BAREKENG: Journal of Mathematics and Its Applications
March 2026 Volume 20 Issue 1 Page 0001–0016
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol20iss1pp0001-0016

STRUCTURAL EQUATION MODELING ANALYSIS ON POVERTY IN WEST KALIMANTAN WITH FINITE MIXTURE IN PARTIAL LEAST SQUARE APPROACH

Muhammad Fauzan[1], Hendra Perdana[2*], Neva Satyahadewi[3]

[1,2,3]*Department of Mathematics, Mathematics and Natural Science Faculty, Universitas Tanjungpura*
Jln. Prof. Dr. Hadari Nawawi, Pontianak, 78124, Indonesia

Corresponding author's e-mail: *** **hendra.perdana@math.untan.ac.id**

Article Info

Article History:

Received: 27th February 2025
Revised: 2nd May 2025
Accepted: 25th July 2025
Available online: 24th November 2025

Keywords:

Finite mixture;
Poverty;
Segments;
West Kalimantan.
.

ABSTRACT

Poverty occurs when individuals or groups lack the necessary resources to fulfill their basic needs. In Indonesia, including West Kalimantan, poverty remains a significant issue influenced by various socio-economic factors. This study aims to identify valid and reliable indicators of poverty and classify regencies/cities in West Kalimantan using the 2023 data from the Central Statistics Agency of West Kalimantan and Indonesia. The analysis applies the Structural Equation Modeling approach with Finite Mixture in Partial Least Squares (FIMIX-PLS). From 19 observed indicators, only 12 were found valid and reliable based on measurement and structural model evaluation. The structural model reveals three significant relationships: the Economy significantly influences Poverty, Health influences Education, and Education influences the Economy. Based on the FIMIX-PLS results, the regencies/cities are segmented into four groups with distinct structural characteristics. Segment 1 reflects the influence of Health on Education, Segment 2 reflects the influence of Health on the Economy, Segment 3 highlights the influence of Economy on Poverty, and Segment 4 captures the influence of Education on the Economy. Detailed interpretations of each segment and their policy implications are presented in the conclusion. The results support the importance of tailored poverty alleviation strategies based on latent regional characteristics and validated model findings.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-ShareAlike 4.0 International License (https://creativecommons.org/licenses/by-sa/4.0/).

How to cite this article:

M. Fauzan, H. Perdana and N. Satyahadewi., "STRUCTURAL EQUATION MODELING ANALYSIS ON POVERTY IN WEST KALIMANTAN WITH FINITE MIXTURE IN PARTIAL LEAST SQUARE APPROACH," *BAREKENG: J. Math. & App.*, vol. 20, iss. 1, pp. 0001-0016, Mar, 2026.

Copyright © 2026 Author(s)
Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article · Open Access

1. INTRODUCTION

Poverty is a multidimensional issue that continues to affect many developing countries. To address this, the United Nations established the United Nations Development Program (UNDP), which focuses on eradicating poverty, promoting inclusive development, and strengthening institutional capacity to ensure sustainable development outcomes [1]. Several international studies emphasize that poverty is influenced by multiple structural factors, including job quality, educational attainment, access to health services, infrastructure development, gender disparities, and geographic location [2].

In various regions of Indonesia, poverty remains a persistent and complex issue due to its multidimensional impacts, such as hampering economic growth and contributing to inflationary pressures. This condition is typically marked by the inability of individuals or households to meet basic needs, such as food, shelter, education, and healthcare, which in turn limits access to essential public services [3]. According to Chambers, poverty is not merely about income deprivation but constitutes an integrated concept comprising five dimensions: (1) powerlessness, (2) dependence, (3) geographical and social isolation, (4) vulnerability to crises, and (5) material deprivation [4]. This perspective emphasizes that living in poverty involves more than just financial hardship; it also includes limited access to education and healthcare, exposure to legal injustice, a heightened risk of crime and exploitation, lack of autonomy, and an inability to shape one's own life path.

In March 2022, the poverty rate in Indonesia dropped to its lowest level, marking a significant achievement as the rate returned to single digits for the first time since the COVID-19 pandemic, namely 9.54% [5]. Since the onset of the pandemic, poverty has evolved beyond a financial and consumption issue, encompassing limited access to employment, health, education, and livelihood activities, factors that impact both the short and long term [6]. This proves that poverty, as a multidimensional phenomenon, cannot be understood through a single lens; hence, a multidimensional approach is essential for effectively reducing poverty rates [7]. West Kalimantan is the province with the highest number of poor people on Kalimantan Island. Although the poverty rate has slightly declined, the reduction has not been significant, and poverty remains highly concentrated in this province [8]. When analyzed by percentage over the past seven years, the poverty rate in West Kalimantan decreased from 8.30% in 2015 to 6.73% in 2022 [9]. However, based on residential classifications, the number of people living in poverty in urban areas increased by 2.35 thousand—from 85.04 thousand in March 2022 to 87.39 thousand in March 2023. Similarly, the rural poor population rose from 265.21 thousand to 265.96 thousand over the same period, marking an increase of 750 people [9].

Poverty is associated with low levels of consumption and income, poor health and education, limited participation in development, and various other issues related to human development. Although the poverty rate in West Kalimantan has declined, further analysis is needed to identify the factors that influence this issue. This study employs Structural Equation Modeling (SEM) with the Finite Mixture Partial Least Squares (FIMIX-PLS) approach. SEM is a multivariate analysis technique used to model the simultaneous linear relationships between unobservable latent variables and observable indicators [10]. Among various SEM techniques, Partial Least Squares (PLS) is particularly suitable for analyzing latent variables, indicators, and measurement errors simultaneously. PLS is often used when the theoretical foundation is weak or when the indicators do not align well with a reflective measurement model [11]. Known as soft modeling, PLS offers flexibility by relaxing several statistical assumptions, such as multivariate normality, specific measurement scales, and large sample size requirements [10].

A Finite Mixture in Partial Least Squares (FIMIX-PLS) is part of the Structural Equation Modeling (SEM) technique, which was introduced to identify unobserved heterogeneity within data groups [12]. This approach considers heterogeneity within the observation units of latent variables, acknowledging that different groups or regions within the dataset may exhibit distinct structural relationships. By applying this method, the research assumes that not all units respond uniformly to the same latent constructs, thereby justifying the use of FIMIX-PLS to capture unobserved segment-specific variations.

This approach allows for a more accurate conclusion by accounting for heterogeneity among latent variables through the segmentation process [13]. The purpose of this study is to analyze valid and reliable indicators that significantly influence poverty, as well as to classify regional groupings in West Kalimantan using the Finite Mixture Partial Least Squares (FIMIX-PLS) method. Although numerous studies have examined poverty determinants, most treat the population as homogeneous, overlooking the possibility of latent heterogeneity across regions. This study addresses that gap by applying a segmentation-based approach that identifies unobserved subpopulations, thereby providing a more nuanced understanding of poverty dynamics. The urgency of this research lies in the persistent concentration of poverty in West Kalimantan despite various development efforts, emphasizing the need for more targeted, evidence-based interventions tailored to specific regional characteristics.

