Mathematics and Statistics 11(1): 1-12, 2023 http://www.hrpub.org
DOI: 10.13189/ms.2023.110101

Development of Nonparametric Structural Equation Modeling on Simulation Data Using Exponential Functions

Tamara Rezti Syafriana[*], Solimun, Ni Wayan Surya Wardhani, Atiek Iriany, Adji Achmad Rinaldo Fernandes

Department of Statistics, Faculty of Mathematics and Natural Sciences, Brawijaya University, Malang, Indonesia

Received August 1, 2022; Revised October 27, 2022; Accepted November 14, 2022

Cite This Paper in the Following Citation Styles

(a): *[1] Tamara Rezti Syafriana, Solimun, Ni Wayan Surya Wardhani, Atiek Iriany, Adji Achmad Rinaldo Fernandes , "Development of Nonparametric Structural Equation Modeling on Simulation Data Using Exponential Functions," Mathematics and Statistics, Vol. 11, No. 1, pp. 1 - 12, 2023. DOI: 10.13189/ms.2023.110101.*

(b): *Tamara Rezti Syafriana, Solimun, Ni Wayan Surya Wardhani, Atiek Iriany, Adji Achmad Rinaldo Fernandes (2023). Development of Nonparametric Structural Equation Modeling on Simulation Data Using Exponential Functions. Mathematics and Statistics, 11(1), 1 - 12. DOI: 10.13189/ms.2023.110101.*

Abstract **Objective:** This study aims to determine the development of nonparametric SEM analysis on simulation data using the exponential function. **Methodology:** This study uses simulation data which is defined as an experimental approach to imitate the behavior of the system using a computer with the appropriate software. This study uses nonparametric structural equation modeling (SEM) analysis. The function used in this study is the exponential function. **Results:** The results showed that with simulation data all relationships have a significant effect on each other which have formative and reflective indicators. Testing the direct effect of Y2 on Y3 produces a structural coefficient value of 0.255 with a p-value <0.001 which means it is significant. The structural coefficient is positive, indicating that the relationship between the two is positive. This means that the higher Y2, the higher Y3. The results of the measurement model get a coefficient of determination of 0.91. It can be explained that 91% of the diversity of variables Y1, Y2, and Y3 can be explained by the X1 variable while 9% is explained by other variables not used in the model. **Novelty:** This study uses simulation data that is made very complex to analyze several related system structures at one time and can use a lot of data to get closer to real conditions to obtain comprehensive results, adjust to the criteria to be studied, and meet the following criteria: nonparametric SEM analysis criteria using the exponential function.

Keywords Structural Equation Model, Linearity, Parametric, Nonparametric

1. Introduction

SEM is the development of regression analysis and path analysis that tests the indicator model simultaneously, which is called structural equation modeling, and tests the relationship between variables [9]. SEM is used to determine the relationship between variables that cannot be measured directly (latent variables) which are complex because they involve a number of predictors and response variables that are interconnected to form a model. SEM is known as a combination of two models, namely the structural model and the measurement model [12]. Structural model is a model that describes the relationships that exist between latent variables. While the measurement model is a model that describes the relationship between latent variables and observed variables (indicators). Based on the measurement process, variables can be categorized into manifest variables and latent variables [10]. In general, latent variables are defined as variables that cannot be measured directly, but these variables must go through indicators that reflect and build on them [9].

The measurement model basically consists of two types, namely reflective and formative, with the main difference being that reflective measurement (based on factor analysis) requires the presence of a common factor because it is an indicator model that reflects variables [13]. Meanwhile, formative measurements (based on Principal Component Analysis/PCA) do not require the presence of a common factor, or in other words, the indicators are not correlated with each other [1].

Before conducting further tests of SEM analysis, there are assumptions that must be met to determine the form of the relationship, namely the assumption of linearity. The assumption of linearity has an influence on the shape of the model, if the assumption of linearity is fulfilled then parametric path analysis is used. However, if the assumption of linearity is not met, then nonparametric path analysis is used [14]. Meanwhile, if the results of the linearity assumption have several linear relationships and there are several non-linear relationships, semiparametric path analysis can be used. The Regression Specification Error Test (RESET) method is one of the methods used to determine the relationship between variables using linearity test.

There are three modeling approaches to path analysis, one of which is a nonparametric approach. The nonparametric approach of the form of the regression function is assumed to be unknown and the linearity assumption is not met [11]. The existence of a relationship with a nonparametric approach between latent variables is a reason for the emergence of nonparametric SEM analysis [8]. When one of the assumptions in SEM is not met and the form of the relationship between variables is not known, the study has been included in nonparametric SEM. In other words, nonparametric SEM is SEM that is caused because the assumption of linearity between latent variables is not met.

One of the nonparametric path analyses is spline. Hidayat, et al. [5] explained that splines are parts of a polynomial that have continuous and segmented (truncated) properties. The advantage of truncated spline path analysis is that it can predict the function when there is a shift in the data pattern. Researchers that use a nonparametric approach include: Efendi, et al. [2] performed linear, quadratic, and cubic nonparametric spline analysis modeling in the timely model of bank credit; Rasyidah, et al. [7] developed a nonparametric regression lined path analysis using the PWLS approach.

This study uses generated data which is defined as an experimental approach to imitating the behavior of a very complex system that is suitable for analyzing several related system structures at one time and can use a lot of data to get closer to real conditions [15]. The simulation is carried out to obtain comprehensive results, adjust to the criteria to be studied, and meet the criteria for nonparametric SEM analysis.

2. Literature Review

2.1. Measurement Model

There are two measurement models that are often used in structural model analysis, namely Principal Component Analysis (PCA) which is commonly used in formative indicator models, and Factor Analysis (FA) which is commonly used in reflective indicator models [16]. The formative indicator model does not require the assumption that all indicators have a common factor. In the formative indicator model, it is not necessary to assume a consistent correlation between indicators. So PCA is used to reduce the number of dimensions of the set of variables consisting of correlated variables into new uncorrelated variables by maintaining as much diversity as possible in the information contained in the set [17]. While factor analysis is used to group several variables which will later be used as one factor that has similarities.

