The Nonparametric Path Function Estimation of Fourier Series at Low Oscillations for Modeling Timely Paying Credit

Sudarshan Bhalshankar

October 13, 2021

Sudarshan Dilip Bhalshankar.
Mechanical Engineering, CAD/CAM.
Sudarshanbhalshankar111@gmail.com

Nonparametric Path Function Estimation of Fourier Series at Low Oscillations for Modeling Timely Paying Credit

Abstract:

Purpose: This research aims to estimate the nonparametric path function of the Fourier series and to describe the lemma and theorem for the analysis of the nonparametric path of the Fourier series at low oscillation levels (K=2,3,4,5).

Method: The analytical method used is a Fourier series nonparametric path analysis with a low level of oscillation. Primary data is obtained from customers at a Bank (Bank X) in Indonesia. The data is in the form of item scores that are used as the average variable so that the average data scale is obtained which is the data of the relevant latent variable.

Findings: The function estimation in nonparametric path analysis using the Fourier series approach is $\hat{\alpha}(\lambda)=\left(n^{-1}X^{'}X+\lambda D\right)^{-1}n^{-1}X^{'}y$. The best nonparametric path model that can describe the 5C variable on Time to Pay through Willingness to Pay is when the oscillation K=4 with R2 is 78%.

Originality: This study applies the Fourier series approach to path analysis in modeling on time to pay credit in the banking sector

Keywords: Path analysis, Fourier series, on-time pay, willingness to pay

1. Introduction

Statistics is a method and science about collecting, processing, presenting, analyzing data, and how to draw general and informative conclusions with a series of procedures both descriptively and inferentially. According to Gujarati and Porter (2012), regression analysis is related to the study of the dependence of one variable, namely the dependent variable, on one or more other variables, namely the independent variable, to estimate the average value (population) of the dependent variable from the value of the dependent variable. known or fixed value of the independent variable (in repeated sampling). Regression analysis is used if you want to know whether the independent variable has a direct influence on the dependent variable. The regression analysis approach can be done in three ways, namely parametric approach, semiparametric approach, and nonparametric approach.

Financial support from financial institutions is very much needed in the context of business development and welfare. The provision of credit can provide convenience for customers in meeting their daily needs. Bank is one of the financial institutions that provide credit products. There are various credit options offered to customers. One of the credit products offered by banks is Home Ownership Loans (KPR). The success of the bank in managing credit is the key to the success of the bank in its operations. If there is congestion in the credit management process, the bank will face problems, one of which is reducing the bank's operating income. The way to overcome the risk of bad credit is that customers must pay credit on time. The 5C principle (Character, Capacity, Capital, Collateral, and Condition) can be used by the bank to decide whether the customer is able to pay the loan on time or not.

Path analysis was first developed by Wright in 1934 (Fernandes et al., 2017). Path analysis is used to test the model of the relationship between variables in the form of cause and effect (Solimun, 2002). Path analysis is a technique that can be used to determine whether there is a causal relationship between exogenous variables and endogenous variables. Path analysis is not only used to determine the direct effect of exogenous variables on endogenous variables, but also explains whether or not there is an indirect effect given by exogenous variables on endogenous variables through mediating endogenous variables.

Path analysis is divided into two, namely parametric-based path analysis and nonparametric-based path analysis. Parametric-based path analysis can be used if the linearity assumption is met, while if the linearity assumption is not met, two possible analyzes can be used, namely nonlinear and/or nonparametric analysis. Solimun (2010) mentions six assumptions that underlie path analysis, namely (1) the relationship between variables is linear and additive, (2) the remainder is normally distributed, (3) the pattern of the relationship between variables is recursive, (4) minimum endogenous variables in the measuring scale. intervals, (5) research variables were measured without error, and (6) the analyzed model was specified based on relevant theories and concepts. The assumption that can make the model change is the assumption of linearity. The assumption of linearity influences the shape of the model. If the linearity assumption is met then the path analysis is parametric, but if the linearity assumption is not met there are 2 possibilities, nonlinear path analysis is used when the nonlinear form is known, but if the nonlinear form is unknown and there is no information about the data pattern, then nonparametric path analysis is used. The relationship between variables can be known using the linearity test, one of which is the Regression Specification Error Test (RESET) method.

