CAUCHY – Jurnal Matematika Murni dan Aplikasi
Volume 11 (1) (2026), Pages 91-105
p-ISSN: 2086-0382; e-ISSN: 2477-3344

Quadratic and Truncated Spline Structural Equation Modeling With Double Bootstrap in The Waste Management Economy

Anggun Fadhila Rizqia[1], Solimun[1*], Nurjannah[1], Kamelia Hidayat[1], and Fachira Haneinanda Junianto[2]

[1]*Department of Statistics, Faculty of Mathematics and Natural Science, Universitas Brawijaya, Indonesia*
[2]*Department of Mathematics, Faculty of Mathematics and Natural Science, Universitas Brawijaya, Indonesia*

Abstract

This study aims to develop and apply a semiparametric Structural Equation Modeling (SEM) approach that integrates quadratic and truncated spline estimation, enhanced with a double bootstrap resampling method. Semiparametric SEM is used to overcome the limitations of conventional SEM, particularly when data complexity and social behavior do not fully satisfy the linearity assumptions. The model was applied to analyze the public mindset and participation in waste management based on the 3R (Reduce, Reuse, Recycle) principle, focusing on the role of waste banks in optimizing the economic value of waste. The truncated spline approach enables flexible modeling of non-linear relationships among latent variables, while the quadratic term captures global curvature effects. Furthermore, the double bootstrap approach yields smaller standard errors and higher relative efficiency in parameter estimation. The simulation and empirical results demonstrate that the semiparametric SEM with double bootstrap produces higher model stability and more accurate parameter estimation compared to the single bootstrap approach. This method provides a robust analytical framework for modeling complex social phenomena, such as community-based waste management.

Keywords: Double Bootstrap; Semiparametric; Structural Equation Modeling; Waste Management Economy.

1 Introduction

Structural equation modeling (SEM) is a statistical approach that allows researchers to examine multiple relationships among variables at the same time, including the links between constructs and their indicators [1]. This method is particularly useful when dealing with latent variables—variables that cannot be observed directly and must be represented through their measured indicators [1]. SEM integrates two interconnected components: the structural model and the measurement model [2], [3]. The structural, or inner model, explains how constructs influence one another, while the measurement, or outer model, specifies how each construct is operationalized through its corresponding indicators [4].

The semiparametric approach is used when some of the relationships between variables are known, while the rest are unknown [5], [6]. One of the estimation methods in the nonparametric

*Corresponding author. E-mail: anggunfadhila24@gmail.com

Submitted: November 18, 2025 Reviewed: December 03, 2025 Accepted: January 02, 2026
DOI: https://doi.org/10.18860/cauchy.v11i1.37591

approach is the spline, specifically the truncated spline [7], [8], [9]. In the semiparametric approach, truncated splines serve as a bridge between the flexibility of the nonparametric approach and the parametric approach. Parametric SEM assumes linear relationships between variables, with an extension in the form of quadratic SEM to capture nonlinear effects (interactions and quadratic) [10]. Conversely, nonparametric SEM is not bound to a specific function form and uses truncated splines to model unknown relationships [11]. The combination of the two gives rise to semiparametric SEM, which is capable of capturing quadratic patterns and local patterns flexibly. Bootstrap resampling is a resampling technique used to estimate distributions, measure uncertainty, and construct confidence intervals. The bootstrap resampling method is used in situations where the data is not normally distributed, the sample size is small, or the model has a complex structure [12]. SEM, especially the semiparametric approach, makes parameter estimation complex.

Waste management within the framework of the circular economy reflects efforts to transform waste into a valuable economic resource. Through reduce, reuse, and recycling processes, waste can be converted into value-added products that generate new business opportunities and reduce waste management costs [13]. The comprehensive economic management of waste from upstream to downstream is aimed at providing economic benefits to the community while ensuring environmental safety [14]. This requires a paradigm shift in waste management, from focusing solely on final disposal to adopting a holistic approach that views waste as an economically valuable resource [15]. One of the most effective community-based initiatives is the waste bank system, where residents can exchange sorted waste for economic incentives. This system has been shown to improve public motivation and participation in proper waste sorting and management practices [16].

The Semiparametric SEM approach was used to analyze people's mindset toward the economic value of waste because complex social data did not fully meet the assumption of linearity. This approach is considered to be more flexible than parametric and nonparametric methods in modeling the relationship between latent variables. Meanwhile, double bootstrap has been introduced in PLS-SEM as an alternative resampling approach, and a prior studies suggest that it may enhance inferential properties compared to the single bootstrap [17]. The SEM Semiparametric approach with double bootstrap was implemented in the case of waste management in Batu City by analyzing the influence of the quality of facilities and infrastructure and the use of waste banks on understanding 3R management and waste economic management.

Nevertheless, the application of semiparametric SEM in the context of waste management and understanding its economic value is still rarely discussed in the literature, especially when the relationship between variables is nonlinear. This gap highlights the necessity for a method which is able to deal with complex relationship phenomena rather than parametric models. This research aims to construct a spline-based semiparametric SEM model with the support of a double bootstrap procedure to better determine the effects of these variables in shaping public understanding of 3R management and the economic value of waste. The key contribution of this study lies in strengthening the methodological aspects by incorporating splines and double bootstrap, which result in more stable estimates while offering a more adaptive analytical approach to understanding public behavior and perceptions within the circular economy framework.