2. RESEARCH METHODS

Partial Least Squares (PLS) is a Structural Equation Modeling (SEM) analysis technique capable of directly analyzing indicator variables, latent variables, and measurement errors. PLS is often used as an alternative method when the available indicators do not meet the requirements of a reflective measurement model, or when the underlying theory is weak [14]. This method integrates principal component analysis and regression in an iterative process, with the primary objective of explaining the variance in the constructs within the model [15]. PLS is considered a powerful analytical tool because it does not require strict assumptions, can be applied to various data scales, and is suitable for small sample sizes [16]. While PLS-SEM is particularly useful for exploratory analysis and early-stage model development, any relationships identified without strong theoretical support should be interpreted with caution. The accurate and meaningful interpretation of these relationships still relies on a solid theoretical foundation [17]. Moreover, PLS can be applied in structural models that involve both formative and reflective indicators.

2.1 Structural Equation Modelling – Partial Least Square (SEM-PLS)

This study employs the Partial Least Squares Structural Equation Modeling (PLS-SEM) approach to analyze relationships between latent constructs and indicators, as well as to assess the predictive contributions among constructs. The evaluation of a PLS-SEM model involves two primary stages: measurement and structural model evaluations.

1. Measurement Model Evaluation

   The evaluation of the measurement model aims to ensure that the indicators used are reliable and valid in measuring the latent constructs. The following aspects were evaluated. First, indicator reliability was assessed through outer loading values. Indicators with loadings above 0.7 are considered to represent the construct adequately, although values between 0.6 and 0.7 can be accepted in exploratory research [16]. Internal consistency reliability was then evaluated using Composite Reliability (CR), which is more appropriate in PLS models than Cronbach's Alpha. CR values above 0.7 indicate satisfactory reliability [16]. Next, convergent validity was tested using Average Variance Extracted (AVE). AVE values above 0.5 confirm that the construct explains more than half of the variance of its indicators [18]. Finally, discriminant validity was evaluated using the Fornell–Larcker criterion and HTMT (Heterotrait-Monotrait Ratio). Discriminant validity is achieved if the square root of AVE is greater than the inter-construct correlations and HTMT values are below 0.9 [18]. These evaluations ensure that each construct in the model is distinct and properly measured by its assigned indicators.

2. Structural Model Evaluation

   Following measurement model validation, the structural model is evaluated to examine the hypothesized relationships among latent variables. The strength and direction of relationships are assessed through path coefficients, with their significance tested using the bootstrapping procedure to obtain p-values and t-values [16]. The coefficient of determination (R^2) is used to measure the proportion of variance in the endogenous variables explained by the exogenous constructs. Higher R^2 values indicate better explanatory power [16]. The effect size (f^2) is calculated to assess the impact of each exogenous variable on the endogenous variable [16]. Predictive relevance (Q^2) is determined via the blindfolding technique, where Q^2 values above zero indicate that the model has predictive relevance [19]. While PLS-SEM does not emphasize global model fit, the Standardized Root Mean Square Residual (SRMR) can be employed to assess model quality; an SRMR below 0.08 suggests acceptable model fit [16].

2.2 Finite Mixture in Partial Least Square

Generally, when estimating a structural equation model, researchers assume that the data collected is homogeneous [13]. However, this assumption is often incorrect and may be considered irrational, especially when data is collected from multiple units or segments [20]. K-means clustering and factor analysis combined with Ordinary Least Squares (OLS) regression are commonly used to address segmentation or grouping problems [14]. However, these methods cannot be applied when the variables involved are unobserved or latent. Although grouping techniques can be used in OLS models to manage heterogeneity, they are not suitable for models that include latent constructs or variables that cannot be directly measured. Therefore, Finite Mixture Partial Least Squares (FIMIX-PLS) is considered a more appropriate method, as it enables the identification of unobserved heterogeneity in models that incorporate latent variables [21].

The FIMIX-PLS estimation assumes that heterogeneity occurs in the structural model with the assumption of η_i Distributed Finite Mixture with a Normal Multivariate Density Function. $f_{i|k.}$ [22]

$$\boldsymbol{\eta}_i \sim \sum_{k=1}^{K} \rho_k f_{i|k} \left(\boldsymbol{\eta}_i | \boldsymbol{\xi}_i, \boldsymbol{B}_k, \boldsymbol{\Gamma}_k, \boldsymbol{\psi}_k\right). \tag{1}$$

The vector $\boldsymbol{\eta}_i$ represents a set of endogenous latent variables that reflect outcome constructs within the inner model of PLS for the i-th observation. These constructs are not directly measured but are estimated from multiple observed indicators through the measurement model. In the context of SEM analysis, $\boldsymbol{\eta}_i$ is considered latent, and its values are determined by the structural relationships defined in the model, including the influence of other constructs via path coefficients. The mixing proportion $\boldsymbol{\rho}_k$ indicates the probability weight of an observation belonging to the k-th latent class in the FIMIX-PLS model. Its value is always positive, and the sum of all mixing proportions across K classes must equal one ($\rho_k > 0$ and $\sum_{k=1}^{K} \rho_k$=1), This implies that each observation probabilistically belongs to one of the latent classes in the mixture model. The magnitude of ρ_k reflects the relative size or dominance of each class in shaping the overall data structure.

The conditional density function $f_{i|k}(\boldsymbol{\eta}_i | \boldsymbol{\xi}_i, \boldsymbol{B}_k, \boldsymbol{\Gamma}_k, \boldsymbol{\psi}_k)$ describes the likelihood of observing the values of the endogenous latent variables $\boldsymbol{\eta}_i$ for the i-th observation, given the exogenous latent variables $\boldsymbol{\xi}_i$ and the class-specific parameters $\boldsymbol{B}_k$, $\boldsymbol{\Gamma}_k$, and $\boldsymbol{\psi}_k$. This function encapsulates the structural information within class k and serves as the foundation for probability estimation in the mixture modeling framework. The matrix $\boldsymbol{B}_k$ contains the path coefficients representing relationships among endogenous latent constructs within class k. Each element in this matrix quantifies the extent to which one construct influences another within the same segment.

The matrix $\boldsymbol{\Gamma}_k$ includes the path coefficients that describe the influence of exogenous latent constructs on endogenous latent constructs, and is also specific to class k. Finally, the matrix $\boldsymbol{\psi}_k$ captures the variances of structural disturbances (i.e., residual errors) in the inner model. Each diagonal element of $\boldsymbol{\psi}_k$ represents the unexplained variance associated with an endogenous latent construction, those portions not accounted for by other constructs or by exogenous variables. This matrix is essential for understanding the stochastic component of the model.

The principle of likelihood is used in the estimation of the FIMIX-PLS model. In FIMIX PLS, the expectation-maximization (EM) algorithm to maximize the likelihood function, which is a combination of the Maximization (M) step and Expectation (E) step. The optimal number of segments in FIMIX-PLS can be seen through the criteria values such as Normed Entropy (EN), Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Consistent AIC (CAIC) [23].