Research involving the selection of oscillation parameters in nonparametric regression analysis of the Fourier Series was carried out by Nurjanah et al. (2015) using a large value oscillation parameter, which is between 60 to 99. The results show that an oscillation of 70 is able to provide a high enough R2 compared to an oscillation of 99 so that many parameters must be estimated as many as 72 parameters. Research conducted by Wisisono et al. (2018) using oscillations with a value between 1 to 18 shows that an oscillation with a value of 16 has been able to provide a fairly high R2 compared to an oscillation of a value of 18 so that many parameters must be estimated as many as 16 parameters. Research conducted by Soliha et al. (2018) using oscillation parameters with values between 3 to 6 shows that oscillations with a value of 3 have been able to provide a high enough R2 compared to oscillations with a value of 6 so that many parameters must be estimated as many as 8 parameters. Based on some of these studies, it can be concluded that a high value oscillation level does not always give a high R2 as well. So that in this study we will compare the oscillation levels of 2, 3, 4 and 5. The choice of the oscillation level to be compared is based on the number of parameters that must be estimated if the oscillation parameter is too large then, in the research conducted by Soliha et al, (2018) proves that the oscillation parameter of less than 10 already gives a fairly high R^2 value.

The Fourier series is a trigonometric polynomial that has flexible properties, so the model can adapt effectively to the local properties of the data. Research on nonparametric regression modeling of the Fourier Series has been carried out by Prahutama (2013) and Sholiha et al. (2018). Prahutama (2013) conducted a study on nonparametric regression analysis of the Fourier Series to analyze the open unemployment rate in East Java. then Sholiha et al. (2018) conducted research on nonparametric regression analysis of the sine and cosine-based Fourier Series in modeling the sales planning of typical Madurese snacks. The Fourier series has the advantage that it is able to overcome data that has a trigonometric distribution, namely sine and cosine (Prahutama, 2013) which can show periodic functions in general (Wisisono et al., 2018). Periodic means that a situation occurs at a fixed time interval (Nurjanah et al., 2015). The data pattern that is suitable for the Fourier Series approach is a repeating data pattern, repetition of the value of endogenous variables for different exogenous variables (Prahutama, 2013). Based on the explanation above, the aim of this research is to estimate the nonparametric path function of the Fourier series and to describe the lemma and theorem for the analysis of the nonparametric path of the Fourier series at low oscillation levels.

2. Literature Review

2.1 Parametric Regression Analysis

Regression analysis according to Kutner et al. (2005) is a statistical methodology that utilizes the relationship between two or more quantitative variables so that the response variable or outcome can be predicted from other variables. Simple linear regression analysis can be used when one predictor variable is used to predict the response variable. The simple linear regression model can be expressed as follows:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \qquad (2.1)$$

with:

Y_i : the value of the response variable on the i-th observation
β_0 : intercept parameter
β_1 : slope parameter
X_i : the value of the predictor variable on the i-th observation
ε_i : error on observation i

If there is more than one predictor variable, multiple linear regression analysis is used. According to Kutner et al. (2005) multiple linear regression analysis is one of the most widely used statistical methods when there is more than one predictor variable used to predict the response variable. Multiple linear regression models can be expressed as follows:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \varepsilon_i \qquad (2.2)$$

The general linear regression model can be expressed as follows:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_p X_{ip-1} + \varepsilon_i \qquad (2.3)$$

with:

Y_i : the value of the response variable on the i-th observation
β_0 : intercept parameter
$\beta_1, \beta_2, \ldots, \beta_p$: slope parameters
$X_{i1}, \ldots, X_{ip-1}$: nilai variabel prediktor dalam pengamatan ke-i.
i : 1, 2, ..., n.
n : number of observations
ε_i : error on observation i

Solving the problem of parameter estimation in multiple linear regression analysis that has more than two predictor variables can be solved by the matrix method. Equation 2.3 is a general equation of the population multiple linear regression model with the number of predictor variables as many as p-1 pieces. If there are n observations and p predictor variables, the regression equation can be written as follows:

$$\begin{aligned}
Y_1 &= \beta_0 + \beta_1 X_{11} + \beta_2 X_{12} + \cdots + \beta_p X_{1p-1} + \varepsilon_1 \\
Y_2 &= \beta_0 + \beta_1 X_{21} + \beta_2 X_{22} + \cdots + \beta_p X_{2p-1} + \varepsilon_2 \\
Y_3 &= \beta_0 + \beta_1 X_{31} + \beta_2 X_{32} + \cdots + \beta_p X_{3p-1} + \varepsilon_3 \qquad (2.4) \\
\vdots \quad & \quad \vdots \quad \vdots \qquad \vdots \qquad\qquad \vdots \qquad \vdots \\
Y_n &= \beta_0 + \beta_1 X_{n1} + \beta_2 X_{n2} + \cdots + \beta_p X_{np-1} + \varepsilon_n
\end{aligned}$$