2 Methods

2.1 Data and Research Variables

The research was conducted using primary data and simulations. The research population consisted of communities from 24 villages and sub-districts in Batu City, with respondents aged 17 years and over. The research sample consisted of residents living in The Batu District, determined using a quota sampling of 100 respondents. The research variables used were Quality of Facilities and Infrastructure (X_1), Waste Banks (X_2), 3R-Based Waste Management (Y_1), and Waste Economic Management (Y_2). All variables were tested for validity and reliability. The results showed that every questionnaire item met the criteria, indicating that all items were valid and reliable, and therefore suitable for further analysis. The variables used in this study are represented through several indicators as in the Table 2.1 below.

Table 1: Research Variables

Variable	Indicator
Quality of Facilities and Infrastructure (X_1)	Facilities and Infrastructure Maintenance ($X_{1.1}$) Quality of Facilities and Infrastructure ($X_{1.2}$) Ease of Use of Facilities and Infrastructure ($X_{1.3}$)
Waste Banks (X_2)	Waste Banks in Residents' Homes ($X_{2.1}$) Effectiveness of Waste Banks ($X_{2.2}$) Operational Efficiency of Waste Banks ($X_{2.3}$)
3R-Based Waste Management (Y_1)	Effectiveness of Reduce ($Y_{1.1}$) Effectiveness of Reuse ($Y_{1.2}$) Effectiveness of Recycle ($Y_{1.3}$)
Waste Management Economics (Y_2)	Efficiency of Waste Management Economics ($Y_{2.1}$) Waste as an Economic Resource ($Y_{2.2}$) Sustainability of Waste Management Economics ($Y_{2.3}$)

2.2 Research Model and Research Steps

The research model used can be seen in Fig. 1.

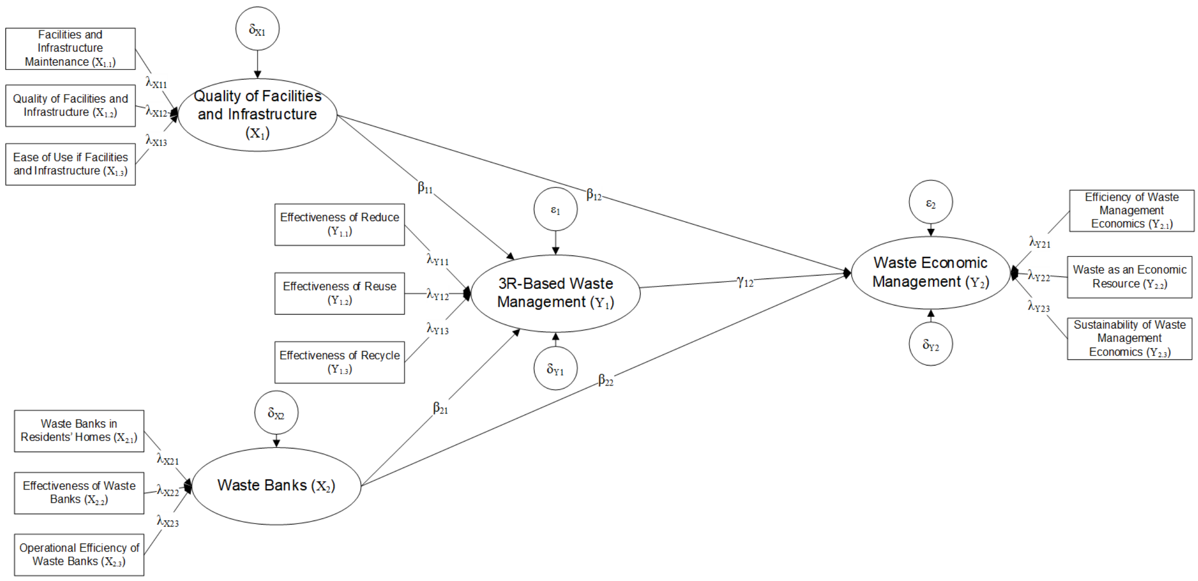

Figure 1: SEM Path Diagram

The model used in Fig. 1 employs semiparametric Structural Equation Modeling (SEM). The analysis was conducted using RStudio with the following steps:

1. Inputting primary data.
2. Defining variables and indicators for the exogenous, mediating endogenous, and pure endogenous constructs.
3. Specifying the structural model (inner model) and the measurement model (outer model), with all latent variable indicators treated as formative.
4. Creating a path diagram and converting it into a system of equations.

5. Assessing linearity using the Ramsey RESET test and the modified quadratic Ramsey RESET test.
6. Estimating the parameters of both the measurement model and the structural functions.
7. Selecting the optimal order and knot for the nonparametric component based on the smallest Generalized Cross-Validation (GCV) value.
8. Performing semiparametric SEM for linear and quadratic specifications, followed by double bootstrap testing for direct and indirect effects.
9. Evaluating the goodness of fit using R^2_{adj}.
10. Conducting simulation studies with repeated replications to assess model stability and performance under different relationship patterns.

2.3 Linearity Assumption

The SEM framework assumes that both the links between latent variables and their indicators, as well as the relationships among latent variables in the inner model, follow a linear, quadratic, or cubic form. To assess this assumption, the Ramsey RESET procedure is applied, and the main steps of the test are outlined in [18].