2.3 Determining the Number of Segments

In FIMIX-PLS, The EN, AIC, BIC, and CAIC criteria are statistical measures used to determine the optimal number of segments. Normed Entropy (EN) ranges from 0 to 1, where a higher value, closer to 1, indicates better segment separation quality. These criteria are calculated using the following formulas [10]:

The Akaike Information Criterion (AIC) for class k is defined as:

$$AIC_k = -2 \ln L + \, c \cdot N_k, \tag{2}$$

where $c = 2$ is a constant and N_k represents the number of independent parameters in the model.

The Bayesian Information Criterion (BIC) for class k is calculated as:

$$BIC_k = -2 \ln L + \ln N \cdot N_k, \tag{3}$$

where N is the number of observations and c is given by $\ln(N)$.

The Consistent Akaike Information Criterion (CAIC) is expressed as:

$$CAIC_k = -2\ln L + (\ln I + 1) \cdot N_k, \tag{4}$$

where I denote the total number of observations, and c is defined as $(\ln I + 1)$.

The Normed Entropy (EN) is calculated using the following formula.

$$EN_k = 1 - \frac{\left[\sum_{i-1}^{I} \sum_{k=1}^{K} -P_{ik} \ln(P_{ik})\right]}{I \ln(K)}, \tag{5}$$

where P_{ik} is the posterior probability that the i-th observation belongs to the k-th class. A higher EN value indicates a clearer and more meaningful separation among segments. In general, better segmentation is associated with a higher NE score, which reflects stronger differentiation and better representation of the data's latent heterogeneity [24].

3. RESULTS AND DISCUSSION

This study uses secondary data from the official websites of the "Badan Pusat Statistik Kalimantan Barat: https://kalbar.bps.go.id/id" and "Badan Pusat Statistik Indonesia: https://www.bps.go.id/id" in 2023. The data has 14 observations according to regencies/cities with 19 indicators. The latent variables used are poverty, health, education, and economy.

Table 1. Latent Variables Indicators

Latent Variables	Code	Indicator
Health (ξ_1)	*SH1* (x_{11})	Life expectancy rate
	SH2 (x_{12})	Poor households who use decent water
	SH3 (x_{13})	Percentage of birth support helper with medical personnel
	SH4 (x_{14})	Percentage of women who are married and use birth control (15-49)
	SH5 (x_{15})	Percentage of population who have BPJS health contribution assistance Recipients
	SH6 (x_{16})	Percentage of households that have their own latrines
Education (η_1)	*PD1* (y_{12})	Literacy rate (percent)
	PD2 (y_{13})	Pure participation rate for senior high school (percent)
	PD3 (y_{14})	School participation rate of 16-18 years old
	PD4 (y_{15})	Open unemployment rate (percent)
	PD5 (y_{16})	Average length of school (year)
Economy (η_2)	*EK1* (y_{21})	Percentage of monthly average per capita expenditure for food
	EK2 (y_{22})	Percentage of poor people aged 15 and over working in the formal sector
	EK3 (y_{23})	Percentage of total labor force
	EK4 (y_{24})	Percentage of households by the main source of lighting
	EK5 (y_{25})	Percentage of poor people aged 15 and over working in the agricultural sector
Poverty (η_3)	*M1* (y_{31})	Number of poor populations
	M2 (y_{32})	Index of poverty depth
	M3 (y_{33})	Index of poverty severity

A path diagram of the constructed poverty structure model in the relationship between endogenous and exogenous variables is illustrated in Fig. 1. This study's structural model design combines three variables: economy, education, and poverty, as endogenous latent variables, and health as exogenous latent variables. The four latent variables are linked through pathways designed to test the hypotheses examined in the study.

3.1 Evaluation of the Measurement Model

The discriminant validity, convergent validity, and composite reliability are used to evaluate the measurement model.

3.1.1 Convergent Validity

The principle of convergent validity requires a good correlation between each indicator and the latent variable that constitutes it. The convergent validity does not meet the standard or is unacceptable if the outer loading value obtained is below 0.6. Furthermore, the re-testing modification is carried out by eliminating invalid indicators so that all indicators that construct each latent variable have valid outer loading values and can be continued for further analysis.

