Received: November 26, 2020. Revised: May 20, 2021. Accepted: June 9, 2021. Published: June 20, 2021.

Modeling of Path Nonparametric Truncated Spline Linear, Quadratic, and Cubic in Model on Time Paying Bank Credit

Erlinda Citra Lucki Efendi[1*], Adji Achmad Rinaldo Fernandes[2], Maria Bernadetha Theresia Mitakda[3]
Department of Statistics, Faculty of Mathematics and Science
Brawijaya University
St. Veteran, Malang 65145 East Java
Indonesia

Abstract: - This study aims to estimate the nonparametric truncated spline path functions of linear, quadratic, and cubic orders at one and two knot points and determine the best model on the variables that affect the timely payment of House Ownership Credit (HOC). In addition, this study aims to test the hypothesis to determine the variables that have a significant effect on punctuality in paying House Ownership Credit (HOC). The data used in this study are primary data. The variables used are service quality and lifestyle as exogenous variables, willingness to pay as mediating variables and on time to pay as endogenous variables. Analysis of the data used in this study is a nonparametric path using R software. The results showed that the best model was obtained on a nonparametric truncated spline linear path model with 2 knot points. The model has the smallest GCV value of 25.9059 and R2 value of 96.96%. In addition, the results of hypothesis testing on function estimation have a significant effect on the relationship between service quality and willingness to pay, the relationship between service quality and on time to pay, the relationship between lifestyle and willingness to pay, and the relationship between lifestyle and on time pay. The novelty of this research is to model and test the hypothesis of nonparametric regression development, namely nonparametric truncated spline paths of linear, quadratic and cubic orders.

Key-Words: - Path Analysis, Path Nonparametric, Truncated Spline, House Ownership Credit.

I. INTRODUCTION

In a study, researchers usually observe a relationship between variables. One method of statistical analysis that can be used is regression analysis, where regression analysis is a method used to observe the relationship between two or more variables and can be used to determine the pattern of the relationship from a model whose form is not yet known [1].

There are three regression analysis approaches that can be done, namely nonparametric, semiparametric, and parametric approaches. In the nonparametric regression approach the form of the regression function is assumed to be unknown and the linearity assumption is not met, the parametric regression approach is assumed to be known and the linearity assumption is met [2], while the semiparametric regression approach is a combination of parametric and nonparametric regression that can be performed. if some of the curves are assumed to be known and some are assumed to be unknown [3].

Many current studies use more than one dependent variable. If there are two dependent variables, regression cannot be performed. Path analysis or path analysis can overcome these weaknesses. According to [4], path analysis is a method that observes the direct and indirect effects of the hypothesized variables. Path analysis is not a causal method (cause and effect), but a method that can be used for causal models formulated on basic knowledge and theories developed by researchers.

Often found in real life, the pattern of the relationship between the dependent and independent variables cannot be known in the shape of the curve. Path analysis based on nonparametric regression is a regression approach or path analysis that is in accordance with the pattern of the relationship between the dependent variable and the independent variable whose curve shape is not known [5]. In addition, nonparametric regression approaches tend to be more free to search for regression patterns so that they are flexible and objective.

House Ownership Credit (HOC) is one of the credit service facilities provided by banks to customers to meet their needs in the form of housing. Banks in providing a credit facility to customers, one of the most considered is the quality service and lifestyle. The quality service to customers can be seen in terms of reliability, responsiveness, assurance, empathy, and tangibles. Meanwhile, the customer's lifestyle towards HOC decisions can be seen in terms of activities, interests, and opinions.

Referring to [6], states that service quality needs to be improved, especially with regard to reliability so that various complaints can be avoided by customers and strive to provide satisfaction to customers regarding the provision of HOC. With the provision of HOC by the bank, there will be possible risks faced, such as willingness to pay credit (willingness to pay) and on time in paying credit. Therefore, to reduce risk, the bank must be more selective and careful in providing loans to customers through an attitude assessment. The assessment is meant by linking the variables of Service Quality and lifestyle with Willingness to Pay (willingness to pay) and Punctuality to Pay. As expected in this study, in order to minimize the HOC risk faced by the bank, by improving the HOC service system, it creates a high willingness of customers to pay for HOC.

Based on the description above, this study will examine the estimation of nonparametric truncated spline path functions with linear, quadratic, and cubic orders at one and two knot points. This study aims to estimate and select the best model on the variables that affect the timely payment of HOC. In addition, this study aims to test the hypothesis to determine the variables that have a significant effect on punctuality in paying HOC. In estimating the truncated spline nonparametric path function, the Least Square (LS) approach is used to determine the variables that affect the punctuality of paying HOC.

II. LITERATURE REVIEW

A. Nonparametric Regression Analysis

Nonparametric regression analysis is performed if the shape of the regression curve is unknown. The nonparametric regression approach is regression modeling that is not tied to regression assumptions. The nonparametric regression model is as follows [7]:

$$Y_i = f(X_i) + \varepsilon_i \quad (1)$$

B. Truncated Spline Nonparametric Regression

Spline is a part of regression analysis, more specifically nonparametric regression, and semiparametric regression. Research in the field of spline, which is independent and characterized, requires a comprehensive process, follows the stages, and is very long. According to [2], spline is a part or piece of polynomials that have segmented and continuous properties (truncated).

If the $f(X_i)$ regression curve is approximated by a spline function of order p with knots point $K_1, K_2, \dots, K_k$ (1the point of the first knot to the k-knot). The following functions are obtained:

$$\hat{f}(X_i) = \hat{\beta}_0 + \sum_{j=1}^{p} \hat{\beta}_j X_i^j + \sum_{k=1}^{K} \delta_k (X_i - K_k)_+^p \quad (2)$$

If equation (2) is substituted into equation (1), then the spline nonparametric regression equation is obtained as follows:

$$Y_i = \beta_0 + \sum_{j=1}^{p} \beta_j X_i^j + \sum_{k=1}^{K} \delta_k (X_i - K_k)_+^p + \varepsilon_i \quad (3)$$

The function is a truncated function given by:

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

where p is the order Spline and K is the selected knot [8].

C. Nonparametric Path Analysis

According to [9], path analysis is an analysis used to identify a causal relationship (cause-effect) between endogenous and exogenous variables. The path analysis model serves to analyze the pattern of relationships between variables. This is done to identify the direct or indirect effects of several exogenous variables on endogenous variables.

The parametric path model is still not able to help when the shape of the regression curve is unknown or unknown and the linearity assumption is not fulfilled. Therefore, an analysis called path analysis or nonparametric path analysis was developed. Nonparametric path analysis is basically an extension of nonparametric regression analysis.

Estimation of the path analysis function can be used with a nonparametric regression approach which shows the relationship between one endogenous variable and more than one exogenous variable (multiexogenous) involving n observations $(X_{1i}, X_{2i}, \ldots, X_{pi}, Y_i)$ which follows the regression model. The following equation is presented [10]:

$$\tilde{y} = \sum_{i=1}^{p} \tilde{f}_i(X_i) + \tilde{\varepsilon}_i;\ i = 1, 2, \ldots, p \quad (4)$$

D. Selection of Optimal Knot Points and Model Goodness Measures

The spline regression model is an analysis model that uses a nonparametric approach in which the least square estimation with the optimal knot point is selected based on the value of the smallest GCV (Generalized Cross Validation) [11]. The GCV method has the advantage of being asymptotically optimal, efficient, and simple in calculations, and does not require information on variance. The selection of knot points using the GCV method will be better if it is used on data that is normally distributed (Gaussian) [12]. In this study, the selection of the optimal knot point was obtained from the smallest GCV value and the measure of the goodness of the model was obtained from the largest coefficient of determination.