1. Determining the first regression equation (linear regression) as in the following equation.

$$Y_i = \beta_0 + \beta_1 X_{1i} + \varepsilon_i \tag{1}$$

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_{1i} \tag{2}$$

Next, calculate the residual standard error R_1^2 denotes the coefficient of determination of the restricted model (linear regression) using Eq. 3 below.

$$R_1^2 = 1 - \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} \tag{3}$$

2. Determine the second regression equation as a comparison model as shown in Eq. 4 and Eq. 5.

$$Y_i = \beta_0^* + \beta_1^* X_{1i} + \beta_2 \hat{Y}_i^{\,2} + \beta_3 \hat{Y}_i^{\,3} + \varepsilon_i \tag{4}$$

$$\hat{Y}_i^* = \hat{\beta}_0^* + \hat{\beta}_1^* X_{1i} + \hat{\beta}_2 \hat{Y}_i^{\,2} + \hat{\beta}_3 \hat{Y}_i^{\,3} \tag{5}$$

Then, the calculation of R_2^2 (denotes the coefficient of determination of the unrestricted model) is performed using the following Eq. 6:

$$R_2^2 = 1 - \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i^*)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} \tag{6}$$

3. Testing whether the relationship is linear or not.
Hypothesis for Ramsey's RESET:

$$H_0 : \beta_2 = \beta_3 = 0 \quad \text{(Linear relationship between variables)},$$
$$H_1 : \text{There is at least one } \beta_j \neq 0,\ j = 2, 3.$$

The test statistic follows an F distribution as in Eq. 7:

$$F = \frac{(R_2^2 - R_1^2)/q}{(1 - R_2^2)/(n - k_{ur})} \sim F_{(q, n-k_{ur})} \tag{7}$$

where:
$q = 2$
$k_{ur} = 4$

k_{ur} denotes the number of parameters in the unrestricted model (including the intercept) The decision is to reject H_0 if the test statistic

$$F > F_\alpha(q, n - k_{ur})$$

or when the p-value < 0.05, which means the relationship between variables is nonlinear.

The subsequent procedure is the modified Ramsey RESET test, which is used when the association between variables is suspected to be nonlinear. This test aims to evaluate whether the relationship follows a quadratic form. The steps for conducting the modified Ramsey RESET test are as follows [18].

1. Determine the first regression equation (quadratic regression) as in the following equation.

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{1i}^2 + \varepsilon_i \tag{8}$$

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{1i}^2 \tag{9}$$

Next, calculate the residuals (R_1^2) denotes the coefficient of determination of the restricted model (quadratic regression) using Eq. 10:

$$R_1^2 = 1 - \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} \tag{10}$$

2. Determine the second regression equation as a comparison model as in Eq. 11 and Eq. 12.

$$Y_i = \beta_0^* + \beta_1^* X_{1i} + \beta_2^* X_{1i}^2 + \beta_3 \hat{Y}_i^2 + \beta_4 \hat{Y}_i^3 + \varepsilon_i \tag{11}$$

$$\hat{Y}_i^* = \hat{\beta}_0^* + \hat{\beta}_1^* X_{1i} + \hat{\beta}_2^* X_{1i}^2 + \hat{\beta}_3 \hat{Y}_i^2 + \hat{\beta}_4 \hat{Y}_i^3 \tag{12}$$

Then, the calculation of R_2^2 (denotes the coefficient of determination of the unrestricted model) is performed using Eq. 13:

$$R_2^2 = 1 - \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i^*)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} \tag{13}$$

3. Testing whether the relationship is linear or not.
Hypothesis for Ramsey's RESET:

$$H_0 : \beta_3 = \beta_4 = 0 \quad \text{(Quadratic relationship between variables)},$$
$$H_1 : \text{There is at least one } \beta_j \neq 0,\ j = 3, 4.$$

The test statistic follows an F distribution as in Eq. 14:

$$F = \frac{(R_2^2 - R_1^2)/q}{(1 - R_2^2)/(n - k_{ur})} \sim F_{(q,\, n-k_{ur})} \tag{14}$$

where:
$q = 2$
$k_{ur} = 5$
k_{ur} denotes the number of parameters in the unrestricted quadratic model (including the intercept)
If the test statistic $F < F_\alpha(q, n - k_{ur})$ or when the p-value > 0.05, then accept H_0, meaning the relationship between variables is quadratic. Conversely, if p-value < 0.05, then reject H_0, meaning the relationship is not quadratic and should be approached using a nonparametric method.

2.4 Structural Equation Modeling

SEM integrates the structural model and the measurement model within a single analytical framework. The structural or inner model represents the connections among latent variables, while the measurement or outer model specifies how each latent variable is linked to its indicators. These two components are more easily interpreted when illustrated through a path diagram. The path diagram derived from the development of both the inner and outer models is presented in Fig. 2.