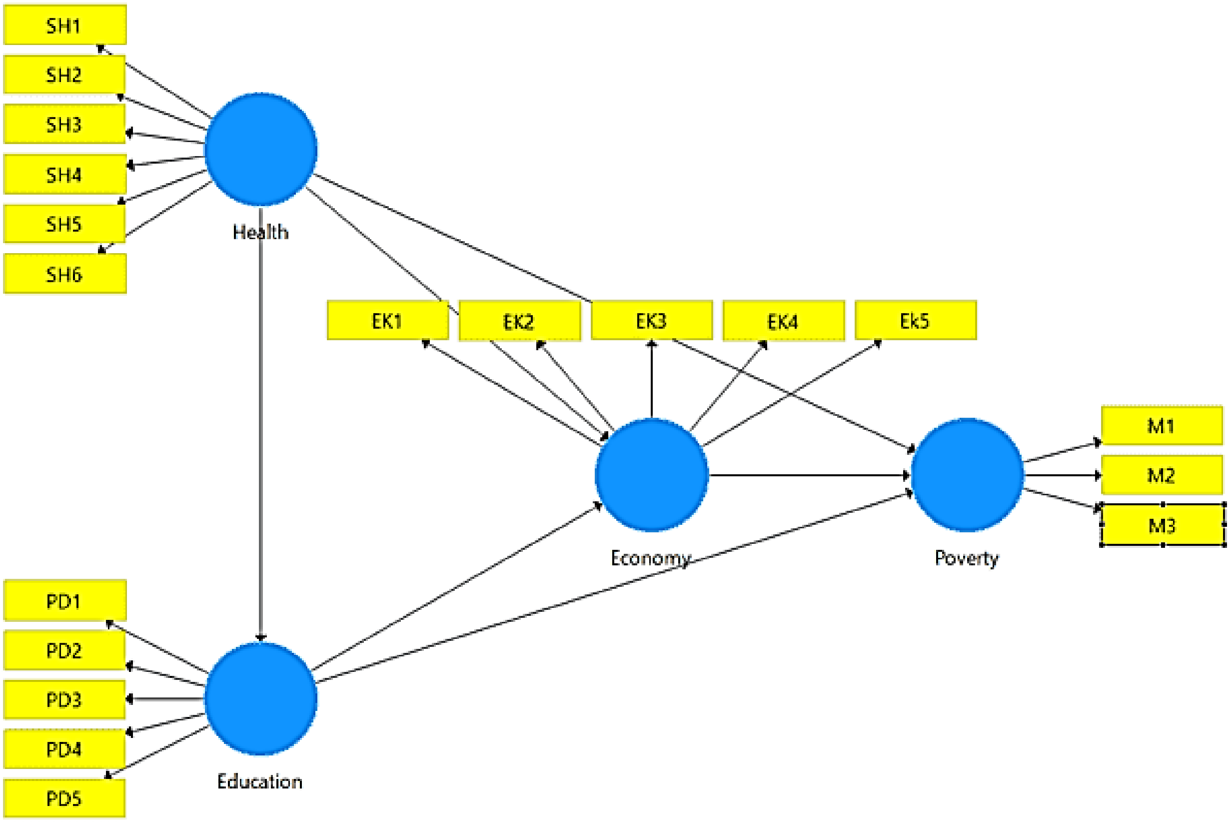

Figure 1. Structure of Poverty

The selection of variables and indicators in this model is based on a comprehensive literature review and adapted from prior empirical studies on multidimensional poverty [25] [10]. The constructions of health, education, and economy are employed as exogenous variables, while poverty serves as the endogenous outcome. Indicators for each construct were selected based on their theoretical relevance and contextual applicability to poverty analysis in Indonesia. The health variable includes six indicators reflecting access to basic healthcare and hygiene practices [26]. Education is measured through five indicators related to schooling and literacy [27]. Economic conditions are represented by indicators such as household expenditure, employment sector participation, and access to basic infrastructure [28]. Lastly, the poverty construct is measured using three key indicators adapted from the Multidimensional Poverty Index (MPI) framework [29] and regional poverty data [9].

For Table 2, there are still indicators with outer loading values below 0.6, so retesting is carried out by issuing invalid variables. Invalid variables include PD1, SH1, SH4, SH5, EK2, EK4, and M3.

Table 2. Outer Loading Value

Latent Variable	Indicator	Outer Loadings	Information
Health	*SH1*	0.034	Invalid
	SH2	0.753	Valid
	SH3	0.953	Valid
	SH4	-0.928	Invalid
	SH5	-0.050	Invalid
	SH6	0.609	Valid
Education	*PD1*	0.403	Invalid
	PD2	0.643	Valid
	PD3	0.750	Valid
	PD4	0.830	Valid
	PD5	0.750	Valid
Economy	*EK1*	0.721	Valid
	EK2	-0.768	Invalid
	EK3	0.920	Valid
	EK4	-0.799	Invalid
	EK5	0.907	Valid
Poverty	*M1*	0.926	Valid
	M2	0.903	Valid
	M3	0.525	Invalid

All indicators in Table 3 are declared appropriate or valid for further analysis because the external load value for each indicator is above 0.6.

Table 3. Outer Loading Value After Retesting

Latent Variable	Indicator	Outer Loadings	Information
Health	*SH2*	0.832	Valid
	SH3	0.900	Valid
	SH6	0.738	Valid
Education	*PD2*	0.644	Valid
	PD3	0.769	Valid
	PD4	0.844	Valid
	PD5	0.723	Valid
Economy	*EK3*	0.923	Valid
	EK5	0.940	Valid
Poverty	*M1*	0.988	Valid
	M2	0.978	Valid

After the convergent validity test, a new path diagram is obtained, which is presented in Fig. 2.

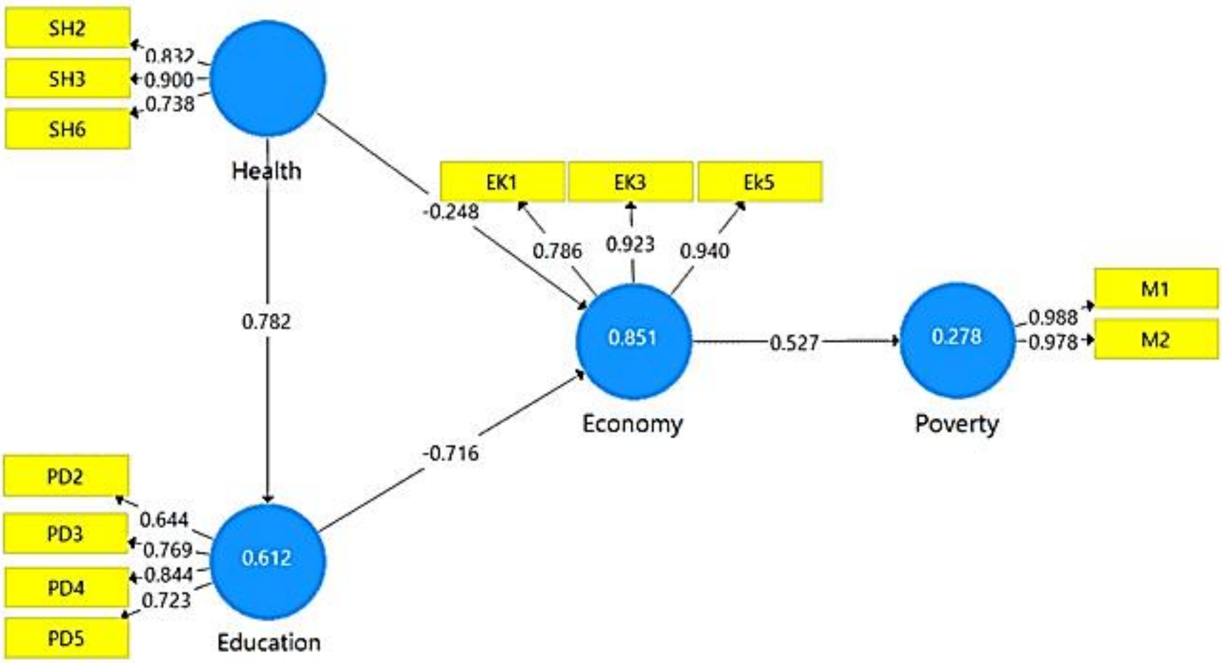

Figure 2. A New Form of Conceptual Model of Poverty Structure

Based on the new model form after a retesting test, valid indicators for the health variables are SH2, SH3, and SH6. For education, valid indicators are PD, PD3, PD4, and PD5. Valid indicators for the economic variables are EK1, EK2, and EK5. Finally, valid indicators for the poverty variables are M1 and M2.

3.1.2 Discriminant Validity

Discriminant validity is done to see how a latent variable differs from other latent variables. It can also ensure how well a construct can explain each indicator that makes it up. Discriminant validity looks at the values obtained on cross-loading, Fornell-Larcker Criterion, and AVE. The reflective indicator is valid based on the loading factor value in the convergent validity evaluation. This is supported by the AVE value for each of its latent variables greater than 0.5, as presented in Table 4.