D.1 Generalized cross validation (GCV)

The best spline estimator is obtained from the optimal knot point. If the optimal knot point is obtained, the best spline function will be obtained. Selection of the best knot point more leads to parsimony or simplicity of the model. The smallest GCV value is the optimal knot point. The following equation is presented [13]:

$$GCV(\boldsymbol{K}) = \frac{MSE(\boldsymbol{K})}{[n^{-1} trace(\boldsymbol{I} - \boldsymbol{A}(\boldsymbol{K}))]^2} \quad (5)$$

Where $MSE(\boldsymbol{K}) = n^{-1}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ and is the knot point and the matrix is obtained:

$\hat{y} = \boldsymbol{A}(\boldsymbol{K})y$

$$A[\mathbf{K}] = \mathbf{X}[\mathbf{K}]\left(\mathbf{X}[\mathbf{K}]^t\,\mathbf{X}[\mathbf{K}]\right)^{-1}\mathbf{X}[\mathbf{K}]^t \quad (6)$$

D.2 Coefficient of determination (R2)

The coefficient of determination is a measure of the contribution of exogenous variables to endogenous variables. The best model can be seen from the value R^2 the biggest. The formula for the coefficient of determination is as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \quad 0 \le R^2 \le 1 \quad (7)$$

In determining the best model, the coefficient of determination has the criteria shown in Table 1 as follows [14].

Table 1. Criteria for Interpretation of the Coefficient of Determination

No.	R2 value	Criteria
1	<0.5	Not good
2	0.5-0.75	Enough
3	0.75>	Good

E. Hypothesis Test

Simultaneous hypothesis testing on the model with the truncated spline approach will be studied, if the following equation is given:

$$\begin{aligned}
f_{1i} &= \beta_{10} + \beta_{11}x_i + \delta_{21}(x_i - k_{11})_+ + \delta_{22}(x_i - k_{12})_+ + \delta_{23}(x_i - k_{13})_+ + \varepsilon_{1i} \\
f_{2i} &= \beta_{20} + \beta_{21}x_{1i} + \delta_{21}(x_i - k_{11})_+ + \delta_{22}(x_i - k_{12})_+ + \delta_{23}(x_i - k_{13})_+ + \beta_{22}y_{1i} \\
&\quad + \gamma_{21}(y_{1i} - k_{21})_+ + \gamma_{22}(y_{1i} - k_{22})_+ + \gamma_{23}(y_{1i} - k_{23})_+ + \varepsilon_{2i} \\
f_{3i} &= \beta_{30} + \beta_{31}x_{1i} + \delta_{31}(x_i - k_{11})_+ + \delta_{32}(x_i - k_{12})_+ + \delta_{33}(x_i - k_{13})_+ + \beta_{32}y_{1i} \\
&\quad + \gamma_{31}(y_{1i} - k_{21})_+ + \gamma_{32}(y_{1i} - k_{22})_+ + \gamma_{33}(y_{1i} - k_{23})_+ + \beta_{33}y_{2i} + \lambda_{31}(y_{2i} - k_{31})_+ \\
&\quad + \lambda_{32}(y_{2i} - k_{32})_+ + \lambda_{33}(y_{2i} - k_{33})_+ + \varepsilon_{3i}
\end{aligned} \quad (8)$$

To test the parameters found in equation (24), simultaneous hypothesis testing can be done as follows:

H_0: $\beta_{10} = \cdots = \beta_{1j} = \delta_{(m+1)1} = \cdots = \delta_{(m+k)1} = \lambda_{(m+1)1} = \cdots = \lambda_{(m+l)1} = 0$, vs

H_1: There is at least one $\beta_{1j} \neq 0$ or $\delta_{(m+k)1} \neq 0$ or $\lambda_{(m+l)1} \neq 0$

Furthermore, the F test statistic is obtained as follows

$$F_{hitung} = \frac{Q}{JKsisa/(n-p)} \quad (5)\ (9)$$

F. Research variable

F.1 Service quality

Service quality is all forms of activities carried out to carry out continuous quality improvements to the processes, products and services produced by the company. According to research by [15] service quality can be measured through reliability, responsiveness, assurance, empathy and tangibles.

F.2 Lifestyle

Lifestyle is a person's pattern of life in everyday life [16]. According to [17], lifestyle is measured through activities, interests and opinions.

F.3 Willingness to pay

Willingness to pay is a value where someone is willing to pay, sacrifice or exchange something to obtain goods or services. According to research by [18] to measure willingness to pay, indicators such as consultation, required documents, method and place of credit payment, payment deadlines and allocation of funds are used to measure the willingness to pay.

F.4 On time pay

According to [19] timeliness is the utilization of information and the level of compliance with the regulations that have been set. On time pay is defined as payments made on time. According to [20] on time paying can be measured by the desire to always pay on time and always pay monthly payments on time.

III. RESEARCH METHOD

A. Research Data

The data used in this study are primary data regarding service quality, lifestyle, willingness to pay, and on time to pay. Data obtained through questionnaires distributed to respondents who have been determined. Respondents in this study were customers of HOC Bank X in Sidoarjo City. The research instrument uses a Likert scale to measure the variables used. The sampling technique used is purposive sampling where the sampling is adjusted on the basis of certain characteristics or conditions that are the same as the population. The sample used in this study is 100 customers of HOC Bank X in 2021. The variables used are quality service (X_1), lifestyle (X_2), willingness to pay (Y_1), and on time to pay (Y_2).

B. Research Model

In this study, two exogenous variables were used, namely variables quality service (X_1), lifestyle (X_2), willingness to pay (Y_1), and on time to pay (Y_2). The path diagram for this study is presented in Figure 1.

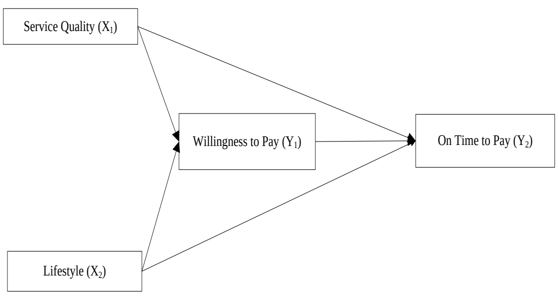


Fig.1. Research Path Diagram

C. Research Procedure

The steps taken in this study are as follows:

1. Make a path diagram based on the predetermined exogenous and endogenous variables.
2. Test the linearity assumption of the relationship between variables using Ramsey's Regression Specification Error Test (RESET) with a calculated F value. If the assumption of linearity is met, then parametric path analysis can be used. However, if the assumption of linearity is not met, it can be continued to the next stage.
3. Estimating the path function with truncated spline with linear, quadratic, and cubic orders with the number of knots, namely 1 point knot, and 2 point knot using the LS method.
4. Obtaining the function estimation results of each degree of the polynomial, namely linear, quadratic, and cubic with the number of knots, namely 1 point knot, and 2 point knot.
5. Selection of the optimum model and the optimum knot point based on the smallest GCV coefficient.
6. Interpret the estimation results of the best truncated spline model nonparametric path function and conclude from the results of the analysis.
7. Perform hypothesis testing on the best model.