2.2. Linearity Assumption

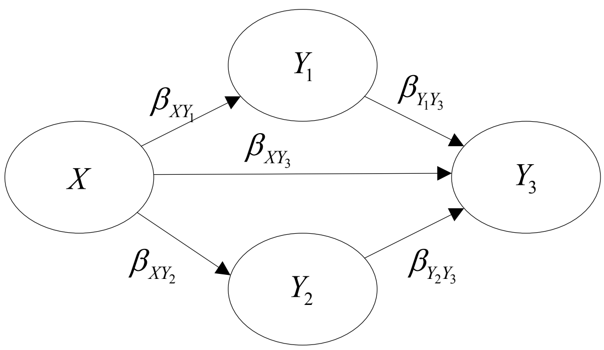
The linearity assumption states that the relationship between the response variable and the predictor variable is correct, which means that the regression curve can be expressed in linear, quadratic, or cubic form. If the assumption of linearity is met, then a parametric approach is used. However, if the linearity assumption is not met, then a nonparametric approach is used. Meanwhile, if the results of the linearity assumption have several linear relationships and there are several non-linear relationships, a semiparametric approach can be used [18]. One method to test the assumption of linearity is the Regression Specification Error Test or RESET which was first introduced in 1969 by Ramsey.

The steps for implementing RESET are [4]:

1) Perform regression analysis using one predictor variable to get the fitted value of the response variable from equation (1).
$$y_i = \beta_0 + \beta_1 X_{i1} + \varepsilon_i \tag{1}$$
2) Perform regression analysis by entering the fitted value obtained from equation (1) as a new predictor variable with the regression equation model in equation (2):
$$y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 y_i^2 + \beta_3 y_i^3 + \varepsilon_i \tag{2}$$
3) Then, we get the value of R_1^2 from equation (1) and the value of R_2^2 from equation (2). After the value R_1^2 and R_2^2 of the two equations are known, then calculate the value of the F test statistic.
4) Based on the value of F, then compare the value of F with the value of F_{table}.

2.3. Path Analysis

Path analysis is a regression development method that involves more than one endogenous variable to determine the direct, indirect, and total effect of exogenous variables on endogenous ones [19]. Path analysis is not a method to find the cause but to find out the magnitude of the causal effect based on the knowledge and theoretical considerations of the researcher. In path analysis, there is at least one exogenous variable (X), two mediating variables ($Y1$) and (Y2), and one endogenous variable (Y3). One form of a path diagram is shown in Figure 1.



Source: Solimun (2010)

Figure 1. Path Diagram

Figure 1, shows that X affects Y_1 and Y_2, Y_1 affects Y_2, while Y_2 is influenced by X and Y_1. From this simple linear path diagram, the following model can be formed:

$$Y_{1i} = \beta_{01} + \beta_{XY1}X_i + \varepsilon_{1i};\ \varepsilon_{1i}\ NIID(0,\sigma^2)$$

$$Y_{2i} = \beta_{02} + \beta_{XY2}X_i + \beta_{Y_1Y_2} + \varepsilon_{1i};\ \varepsilon_{1i}\ NIID(0,\sigma^2) \quad (3)$$

There are three types of path analysis effects, namely [10]:

1) Direct Effect
2) Indirect Effect
3) Total Effect

2.4. Exponential Function

One of the distributions that are widely used in statistics, especially stochastic processes, is the exponential distribution. The exponential distribution is a special case of the gamma distribution. This distribution is widely used as a model in engineering and science [20]. The continuous random variable X has an exponential distribution with the parameter β, if the solid function is given with β > 0, the exponential distribution is also called the gamma distribution with the exponential distribution being a distribution that is useful for finding the time difference that occurs in a certain opportunity. In this exponential distribution, search or data processing is used using random variables. Where the random variable itself is a variable in the form of a value or number which is the outcome of a random experiment. A random variable is discrete if only a certain value can be calculated. However, a random variable is continuous if it is any value in an interval.

2.5. Nonparametric SEM

The existence of a relationship with a nonparametric approach between latent variables is a reason for the emergence of nonparametric SEM analysis [8]. When one of the assumptions in SEM is not met and the form of the relationship between variables is not known, the study has been included in nonparametric SEM [21]. In other words, nonparametric SEM is SEM that is caused because the assumption of linearity between latent variables is not met. One of the initial identifications of the existence of a relationship with a nonparametric approach between latent variables is a scatter plot between exogenous latent variables and nonlinear endogenous latent variables. Based on the development of regression analysis and nonparametric path analysis, a nonparametric structural model equation can be formed.

2.6. Model Suitability

The appropriate model is a model that has a good size at the evaluation stage. In this study, two measures are used, namely the Root of Mean Square Error (RMSE) and coefficient of determination (R^2). RMSE is a measure used to evaluate the model by calculating the difference between the observed value y_{kit} and the estimated value $\hat{f}_{kit}$ for each response, so that it is obtained[3]:

$$RMSE_{y_k} = \sqrt{\frac{\sum_{i=1}^{N}\left(y_{ki}-\hat{f}_{ki}\right)^2}{N}}, k = 1,2. \quad (4)$$

The value of R^2 or the coefficient of determination measures how big the ratio between the diversity of data that can be explained by the model compared to the total diversity of the data.

$$R_{y_k}^2 = 1 - \frac{\sum_{i=1}^{N}\left(y_{ki}-\hat{f}_{ki}\right)^2}{\sum_{i=1}^{N}(y_{ki}-\bar{y}_k)^2}, k = 1,2. \quad (5)$$

The best/suitable model is the model that has the smallest value of $RMSE_{y_1}$ and $RMSE_{y_2}$, the largest values of $R_{y_1}^2$ and $R_{y_2}^2$.

2.7. Simulation Data

Simulation is the process of planning a mathematical or logical model of a real system and conducting experiments on the model using a computer to describe, explain and predict the behavior of the system [6]. The simulation method is the design of a model of a real system and the execution of experiments with this model to understand the behavior of the system or to develop a strategy (within a limit determined by one or more criteria) about the operation of the system. The simulation method can explain the behavior of a mathematical model that is made according to the character of the original system so that an analyst can conclude about the behavior of the real-world system.

3. Methodology

The data used in this research is simulation data on various experimental functions, namely the exponential function. Analysis tools using R Software.