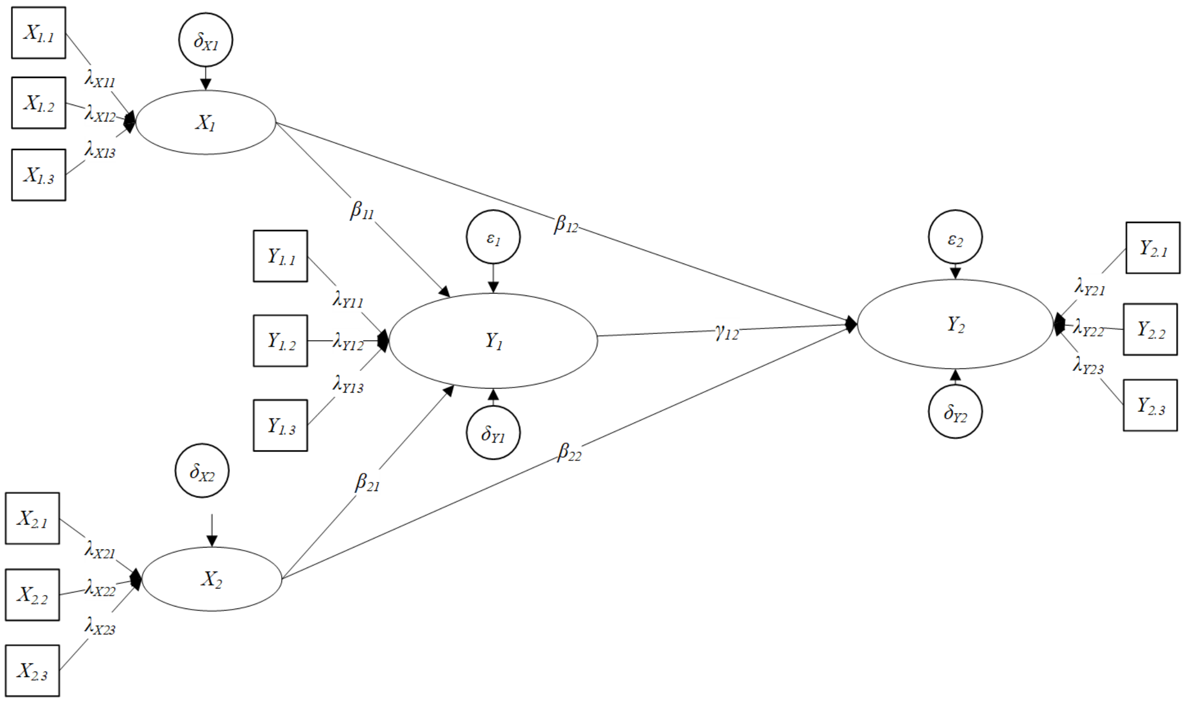

Figure 2: Path Diagram

The measurement model, or outer model, explains how latent variables are linked to their respective indicators. In constructing an outer model, researchers must decide whether each latent variable is measured reflectively or formatively [1]. This determination can be guided by theoretical foundations, findings from prior studies, empirical considerations, as well as the researcher's intuition and logical reasoning [1]. In a formative indicator model, the latent variable is illustrated as an ellipse with arrows pointing toward it from its indicators. These arrows signify that the indicators collectively construct the latent variable. The corresponding path analysis in the outer model can be expressed through the following equation.

Exogenous latent variables with formative properties

$$x_{1i} = \lambda_{x11}X_{11i} + \lambda_{x12}X_{12i} + \lambda_{x13}X_{13i} + \delta_{x1i} \tag{15}$$

$$x_{2i} = \lambda_{x21}X_{21i} + \lambda_{x22}X_{22i} + \lambda_{x23}X_{23i} + \delta_{x2i} \tag{16}$$

Endogenous latent variables with formative properties

$$y_{1i} = \lambda_{y11}Y_{11i} + \lambda_{y12}Y_{12i} + \lambda_{y13}Y_{13i} + \delta_{y1i} \tag{17}$$

$$y_{2i} = \lambda_{y21}Y_{21i} + \lambda_{y22}Y_{22i} + \lambda_{y23}Y_{23i} + \delta_{y2i} \tag{18}$$

In SEM analysis, it is possible to have relationships between latent variables in the form of propositions, so this analysis can be used in exploratory relationship analysis. The inner model linear equation can be written in the following equation.

$$Y_{1i} = \beta_{11}X_{1i} + \beta_{21}X_{2i} + \varepsilon_{1i} \tag{19}$$

$$Y_{2i} = \beta_{12}X_{1i} + \beta_{22}X_{2i} + \gamma_{12}Y_{1i} + \varepsilon_{2i} \tag{20}$$

2.5 Semiparametric Structural Equation Modeling (SEM)

If it is assumed that the entire inner model is nonparametric, then a nonparametric structural model equation can be formed as in the following equation.

$$Y_{1i} = f_1(X_{1i}) + \varepsilon_{1i} \tag{21}$$

$$Y_{2i} = f_2(X_{1i}, Y_{1i}) + \varepsilon_{2i} \tag{22}$$

The nonparametric truncated linear spline (order $m = 1$) SEM equation model using one knot can be written in the following equation:

$$\hat{f}(X_{1i}) = \hat{\beta}_0 + \sum_{j=1}^{p} \hat{\beta}_j X_{ji} + \sum_{k=1}^{K} \hat{\delta}_k (X_{1i} - K_k)_+^p \tag{23}$$

With the truncated function:

$$(X_i - K_k)_+^p = \begin{cases} (X_i - K_k)^p, & X_i \geq K_k, \\ 0, & X_i < K_k \end{cases} \tag{24}$$

Semiparametric truncated spline Structural Equation Modeling extends the principles of semiparametric regression and semiparametric path analysis. This approach is particularly useful when certain relationship structures in the model are known, while others remain unspecified. Despite its potential, the application of semiparametric SEM using truncated spline techniques has been relatively limited in existing research. In semiparametric SEM, the optimal knot configuration is determined by identifying the model with the lowest Generalized Cross Validation (GCV) value. GCV serves as an internal evaluation method used to choose both the number and position of knots in each model. The most appropriate spline estimator is the one associated with this optimal knot selection. The GCV expression is presented below.

$$GCV(K) = \frac{MSE(K)}{\left[n^{-1}\,\text{trace}(I - A(K))\right]^2} \tag{25}$$

where

$$MSE(K) = n^{-1} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2,$$

K is the knot, and the matrix $A(K)$ is obtained from:

$$A[K] = X[K] \left(X[K]^T \hat{\Sigma}^{-1} X[K] \right)^{-1} X[K]^T \hat{\Sigma}^{-1} \tag{26}$$

Eq. 26 is a function of the knot, and $K = (K_1, K_2, \ldots, K_k)^T$ is the knot value.

2.6 Resampling Double Bootstrap

Bootstrap resampling is a technique that generates multiple random samples using a "resampling with replacement" procedure, meaning each selected observation is returned to the dataset before the next draw. In this method, the number of resampled observations does not exceed the size of the original dataset. The steps for estimating standard errors through the bootstrap approach are outlined as follows.