IV. RESEARCH RESULT

A. Structural model development

Lemma 1 Forms of the linear truncated spline nonparametric path model

If given paired data $(X_{1i}, X_{2i}, Y_{1i}, Y_{2i})$ with $i = 1,2,\dots,n$ those following the nonparametric path analysis model, then the form of the linear nonparametric path analysis function is obtained as presented in equation (10) and the model in (11).

$$Y_{1i} = f_1(X_{1i}, X_{2i}) + \varepsilon_{1i} \qquad (10)$$

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

$$f_{1i} = \beta_{10} + \beta_{11}X_{1i} + \delta_{11}(X_{1i} - K_{11})_+ + \beta_{12}X_{2i} + \delta_{12}(X_{2i} - K_{21})_+$$

$$f_{2i} = \beta_{20} + \beta_{21}X_{1i} + \delta_{21}(X_{1i} - K_{11})_+ + \beta_{22}X_{2i} + \delta_{22}(X_{2i} - K_{21})_+ + \beta_{23}Y_{1i} + \delta_{23}(f_{1i} - K_{31})_+ \qquad (11)$$

Proof

Before obtaining a model in the linear nonparametric path analysis, the first model is obtained from (a) multiple linear regression analysis; (b) simple linear path analysis; and (c) nonparametric regression analysis as follows:

First part:

It is known that the multiple linear regression model with equation (12) and the model in (13)

$$Y_i = f(X_1, X_2) + \varepsilon_i \qquad (12)$$

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon_i \qquad (13)$$

The second part:

It is known that the simple path analysis model is presented in equation (14) and the model in (15)

$$Y_{1i} = f_1(X_{1i}, X_{2i}) + \varepsilon_{1i}$$
$$Y_{2i} = f_2(X_{1i}, X_{2i}, Y_{1i}) + \varepsilon_{2i} \qquad (14)$$

$$Y_{1i} = \beta_{10} + \beta_{11} X_1 + \beta_{12} X_2 + \varepsilon_{1i}$$
$$Y_{2i} = \beta_{20} + \beta_{21} X_1 + \beta_{22} X_2 + \beta_{23} Y_1 + \varepsilon_{2i} \qquad (15)$$

Part three:

After knowing the equation and multiple linear regression models, a nonparametric regression model can be made as presented in Equations (16) and (17).

$$Y_{1i} = f_1(X_{1i}, X_{2i}) + \varepsilon_{1i} \qquad (16)$$

$$f_{1i} = \beta_{10} + \beta_{11} X_{1i} + \delta_{11}(X_{1i} - K_{11})_+ + \beta_{12} X_{2i} + \delta_{12}(X_{2i} - K_{21})_+ \qquad (17)$$

From the equations in the simple linear regression analysis model, simple path analysis, and nonparametric regression analysis that has been described, a function that is formed as in equations (10) and (11) can be obtained, so that the following matrix is obtained:

$$\underset{\sim}{f}_{2nx1} = \boldsymbol{X}_{2nx12} \underset{\sim}{\beta}_{12x1}$$

$$\begin{bmatrix} f(X_{11}) \\ f(X_{12}) \\ \vdots \\ f(X_{1n}) \\ f(X_{21}) \\ f(X_{22}) \\ \vdots \\ f(X_{2n}) \end{bmatrix} = \begin{bmatrix} X_X & \underset{\sim}{0}_{nx7} \\ \underset{\sim}{0}_{nx5} & X_{XY} \end{bmatrix} \begin{bmatrix} \beta_{10} \\ \beta_{11} \\ \beta_{12} \\ \beta_{20} \\ \beta_{21} \\ \beta_{22} \\ \beta_{23} \\ \delta_{11} \\ \delta_{12} \\ \delta_{21} \\ \delta_{22} \\ \delta_{31} \end{bmatrix} \quad \text{dimana } X_X = \begin{bmatrix} 1 & X_{11} & (X_{11}-K_{11})_+ & X_{21} & (X_{21}-K_{21})_+ \\ 1 & X_{12} & (X_{12}-K_{11})_+ & X_{22} & (X_{22}-K_{21})_+ \\ & & \vdots & & \\ 1 & X_{1n} & (X_{1n}-K_{11})_+ & X_{2n} & (X_{2n}-K_{21})_+ \end{bmatrix}; X_{XY} = \begin{bmatrix} 1 & X_{11} & (X_{11}-K_{11})_+ & X_{21} & (X_{21}-K_{21})_+ & Y_{11} & (Y_{11}-K_{31})_+ \\ 1 & X_{12} & (X_{12}-K_{11})_+ & X_{22} & (X_{22}-K_{21})_+ & Y_{12} & (Y_{12}-K_{31})_+ \\ & & & \vdots & & & \\ 1 & X_{1n} & (X_{1n}-K_{11})_+ & X_{2n} & (X_{2n}-K_{21})_+ & Y_{1n} & (Y_{1n}-K_{31})_+ \end{bmatrix}$$

Lemma 2 Form of quadratic truncated spline nonparametric path model

If given paired data $(X_{1i}, X_{2i}, Y_{1i}, Y_{2i})$ with $i = 1,2,\ldots,n$ those that follow the nonparametric path analysis model, then the quadratic nonparametric path analysis function is obtained as presented in equation (18) and the model in (19).

$$Y_{1i} = f_1(X_{1i}, X_{1i}^2, X_{2i}, X_{2i}^2) + \varepsilon_{1i} \qquad (18)$$
$$Y_{2i} = f_2(X_{1i}, X_{1i}^2, X_{2i}, X_{2i}^2, Y_{1i}, Y_{1i}^2) + \varepsilon_{2i}$$

$$f_{1i} = \beta_{10} + \beta_{11} X_{1i} + \beta_{12} X_{1i}^2 + \delta_{11}(X_{1i} - K_{11})_+^2 + \beta_{13} X_{2i} + \beta_{14} X_{2i}^2 + \delta_{12}(X_{2i} - K_{21})_+^2$$
$$f_{2i} = \beta_{20} + \beta_{21} X_{1i} + \beta_{22} X_{1i}^2 + \delta_{21}(X_{1i} - K_{11})_+^2 + \beta_{23} X_{2i} + \beta_{24} X_{2i}^2 + \delta_{22}(X_{2i} - K_{21})_+^2 \qquad (19)$$

Proof

Before obtaining a model in quadratic nonparametric path analysis, the first model is obtained from (a) quadratic path analysis; and (b) linear nonparametric path analysis as follows:

First part:

It is known that the quadratic path analysis model with equation (20) and the model in (21)

$$Y_{1i} = f(X_1, X_1^2, X_2, X_2^2) + \varepsilon_{1i} \qquad (20)$$
$$Y_{2i} = f(X_1, X_1^2, X_2, X_2^2, Y_1, Y_1^2) + \varepsilon_{1i}$$

$$Y_{1i} = \beta_{10} + \beta_{11} X_1 + \beta_{12} X_1^2 + \beta_{13} X_2 + \beta_{14} X_2^2 + \varepsilon_{1i}$$
$$Y_{2i} = \beta_{20} + \beta_{21} X_1 + \beta_{22} X_1^2 + \beta_{23} X_2 + \beta_{24} X_2^2 + \beta_{25} Y_1 + \beta_{26} Y_1^2 + \varepsilon_{2i} \qquad (21)$$

The second part:

Obtained from the results of Lemma 1 which is a form of linear truncated spline nonparametric path analysis presented in equations (10) and (11) with the following matrix form:

$$\underset{\sim}{f}_{2nx1} = \boldsymbol{X}_{2nx12} \underset{\sim}{\beta}_{12x1} \qquad (16)$$

$$\begin{bmatrix} f(X_{11}) \\ f(X_{12}) \\ \vdots \\ f(X_{1n}) \\ f(X_{21}) \\ f(X_{22}) \\ \vdots \\ f(X_{2n}) \end{bmatrix} = \begin{bmatrix} X_X & \underset{\sim}{0}_{nx7} \\ \underset{\sim}{0}_{nx5} & X_{XY} \end{bmatrix} \begin{bmatrix} \beta_{10} \\ \beta_{11} \\ \beta_{12} \\ \beta_{20} \\ \beta_{21} \\ \beta_{22} \\ \beta_{23} \\ \delta_{11} \\ \delta_{12} \\ \delta_{21} \\ \delta_{22} \\ \delta_{31} \end{bmatrix} \quad \text{dimana } X_X = \begin{bmatrix} 1 & X_{11} & (X_{11}-K_{11})_+ & X_{21} & (X_{21}-K_{21})_+ \\ 1 & X_{12} & (X_{12}-K_{11})_+ & X_{22} & (X_{22}-K_{21})_+ \\ & & \vdots & & \\ 1 & X_{1n} & (X_{1n}-K_{11})_+ & X_{2n} & (X_{2n}-K_{21})_+ \end{bmatrix}; X_{XY} = \begin{bmatrix} 1 & X_{11} & (X_{11}-K_{11})_+ & X_{21} & (X_{21}-K_{21})_+ & Y_{11} & (Y_{11}-K_{31})_+ \\ 1 & X_{12} & (X_{12}-K_{11})_+ & X_{22} & (X_{22}-K_{21})_+ & Y_{12} & (Y_{12}-K_{31})_+ \\ & & & \vdots & & & \\ 1 & X_{1n} & (X_{1n}-K_{11})_+ & X_{2n} & (X_{2n}-K_{21})_+ & Y_{1n} & (Y_{1n}-K_{31})_+ \end{bmatrix}$$