3.1. Research Steps

In order to compare the ability of the truncated spline estimator in nonparametric SEM on the variance-covariance matrix which considers the correlation between the response variables with the variance-covariance matrix at the correlation level ($|\rho|$= 0.3), with the following steps:

1) Set a nonparametric SEM model, and N=100 (sample size)
2) Set value of $x_i, i = 1,2,\ldots,N$, $x_i \in [0.1]$ or $x_i = \frac{2i-1}{2N}$.
3) Generate random error $\underset{\sim}{\varepsilon}$ N-variat distributed (=2N) with $E(\underset{\sim}{\varepsilon}) = \underset{\sim}{0}$ and $\mathrm{Var}(\underset{\sim}{\varepsilon}) = \boldsymbol{\Sigma}$ under the following two conditions:
4) Condition 1: The random error variance-covariance matrix $\boldsymbol{\Sigma}_{\mathrm{MK}}$ considers the correlation between responses as follows:
$$\boldsymbol{\Sigma}_{\mathrm{MK}} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}$$
5) Condition 2: The random error variance-covariance matrix $\boldsymbol{\Sigma}_{\mathrm{MK}}$ considers the correlation between responses as follows:
$$\boldsymbol{\Sigma}_{\mathrm{MK}} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{22} \end{bmatrix}$$
6) Variation of correlation value sets $|\rho_i|$ = 0.3 on the sub-matrix $\boldsymbol{\Sigma}_{12}$.
7) Setting variables and indicators on exogenous, mediating, and endogenous variables in Figure 1
8) Designing a structural model (inner model) by determining the relationship between latent variables.
9) Designing a measurement model (outer model) by determining the indicator properties of the latent variables, namely formative and reflective.
10) Construct a path diagram based on the design of the inner model and outer model.
11) Convert the path diagram into a system of equations
12) Performing PCA and FA analysis of indicators of endogenous intervening variables that have a reflective formative measurement model
13) Perform linearity assumptions to determine the form of the relationship between variables
14) Perform nonparametric SEM analysis.
15) Determination of the best model
16) Perform hypothesis testing for the formed model.
17) Checking the validity of the model
18) Interpret the magnitude of the path analysis coefficient and the effect in the formed model

The measurement and structural model in this study can be seen in Figure 2.

Description:

Xj: exogenous latent variable j=1,2

Ym: endogenous latent variable m=1,2

$\boldsymbol{\lambda}_{xij}$: exogenous variable loading factor

$\boldsymbol{\lambda}_{yim}$: endogenous variable loading factor

$\boldsymbol{\beta}$: coefficient of influence of latent variable

ε: error model

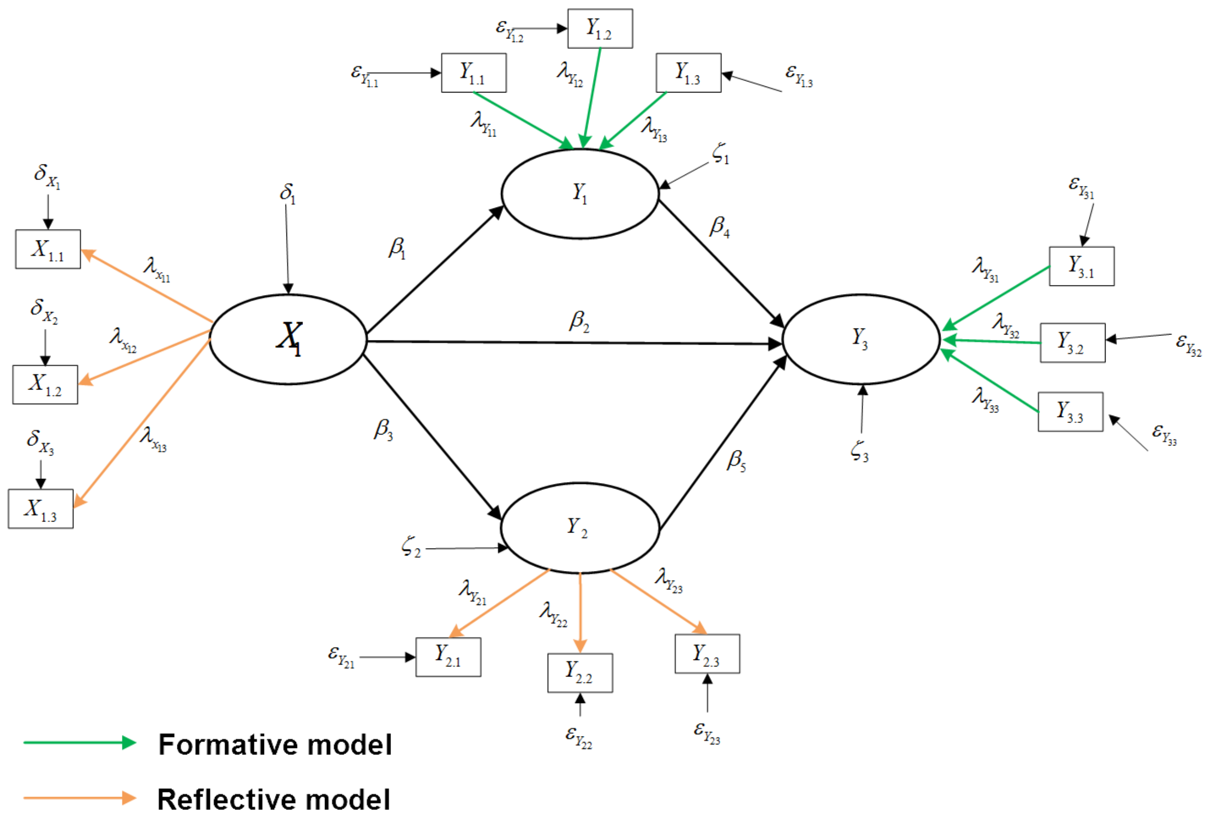
δ: measurement error in manifest variable for exogenous latent variable

ζ : measurement error in manifest variables for endogenous latent variables

4. Results and Discussion

The study used simulation data by setting n=100 and (|ρ| = 0.3). There are 4 variables used in this study, namely one exogenous variable and three endogenous variables. The results of PCA and FA analysis will form data based on the variables shown in Table 1.

Based on Table 1, the results of PCA and FA analysis as a reference in the simulation process. To determine the relationship between variables, a linearity test was performed. The results of the linearity test are in Table 2.