Table 4. Cross Loading Values

Indicator	Economy	Poverty	Health	Education
SH2	-0.596	-0.572	0.832	0.513
SH3	-0.788	-0.764	0.900	0.856
SH6	-0.580	-0.522	0.738	0.487
PD2	-0.348	-0.219	0.396	0.644
PD3	-0.605	-0.326	0.601	0.769
PD4	-0.923	-0.582	0.816	0.844
PD5	-0.674	-0.328	0.389	0.723
EK1	0.789	0.356	-0.619	-0.645
EK3	0.923	0.582	-0.816	-0.844
EK5	0.940	0.439	-0.696	-0.903
M1	0.850	0.988	-0.763	-0.540
M2	0.436	0.978	-0.744	-0.488

In Table 4, the cross-loading value of each indicator on each construct is higher than that of the other constructs. Thus, it was concluded that it had good discriminant validity for all indicators. Apart from the cross-loading value, discriminant validity can be seen through the Fornell-Larcker Criterion and AVE criteria in Tables 5 and 6.

Table 5. Fornell-Larcker Criterion Value

Latent Variable	Economy	Poverty	Health	Education
Economy	0.886	-	-	-
Poverty	0.527	0.983	-	-
Health	-0.808	-0.767	0.826	-
Education	-0.909	-0.526	0.782	0.748

The Fornell-Larcker Criterion is considered satisfactory if the latent variable exhibits a higher value than it is correlated with other latent variables in the same column [18]. In Table 5, Economics has an AVE root value (√AVE) of 0.886 for the latent variable. This value is the largest compared to the correlation values of Economy with Poverty, Economy with Health, and Economy with Education. This also applies to the value of the latent variable √AVE of Poverty with Poverty of 0.983 and Health to Health of 0.826. However, the variable Education to Education is 0.748. for this reason, it is necessary to continue to see the AVE value listed in Table 6.

Table 6 shows that the AVE value in the Economic variables is 0.789, Poverty is 0.966, Health is 0.682, and Education is 0.560, indicating that all have good convergent validity values. Although it can be seen that the value in the Fornell-Larcker Criterion table for the relationship between the education variable and itself is not greater than the relationship with the health variable, this is still acceptable because there is enough evidence that the value of cross-loading and AVE is fulfilled.

Table 6. Average Variance Extracted Value of Each Latent Variable

Latent Variable	Average Variance Extracted (AVE)
Economy	0.784
Poverty	0.966
Health	0.682
Education	0.560

3.1.3 Composite Reliability

Study can be reliable if it has similar data at different times. If the composite reliability value is greater than or equal to 0.7, a variable can be said to meet the reliability [30].

Table 7. Composite Reliability and Cronbach's Alpha Values

Latent Variable	Cronbach's Alpha	Composite Reliability
Economy	0.861	0.916
Poverty	0.966	0.983
Health	0.769	0.865
Education	0.749	0.835

Table 7 shows that this study's reliability is acceptable, and all latent variables meet the required reliability criteria.

3.1.4 Equation for Measurement Model (Outer Model)

Based on Figure 2, the equation for the measurement model is obtained as follows.

1. Health Variable Measurement Model
   $x_{12} = 0.832\xi_1; x_{13} = 0.900\xi_1; x_{16} = 0.738\xi_1$
2. Educational Variable Measurement Model
   $y_{12} = 0.644\eta_1; y_{13} = 0.769\eta_1; y_{14} = 0.844\eta_1; y_{15} = 0.723\eta_1$
3. Economic Variable Measurement Model
   $y_{21} = 0.786\eta_2; y_{23} = 0.923\eta_2; y_{25} = 0.940\eta_2$
4. Poverty Variable Measurement Model
   $y_{31} = 0.988\eta_3; y_{32} = 0.978\eta_3$

3.2 Evaluation of the Structural Model

The evaluation of the structural model ensures that the relationship between the latent variables is significant and follows the hypothesis [19]. In SEM-PLS, structural evaluation is conducted by assessing the model's explanatory power through the coefficient of determination (R^2), calculated using the formula $R^2 = 1 -$ (Residual Sum of Squares/Total Sum of Squares). This value indicates the proportion of variance in the endogenous latent variable that is explained by the exogenous constructs. Additionally, to examine the significance of the relationships between constructs within the structural model, the path coefficient values are analyzed [31].

The R-square value in Table 8 for economics is 0.851, indicating moderate model [31]. This means that the 85.1% variation in the health and education variables can explain the economic variable. The education variable indicates that it can explain 61.2% of the variation in the health variable. The last variable is the poverty variable, meaning that 27.8% of the variation in the economic variable can be explained.

Table 8. Determination Coefficient Model Structural

Endogenous Latent Variables	R^2
Economy	0.851
Education	0.612
Poverty	0.278

The R-square value for the latent variable Poverty is low because several additional factors, such as economic influences and other unaccounted variables, contribute to Poverty. In SEM-PLS, the structural model was evaluated using bootstrapping to assess the path coefficients between latent variables.

Table 9. Path Coefficient Values Indirect Effects Using Bootstrap

Inner Model	Original Sample	t-statistic	p-value
Health -> Education -> Economy -> Poverty	-0295	2.101	0.036
Education -> Economy -> Poverty	-0.377	2.333	0.020
Health -> Education -> Economy	-0.560	2.714	0.007
Health -> Economy -> Poverty	-0.131	0.819	0.413

In Table 9, the p-value of the indirect effect of health on poverty through education and economics is significant, meaning that the education and economic variables significantly mediate the health variable on poverty. The p-value of education on poverty through the economy is also significant. Then, the p-value on the variable health towards the economy through education is also significant, which means that education significantly mediates between health and the economy. This result is reported as part of the overall model interpretation, and no model re-specification was conducted, as the focus remains on theoretically justified pathways and the model as a whole meets the required standards of validity and reliability. The insignificant pathway is acknowledged as a reflection of empirical findings, which can still offer valuable insight into the complexity of inter-variable relationships without altering the core model structure.

Table 10. Path Coefficient Values Direct Effects Using Bootstrap

Inner Model	Original Sample	t-Statistic	p-Value
Economy -> Poverty	0.527	2.812	0.005
Health -> Economy	-0.248	0.989	0.323
Health -> Education	0.782	14.023	0.000
Education -> Economy	-0.716	3.085	0.002

Results in Table 10 show that, based on the value of direct effects, three pathways significantly impact the latent variables of the economy on poverty, health on education, and education on the economy. The equation of the inner model or structural model can be expressed as follows.

1. Structural Model of Educational Variables
   $\eta_1 = 0.782\xi_1 + \zeta_1;$
2. Structural Model of Economy Variables
   $\eta_2 = -0.248\xi_1 - 0.716\eta_1 + \zeta_2;$
3. Structural Model of Poverty Variables
   $\eta_3 = 0.527\eta_2 + \zeta_3.$

Based on Table 10, three structural pathways show significant effects. First, health has a strong and significant positive effect on education ($\beta = 0.782$; $p < 0.001$), indicating that improvements in health contribute positively to educational outcomes. Second, education has a significant negative effect on the economy ($\beta = -0.716$; $p = 0.002$), suggesting that in this context, higher education levels may not directly translate into improved economic conditions, possibly due to labor market mismatch or structural economic limitations. Third, the economy has a significant positive effect on reducing poverty ($\beta = 0.527$; $p = 0.005$), meaning that better economic conditions are associated with lower poverty levels. The path from health to the economy ($\beta = -0.248$; $p = 0.323$) is not statistically significant and thus is not interpreted further in the structural model.