From the equation in the quadratic path analysis model, and the results in Lemma 1, a function is formed as in equations (18) and (19), so that the following matrix is obtained:

$$\underset{\sim}{f}_{2nx1} = \boldsymbol{X}_{2nx17} \underset{\sim}{\beta}_{17x1}$$

$$\begin{bmatrix} f(X_{11}) \\ f(X_{12}) \\ \vdots \\ f(X_{1n}) \\ f(X_{21}) \\ f(X_{22}) \\ \vdots \\ f(X_{2n}) \end{bmatrix} = \begin{bmatrix} X_{XX^2} & \underset{\sim}{0}_{nx10} \\ \underset{\sim}{0}_{nx7} & X_{XX^2YY^2} \end{bmatrix} \begin{bmatrix} \beta_{10} \\ \beta_{11} \\ \beta_{12} \\ \beta_{13} \\ \beta_{14} \\ \beta_{20} \\ \beta_{21} \\ \beta_{22} \\ \beta_{23} \\ \beta_{24} \\ \beta_{25} \\ \beta_{26} \\ \delta_{11} \\ \delta_{12} \\ \delta_{21} \\ \delta_{22} \\ \delta_{31} \end{bmatrix} \quad \text{dimana } X_{XX^2} = \begin{bmatrix} 1 & X_{11} & X_{11}^2 & (X_{11}-K_{11})_+^2 & X_{21} & X_{21}^2 & (X_{21}-K_{21})_+^2 \\ 1 & X_{12} & X_{11}^2 & (X_{11}-K_{11})_+^2 & X_{21} & X_{21}^2 & (X_{21}-K_{21})_+^2 \\ & & & \vdots & & & \\ 1 & X_{1n} & X_{1n}^2 & (X_{11}-K_{11})_+^2 & X_{2n} & X_{2n}^2 & (X_{21}-K_{21})_+^2 \end{bmatrix}; X_{XX^2YY^2} = \begin{bmatrix} 1 & X_{11} & X_{11}^2 & (X_{11}-K_{11})_+^2 & X_{21} & X_{21}^2 & (X_{21}-K_{21})_+^2 & Y_{11} & Y_{11}^2 & (Y_{11}-K_{31})_+^2 \\ 1 & X_{12} & X_{12}^2 & (X_{12}-K_{11})_+^2 & X_{22} & X_{22}^2 & (X_{22}-K_{21})_+^2 & Y_{12} & Y_{12}^2 & (Y_{12}-K_{31})_+^2 \\ & & & & \vdots & & & & & \\ 1 & X_{1n} & X_{1n}^2 & (X_{1n}-K_{11})_+^2 & X_{2n} & X_{2n}^2 & (X_{2n}-K_{21})_+^2 & Y_{1n} & Y_{1n}^2 & (Y_{1n}-K_{31})_+^2 \end{bmatrix}$$

Lemma 3 Forms of cubic truncated spline nonparametric path analysis model

If given paired data $(X_{1i}, X_{2i}, Y_{1i}, Y_{2i})$ with $i = 1,2,\ldots,n$ those that follow the nonparametric path analysis model, then the form of the cubic nonparametric path analysis function is obtained as presented in equation (22) and the model in (23).

$$Y_{1i} = f_1(X_{1i}, X_{1i}^2, X_{1i}^3, X_{2i}, X_{2i}^2, X_{2i}^3) + \varepsilon_{1i}$$
$$Y_{2i} = f_2(X_{1i}, X_{1i}^2, X_{1i}^3, X_{2i}, X_{2i}^2, X_{2i}^3, Y_{1i}, Y_{1i}^2, Y_{1i}^3) + \varepsilon_{2i} \qquad (22)$$

$$f_{1i} = \beta_{10} + \beta_{11} X_{1i} + \beta_{12} X_{1i}^2 + \beta_{13} X_{1i}^3 + \delta_{11}(X_{1i} - K_{11})_+^3 + \beta_{14} X_{2i} + \beta_{15} X_{2i}^2 + \beta_{16} X_{2i}^3 + \delta_{12}(X_{2i} - K_{21})_+^3$$
$$f_{2i} = \beta_{20} + \beta_{21} X_{1i} + \beta_{22} X_{1i}^2 + \beta_{23} X_{1i}^3 + \delta_{21}(X_{1i} - K_{11})_+^3 + \beta_{24} X_{2i} + \beta_{25} X_{2i}^2 + \beta_{26} X_{2i}^3 + \delta_{22}(X_{2i} - K_{21})_+^3$$
$$+ \beta_{27} Y_{1i} + \beta_{28} Y_{1i}^2 + \beta_{29} Y_{1i}^3 + \delta_{23}(f_{1i} - K_{31})_+^3 \qquad (23)$$

Proof

From the results in Lemma 1 and Lemma 2, an expansion can be made by adding the cubic order, so we can obtain a function that is formed as in equations (22) and (23), the following matrix is obtained:

$$\underset{\sim}{f}_{2nx1} = \boldsymbol{X}_{2nx22}\underset{\sim}{\beta}_{22x1}$$

$$\begin{bmatrix} f(X_{11}) \\ f(X_{12}) \\ \vdots \\ f(X_{1n}) \\ f(X_{21}) \\ f(X_{22}) \\ \vdots \\ f(X_{2n}) \end{bmatrix} = \begin{bmatrix} X_{XX^3} & \underset{\sim}{0}_{nx13} \\ \underset{\sim}{0}_{nx9} & X_{XX^3YY^3} \end{bmatrix} \begin{bmatrix} \beta_{10} \\ \beta_{11} \\ \beta_{12} \\ \beta_{13} \\ \beta_{14} \\ \beta_{15} \\ \beta_{16} \\ \beta_{20} \\ \beta_{21} \\ \beta_{22} \\ \beta_{23} \\ \beta_{24} \\ \beta_{25} \\ \beta_{26} \\ \beta_{27} \\ \beta_{28} \\ \beta_{29} \\ \delta_{11} \\ \delta_{12} \\ \delta_{21} \\ \delta_{22} \\ \delta_{31} \end{bmatrix} \text{ dimana } X_{XX^2} = \begin{bmatrix} 1 & X_{11} & X_{11}^2 & X_{11}^3 & (X_{11}-K_{11})_+^3 & X_{21} & X_{21}^2 & X_{21}^3 & (X_{21}-K_{21})_+^3 \\ 1 & X_{12} & X_{12}^2 & X_{12}^3 & (X_{12}-K_{11})_+^3 & X_{22} & X_{22}^2 & X_{22}^3 & (X_{22}-K_{21})_+^3 \\ & & & & \vdots & & & & \\ 1 & X_{1n} & X_{1n}^2 & X_{1n}^3 & (X_{1n}-K_{11})_+^3 & X_{2n} & X_{2n}^2 & X_{2n}^3 & (X_{2n}-K_{21})_+^3 \end{bmatrix};$$

$$X_{XX^2YY^2} = \begin{bmatrix} 1 & X_{11} & X_{11}^2 & X_{11}^3 & (X_{11}-K_{11})_+^3 & X_{21} & X_{21}^2 & X_{21}^3 & (X_{21}-K_{21})_+^3 & Y_{11} & Y_{11}^2 & Y_{11}^3 & (Y_{11}-K_{21})_+^3 \\ 1 & X_{12} & X_{12}^2 & X_{12}^3 & (X_{12}-K_{11})_+^3 & X_{22} & X_{22}^2 & X_{22}^3 & (X_{22}-K_{21})_+^3 & Y_{12} & Y_{12}^2 & Y_{12}^3 & (Y_{12}-K_{21})_+^3 \\ & & & & & & \vdots & & & & & & \\ 1 & X_{1n} & X_{1n}^2 & X_{1n}^3 & (X_{1n}-K_{11})_+^3 & X_{2n} & X_{2n}^2 & X_{2n}^3 & (X_{2n}-K_{21})_+^3 & Y_{1n} & Y_{1n}^2 & Y_{1n}^3 & (Y_{1n}-K_{21})_+^3 \end{bmatrix}$$