Source: Solimun, et. al. (2017)

Figure 2. Measurement Model and Structural Model

Table 1. PCA and FA Results

n	X1	Y1	Y2	Y3
1	-0.9382	5.9596	-0.0694	3.4389
2	-0.2620	7.1263	-0.1836	-0.1163
3	2.3001	6.3257	1.5889	1.5289
:	:	:	:	:
99	-2.5114	5.1303	2.9387	4.1271
100	3.1846	5.1510	1.875	4.1271

Table 2. Linearity Test

Relationship	*p*-value	Decision
X1~Y1	<0.001	Non-Linear
X1~Y2	<0.001	Non-Linear
X1~Y3	<0.001	Non-Linear
Y1~Y3	0.6326	Linear
Y2~Y3	0.7063	Linear

Based on the results of Table 2, it can be explained that there is a non-linear relationship, namely the relationship between X1~Y1, X1~Y2, and X1~Y3 because it accepts H0. P-Value is less than 0.05. In addition, the relationship between Y1~Y3 and Y2~Y3 has a p-value greater than 0.05 so it rejects the hypothesis which means it is linear. The results of secondary data have two forms, namely, there is a linear relationship and there is a non-linear relationship. The results of Table 2 are presented against the data pattern in Figure 3.

In SEM analysis there are two models, namely the outer model and the inner model. The first stage in this research is the outer model (measurement model). There are two outer models of WarpPLS SEM, namely the reflective model and the formative model. In the outer model, there are values for outer loading (for reflexive indicators) and outer weight (for formative indicators) which show the weight of each indicator as a measure of each latent variable. The indicator with the largest outer loading or outer weight indicates that the indicator is the strongest (dominant) variable measure.

1) Measurement Model of X1

Variable X1 has three indicators based on Table 3 to determine the measurement model, the value of the weight of the measurement, and the p-value of each indicator of the attitude variable which is presented in Table 3.

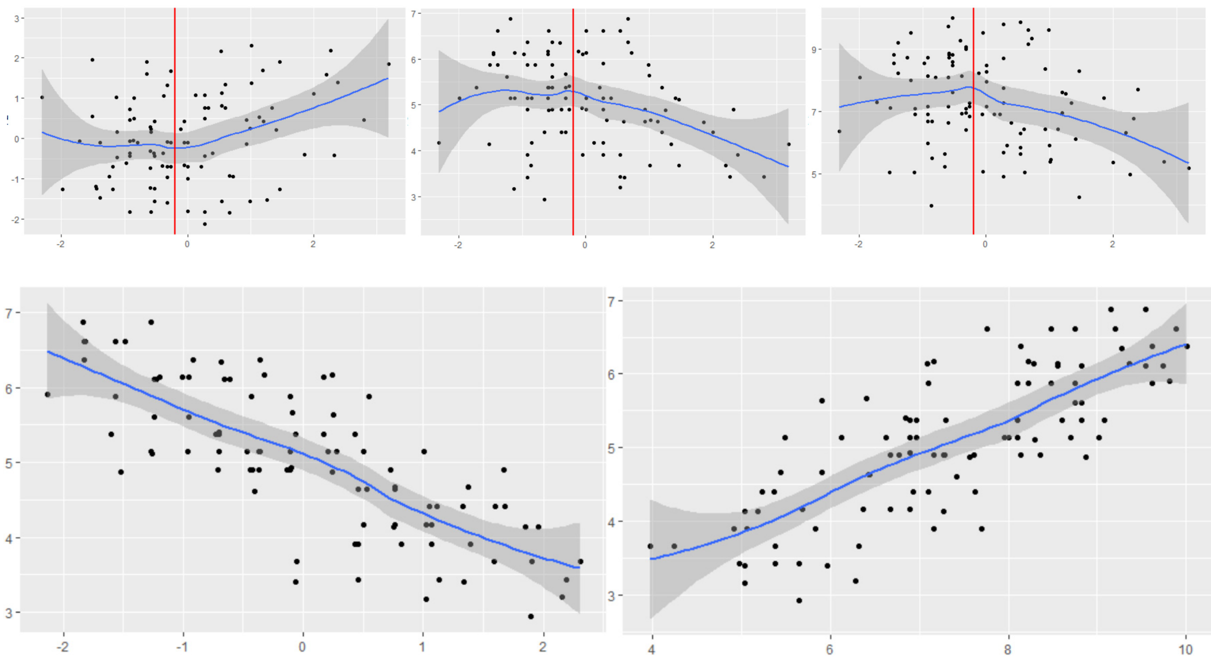
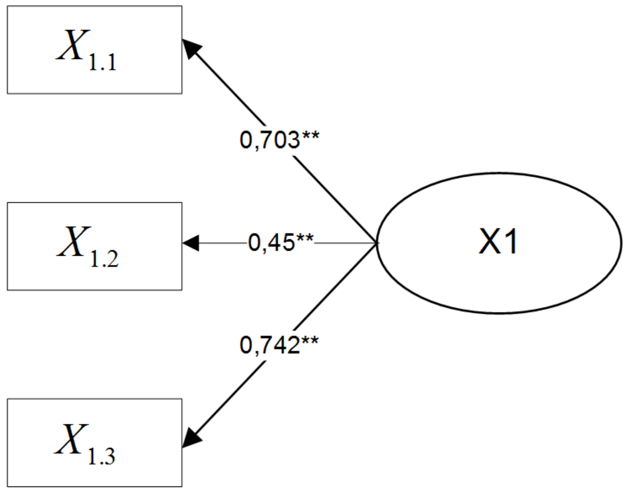


Figure 3. Relationship Curves Between Variables

Table 3. Variable Loading Factor Results (X1)

Variable	Indicator	Measurement Model	Outer Loading	p-value
X_1	X_{11}	Reflective	0.703	<0.001
	X_{12}	Reflective	0.45	<0.001
	X_{13}	Reflective	0.742	<0.001



Source: Researcher (2022)

Figure 4. Measurement Model of X1

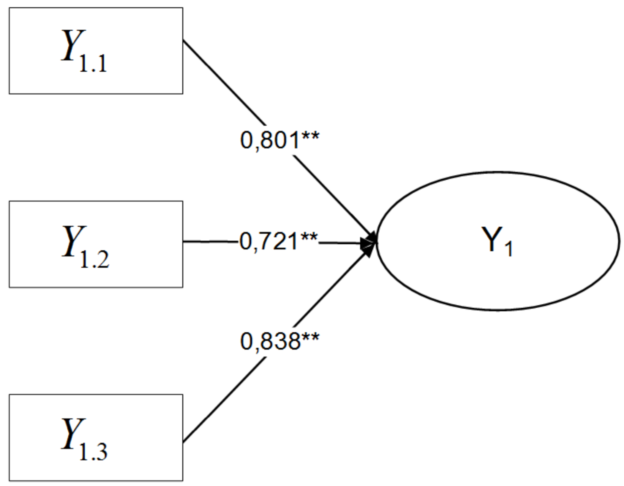
Based on the results of Table 3, the value of the outer loading indicator variable X1 can be presented in Figure 4.