3.3 Group Determination with FIMIX-PLS

FIMIX-PLS produces several groups based on predetermined statistical criteria: BIC, AIC, CAIC, and EN values. Table 11 presents all these criterion values.

Table 11. AIC, BIC, CAIC, and EN Values in Each Segment

Criteria	$k = 2$	$k = 3$	$k = 4$	$k = 5$
AIC	83.953	83.053	**70.354**	70.737
BIC	93.539	97.751	**90.165**	95.660
CAIC	108.539	120.751	**121.165**	154.660
EN	0.879	0.937	**0.987**	0.971

Based on Table 11, $k = 4$ is the best segmentation because it gives the highest EN value of 0.987. Meanwhile, the AIC, BIC, and CAIC values were smaller than those of other segmentations. So, it can be concluded that four segments were formed. Each group formed tends to influence each other between latent variables.

Table 12 below describes the provincial grouping results in West Kalimantan based on the FIMIX-PLS results.

Table 12. Grouping of Regency/City Based on FIMIX-PLS

Segment	Count	Regency/City
1	5	Landak, Sanggau, Sekadau, Kayong Utara, Singkawang City
2	4	Kapuas Hulu, Melawi, Mempawah, Kubu Raya
3	3	Sambas, Sintang, Pontianak City
4	2	Bengkayang, Ketapang

The area included in Group 1 consists of five regencies/cities: Landak, Sanggau, Sekadau, Kayong Utara, and Singkawang City. The area included in Group 2 consists of four regencies/cities: Kapuas Hulu, Melawi, Mempawah, and Kubu Raya. The area included in Group 3 consists of three regencies/cities: Sambas, Sintang, and Pontianak City. The area included in Group 4 consists of two regencies/cities: Bengkayang and Ketapang. Table 13 summarizes the effect of each latent variable on other latent variables.

Table 13. Path Coefficient FIMIX-PLS on The Latent Variable of Each Segment

Path	Segment			
	1	2	3	4
Economy Poverty	0.414	0.906	**0.999**	-0.985
Health → Economy	-0.871	**-1.280**	0.176	1.493
Health → Education	**0.842**	0.908	0.853	0.786
Education → Economy	-0.150	0.394	-1.145	**10.787**

Table 13 shows that group 1, consisting of 5 districts/cities, has the characteristics of the influence of health on education greater than the influence of health on the economy, economy on poverty, and education on the economy. A positive score on the path coefficient indicates that improving health will increase the percentage of education in districts/cities in group 1. Group 2 consists of 4 regencies/cities with the characteristics of the influence of health on the economy.

Group 3 consists of 3 regencies/cities that have more significant characteristics of economic influence on poverty than groups 1, 2, and 4. A negative value on the path coefficient indicates that increasing education will reduce the percentage of the economy in districts/cities in group 3. Group 4 consists of 2 towns/regencies with characteristics of education's influence on the economy more significant than in groups 1, 2, and 3. A positive value in the path coefficient indicates that increasing education will increase the percentage of Economy in districts/cities in group 4. Based on the analysis results from FIMIX-PLS, Figure 3 is a map of the division of regions into four segments using ArcGIS software.

In Fig. 3, segment 1 represents regions with the influence of health on education; segment 2 represents areas with the characteristics of the impact of health on the economy. Segment 3 has the characteristics of the economy's influence on poverty, and Segment 4 shows the influence of education on the economy.

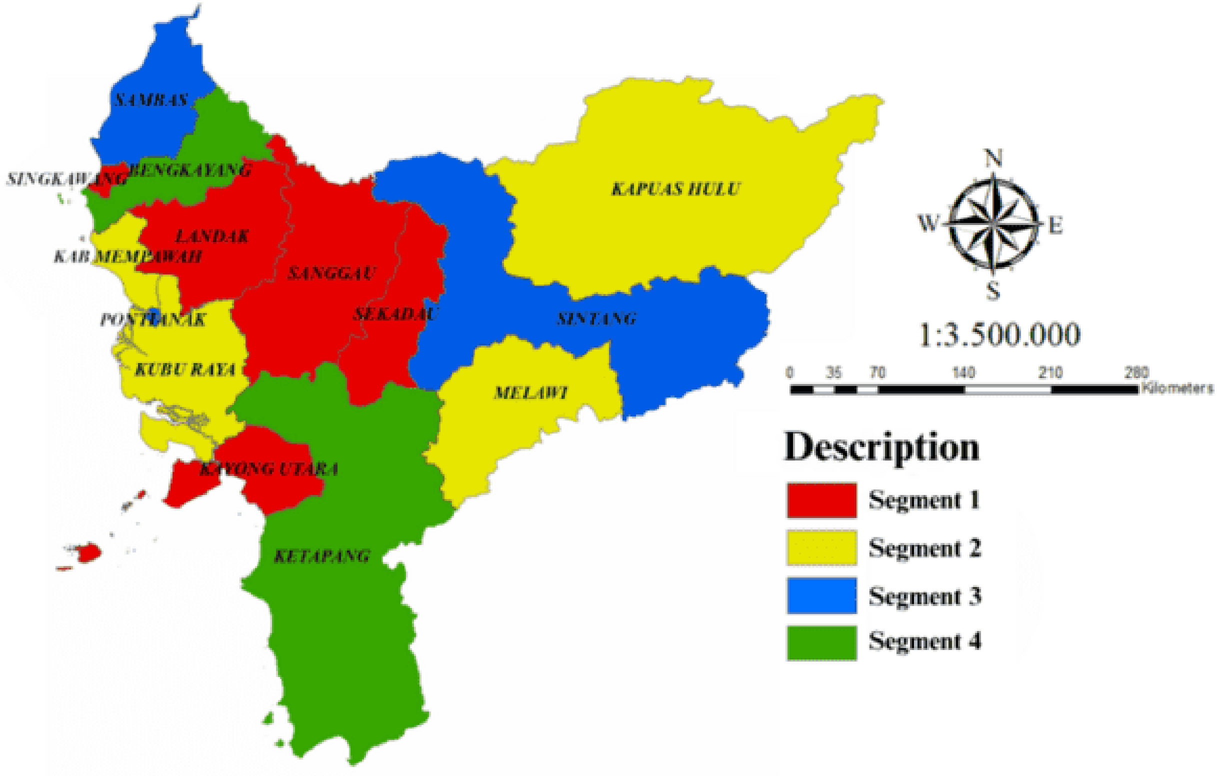

Figure 1. Map of Regency/City Segmentation Results Based on FIMIX-PLS

By looking at the average results of each segment of all indicators, what needs to be of concern to the government based on indicators with shortcomings in each segment described in Table 14.