From the above equation, it can be written in matrix form as follows:

$$\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \vdots \\ Y_n \end{pmatrix} = \begin{pmatrix} 1 & X_{11} & X_{12} & \cdots & X_{1p-1} \\ 1 & X_{21} & X_{22} & \cdots & X_{2p-1} \\ 1 & X_{31} & X_{32} & \cdots & X_{3p-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots & X_{np-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_n \end{pmatrix} \qquad (2.5)$$

The linear regression model in matrix form can be written as follows:

$$\underset{\sim}{Y} = \mathbf{X}\underset{\sim}{\beta} + \underset{\sim}{\varepsilon} \quad (2.6)$$

Parameter estimation in the parametric regression model with a matrix approach is done by minimizing the number of squares of errors

$$\underset{\sim}{\varepsilon} = \underset{\sim}{Y} - \mathbf{X}\underset{\sim}{\beta} \quad (2.7)$$

$$\underset{\sim}{\varepsilon}'\underset{\sim}{\varepsilon} = \left(\underset{\sim}{Y} - \mathbf{X}\underset{\sim}{\beta}\right)'\left(\underset{\sim}{Y} - \mathbf{X}\underset{\sim}{\beta}\right)$$

Then it is derived to the parameter, namely β and equates to zero

$$\underset{\sim}{\varepsilon}'\underset{\sim}{\varepsilon} = \underset{\sim}{Y}'\underset{\sim}{Y} - \underset{\sim}{Y}'\underset{\sim}{\beta}\mathbf{X} - \underset{\sim}{\beta}'\mathbf{X}'Y + \underset{\sim}{\beta}'\mathbf{X}'\mathbf{X}\underset{\sim}{\beta}$$

$$\underset{\sim}{\varepsilon}'\underset{\sim}{\varepsilon} = \underset{\sim}{Y}'\underset{\sim}{Y} - 2\underset{\sim}{\beta}'\mathbf{X}'Y + \left(\mathbf{X}'\mathbf{X}\right)\underset{\sim}{\beta}^2$$

$$\frac{\partial\underset{\sim}{\varepsilon}'\underset{\sim}{\varepsilon}}{\partial\beta} = \frac{\partial\underset{\sim}{Y}'\underset{\sim}{Y} - 2\underset{\sim}{\beta}'\mathbf{X}'\underset{\sim}{Y} + \left(\mathbf{X}'\mathbf{X}\right)\underset{\sim}{\beta}^2}{\partial\beta}$$

$$\frac{\partial\underset{\sim}{\varepsilon}'\underset{\sim}{\varepsilon}}{\partial\beta} = 0 - 2\mathbf{X}'\underset{\sim}{Y} + 2\left(\mathbf{X}'\mathbf{X}\right)\underset{\sim}{\beta}$$

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

$$2\left(\mathbf{X}'\mathbf{X}\right)\underset{\sim}{\hat{\beta}} = 2\mathbf{X}'\underset{\sim}{Y}$$

$$\left(\mathbf{X}'\mathbf{X}\right)\underset{\sim}{\hat{\beta}} = \mathbf{X}'\underset{\sim}{Y}$$

$$\left(\mathbf{X}'\mathbf{X}\right)^{-1}\left(\mathbf{X}'\mathbf{X}\right)\underset{\sim}{\hat{\beta}} = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\underset{\sim}{Y}$$

$$\mathbf{I}\underset{\sim}{\hat{\beta}} = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\underset{\sim}{Y}$$

So that the parameter estimation for multiple linear regression with a matrix approach is as follows:

$$\underset{\sim}{\hat{\beta}} = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\underset{\sim}{Y} \quad (2.8)$$

2.2 Nonparametric Regression Analysis

Nonparametric regression is a very flexible regression in modeling data patterns so that the subjectivity of the researcher can be minimized (Soliha et al., 2018). Nonparametric regression analysis is used if the classical assumptions in parametric regression analysis (assumptions of normality, nonmulticollinearity, and homoscedasticity) are not met. This method is best suited for concluding situations where there is little or no prior information available about the regression curve or data pattern (Eubank, 1999).