Based on the study in Figure 4, it is explained that the X1 variable is the strongest indicator, namely X11 with an outer loading value of 0.703. The second strongest indicator is X12 with an outer loading value of 0.45 and followed by an indicator in the third position, namely the X13 component with an indicator value of 0.742. Referring to Table 4, the p-value of the three indicators is <0.001 where the three indicators are significant and can reflect the X1.

2) Measurement Model of Y1

The Y1 variable has three indicators based on Table 4 to determine the measurement model, the measurement weight value, and the p-value of each Y1 variable indicator which are presented in Table 4.

Based on the results of Table 4, the value of the Y1 variable outer weight indicator can be presented in Figure 5.



Source: Researcher (2022)

Figure 5. Measurement Model of Y1

Based on the study in Figure 5, it is explained that in the Y1 variable, the strongest indicator is Y13 with an external weight value of 0.838. The second strongest indicator is Y11 with an outer weight value of 0.801 and followed by an indicator in the third position, namely Y12 with an outer weight value of 0.721.

Referring to Table 4, the p-value of the three indicators is < 0.001 where the three indicators are significant and can form the Y1 variable.

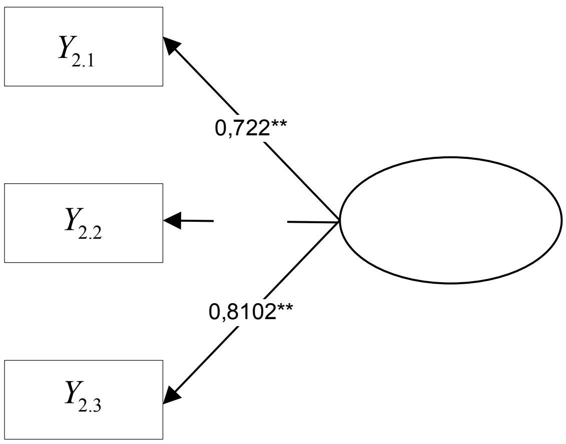
Table 4. Variable Outer Weight Results (Y1)

Variable	Indicator	Measurement Model	Outer Weight	p-value
Y_1	Y_{11}	Formative	0.801	<0.001
	Y_{12}	Formative	0,721	<0.001
	Y_{13}	Formative	0.838	<0.001

3) Measurement Model of Y2

Variable Y2 has three indicators based on Table 5 to determine the measurement model, the value of the measurement weight, and the p-value of each indicator variable Y2 which is presented in Table 5.

Based on the results of Table 5, the value of the outer loading variable Y2 can be presented in Figure 6.



Source: Researcher (2022)

Figure 6. Measurement Model of Y2

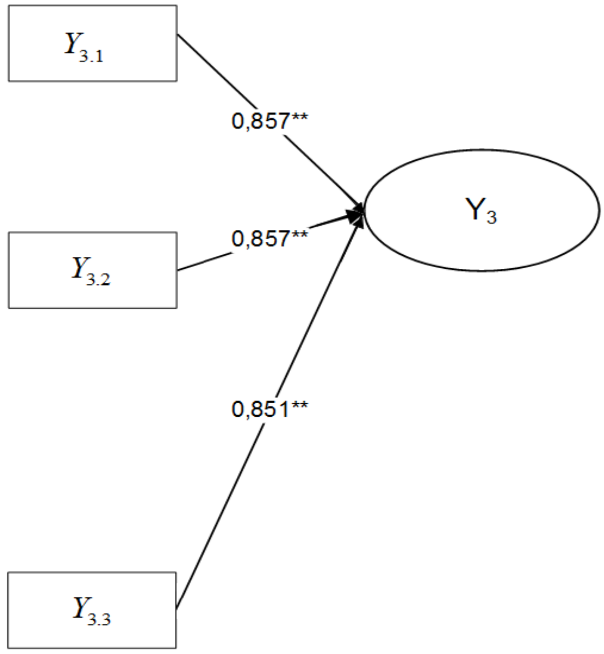
Referring to Table 6, the p value of the three indicators is < 0.001 where the three indicators are significant which can reflect the Y2 variable.

4) Measurement Model of Y3

The Y3 variable has two indicators based on Table 6 to determine the measurement model, the value of the measurement weight, and the p-value of each Y3 variable indicator which are presented in Table 6.

Based on the results of Table 6, the value of the outer weight indicator variable on Variable Y3 can be presented in Figure 7.

Based on the analysis of Figure 7, it is explained that the Y3 variable has the three strongest indicators, namely Y31 with an external weight value of 0.857. Another strongest indicator is Y32 with an outer weight value of 0.857. While the second strongest indicator is Y33 with an external weight value of 0.851. Referring to Table 7, the p-value of the three indicators is <0.001 where the three significant indicators can form the Y3 variable.