1. Determine the number of B iterations in the bootstrap sample $(xy_1^*, xy_2^*, \ldots, xy_B^*)$ obtained by randomly sampling with replacement n elements from the original sample.
2. Calculate the bootstrap replication as in the following equation.

$$\hat{\theta}_{(b)}^* = s(xy_{(b)}^*), \qquad b = 1, 2, \ldots, B \tag{27}$$

3. Estimate the standard error as in the following equation.

$$SE_{\hat{\theta}^*} = \sqrt{\frac{\sum_{b=1}^{B}\left(\hat{\theta}^*_{(b)} - \bar{\hat{\theta}}^*_{(.)}\right)^2}{B}} \tag{28}$$

where

$$\bar{\hat{\theta}}^*_{(.)} = \frac{1}{B}\sum_{b=1}^{B}\hat{\theta}^*_{(b)}$$

The double bootstrap procedure introduced in 1988, which is a bootstrap technique that uses two levels of resampling [19]. The procedure begins by performing bootstrap B_1 replications. The procedure then continues by performing bootstrap B_2 replications. Therefore, the total number of test statistics calculations that must be performed is $B_1 + B_1B_2$. After obtaining the first bootstrap data set $xy^*_{(b)}$ and the test statistic value $\hat{\tau}^*_b$, with $b = 1, 2, \dots, B_1$, for each first bootstrap sample, a second bootstrap data set is generated, denoted by $xy^*_{(bj)}$ and the test statistic value $\hat{\tau}^*_{bj}$, with $b = 1, 2, \dots, B_1$ and $j = 1, 2, \dots, B_2$.

The evaluation of resampling methods can be conducted by comparing their relative efficiency values. Relative efficiency expresses the ratio of variance between two estimators, indicating which method provides more stable estimates [19], [20]. It can be formulated as Eq. 29.

$$Eff(\hat{\beta}_{DB}, \hat{\beta}_{SB}) = \frac{V(\hat{\beta}_{SB})}{V(\hat{\beta}_{DB})} \tag{29}$$

Where:

$V(\hat{\beta}_{SB})$: Variance of the parameter estimator obtained using the Single Bootstrap technique

$V(\hat{\beta}_{DB})$: Variance of the parameter estimator obtained using the Double Bootstrap technique

2.7 Simulation Study Design

A simulation study was conducted to evaluate the robustness and flexibility of the proposed semiparametric SEM framework under different functional relationship patterns. The simulation employed a sample size of $n = 100$, consistent with the empirical application, to maintain comparability between simulated and observed data.

The data generating process (DGP) was specified to reflect the structural configuration of the proposed SEM, consisting of two exogenous variables (X_1, X_2), one moderating variable (Y_1), and one endogenous variable (Y_2). The structural equations were generated as Eq. 19 and Eq. 20, with functional forms specified as linear, quadratic, or nonparametric depending on the simulation scenario. Linear scenarios used the original forms of Eq. 19 and Eq. 20, quadratic scenarios incorporated squared terms in the relevant paths, and nonparametric scenarios modeled the nonlinear components using truncated spline functions with a predifined knot. The error terms were independently generated from a normal distribution, $\varepsilon \sim \mathcal{N}(0, \sigma^2)$, with constat variance across replications. All true parameter values were fixed across replications to ensure comparability of model performance across scenarios.

Six simulation scenarios were constructed by varying the functional form (linear, quadratic, truncated spline) and their positions in the structural model. Each scenario was replicated 100 times. Model performance was evaluated using the coefficient of determination R^2, wich was computed for each replication and subsequently averaged across replications to obtain the reported values.

3 Results and Discussion

3.1 Linearity Test

The results of the test using Ramsey's RESET test can be seen in the following table.

Table 2: Linearity Test Results

Relationship	Linear		Quadratic		Cubic	
	p-value	Result	p-value	Result	p-value	Result
$X_1 \to Y_1$	0.423	Linear	0.017	Not Quadratic	0.017	Not Cubic
$X_1 \to Y_2$	0.387	Linear	0.038	Not Quadratic	0.021	Not Cubic
$X_2 \to Y_1$	0.041	Not Linear	0.013	Not Quadratic	0.008	Not Cubic
$X_2 \to Y_2$	0.046	Not Linear	0.212	Quadratic	0.040	Not Cubic
$Y_1 \to Y_2$	0.451	Linear	0.018	Not Quadratic	0.015	Not Cubic

Based on Table 2, it can be seen that the results of linearity testing between each variable as follows:

1. The p-value for the relationship between Quality of Facilities and Infrastructure (X_1) and 3R-Based Waste Management (Y_1) exceeds $\alpha = 0.05$, indicating that this relationship is linear.
2. The relationship between Quality of Facilities and Infrastructure (X_1) and Waste Management Economics (Y_2) also shows a p-value greater than α, confirming a linear form.
3. For the relationship between Waste Banks (X_2) and 3R-Based Waste Management (Y_1), the p-value is below α, suggesting nonlinearity. A modified Ramsey RESET test was therefore conducted, but the form of the relationship could not be determined, leading to the use of a nonparametric approach.
4. The relationship between Waste Banks (X_2) and Waste Management Economics (Y_2) likewise yields a p-value below α, indicating nonlinearity. The modified Ramsey RESET test identified this relationship as quadratic.
5. Meanwhile, the p-value for the relationship between 3R-Based Waste Management (Y_1) and Waste Management Economics (Y_2) is above α, allowing the conclusion that the relationship is linear.