Theorem

If paired data is given following the nonparametric path model in the crosssection data that fulfill the form of the nonparametric path function for data crosssection as presented in Lemma 1, Lemma 2, and Lemma 3, the truncated spline estimator method that minimizes the number of squares of error using the Ordinary Least Square method is presented. as follows:

$$\underset{\underset{\sim}{\beta} \in R^{p+1+k}}{Min}\left\{\underset{\sim}{\varepsilon}^t\underset{\sim}{\varepsilon}\right\} = \underset{\underset{\sim}{\beta} \in R^{p+1+k}}{Min}\left\{\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}^t \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}\right\}$$

lowered against $\hat{\underset{\sim}{\beta}}$ and then equated with zero, then:

$$\frac{\partial\left(\underset{\sim}{\varepsilon}^{\mathrm{T}}\underset{\sim}{\varepsilon}\right)}{\partial\left(\hat{\underset{\sim}{\beta}}\right)} = 0$$

$$\frac{\partial\left(\underset{\sim}{Y}^{\mathrm{T}}\underset{\sim}{Y} - 2\underset{\sim}{\theta}^{\mathrm{T}}\mathbf{X}(\underset{\sim}{K})^{\mathrm{T}}\underset{\sim}{Y} + \mathbf{X}(\underset{\sim}{K})^{\mathrm{T}}\underset{\sim}{\beta}^{\mathrm{T}}\mathbf{X}(\underset{\sim}{K})\underset{\sim}{\beta}\right)}{\partial\left(\hat{\underset{\sim}{\beta}}\right)} = 0$$

$$-2\mathbf{X}(\underset{\sim}{K})^{\mathrm{T}}\underset{\sim}{Y} + 2\mathbf{X}(\underset{\sim}{K})^{\mathrm{T}}\mathbf{X}(\underset{\sim}{K})\hat{\underset{\sim}{\beta}} = 0$$

$$-2\left(\mathbf{X}(\underset{\sim}{K})^{\mathrm{T}}\underset{\sim}{Y} - \mathbf{X}(\underset{\sim}{K})^{\mathrm{T}}\mathbf{X}(\underset{\sim}{K})\hat{\underset{\sim}{\beta}}\right) = 0$$

$$\mathbf{X}(\underset{\sim}{K})^{\mathrm{T}}\underset{\sim}{Y} - \mathbf{X}(\underset{\sim}{K})^{\mathrm{T}}\mathbf{X}(\underset{\sim}{K})\hat{\underset{\sim}{\beta}} = 0$$

$$\mathbf{X}(\underset{\sim}{K})^{\mathrm{T}}\mathbf{X}(\underset{\sim}{K})\hat{\underset{\sim}{\beta}} = \mathbf{X}(\underset{\sim}{K})^{\mathrm{T}}\underset{\sim}{Y}$$

$$\hat{\underset{\sim}{\beta}} = \left[\mathbf{X}(\underset{\sim}{K})^{\mathrm{T}}\mathbf{X}(\underset{\sim}{K})\right]^{-1}\left[\mathbf{X}(\underset{\sim}{K})^{\mathrm{T}}\underset{\sim}{Y}\right] \qquad (24)$$

Proof

In the form of a matrix, the nonparametric path is expressed as follows:

$$\begin{bmatrix} Y_{11} \\ Y_{12} \\ \vdots \\ Y_{1n} \\ Y_{21} \\ Y_{22} \\ \vdots \\ Y_{2n} \end{bmatrix}_{nx1} = \begin{bmatrix} f(X_{11}) \\ f(X_{12}) \\ \vdots \\ f(X_{1n}) \\ f(X_{21}) \\ f(X_{22}) \\ \vdots \\ f(X_{2n}) \end{bmatrix}_{nx1} + \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \vdots \\ \varepsilon_{1n} \\ \varepsilon_{21} \\ \varepsilon_{22} \\ \vdots \\ \varepsilon_{2n} \end{bmatrix}_{nx1} \qquad (25)$$

Truncated spline nonparametric regression curve equation for $i = 1, 2, \ldots, n$ written in the form of an equation as follows:

$$\begin{bmatrix} f(X_1) \\ f(X_2) \\ \vdots \\ f(X_n) \end{bmatrix} = \begin{bmatrix} 1 & X_1 & \cdots & X_1^p & (x_1-K_1)_+^m & \cdots & (x_1-K_r)_+^m \\ 1 & X_2 & \cdots & X_2^p & (x_2-K_1)_+^m & \cdots & (x_2-K_r)_+^m \\ & & & \vdots & & & \\ 1 & X_n & \cdots & x_n^p & (x_n-K_1)_+^m & \cdots & (x_n-K_r)_+^m \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_m \\ \delta_1 \\ \delta_2 \\ \vdots \\ \delta_r \end{bmatrix} \qquad (26)$$

It can also be written as:

$$\underset{\sim}{f}(X) = \mathbf{X}(\underset{\sim}{K})\underset{\sim}{\beta} \qquad (27)$$

Based on the nonparametric regression model, the truncated spline function of order m with knots points K_1, K_2, …, K_h. assumed to be normally distributed, independent of each other with a mean value of zero ε_i and a variance of σ^2, so that $\varepsilon_i \sim N\left(0, \sigma^2\right)$. Nonparametric regression curve estimation can be done by using the Ordinary Least Square (OLS) method. In its estimation, OLS minimizes the number of squares of errors, where the error is obtained from equation (24).

$$\varepsilon_i = Y_i - f(X_i) \qquad (28)$$

Equation (28) can be expressed in matrix notation as follows:

$$\underset{\sim}{\varepsilon} = \underset{\sim}{Y} - \underset{\sim}{f}(X) \qquad (29)$$

The parameter estimates obtained by the OLS method are as follows:

$$\begin{aligned} \underset{\sim}{\varepsilon}^{\mathrm{T}}\underset{\sim}{\varepsilon} &= \left(\underset{\sim}{Y} - \underset{\sim}{f}(X)\right)^{\mathrm{T}}\left(\underset{\sim}{Y} - \underset{\sim}{f}(X)\right) \\ &= \left(\underset{\sim}{Y} - \mathbf{X}(\underset{\sim}{K})\underset{\sim}{\beta}\right)^{\mathrm{T}}\left(\underset{\sim}{Y} - \mathbf{X}(\underset{\sim}{K})\underset{\sim}{\beta}\right) \end{aligned}$$

$$= \left(\underset{\sim}{Y}^{\mathrm{T}} - \mathbf{X}\left(\underset{\sim}{K}\right)^{\mathrm{T}} \underset{\sim}{\beta}^{\mathrm{T}}\right)\left(\underset{\sim}{Y} - \mathbf{X}\left(\underset{\sim}{K}\right)\underset{\sim}{\beta}\right)$$

$$= \underset{\sim}{Y}^{\mathrm{T}}\underset{\sim}{Y} - \underset{\sim}{Y}^{\mathrm{T}}\mathbf{X}\left(\underset{\sim}{K}\right)\underset{\sim}{\beta} - \underset{\sim}{\beta}^{\mathrm{T}}\mathbf{X}\left(\underset{\sim}{K}\right)^{\mathrm{T}}\underset{\sim}{Y} + \mathbf{X}\left(\underset{\sim}{K}\right)^{\mathrm{T}}\underset{\sim}{\beta}^{\mathrm{T}}\mathbf{X}\left(\underset{\sim}{K}\right)\underset{\sim}{\beta}$$

$$= \underset{\sim}{Y}^{\mathrm{T}}\underset{\sim}{Y} - 2\underset{\sim}{\beta}^{\mathrm{T}}\mathbf{X}\left(\underset{\sim}{K}\right)^{\mathrm{T}}\underset{\sim}{Y} + \mathbf{X}\left(\underset{\sim}{K}\right)^{\mathrm{T}}\underset{\sim}{\beta}^{\mathrm{T}}\mathbf{X}\left(\underset{\sim}{K}\right)\underset{\sim}{\beta} \tag{30}$$