Source: Researcher (2022)

Figure 7. Measurement Model of Y3

Testing the semiparametric SEM structural model based on the shape of Figure 5 which describes the X1~Y1 relationship; X1~Y2; and X1~Y3 have a non-linear relationship. While the form of the relationship between Y1~Y3 and Y2~Y3 has a non-linear relationship. The process of analyzing semiparametric SEM data is limited to linear form (order = 1) and 1 node point, where the equation can be formed as follows:

$$\hat{f}_1 = \hat{\beta}_{01} + \hat{\beta}_{11}x_{1i} + \delta_{11}(x_{1i} - K_{11})_+$$

$$\hat{f}_2 = \hat{\beta}_{02} + \hat{\beta}_{12}x_{1i} + \delta_{12}(x_{1i} - K_{11})_+$$

$$\hat{f}_3 = \hat{\beta}_{03} + \hat{\beta}_{13}x_{1i} + \delta_{12}(x_{1i} - K_{11})_+ + \hat{\beta}_{23}y_{1i} + \hat{\beta}_{43}y_{2i}$$

Table 5. Variable Loading Factor Results (Y2)

Variable	Indicator	Measurement Model	Outer Loading	p-value
Y_2	Y_{21}	Reflective	0.722	<0.001
	Y_{22}	Reflective	0.853	<0.001
	Y_{23}	Reflective	0.8102	<0.001

Table 6. Variable Outer Weight Results Y3

Variable	Indicator	Measurement Model	Outer Weight	p-value
Y_3	Y_{31}	Formative	0.857	<0.001
	Y_{32}	Formative	0.857	<0.001
	Y_{33}	Formative	0.851	<0.001

Structural model testing is a hypothesis testing in research. Hypothesis testing is done partially by t-test (T-Statistics) on each direct influence path. Table 7 presents the results of testing the direct influence hypothesis on secondary data.

The results of Table 7 testing the structural model for the following significant effects:

1 The Direct Effect of X_1 on Y_1

Testing the direct effect of X1 on Y1 Figure 8, there are two different areas on the X1 variable. The first regime (region) is when X1 is less than 1 knot (K1) of -0.1192 (from the measurement model the range is 43.0%). In the second regime (region), when the attitude variable (X1) is more than 1 knot. The results of the analysis show that the two direct influence coefficients are negative and positive with p-value < 0.001 which means significant. This shows that there is a significant effect between X1 and Y1. When you have a low attitude (0%) to a moderate attitude (43%), it will result in a decrease in Y1. When X1 is more than 43% then Y1 has increased.

2 The Direct Effect of X_1 on Y_2

Testing the direct effect of X1 on Y2 Figure 9, there are two different areas on the X1 variable. The first regime (region) is when X1 is less than 2 knots (K2) by -0.3421 (from the measurement model the range is 37.0%). In the second regime (region), when the attitude variable (X1) is more than 2 knots. The results of the analysis show that the two direct influence coefficients are negative and positive with p-value < 0.001 which means significant. This shows that there is a significant effect between X1 and Y2. When you have a low attitude (0%) to a moderate attitude (37%), it will result in an increase in Y2. When X1 is more than 37% then Y2 has decreased.

Table 7. Structural Model

Relationship	Coefficient	Estimation	p-value	Result
X_1 to Y_1	$\beta_1 x_1$	0.357	<0.001	Significant
	$\beta_1(x_1 - k_1)$	-0.824	<0.001	
X_1 to Y_2	$\beta_1 x_1$	-0.213	<0.001	Significant
	$\beta_1(x_1 - k_2)$	0.421	<0.001	
X_1 to Y_3	$\beta_1 x_1$	0.152	<0.001	Significant
	$\beta_1(x_1 - k_3)$	-0.381	<0.001	
Y_1 ke Y_3	$\beta_1 y_1$	0.234	<0.001	Significant
Y_2 ke Y_3	$\beta_1 y_2$	-0.284	<0.001	Significant