3.2 Optimal Knot

Generalized Cross Validation (GCV) is used to determine the location of the optimal knot point. The optimal knot point is determined based on the smallest GCV value. The knot and GCV values are shown in Table 3.

Table 3: Optimal Knot Point

No	Knot	GCV	$\mathbf{R^2_{adj}}$
1	0.420	0.625	0.812
2	-1.310	0.642	0.798

Based on the GCV calculation results in Table 3, it can be seen that the smallest GCV value is obtained at 1 knot point, $K = 0.420$, with GCV $= 0.625$ and $R^2_{adj} = 0.812$. Therefore, the knot point at $X_2 = 0.420$ was selected as the optimal knot for the nonparametric function $f_1(X_2)$. The spline specification consequently distinguishes two regimes in the predictor domain, namely for $X_2 < 0.420$ and $X_2 > 0.420$, each indicating different functional patterns in the relationship between X_2 and Y_1.

3.3 Measurement Model

The results of the measurement model for each variable indicator can be seen based on the outer weight. The outer weight is presented in Table 4.

Table 4: Outer Weight

Variable	Indicator	Coefficients	Standard Error	t-test Statistics	p-value
X_1	$X_{1.1}$	0.812	0.068	11.941	0.002
	$X_{1.2}$	0.743	0.074	10.040	0.003
	$X_{1.3}$	0.879	0.063	13.952	0.001
X_2	$X_{2.1}$	0.691	0.077	8.974	0.004
	$X_{2.2}$	0.755	0.069	10.942	0.003
	$X_{2.3}$	0.720	0.090	8.000	0.003
Y_1	$Y_{1.1}$	0.834	0.064	13.031	0.001
	$Y_{1.2}$	0.788	0.066	11.939	0.002
	$Y_{1.3}$	0.665	0.095	7.000	0.004
Y_2	$Y_{2.1}$	0.702	0.070	10.029	0.003
	$Y_{2.2}$	0.817	0.054	15.130	0.001
	$Y_{2.3}$	0.760	0.069	11.014	0.002

As the formative measurement results indicate in Table 4, all indicators are significantly related to their constructs at p-value < 0.01 in all domains. Regarding the Quality of Facilities and Infrastructure construct (X_1), each of maintenance, facility quality, and ease of use are significant (p-value < 0.001) further suggesting to us that each factor has a relevant influence on X_1. Ease of use offers the most significant incremental contribution of these aspects, implying that user-friendly facilities matter the most in shaping the construct and maintenance and quality are important auxiliary components.

In the case of X_2 for the Waste Banks construct, all indicators of the existence of the waste banks in households, perceived effectiveness, and operational efficiency are significant (p-value $<$ 0.01). The effectiveness indicator has the strongest contribution to X_2, reflecting that community perceptions of effective waste banks become the prime instrument to inform this construct. The relevance of the remaining indicators, for now, also indicates that household participation and operational performance contribute to the construct's conceptualization, respectively.

All the indicators for 3R-Based Waste Management (Y_1) and Waste Management Economics (Y_2) have significant contributions (p-value < 0.01). The Reduce dimension in Y_1 is the strongest contributor, illustrating its importance as part of the 3R framework, whereas Reuse and Recycle are complementary and important elements. The perception of waste as an economic resource dominates the construct for Y_2, indicating that economic valuation drives the construct, whereas efficiency and sustainability also contribute substantially, having statistically reliable results.

In addition to evaluating the significance of the outer weights, collinearity among formative indicators was assessed to ensure the adequacy of the formative measurement model. The outer weight is presented in table below.

Table 5: Collinearity Diagnostics of Formative Indicators (VIF)

Construct	Indicator	VIF
X_1	$X_{1.1}$	1.82
	$X_{1.2}$	2.10
	$X_{1.3}$	2.35
X_2	$X_{2.1}$	1.74
	$X_{2.2}$	2.05
	$X_{2.3}$	2.48
Y_1	$Y_{1.1}$	1.69
	$Y_{1.2}$	2.12
	$Y_{1.3}$	2.90
Y_2	$Y_{2.1}$	1.88
	$Y_{2.2}$	2.20
	$Y_{2.3}$	2.41

Based on the table, all VIF are below 10, which is commonly considered acceptable in formative measurement models.

3.4 Structural Model using Double Bootstrap

From the semiparametric SEM structural model, after verifying the linearity assumption, three relationships are identified as linear, one as quadratic, and one as nonparametric. Using this established model, the next step involves examining the direct effects among the variables. To ensure that the parameter estimates are more reliable and stable, the analysis is extended by applying both single and double bootstrap techniques. The results of the direct effect testing obtained through these bootstrap procedures are presented in Table 6.

Table 6: Direct Effect using Single and Double Bootstrap

Relationship	$\hat{\beta}_i$	Coef.	Single Bootstrap		Double Bootstrap		Relative Efficiency
			SE	p-value	SE	p-value	
$X_1 \to Y_1$	$\hat{\beta}_1 X_1$	0.372	0.096	<0.001	0.063	<0.001	2.324
$X_1 \to Y_2$	$\hat{\beta}_2 X_1$	0.286	0.077	<0.001	0.050	<0.001	2.370
$X_2 \to Y_1$	$\hat{\beta}_3 X_2$	0.198	0.080	0.013	0.052	<0.001	2.384
	$\hat{\beta}_4 (X_2 - K_{21})_+$	-1.153	0.104	<0.001	0.077	<0.001	1.806
$X_2 \to Y_2$	$\hat{\beta}_5 X_2$	0.415	0.137	0.002	0.107	<0.001	1.629
	$\hat{\beta}_6 X_2^2$	0.094	0.057	0.099	0.042	0.024	1.882
$Y_1 \to Y_2$	$\hat{\beta}_7 Y_1$	0.527	0.056	<0.001	0.040	<0.001	1.968