The parameter estimate is obtained by forming the normal equation from equation (30) as follows:

$$\frac{\partial\left(\underset{\sim}{\varepsilon}^{\mathrm{T}}\underset{\sim}{\varepsilon}\right)}{\partial\left(\hat{\underset{\sim}{\beta}}\right)} = 0$$

$$\frac{\partial\left(\underset{\sim}{Y}^{\mathrm{T}}\underset{\sim}{Y} - 2\underset{\sim}{\theta}^{\mathrm{T}}\mathbf{X}\left(\underset{\sim}{K}\right)^{\mathrm{T}}\underset{\sim}{Y} + \mathbf{X}\left(\underset{\sim}{K}\right)^{\mathrm{T}}\underset{\sim}{\beta}^{\mathrm{T}}\mathbf{X}\left(\underset{\sim}{K}\right)\underset{\sim}{\beta}\right)}{\partial\left(\hat{\underset{\sim}{\beta}}\right)} = 0$$

$$-2\mathbf{X}\left(\underset{\sim}{K}\right)^{\mathrm{T}}\underset{\sim}{Y} + 2\mathbf{X}\left(\underset{\sim}{K}\right)^{\mathrm{T}}\mathbf{X}\left(\underset{\sim}{K}\right)\hat{\underset{\sim}{\beta}} = 0$$

$$-2\left(\mathbf{X}\left(\underset{\sim}{K}\right)^{\mathrm{T}}\underset{\sim}{Y} - \mathbf{X}\left(\underset{\sim}{K}\right)^{\mathrm{T}}\mathbf{X}\left(\underset{\sim}{K}\right)\hat{\underset{\sim}{\beta}}\right) = 0$$

$$\mathbf{X}\left(\underset{\sim}{K}\right)^{\mathrm{T}}\underset{\sim}{Y} - \mathbf{X}\left(\underset{\sim}{K}\right)^{\mathrm{T}}\mathbf{X}\left(\underset{\sim}{K}\right)\hat{\underset{\sim}{\beta}} = 0$$

$$\mathbf{X}\left(\underset{\sim}{K}\right)^{\mathrm{T}}\mathbf{X}\left(\underset{\sim}{K}\right)\hat{\underset{\sim}{\beta}} = \mathbf{X}\left(\underset{\sim}{K}\right)^{\mathrm{T}}\underset{\sim}{Y} \tag{31}$$

Based on equation (24), the nonparametric path estimation is obtained based on the matrix notation in equation (27) which is as follows:

$$\hat{\underset{\sim}{f}}(X) = \mathbf{X}\left(\underset{\sim}{K}\right)\hat{\underset{\sim}{\beta}}$$

$$\hat{\underset{\sim}{f}}(X) = \mathbf{X}\left(\underset{\sim}{K}\right)\left[\mathbf{X}\left(\underset{\sim}{K}\right)^{\mathrm{T}}\mathbf{X}\left(\underset{\sim}{K}\right)\right]^{-1}\left[\mathbf{X}\left(\underset{\sim}{K}\right)^{\mathrm{T}}\underset{\sim}{Y}\right] \tag{32}$$

So, equation (32) shows the curve estimate for the nonparametric path model with the truncated spline approach.

B. Linearity testing

In statistical modeling, information is needed about the pattern of relationships between variables to determine whether a method is used through a parametric or nonparametric, or semiparametric approach. The results of linearity testing with the Ramsey RESET are as follows:

Table 2. Ramsey RESET Linearity Test Results.

Variable	P-value	Relationship
X1 to Y1	<0.0001	Nonlinear
X2 to Y1	0.0263	Nonlinear
X1 to Y2	0.0009	Nonlinear
X2 to Y2	0.0061	Nonlinear
Y1 to Y2	0.0158	Nonlinear

Based on Table 2, the test results show that p-value$<\alpha$ (0.05), then it can be decided to reject H_0. Therefore, the relationship between variables indicates that the data does not meet the linearity assumption.

C. Truncated spline nonparametric Path model degree of linear polynomial (order p=1)

The truncated spline nonparametric path model at the moment is linear (order p = 1) with 1 knot for two exogenous variables and two endogenous variables as follows:

$$\hat{f}_{1\mathrm{i}} = 96.7248 + 0.1263X_{1i} - 0.0729(X_{1i} - 65.76)_{+} - 0.2559X_{2i} + 0.3827(X_{2i} - 35.85)_{+}$$

$$\hat{f}_{2\mathrm{i}} = 60.8009 + 0.1388X_{1i} - 0.2725(X_{1i} - 65.76)_{+} + 0.2390X_{2i} - 0.4989(X_{2i} - 35.85)_{+} - 0.6755f_{1\mathrm{i}} + 0.2728(f_{1i} - 95.45)_{+} \tag{33}$$

The truncated spline nonparametric path model at the moment is linear (order p = 1) with 2 knots for two exogenous variables and two endogenous variables as follows:

$$\hat{f}_{1\mathrm{i}} = 96.8497 + 0.2116X_{1i} - 0.2917(X_{1i} - 60.90)_{+} + 0.1723(X_{1i} - 68.22)_{+} - 0.4131X_{2i} + 1.0052(X_{2i} - 35.85)_{+} - 0.6783(X_{2i} - 40.47)_{+}$$

$$\hat{f}_{2\mathrm{i}} = -0.2406 + 0.1286X_{1i} - 0.1038(X_{1i} - 60.90)_{+} - 0.1813(X_{1i} - 68.22)_{+} + 0.5313X_{2i} - 1.1000(X_{2i} - 35.85)_{+} + 0.3198(X_{2i} - 40.47)_{+} - 0.1168f_{1\mathrm{i}} - 0.3452(f_{1i} - 95.45)_{+} - 4.9569(f_{1i} - 99.72)_{+} \tag{34}$$

D. Truncated spline nonparametric path model degree of quadratic polynomial (order p = 2)

The truncated spline nonparametric path model at the quadratic time (order p = 2) with 1 knot for two exogenous variables and two endogenous variables as follows:

$$\hat{f}_{1\mathrm{i}} = 0.0842 + 2.4051X_{1i} - 0.0205X_{1i}^{2} + 0.0248(X_{1i} - 60.90)_{+}^{2} + 1.4117X_{2i} - 0.0189X_{2i}^{2} + 0.0172(X_{2i} - 35.85)_{+}^{2}$$

$$\hat{f}_{2\mathrm{i}} = 0.0130 + 0.3659X_{1i} - 0.0024X_{1i}^{2} - 0.0026(X_{1i} - 60.90)_{+}^{2} + 0.2280X_{2i} - 0.0036X_{2i}^{2} - 0.0068(X_{2i} - 35.85)_{+}^{2} + 0.5983f_{1\mathrm{i}} - 0.0068f_{1i}^{2} + 0.0547(f_{1i} - 95.45)_{+}^{2} \tag{35}$$

