 to -0.07	0.384	0.28-0.54
ED→POV	-0.267	-0.39 to -0.16	0.311	0.21-0.44

Table 5 presents the estimate of average path coefficients, which are statistically significant at the 95% confidence interval. This means that the relationship between latent variables is significant. In addition, Table 5 also shows the standard deviation of path coefficients statistically significant at the 95% confidence interval. This means that there are substantial differences in the estimated path coefficients in each region.

Evaluation of the structural model can also be seen through the value of the critical ratio. The Critical Ratio value can be used to see whether an exogenous latent variable significantly influences the endogenous latent variable or not. By using equation (18), the following results are obtained:

Table 6. Estimate of Average Path Coefficients, Standard Error and Critical Ratio

Latent	Estimate	Std. Error	Critical Ratio
HE→ED	0.497	0.035	14.20
HE→POV	-0.158	0.059	2.68
ED→POV	-0.267	0.060	4.45

At the 5% significance level, the value of $t_{\alpha,n-k}$=1.964, where n = 514 and k = 33. Based on the value of t table, it can be concluded:

1. Health quality has a negative and significant effect on poverty (Critical Ratio ≥ 1.964). This means that if the quality of health is getting better, it will reduce poverty.
2. Education has a negative and significant effect on poverty (Critical Ratio ≥ 1.964). This means that if the education is getting better, it will reduce poverty.
3. Health quality has a positive and significant effect on education (Critical Ratio ≥ 1.964). This means that the better the quality of health, the higher education will be.

FIT value used to evaluate overall goodness of fit model. Based on equation (19), the FIT value is 0.622 which means the model is able to explain about 62.2% datas's variation. The results aligns with Amanah & Rahmawati (2023) and Zebua & Harefa (2022), who studied factors influencing poverty in East Java and Sumatra, focusing on a single region. However, multilevel GSCA's advantage lies in its capability to perform within-group and between-group analysis, providing a comprehensive understanding of poverty in Indonesia by analyzing regional differences. This method enables analysis of poverty factors at the regional level.

As previously explained, the multilevel GSCA expresses group-dependent loading factors and path coefficients as the sum of mean and standard deviation, allowing estimation of values for each group. The estimated enable to see differences in loading factors and path coefficients in certain groups. For instance, Table 7 displays the estimated path coefficient values for each region.

Table 7. Value of Path Coefficient for 6 Regions and Their Critical Ratio

Region	Latent	Estimate	Std. Error	Critical Ratio
Sumatera	HE→ED	0.480	0.073	6.575
	HE→POV	0.083	0.061	1.361*
	ED→POV	-0.218	0.059	3.695
Jawa	HE→ED	0.287	0.066	4.348
	HE→POV	0.378	0.068	5.559
	ED→POV	-0.681	0.071	9.592
Bali & Nusa Tenggara	HE→ED	0.669	0.107	6.252
	HE→POV	-0.846	0.199	4.251
	ED→POV	-0.369	0.214	1.724*

Kalimantan	HE→ED	0.439	0.123	3.569
	HE→POV	0.004	0.195	0.021*
	ED→POV	0.312	0.182	1.714*
Sulawesi	HE→ED	0.470	0.128	3.672
	HE→POV	-0.225	0.099	2.273
	ED→POV	-0.154	0.093	1.656*
Maluku & Papua	HE→ED	0.635	0.088	7.216
	HE→POV	-0.343	0.126	2.722
	ED→POV	-0.491	0.129	3.806

*Not significant

Table 7 provides the estimated path coefficient values for each region as follows:

1. Health quality has a positive and significant effect on education in all regions (Critical Ratio ≥ 1.964). This means that the better the quality of health in all regions, the higher education will be.
2. Health quality has a negative and significant effect on poverty in Bali & Nusa Tenggara, Sulawesi, as well as Maluku & Papua (Critical Ratio ≥ 1.964). This means that if the quality of health in the region is getting better, it will reduce poverty.
3. Education has a negative and significant effect on poverty in Sumatra, Java, and Maluku & Papua (Critical Ratio ≥ 1.964). This means that if the education in the region is getting better, it will reduce poverty.

Table 7 reveals diverse findings. In Java, the quality of health positively and significantly impacts poverty, suggesting that high health quality is associated with high poverty rates. Conversely, in Sumatra and Kalimantan, health quality shows no significant effect on poverty. Similarly, education shows no significant effect in Bali & Nusa Tenggara, Kalimantan and Sulawesi. This implies that high education level does not guarantee low poverty. Despite Table 6 indicating a negative and significant effect of health quality and education on poverty in Indonesia, but regional results present different results. But, these results are consistent with Table 5, which show significant standard deviations, suggesting variations between groups of health and education on poverty.

Based on these results, the multilevel GSCA method can be a reference when using multilevel data because it provides comprehensive results by analyzing relationships within-group and between-group simultaneously. Multilevel GSCA enables deeper understanding of complex phenomena like poverty across various regions by considering variations at both individual and group levels. So that multilevel GSCA aids in adjusting strategies to effectively address regional disparities.