Table 14. Average Results of All Indicators in Each Segment

Indicator	Segment 1	Segment 2	Segment 3	Segment 4
SH2	79.34	**75.65**	89.867	75.95
SH3	79.812	82.558	88.433	**78.07**
SH6	86.234	**84.67**	91.37	87.99
PD2	52.936	54.07	61.1	**52.15**
PD3	69.256	69.515	70.187	**69.125**
PD4	3.988	4.685	**5.627**	4.745
PD5	**7.274**	7.37	8.28	7.385
EK1	59.606	61.356	58.24	**58.215**
EK3	96.012	95.315	**94.373**	95.255
EK5	54.238	44.988	**39.717**	51.31
M1	6.898	7.18	6.57	**7.765**
M2	1.046	1.26	0.913	**1.4**

Based on Table 14, segment one needs to consider the percentage of the population with BPJS Contribution Assistance Recipients (PD5). Segment 2, namely, poor households that use decent water (SH2) and the percentage of poor people who use their latrines (SH6), also needs to be considered.

In segment 3, the indicators with the lowest average values are the Open Unemployment Rate (PD4), the percentage of employment relative to the labor force (EK3), and the percentage of low-income individuals working in the agricultural sector (EK5). In segment 4, the indicators with the smallest average values are the percentage of mothers receiving medical assistance during childbirth (SH3), the Pure Participation Rate of High School/Equivalent (PD2), the School Participation Rate 16-18 years old (PD3), the percentage of per capita expenditure on food (EK1), the percentage of the poor population (M1), and the poverty depth index (M2).

4. CONCLUSION

There are 12 valid and reliable indicators in poverty modeling in West Kalimantan. From the direct influence analysis results, some paths have a significant influence, namely the latent variables of the economy on poverty, health on education, and education on the economy. The indirect effects of economic and educational variables significantly mediate the relationship between health and poverty. Economics also significantly mediates between education and poverty, which further mediates the link between health and economics.

The resulting segmentation consists of four distinct groups. Segment 1, characterized by the influence of health and education, includes 5 regions: Landak, Sanggau, Sekadau, Kayong Utara, and Singkawang. Segment 2, which reflects the effect of health on the economy, includes Kapuas Hulu, Melawi, Mempawah, and Kubu Raya. Segment 3, dominated by the effect of the economy on poverty, includes Sambas, Sintang, and Pontianak City. Lastly, Segment 4, defined by the effect of education on the economy, includes Bengkayang and Ketapang.

These segmentation results are relevant for formulating area-specific poverty alleviation policies. As shown in Table 14, each segment has unique weaknesses in key indicators. For instance, Segment 1 should focus on improving access to BPJS assistance (PD5), while Segment 2 needs targeted intervention in clean water access (SH2) and sanitation (SH6). Segment 3 requires employment-based interventions (PD4, EK3, EK5), and Segment 4 demands educational support (PD2, PD3) and nutritional poverty-related assistance (EK1, M1, M2). Policy implementation should be tailored to each region’s segment profile, prioritizing the most vulnerable indicators to effectively reduce multidimensional poverty. The segmentation output provides actionable insights and supports the relevance of differentiated policy strategies across districts in West Kalimantan.

Author Contributions

Muhammad Fauzan: Conceptualization, Formal Analysis, Methodology, Writing - Original Draft. Hendra Perdana: Data curation, Project Administration, Software, Writing - Review and Editing. Neva Satyahadewi: Formal Analysis, Supervision, Validation, Visualization, Writing - Original Draft. All authors discussed the results and contributed to the final manuscript.

Funding Statement

The authors received no financial support for the research, authorship, and/or publication of this article.

Acknowledgment

The authors would like to express their gratitude and appreciation to all those who have already supported the completion of this work.

Declarations

The authors declare no conflicts of interest to report study.

REFERENCES

[1] United Nations Development Programme (UNDP), “ENDING POVERTY BY 2030: UNDP’S PERSPECTIVE AND ROLE,” New York, NY, US, 2016. [Online]. Available:: https://www.undp.org/sites/g/files/zskgke326/files/publications/ISSUE_BRIEF_Ending_Poverty_by_2030.pdf

[2] World Health Organization, “SOCIAL DETERMINANTS OF HEALTH - OVERVIEW,” 2024.

[3] A. Sumner, C. Hoy, and E. Ortiz-Juarez, “ESTIMATES OF THE IMPACT OF COVID-19 ON GLOBAL POVERTY,” UNU-WIDER, Helsinki, Finland, Apr. . Doi: https://doi.org/10.35188/UNU-WIDER/2020/800-9

[4] S. Gwariro et al., “POVERTY: WHAT IS IT AND WHAT IS IT NOT? A CONCEPT ANALYSIS,” *Int. J. Heal. Sci. Res.*, vol. 7, no. 5, pp. 345–353, 2017.

[5] Perkumpulan Prakarsa, “SATU DEKADE INDEKS KEMISKINAN MULTIDIMENSI INDONESIA 2012-2021,” Jakarta,

2023.

[6] S. Alkire, *THE EFFECT OF COVID-19 ON MULTIDIMENSIONAL POVERTY*, no. 10. 2020. [Online]. Available: www.mppn.org

[7] F. Bourguignon and S. Chakravarty, *THE MEASUREMENT OF MULTIDIMENSIONAL POVERTY*. Springer, 2019.doi: https://doi.org/10.1007/978-981-13-3432-0_7

[8] M. Zainal, M. U. Burhan, and I. Noor, "POVERTY ANALYSIS: POVERTY DETERMINANTS IN WEST KALIMANTAN PROVINCE 2010-2018," *South East Asia J. Contemp. Business, Econ. Law*, vol. 21, no. 5, pp. 78–86, 2020.

[9] Badan Pusat Statistik Kalimantan Barat, "PROFIL KEMISKINAN DI KALIMANTAN BARAT SEPTEMBER 2023," *Ber. Resmi Stat.*, 2023.

[10] E. D. Anggita, A. Hoyyi, and A. Rusgiyono, "ANALISIS STRUCTURAL EQUATION MODELLING PENDEKATAN PARTIAL LEAST SQUARE DAN PENGELOMPOKAN DENGAN FINITE MIXTURE PLS (FIMIX-PLS)," *J. Gaussian*, vol. 8, no. 1, pp. 35–45, 2019.doi: https://doi.org/10.14710/j.gauss.v8i1.26620

[11] R. N. Meilita, Y. N. Nasution, and M. N. Hayati, "STRUCTURAL EQUATION MODELLING DENGAN PENDEKATAN PARTIAL LEAST SQUARE (STUDI KASUS: PENGARUH LOCUS OF CONTROL, SELF EFFICACY, DAN LINGKUNGAN KERJA TERHADAP KINERJA KARYAWAN KALTIM POST SAMARINDA)," *Pros. Semin. Sains dan Teknol. FMIPA Unmul*, vol. 1, no. 1, pp. 41–45, 2016.