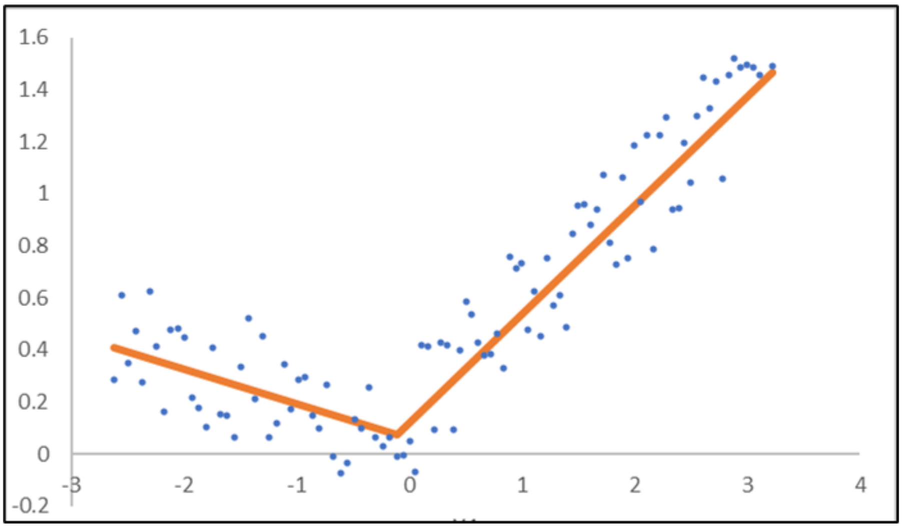


Figure 8. The Direct Effect of X_1 on Y_1

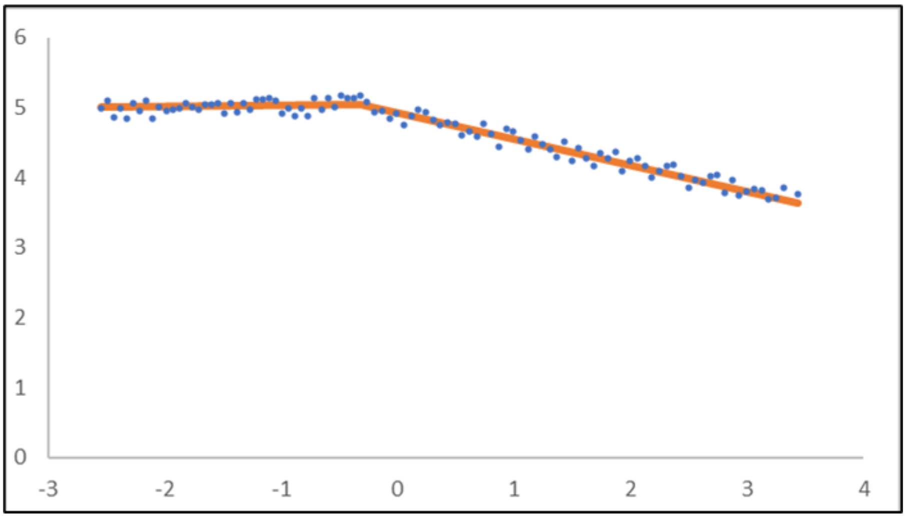


Figure 9. The Direct Effect of X_1 on Y_2

3 The Direct Effect of X_1 on Y_3

Testing the direct effect of X1 on Y3 Figure 10, there are two different areas on the X1 variable. The first regime (region) is when X1 is less than 3 knots (K3) of -0.2461 (from the measurement model the range is 35.0%). In the second regime (region), when the attitude variable (X1) is more than 3 knots. The results of the analysis show that the two direct influence coefficients are negative and positive with p-value <0.001 which means significant. This shows that there is a significant effect between X1 and Y3. When you have a low attitude (0%) to a moderate attitude (35%), it will result in an increase in Y2. When X1 is more than 35% then Y2 has decreased.

4 The Direct Effect of Y_1 on Y_3

Testing the direct effect of Y1 on Y3 Figure 11 produces a structural coefficient value of -0.265 with a p-value <0.001 which means it is significant. The structural coefficient is negative, indicating that the relationship between the two is negative. This means that the higher Y1 then Y3 has decreased.

5 The Direct Effect of Y_2 onY_3

Testing the direct effect of Y2 on Y3 Figure 12 produces a structural coefficient value of 0.255 with a p-value <0.001 which means it is significant. The structural coefficient is positive, indicating that the relationship between the two is positive. This means that the higher Y2, the higher Y3.

The results of the direct effect are in Table 7 if written using the following equation:

$$\hat{f}_1 = 7.722 + 0.357x_{1i} - 0.824(x_{1i} + 0.119)_+$$

$$\hat{f}_2 = 0.560 - 0.213x_{1i} + 0.421(x_{1i} + 0.342)_+$$

$$\hat{f}_3 = 6{,}452 + 0.152 - 0.381(x_{1i} + 0.246)_+ + 0.234y_{1i} - 0.284y_{2i}$$

The results of the measurement model get a coefficient of determination of 0.91. It can be explained that 91% of the diversity of variables Y1, Y2, and Y3 can be explained by the X1 variable while 9% is explained by other variables not used in the model.

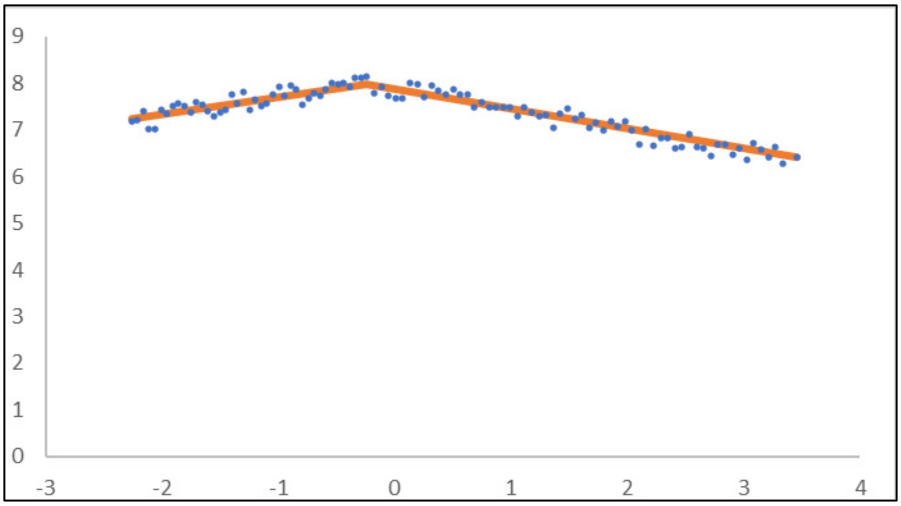


Figure 10. The Direct Effect of X_1 on Y_3

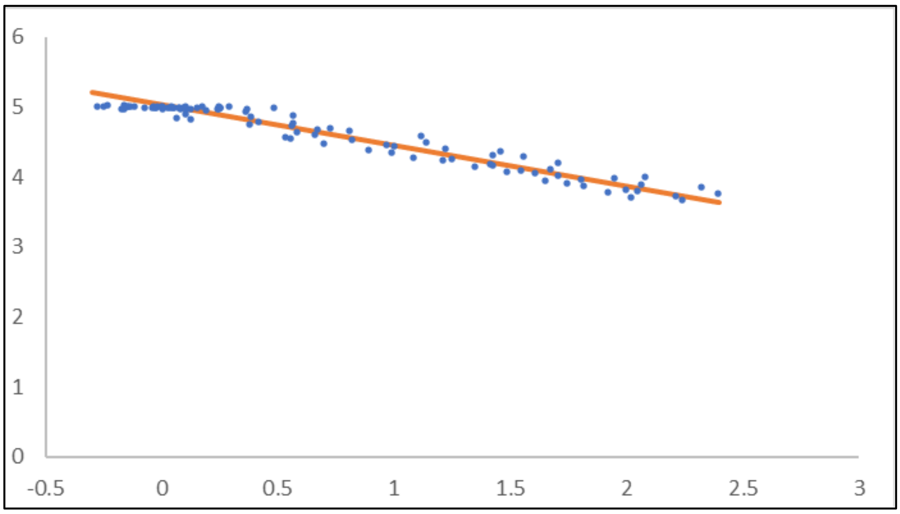


Figure 11. The Direct Effect of Y_1 on Y_3

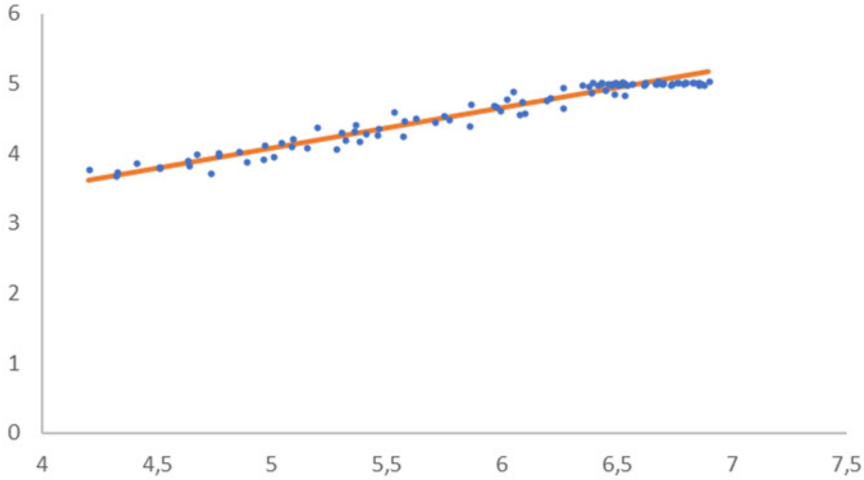


Figure 12. The Direct Effect of Y_2 on Y_3

5. Conclusion

1. The results of the direct effect of semiparametric SEM state that all relationships have a significant effect with the following function equation.