5. CONCLUSION

Based on the results of a case study of the factors that affect poverty, it is found that all indicators are significant for each of the latent variables. In the structural model, it is found that health quality has a negative and significant effect on poverty, education has a negative and significant effect on poverty and health quality has a positive and significant effect on education. The overall goodness of fit (FIT) value is 0.622, meaning that the model is able to explain about 62.2% data variation.

REFERENCES

Abell, N., Springer, D. W., & Kamata, A. (2009). *Developing and Validating Rapid Assessment Intrument* (1 ed). New York: Oxford University Press.

Amanah, F., & Rahmawati, F. (2023). Study of Regularized Generalized Structured Component Analysis to Overcome Multicollinearity in Component-Based SEM. *Jurnal Matematika, Statistika dan Komputasi*, 20(1), 281–293. https://doi.org/10.20956/j.v20i1.28069

Anggita, E. D., Hoyyi, A., & 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). *Jurnal Gaussian*, 8(1), 35–45. https://doi.org/10.14710/j.gauss.v8i1.26620

Badan Pusat Statistik (BPS). (2023). *Data dan Informasi Kemiskinan Kabupaten/Kota Tahun 2023* (Vol. 15). Jakarta: Badan Pusat Statistik.

Burić, I., & Kim, L. E. (2020). Teacher Self-Efficacy, Instructional Quality, and Student Motivational Beliefs: An Analysis Using Multilevel Structural Equation Modeling. *Learning and Instruction*, 66, 101302. https://doi.org/10.1016/j.learninstruc.2019.101302

Chin, W. W. (1998). The Partial Least Squares Approach to Structural Equation Modeling. In *Modern Methods for Business Research* (pp. 295–336). London: Lawrence Erlbaum Associates Publishers.

Gallego-Sánchez, M. D. C., De-Pablos-Heredero, C., Medina-Merodio, J.-A., Robina-Ramírez, R., & Fernandez-Sanz, L. (2021). Relationships Among Relational Coordination Dimensions: Impact on The Quality of Education Online with A Structural Equations Model. *Technological Forecasting and Social Change*, 166, 120608. https://doi.org/10.1016/j.techfore.2021.120608

Ghozali, I. (2008). *Generalized Structured Component Analysis (GSCA): Model Persamaan Structural Berbasis Komponen*. Semarang: Badan Penerbit Universitas Diponegoro.

Hwang, H., & Takane, Y. (2014). *Generalized Structured Component Analysis: A Component-Based Approach to Structural Equation Modeling*. Boca Raton: Chapman & Hall/CRC Press.

Javadi, M. H., Khoshnami, M. S., Noruzi, S., & Rahmani, R. (2022). Health Anxiety and Social Health Among Health Care Workers and Health Volunteers Exposed to Coronavirus Disease in Iran (2020): A Structural Equation Modeling. *Journal of Affective Disorders Reports*, 8, 100321. https://doi.org/10.1016/j.jadr.2022.100321

Long, R., Wang, J., Chen, H., Li, Q., Wu, M., & Tan-Soo, J.-S. (2023). Applying Multilevel Structural Equation Modeling to Energy-Saving Behavior: The Interaction of Individual and City-Level Factors. *Energy Policy*, 174, 113423. https://doi.org/10.1016/j.enpol.2023.113423

Rabe-Hesketh, S., Skrondal, A., & Zheng, X. (2007). Multilevel Structural Equation Modeling. In *Handbook of Latent Variable and Related Models* (pp. 209–227). Elsevier. https://doi.org/10.1016/B978-044452044-9/50013-6

Steenbergen, M. R., & Jones, B. S. (2002). Modeling Multilevel Data Structures. *American Journal of Political Science*, 46(1), 218-237. https://doi.org/10.2307/3088424

Suhriani, I. F., & Abdurakhman. (2019). Pendekatan SEM Berbasis Komponen Menggunakan Generalised Structured Component Analysis (GSCA). *Jurnal Sains Matematika dan Statistika*, 5(2). DOI : 10.24014/jsms.v5i2.7629

Supandi, E. D. (2020). Structural Equation Modeling With Generalized Structured Component Analysis on The Relationship Between Renumeration and Motivation on Employee Performance at UIN Sunan Kalijaga Yogyakarta. *Media Statistika*, 13(2), 136–148. https://doi.org/10.14710/medstat.13.2.136-148

Tenenhaus, M. (2008). Component-based Structural Equation Modelling. *Total Quality Management & Business Excellence*, 19(7–8), 871–886. https://doi.org/10.1080/14783360802159543

Zebua, H. I., & Harefa, G. A. (2022). Structural Equation Model (SEM) dalam Pemodelan Kemiskinan di Pulau Sumatera. *Indonesian Journal of Applied Statistics*, 5(1), 48-57. https://doi.org/10.13057/ijas.v5i1.50493