Table 6 presents the results of the direct effect testing using both Single and Double Bootstrap methods. All relationships show significant p-values < 0.05, indicating that the path coefficients are significant. The Standard Errors (SE) of the Double Bootstrap method are constantly smaller than those of the Single Bootstrap, implying that the Double Bootstrap provides more stable and precise parameter estimates. This pattern is further confirmed by the relative efficiency values, all of which exceed one. The results of Double Bootstrap making it a more reliable resampling approach for estimating the parameters in the semiparametric SEM model. The result of the structural model estimation as follows.

$$\hat{f}_{1i} = 0.428 + 0.372\, X_1 + 0.198\, X_2 - 1.153\,(X_2 - K_{21})_+ \tag{30}$$

$$\hat{f}_{2i} = -0.317 + 0.286\, X_1 + 0.415\, X_2 + 0.094\, X_2^2 + 0.527\, Y_1 \tag{31}$$

3.5 Simulation Study Results

The simulation study was carried out to observe how the proposed semiparametric SEM framework behaves under different data conditions. The simulation varies both the form of the relationship and their positions within the model. Specifically, the structural paths $X \to Y_1$, $X \to Y_2$, and $Y_1 \to Y_2$ are specified as linear, quadratic, or nonparametric using truncated spline, and are evaluated at the exogenous-mediator, exogenous-endogenous, and mediator-endogenous positions.

Six simulation scenarios are formed by combining these relationship types across the three structural paths. For each scenario, data are generated with additive error terms having zero mean and constant variance, and the simulation is repeated 100 times. Model performance is assessed using the coefficient of determination (R^2), which is calculated for each replication ant the averaged across replications. The simulation results are summarized in Table 7.

Table 7: Simulation Performance Based on R^2

Relationship	Position	$\mathbf{R^2}$
Linier	Exogenous–Mediator	0.950
Linier	Exogenous–Endogenous	0.792
Linier	Mediator–Endogenous	0.781
Quadratic	Exogenous–Mediator	0.752
Quadratic	Exogenous–Endogenous	0.882
Quadratic	Mediator–Endogenous	0.917
Truncated Spline	Exogenous–Mediator	0.849
Truncated Spline	Exogenous–Endogenous	0.881
Truncated Spline	Mediator–Endogenous	0.925

Based on Table 7, the highest $R^2 = 0.950$ is obtained in the linear relationship at the exogenous-mediator position, indicating strong model performance under this configuration. Among nonlinear patterns, the quadratic and nonparametric relationships at the mediator-endogenous position also show high explanatory power (0.917 and 0.925, respectively). These results confirm that the semiparametric SEM model can flexibly capture both linear and nonlinear relationships across different structural positions.

3.6 Discussion

The results suggest that the better facilities and infrastructure consistently encourage 3R-based waste management practices. When waste facilities are easier to access, well maintained, and user-friendly, people are more likely to separate waste and participate in reduce, reuse, and recycle activities. The linear nature of this relationship indicates that gradual improvements in infrastructure are directly followed by improvements in behavior, without the need for a specific threshold to be reached. This highlights the importance of continuous and sustained investment in waste management infrastructure.

In contrast, the role of waste banks appears to be more complex. Their performance shows nonlinear relationships with both 3R-based waste management and waste management economics, suggesting that waste banks only become truly effective when operate at a sufficient level of functionality. At low performance levels, improvements may have limited impact, but once operational effectiveness and public trust are established, waste banks can significantly influence recycling behavior and economic outcomes. From a policy perspective, this implies that simply establishing waste banks is not enough; strengthening their management, efficiency, and credibility is crucial to ensure that 3R- practices lead to tangible economic benefits.

4 Conclusion

This study results that the semiparametric Structural Equation Modeling (SEM) approach effectively captures both linear and nonlinear relationships between variables in the waste management economy framework. The linearity test results indicate three linear relationships, one quadratic relationship, and one nonparametric relationship, confirming the presence of diverse functional patterns in the model. The application of single and double bootstrap methods shows that all path coefficients are significant, with the double bootstrap producing smaller standard errors and higher relative efficiency values, indicating more stable and accurate parameter estimation. Simulation results further validation the robustness of the model, where the highest explanatory power ($R^2 = 0.950$) is observed in the linear exogenous-mediator relationship. Then, the quadratic and truncated spline forms at the mediator-endogenous level also yield high R^2 values (0.917 and 0.925). Overall, these findings confirm that the semiparametric SEM with double bootstrap approach provides a flexible and reliable framework for analyzing complex structural relationships, particularly when linear and nonlinear patterns coexist within the same system.

CRediT Authorship Contribution Statement

Anggun Fadhila Rizqia: Conceptualization, Methodology, Analysis, Writing–Original Draft. **Solimun:** Supervisor, Methodology, Review Draft. **Nurjannah:** Co-Supervisor, Analysis, Review Draft. **Kamelia Hidayat:** Literature Review, Data Preparation, Visualization. **Fachira Haneinanda Junianto:** Data Preparation, Analysis, Formatting.

Declaration of Generative AI and AI-assisted technologies

Generative AI tools (ChatGPT, OpenAI GPT-5.1) were used to assist in formatting LaTeX code, improving grammar, and clarifying explanations. All analytical decisions, interpretations, and conclusions were made by the authors.

Declaration of Competing Interest

The authors declare no competing interests.

Funding and Acknowledgments

This research received no external funding.

Data and Code Availability

The data used in this study are primary data collected for a continuing research project and cannot be made publicly available at this stage. The dataset can be obtained from the corresponding author upon reasonable request.