[12] F. Hahn, C. Johnson, M. D. Herrmann, and A. Huber, "CAPTURING CUSTOMER HETEROGENEITY USING A FINITE MIXTURE PLS APPROACH," *Schmalenbach Bus. Rev.*, vol. 54, no. 3, pp. 243–269, 2002.doi: https://doi.org/10.1007/BF03396655

[13] M. Fordellone and M. Vichi, "STRUCTURAL EQUATION MODELING AND SIMULTANEOUS CLUSTERING THROUGH THE PARTIAL LEAST SQUARES ALGORITHM," pp. 1–32, 2018.

[14] J. Henseler, "PARTIAL LEAST SQUARES PATH MODELING: QUO VADIS?," *Qual. Quant.*, vol. 52, no. 1, pp. 1–8, 2018, doi: https://doi.org/10.1007/s11135-018-0689-6.

[15] H. Joshi and D. Chawla, "IDENTIFYING UNOBSERVED HETEROGENEITY IN MOBILE WALLET ADOPTION – A FIMIX-PLS APPROACH FOR USER SEGMENTATION," *Int. J. Bank Mark.*, vol. 41, no. 1, pp. 210–236, 2023, doi: https://doi.org/10.1108/IJBM-03-2022-0132.

[16] J. F. Hair, G. T. Hult, C. Ringle, and M. Sarstedt, *A PRIMER ON PARTIAL LEAST SQUARES STRUCTURAL EQUATION MODELING (PLS-SEM)*. 2017.

[17] Z. Hikmah, H. Wijayanto, and M. N. Aidi, "COMPARISON OF THE CB-SEM AND PLS-SEM MODELS IN THE INFORMATION AND COMMUNICATIONS TECHNOLOGY INDUSTRY IN INDONESIA," *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 17, no. 2, pp. 1151–1164, 2023.doi: https://doi.org/10.30598/barekengvol17iss2pp1149-1162

[18] M. R. Ab Hamid, W. Sami, and M. H. Mohmad Sidek, "DISCRIMINANT VALIDITY ASSESSMENT: USE OF FORNELL & LARCKER CRITERION VERSUS HTMT CRITERION," *J. Phys. Conf. Ser.*, vol. 890, no. 1, 2017, doi: https://doi.org/10.1088/1742-6596/890/1/012163.

[19] J. F. Hair, G. T. M. Hult, C. M. Ringle, M. Sarstedt, N. P. Danks, and S. Ray, *PARTIAL LEAST SQUARES STRUCTURAL EQUATION MODELING (PLS-SEM) USING R: A WORKBOOK*. Springer International Publishing, 2021. doi: https://doi.org/10.1007/978-3-030-80519-7_5.

[20] V. Esposito Vinzi, L. Trinchera, S. Squillacciotti, and M. Tenenhaus, "REBUS-PLS: A RESPONSE-BASED PROCEDURE FOR DETECTING UNIT SEGMENTS IN PLS PATH MODELLING," *Appl. Stoch. Model. Bus. Ind.*, vol. 24, pp. 439–458, Sep. 2008, doi: https://doi.org/10.1002/asmb.728.

[21] M. A. Mukid, B. W. Otok, and A. Suharsono, "SEGMENTATION IN STRUCTURAL EQUATION MODELING USING A COMBINATION OF PARTIAL LEAST SQUARES AND MODIFIED FUZZY CLUSTERING," *Symmetry (Basel).*, vol. 14, no. 11, 2022, doi: https://doi.org/10.3390/sym14112431.

[22] C. M. Ringle, A. Will, and S. Wende, *THE FINITE MIXTURE PARTIAL LEAST SQUARES APPROACH - METHODOLOGY AND APPLICATION*. 2010.doi: https://doi.org/10.1007/978-3-540-32827-8_9

[23] H. Nassiri and S. I. Mohammadpour, "INVESTIGATING SPEED-SAFETY ASSOCIATION: CONSIDERING THE UNOBSERVED HETEROGENEITY AND HUMAN FACTORS MEDIATION EFFECTS," *PLoS One*, vol. 18, no. 2 February, pp. 1–26, 2023, doi: https://doi.org/10.1371/journal.pone.0281951.

[24] A. Kuckertz, C. Bulut, and L. Brändle, "UNOBSERVED HETEROGENEITY IN FIRM PERFORMANCE: THE ALIGNMENT OF ENTREPRENEURIAL ORIENTATION AND ORGANIZATIONAL ERROR MANAGEMENT CULTURE," *J. Bus. Res.*, vol. 179, no. December 2022, 2024, doi: https://doi.org/10.1016/j.jbusres.2024.114701.

[25] R. R. Nuryanti and T. Soebagijo, "SEM-PLS UNTUK ANALISIS STRUKTUR KEMISKINAN PADA MASA PANDEMI COVID-19," *Semin. Nas. Off. Stat.*, vol. 2021, no. 1, pp. 195–203, 2021, doi: https://doi.org/10.34123/semnasoffstat.v2021i1.836.

[26] B. Adhitya, A. Prabawa, and H. Kencana, "ANALISIS PENGARUH PENDIDIKAN, KESEHATAN, SANITASI DAN RATA-RATA JUMLAH ANGGOTA KELUARGA PER RUMAH TANGGA TERHADAP KEMISKINAN DI INDONESIA," *Ekon. J. Econ. Bus.*, vol. 6, no. 1, p. 288, 2022, doi: https://doi.org/10.33087/ekonomis.v6i1.501.

[27] A. R. Hakim, "KONSEP LANDASAN DASAR PENDIDIKAN KARAKTER DI INDONESIA," *J. Educ.*, vol. 6, no. 1, pp. 2361–2373, 2023, doi: https://doi.org/10.31004/joe.v6i1.3258.

[28] E. P. Sari and Novianti, "PENGARUH PDRB TERHADAP KEMISKINAN DI KALIMANTAN BARAT TAHUN 2017–2022," *EKODESTINASI J. Ekon. Bisnis, dan Pariwisata*, vol. 2, no. 1, pp. 36–56, 2024, doi:

https://doi.org/10.59996/ekodestinasi.v2i1.409.
[29] Alkire, S., Kanagaratnam, U., and N. Suppa, "THE GLOBAL MULTIDIMENSIONAL POVERTY INDEX (MPI) 2020," 2020.
[30] A. Maru, W. Jekale, and B. Asteray, "INVESTIGATING THE IMPACT OF LEAN CONSTRUCTION PRINCIPLES ON CONTRACTORS' PROJECT PERFORMANCE IN ETHIOPIA USING PLS-SEM," *J. Proj. Manag.*, vol. 9, no. 3, pp. 227–238, 2024, doi: https://doi.org/10.5267/j.jpm.2024.4.004.
[31] J. Hair and A. Alamer, "PARTIAL LEAST SQUARES STRUCTURAL EQUATION MODELING (PLS-SEM) IN SECOND LANGUAGE AND EDUCATION RESEARCH: GUIDELINES USING AN APPLIED EXAMPLE," *Res. Methods Appl. Linguist.*, vol. 1, no. 3, 2022, doi: https://doi.org/10.1016/j.rmal.2022.100027.