The truncated spline nonparametric path model at quadratic time (order p = 2) with 2 knots for two exogenous variables and two endogenous variables as follows:

$$\hat{f}_{1i} = 0.0932 + 2.5214X_{1i} - 0.0221X_{1i}^2 + 0.0397(X_{1i} - 60.90)_+^2 - 0.0175(X_{1i} - 68.22)_+^2 + 1.4981X_{2i} - 0.0228X_{2i}^2 + 0.0674(X_{2i} - 35.85)_+^2 - 0.0609(X_{2i} - 40.47)_+^2$$

$$\hat{f}_{2i} = 0.0122 + 0.3313X_{1i} - 0.0018X_{1i}^2 - 0.0092(X_{1i} - 60.90)_+^2 + 0.0082(X_{1i} - 68.22)_+^2 + 0.2021X_{2i} - 0.0018X_{2i}^2 - 0.0337(X_{2i} - 35.85)_+^2 + 0.0332(X_{2i} - 40.47)_+^2 + 0.5631f_{1i} - 0.0065f_{1i}^2 + 0.0486(f_{1i} - 95.45)_+^2 + 0.0000(f_{1i} - 100)_+^2 \quad (36)$$

E. Truncated spline nonparametric path model degrees of cubic polynomials (order p=3)

The truncated spline nonparametric path model at cubic time (order p = 3) with 1 knot for two exogenous variables and two endogenous variables as follows:

$$\hat{f}_{1i} = 0.0001 + 0.0019X_{1i} + 0.0621X_{1i}^2 - 0.0006X_{1i}^3 + 0.0020(X_{1i} - 60.90)_+^3 + 0.0017X_{2i} + 0.0225X_{2i}^2 - 0.0003X_{2i}^3 - 0.0002(X_{2i} - 43.51)_+^3$$

$$\hat{f}_{2i} = 0.0001 + 0.0001X_{1i} + 0.0036X_{1i}^2 - 0.0001X_{1i}^3 - 0.0001(X_{1i} - 60.90)_+^3 + 0.0001X_{2i} + 0.0015X_{2i}^2 - 0.0001X_{2i}^3 + 0{,}0001(X_{2i} - 43.51)_+^3 + 0.0002f_{1i} + 0.0088f_{1i}^2 - 0.0001f_{1i}^3 + 0.0000(f_{1i} - 98.65)_+^3 \quad (37)$$

The truncated spline nonparametric path model at cubic time (order p = 3) with 2 knots for two exogenous variables and two endogenous variables as follows:

$$\hat{f}_{1i} = 0.0001 + 0.0025X_{1i} + 0.0711X_{1i}^2 - 0.0008X_{1i}^3 + 0.0071(X_{1i} - 60.90)_+^3 - 0.0078(X_{1i} - 68.22)_+^3 + 0.0015X_{2i} + 0.0274X_{2i}^2 - 0.0004X_{2i}^3 - 0.0041(X_{2i} - 36.67)_+^3 + 0.0762(X_{2i} - 40.58)_+^3$$

$$\hat{f}_{2i} = 0.0001 + 0.0001X_{1i} + 0.0029X_{1i}^2 - 0.0001X_{1i}^3 - 0.0001(X_{1i} - 60.90)_+^3 + 0.0001(X_{1i} - 68.22)_+^3 + 0.0001X_{2i} + 0.0014X_{2i}^2 - 0.0001X_{2i}^3 - 0.0004(X_{2i} - 36.67)_+^3 + 0.0005(X_{2i} - 40.58)_+^3 + 0.0002f_{1i} - 0.0087f_{1i}^2 - 0.0001f_{1i}^3 - 0.0006(f_{1i} - 95.57)_+^3 - 0.0001(f_{1i} - 97.88)_+^3 \quad (38)$$

F. Best order polynomial truncated spline nonparametric model

The best truncated spline nonparametric path model is obtained if the knot point is optimal. Meanwhile, to obtain the optimal knot point, you must find the smallest GCV value. The estimation of the nonparametric truncated spline path model for the linear, quadratic, and cubic orders with 1 knot point and 2 best knot point is as follows:

Table 3. Estimation of Best Truncated Spline Nonparametric Path Functions.

Order	Knots	R^2	GCV
Linear	1	0.9621	25.9720
Linear	2	0.9696	25.9059
Quadratic	1	0.9609	27.1042
Quadratic	2	0.9613	26.8356
Cubic	1	0.9587	38.1283
Cubic	2	0.9595	32.5142

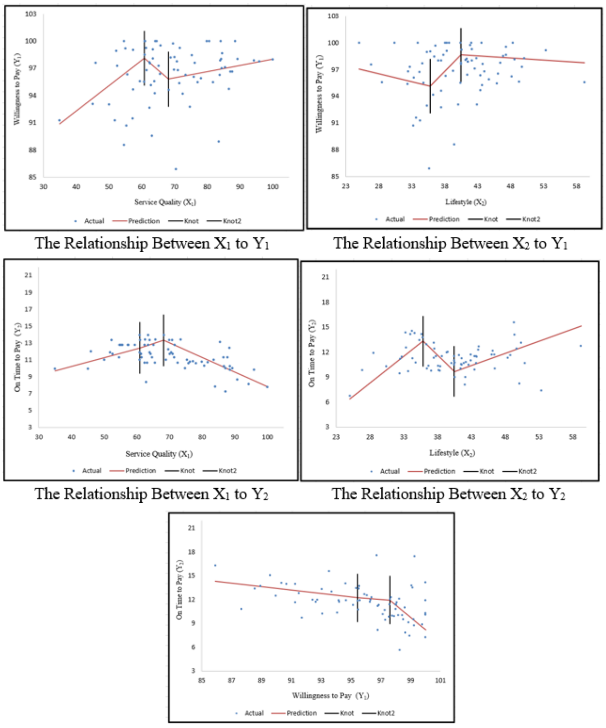
Based on Table 3, it is known that the smallest GCV value of 25.9059 and the largest R^2 value of 0.9696 is obtained in the nonparametric path model truncated spline degree of linear polynomial (order p = 2) with 2 knots. The effect of the best model estimation results is obtained as follows:

$$\hat{f}_{1i} = 96.8497 + 0.2116X_{1i} - 0.2917(X_{1i} - 60.90)_+ + 0.1723(X_{1i} - 68.22)_+ - 0.4131X_{2i} + 1.0052(X_{2i} - 35.85)_+ - 0.6783(X_{2i} - 40.47)_+$$

$$\hat{f}_{2i} = -0.2406 + 0.1285X_{1i} - 0.1038(X_{1i} - 60.90)_+ - 0.1813(X_{1i} - 68.22)_+ + 0.5313X_{2i} - 1.1000(X_{2i} - 35.85)_+ + 0.3198(X_{2i} - 40.47)_+ - 0.1169Y_{1i} - 0.3452(f_{1i} - 95.45)_+ - 4.9569(f_{1i} - 99.72)_+ \quad (39)(37)$$

The nonparametric path model truncated spline degree of linear polynomial (order p = 1) with 2 point knots has several relationships. In the form of an image, the nonparametric path model is cut by a degree of linear polynomial spline (order p = 1) with 2 point knots presented as follows:

The nonparametric path model truncated spline degree of linear polynomial (order p = 1) with 2 point knots has several relationships. The relationship between the variable quality service (X_1) and the variable willingness to pay (Y_1), the relationship between the variable lifestyle (X_2) and the variable willingness to pay (Y_1), the relationship between the variable quality service (X_1) with the variable on time to pay (Y_2), the relationship between the variable lifestyle (X_2) and the variable on time to pay (Y_2), and the relationship between the variable willingness to pay (Y_1) and the variable on time to pay (Y_2). In the form of an image, only the relationship between one exogenous variable and one endogenous variable is presented to facilitate discussion. In the form of an image, the nonparametric path model is cut by a degree of linear polynomial spline (order p = 1) with 2 point knots presented as follows:



The Relationship Between X_1 to Y_1 The Relationship Between X_2 to Y_1

The Relationship Between X_1 to Y_2 The Relationship Between X_2 to Y_2

The Relationship Between Y_1 to Y_2

Fig.2. Functions of Truncated Spline Linear Nonparametric Path Model 2 Knots

The truncated spline linear 2 points knot nonparametric path model is as follows:

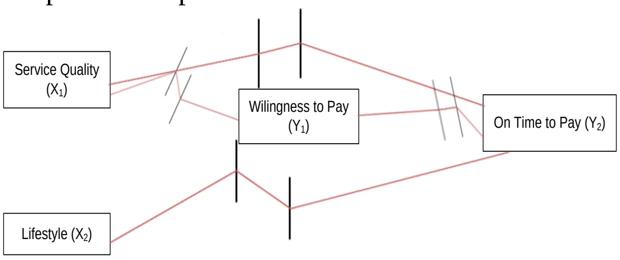


Fig.3. Truncated Spline Linear 2 Knot Nonparametric Path Model

G. Hypothesis Test

The linear hypothesis testing of parameters in the best model of path truncated spline analysis is used to determine whether there is an effect of exogenous variables on endogenous variables. The test was carried out on the best model, namely the linear order truncated spline path model with 2 knots point. Analysis results of hypothesis testing for the truncated spline path model, the degree of two knots linear polynomials is presented in Table 4.

Table 4. Results of Testing the Linear Parameter Hypothesis.

Function	***P-value***	**Decision**	**Conclusion**
X1 to Y1	<0.0001	Reject H_0	Significant
X1 to Y2	0.0003	Reject H_0	Significant
X2 to Y1	<0.0001	Reject H_0	Significant
X2 to Y2	0.0051	Reject H_0	Significant
Y1 to Y2	0.9999	Accept H_0	Not significant

Based on Table 4, the results of hypothesis testing on function estimates that have a significant effect on the relationship between service quality and willingness to pay, the relationship between service quality and on time to pay, the relationship between lifestyle and willingness to pay, and the relationship between lifestyle and on time pay. While testing the hypothesis on the estimation of the relationship function which is not significant when the relationship between willingness to pay and on time to pay.

V. CONCLUSION AND SUGGESTION

A. Conclusion

Based on the results of the analysis and discussion that has been done, it can be concluded that:

1) Estimation of nonparametric truncated spline function of linear, quadratic, and cubic 1 and 2 knot points on the variables of service quality, lifestyle, and willingness to pay that affect the variable on time to pay. The best model is obtained on a nonparametric truncated spline linear path model with 2 knot points. This model has the smallest GCV value compared to other models, which is 25.9059 and the R^2 value is 96.96%.
2) The results of hypothesis testing indicate that there are four relationships that have a significant effect, namely the relationship between service quality and willingness to pay, the relationship between service quality and punctuality in paying, the relationship between lifestyle and willingness to pay, and the relationship between lifestyle and punctuality in paying. Meanwhile, the estimated function that has no significant effect on the relationship between willingness to pay and on time to pay.

B. Suggestion

Based on the results of the analysis and the conclusions obtained in this study, the suggestion that can be given is that this study only estimates the function of linear, quadratic, cubic with one knot and point. For further research, it can be developed by estimating functions in high order with high knot points. In addition to the estimation of the nonparametric path function that has been carried out, further research it can be developed by estimating the semiparametric path function.

REFERENCES

[1] Uyanık, G. K., & Güler, N., A Study on Multiple Linear Regression Analysis, *Procedia - Social and Behavioral Sciences*, Vol.106, 2013, pp. 234–240.

[2] Hidayat, R., Yuliani, Y., & Sam, M., Nonparametric Regression Model with Truncated Spline Approach, *Proceedings*, Vol.3, No.1, 2017.

[3] Suparti, S., & Prahutama, A., Nonparametric Regression Modeling Using Local Polynomial Approach on Electrical Load in Semarang City, *Media Statistics*, Vol.9, No.2, 2016, pp. 85-93.

[4] Sudaryono, S., Application of Analysis (Path Analysis) Based on the Order of Placement of Variables in Research, *Journal of Education and Culture*, Vol.17, No.4, 2011, pp. 391-403.

[5] Eubank, R. L., Huang, C., Maldonado, Y. M., Wang, N., Wang, S., & Buchanan, R. J., Smoothing spline estimation in varying- coefficient models, *Journal of the Royal Statistical Society: Series B, Statistical Methodology*, Vol.66, No.3, 2004, pp. 653-667.

[6] Husnawaty, H., Quality of Credit Provisions on Increasing Home Ownership Loans (HOC), *Bongaya Journal for Research in Accounting (BJRA)*, Vol.3, No.1, 2020, pp. 65-75.

[7] Prahutama, A., Nonparametric Regression Model with Fourier Series Approach in the Case of the Open Unemployment Rate in East Java, *In Proceedings of the UNDIP National Statistics Seminar*, 2011.

[8] Bintariningrum, M. F., & Budiantara, I. N., Truncated spline nonparametric regression modeling and its application to the crude birth rate in Surabaya, *ITS Journal of Science and Arts*, Vol.3, No.1, 2014, pp. D7-D12.

[9] Kadir, A. *Applied Statistics: Concepts, Examples, and Data Analysis with the SPSS/Lisrel Program in Research. Third edition*, Fourth Printing, PT Raja Grafindo Persada, 2018.

[10] Fernandes, A. A. R., & Suhartono, Reproducing Kernel Hilbert space for penalized regression multi-predictors: Case in longitudinal data, *International Journal of Mathematical Analysis*, Vol.8, No.40, 2014, pp. 1951-1961.

[11] Puspitasari, E. & Sutanto, H. T., Optimal Spline Knot Regression Model to Determine the Factors Affecting the Number of Infant Mortality in East Java, *MATHunesa Journal of Scientific Mathematics*, Vo.2, No.1, 2013.

[12] Nurmalisa, W., & Nur, I. M., Nonparametric Spline Regression Modeling with Optimal Knot Point Selection Using Unbiased Risk and Generalized Cross Validation Methods in the Case of the Human Development Index in Central Java, *In Proceedings of the Unimus Student National Seminar 2*, 2019.

[13] Wulandari, I. D. A. M. I., & Budiantara, I. N., Analysis of Factors Affecting Percentage of Poor Population and Per capita Food Expenditure in East Java Using Biendogenous Spline Nonparametric Regression, *ITS Journal of Science and Arts*, Vol.3, No.1, 2014, pp. D30-D35.

[14] Hair, J. F., Ringle, C. M., & Sarstedt, M., PLS-SEM: Indeed a silver bullet, *Journal of Marketing theory and Practice*, Vol.19, No.2, 2011, pp. 139-152.

[15] Zeithaml, V. A., Bitner, M. J., & Gremler, D. D., *Service marketing (5th ed.)*, The McGraw-Hill Companies Inc, 2009.

[16] Kotler, *Marketing Management in Indonesia. First Edition*, Salemba Empat Publisher, 2011.

[17] Kaharu, D., & Budiarti, A., Effect of Lifestyle, Promotion, and Product Quality on Purchase Decisions at Cosmic, *Journal of Management Science and Research (JIRM)*, Vol.5, No.3, 2016.

[18] Permadi, T., Nasir, A., & Anisma, Y., Study of willingness to pay taxes on individual taxpayers who do independent work (case in KPP Pratama Tampan Pekanbaru), *Journal of Economics*, Vol.21, No.2, 2013.

[19] Kadir, A., Factors that affect the timeliness of financial reporting empirical studies on manufacturing companies on the Jakarta stock exchange, *Journal of Management and Accounting*, Vol.12, No.1, 2011.

[20] Ukago, K., Factors Affecting Timeliness of Financial Reporting: Empirical Evidence on the Jakarta Stock Exchange, *Maksi Journal*, Vol.5, No.1, 2005, pp. 13-33.