References

[1] Solimun, A. A. R. Fernandes, and Nurjannah, *Metode Statistika Multivariat: Pemodelan Persamaan Struktural (SEM) Pendekatan WarpPLS*. UB Press, 2017.

[2] T. R. Syafriana, Solimun, N. W. S. Wardhani, A. Iriany, and A. A. R. 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.

[3] 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*. Springer, 2021.

[4] F. Magno, F. Cassia, and C. M. Ringle, "A brief review of partial least squares structural equation modeling (pls-sem) use in quality management studies," *The TQM Journal*, vol. 36, no. 5, pp. 1242–1251, 2024. DOI: 10.1108/TQM-06-2022-0197.

[5] A. A. R. Fernandes and Solimun, *Analisis Regresi dalam Pendekatan Fleksibel*. UB Press, 2021.

[6] S. Xavier and T. L. Alexander, "Interaction and quadratic effect on cardiovascular disease: A nonlinear structural equation modeling," *International Journal of Advanced Research*, vol. 3, no. 10, pp. 832–841, 2015. Available online.

[7] A. S. Suriaslan, I. N. Budiantara, and V. Ratnasari, "Truncated spline regression for binary response: A comparative study of nonparametric and semiparametric approaches," *Communications in Mathematical Biology and Neuroscience*, pp. 1–22, 2025. DOI: 10.28919/cmbn/9209.

[8] D. V. Amanda, A. Iriany, A. A. R. Fernandes, and Solimun, "Analysis of path nonparametric truncated spline maximum cubic order in banking credit of risk behavior model," *BAREKENG: Journal of Mathematics and Its Applications*, vol. 18, no. 4, pp. 2639–2652, 2024. DOI: 10.30598/barekengvol18iss4pp2639-2652.

[9] F. H. Junianto, A. A. R. Fernandes, Solimun, and R. B. Hamdan, "Development of semiparametric path analysis modeling truncated spline: Determinants of increased regional economic growth," *BAREKENG: Journal of Mathematics and Its Applications*, vol. 18, no. 4, pp. 949–960, 2024. DOI: 10.30598/barekengvol18iss2pp0949-0960.

[10] F. H. Junianto, A. A. R. Fernandes, Solimun, and R. B. Hamdan, "Structural equation modeling semiparametric in modeling the accuracy of payment time for customers of credit bank in indonesia," *JTAM (Jurnal Teori dan Aplikasi Matematika)*, vol. 8, no. 4, pp. 1082–1095, 2024. DOI: 10.31764/jtam.v8i4.23668.

[11] J. J. Faraway, *Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models*, 2nd ed. CRC Press, 2016.

[12] H. L. Shang, "Double bootstrapping for visualising the distribution of descriptive statistics of functional data," *Journal of Statistical Computation and Simulation*, vol. 91, no. 10, pp. 1–22, 2017. DOI: 10.48550/arXiv.2102.01144.

[13] D. Mandasari, B. Wirjodirdjo, and M. Anityasari, "Peningkatan fasilitas bank sampah sebagai upaya pengurangan timbunan sampah perkotaan di tps surabaya," *Jurnal Teknik ITS*, vol. 9, no. 2, pp. 322–327, 2020. DOI: 10.12962/j23373539.v9i2.56123.

[14] N. A. Jauhariyah, N. I. Susanti, Mahmudah, F. N. S. Indana, and M. K. Qohar, "Pengembangan pemberdayaan ekonomi pesantren melalui pengelolaan sampah secara berkelanjutan," *Loyalitas: Jurnal Pengabdian Kepada Masyarakat*, vol. 6, no. 1, pp. 116–127, 2023. Available online.

[15] E. K. Purwendah, Rusito, and A. Periani, "Kewajiban masyarakat dalam pemeliharaan kelestarian lingkungan hidup melalui pengelolaan sampah berbasis masyarakat," *Jurnal Locus Delicti*, vol. 3, no. 2, pp. 121–134, 2022. DOI: 10.23887/jld.v3i2.1609.

[16] D. Ariefahnoor, N. Hasanah, and A. Surya, "Pengelolaan sampah desa gudang tengah melalui manajemen bank sampah," *Jurnal Kacapuri: Jurnal Keilmuan Teknik Sipil*, vol. 3, no. 1, pp. 14–30, 2020. DOI: 10.31602/jk.v3i1.3594.

[17] N. F. Richter, S. Schubring, S. Hauff, C. M. Ringle, and M. Sarstedt, "When predictors of outcomes are necessary: Guidelines for the combined use of pls-sem and nca," *Industrial Management & Data Systems*, vol. 120, no. 12, pp. 2243–2267, 2020. DOI: 10.1108/IMDS-11-2019-0638.

[18] Solimun, A. A. R. Fernandes, I. Rahmawati, L. Muflikhah, F. Ubaidillah, and A. A. Rohimi, "Research on structural flexibility and acceptance model (sfam) reconstruction based on disruption innovation in the social humanities and education sector," *WSEAS Transactions on Mathematics*, vol. 20, pp. 657–675, 2021. DOI: 10.37394/23206.2021.20.70.

[19] R. Beran, "Prepivoting: A bootstrap view of test statistics–asymptotic refinements," *Journal of the American Statistical Association*, vol. 82, no. 403, pp. 687–697, 1988. Available online.

[20] M. F. Papalia, Solimun, and Nurjannah, "Comparison of resampling efficiency levels of jackknife and double jackknife in path analysis," *BAREKENG: Journal of Mathematics and Its Applications*, vol. 17, no. 2, pp. 807–818, 2023. DOI: 10.30598/barekengvol17iss2pp0807-0818.