$$\hat{f}_1 = 7.722 + 0.357x_{1i} - 0.824(x_{1i} + 0.119)_+$$

$$\hat{f}_2 = 0.560 - 0.213x_{1i} + 0.421(x_{1i} + 0.342)_+$$

$$\hat{f}_3 = 6.452 + 0.152 - 0.381(x_{1i} + 0.246)_+$$

$$+0.234y_{1i} - 0.284y_{2i}$$

2. Function estimation with semiparametric Structural Equation Modeling (SEM) can be done on the simulation data, the result is a coefficient of determination of 0.91. It means that 91% of the data diversity can be captured by the function, while the remaining 9% is explained by other variables and errors.

REFERENCES

[1] Chen H., Wang Y, "A Penalized Spline Approach to Functional Mixed Effects Model Analysis," Biometrics, vol. 67, no. 1, pp. 861-870, 2011. DOI: 10.1111/j.1541-0420.2010.01524.x

[2] Efendi E.C.L., Fernandes A.A.R., Mitakda M. B. T, "Modeling of path nonparametric truncated spline linear, quadratic, and cubic in model on time paying bank credit," Mathematics and Statistics, https://doi.org/10.13189/ms.2021.090611, vol. 9, no. 6, pp. 947-957, 2021.

[3] Greene J.C, "Mixed Methods In Social Inquiry," in Mixed Methods In Social Inquiry, 1st ed, Jossey-Bass, 2007.

[4] Gujarati D. N, "Basic Econometrics," in Basic Econometrics, 4th ed, McGraw-Hill, 2003.

[5] Hidayat R., Yuliani., Sam M, "Model regresi nonparametrik dengan pendekatan spline truncated," Prosiding Seminar Nasional, vol. 3, no. 1, pp. 203–210, 2017.

[6] Hoover S.V., Perry R.F, "*Simulation; A Problem-Solving Approach*". Addison Wesley: USA, 1990.

[7] Rasyidah F.L.N., Fernandes A.A.R., Iriany A., Solimun., Wardhani N. W. S, "Development Of Path Analysis Based On Nonparametric Regression," Journal of Theoretical and Applied Information Technology, vol. 99, no. 24, pp. 5602–5612, 2021.

[8] Ruliana, "Structural Equation Modeling (Sem) Nonlinier Dengan Pendekatan Polynomial Truncated Spline," 2018.

[9] Solimun S., Fernandes A. A. R., Nurjannah, "Metode Statistika Multivariat. Pemodelan Persamaan Struktural (SEM)," UB Press, 2017.

[10] Solimun, "Analisis Multivariat Pemodelan Struktural Metode Partial Least Square- PLS," CV. Citra, 2010.

[11] Lestari B., Budiantara I. N, "Spline estimator and its asymptotic properties in multiresponse nonparametric regression model," Songklanakarin Journal of Science & Technology, vol. 42, no. 3, pp 533-548, 2020.

[12] Raharjo, K., Nurjannah, N., Solimun, S., & Fernandes, A. A. R. (2018). The influence of organizational culture and job design on job commitment and human resource performance. Journal of Organizational Change Management.

[13] Fernandes, S., & Rinaldo, A. A. R. A. A. (2018). The mediating effect of service quality and organizational commitment on the effect of management process alignment on higher education performance in Makassar, Indonesia. Journal of Organizational Change Management.

[14] Fernandes, A. A. R. (2017). Moderating effects orientation and innovation strategy on the effect of uncertainty on the performance of business environment. International Journal of Law and Management.

[15] Fernandes, A. A. R., Budiantara, I. N. I., Otok, B. W., & Suhartono. (2014). Reproducing Kernel Hilbert space for penalized regression multi-predictors: Case in longitudinal data. International Journal of Mathematical Analysis, 8(40), 1951-1961.

[16] Purbawangsa, I. B. A., Solimun, S., Fernandes, A. A. R., & Rahayu, S. M. (2019). Corporate governance, corporate profitability toward corporate social responsibility disclosure and corporate value (comparative study in Indonesia, China and India stock exchange in 2013-2016). Social Responsibility Journal, 16(7), 983-999.

[17] Fernandes, A. A. R., Hutahayan, B., Arisoesilaningsih, E., Yanti, I., Astuti, A. B., & Amaliana, L. (2019, June). Comparison of curve estimation of the smoothing spline nonparametric function path based on PLS and PWLS in various levels of heteroscedasticity. In IOP Conference Series: Materials Science and Engineering (Vol. 546, No. 5, p. 052024). IOP Publishing.

[18] Fernandes, A. A. R., & Taba, I. M. (2018). Welding technology as the moderation variable in the relationships between government policy and quality of human resources and workforce competitiveness. Journal of Science and Technology Policy Management.

[19] Fernandes, A. A. R. (2017). Investigation of instrument validity: Investigate the consistency between criterion and unidimensional in instrument validity (case study in management research). International journal of law and management.

[20] Sumardi, S., & Fernandes, A. A. R. (2020). The influence of quality management on organization performance: service quality and product characteristics as a medium. Property Management, 38(3), 383-403.

[21] Benny Hutahayan, A. A. R. F. S. N. (2020). Comparison of use of Linkage in Integrated Cluster with Discriminal Analysis Approach. International Journal of Advanced Science and Technology, 29(3), 5654 - 5668. Retrieved from http://sersc.org/journals/index.php/IJAST/article/view/6191