The use of parametric regression can pose a risk if it is forced on data whose data pattern is not known, that is, it can create an unrepresentative regression model so that decisions made through hypothesis testing are inaccurate. To find out the pattern of the relationship between the predictor variable (X) and the response variable (Y) whose curve shape is not yet known, a nonparametric regression model can be used as follows (Soliha et al., 2018):

$$Y_i = \hat{f}\left(x_i\right) + \varepsilon_i \quad (2.9)$$

If the linearity assumption is met, then the analysis is continued by using parametric path analysis by fulfilling the following assumptions:

$$\varepsilon_i \sim \mathrm{N}(0, \sigma^2)$$

But if the assumption of linearity is not met then the analysis uses a nonlinear and or nonparametric path.

dimana:

Y_i : the value of the response variable on the i-th observation

x_i : the value of the predictor variable in the i-th observation.

$\hat{f}$: regression curve.

i : 1, 2, ..., n.
n : number of observations
ε_i : error on observation i

2.3 Path Analysis

Path analysis was first developed in 1934 by a geneticist, Sewall Wright. Path analysis is a technique for estimating the effect of a set of independent variables on the dependent variable from a series of observed correlations. The purpose of path analysis is to measure the direct effect on each separate path in the system thereby finding out the extent to which the variation of a given effect can be determined by each cause. Solimun (2010) describes six assumptions that underlie path analysis, namely:

1) The relationship between variables is linear and additive. The assumption of linearity can be checked with a scatter plot, but the results will be subjective. Another way of checking the assumption of linearity is with the Regression Specification Error Test (RESET) introduced by Ramsey in 1969.
2) The error is normally distributed (remaining normality). The method for testing the normality of the residuals is the Kolmogorov-Smirnov. According to Widarjono (2005), the test for the effect of the predictor variable on the response variable is valid if the residuals obtained have a normal distribution.
3) The pattern of relationship between variables is recursive (one-way causal flow system). The characteristics of the recursive model are:
   a. Between ε_i are mutually free.
   b. Betwee ε_i and X_i are mutually free.
4) Minimum endogenous variable in interval measuring scale.
5) Research variables were measured without error (valid and reliable research instrument).
6) The analyzed model is specified based on the relevant theories and concepts.

The assumption that can make the model change is the assumption of linearity. The assumption of linearity has an influence on the shape of the model. If the linearity assumption is met, then the path analysis is parametric, but if the linearity assumption is not met there are 2 possibilities, if the nonlinear form is known, then use nonlinear path analysis, if the nonlinear form is unknown and there is no information about the data pattern, then use nonparametric path analysis.

2.3 Nonparametric Path Analysis Fourier Series

One approach that can be used in nonparametric path analysis is Fourier series path analysis. The Fourier series is a trigonometric polynomial that has flexibility, so it can adapt effectively to the local nature of the data (Wisisono et al., 2018). The Fourier series has the advantage that it is able to overcome data that has a trigonometric distribution (sine and cosine) (Prahutama, 2013). The data pattern that is suitable for the Fourier Series approach is a repeating data pattern, repetition of the value of the dependent variable for different independent variables. The estimator for the nonparametric path function of the Fourier series is as follows (Tripena, 2009):

$$\hat{f}_\lambda(x_i) = \hat{b}(\lambda)x_i + \frac{1}{2}\hat{a}_0(\lambda) + \sum_{k=1}^{K}\hat{a}_k(\lambda)\cos kx_i \qquad (2.10$$

3. Methodology

The data used in this study is primary data. Primary data was obtained from one of the Banks (Bank X) in Indonesia with a Likert scale with a second order measurement model. The seizure data consisted of 5 exogenous variables, 1 intervening endogenous variable, and 1 pure endogenous variable. Measurement of variables using the average score of each item. This method uses the average scale of all indicators on each variable so that the average scale data is obtained which is the data of the relevant latent variable.

4. Result And Discusion

4.1. The Lemma and Theorem of Nonparametric Path Analysis Fourier Series

Lemma 4.1 Form of Fourier Series Nonparametric Path Model at Oscillation Rate = 2

If given paired data $(X_{1i}, X_{2i}, Y_{1i}, Y_{2i})$ with $i = 1,2,\dots,n$; which follows the non-parametric path analysis model, the form of the nonparametric path analysis function is obtained as presented in equation (4.1)

$$\hat{\underset{\sim}{f}} = \mathbf{X}\underset{\sim}{\alpha} \tag{4.1}$$

If the function is described for each response variable, it is presented in equation (4.2)

$$\begin{aligned} Y_{1i} &= f_1(X_{1i}, X_{2i}) + \varepsilon_{1i} \\ Y_{2i} &= f_2(X_{1i}, X_{2i}, Y_{1i}) + \varepsilon_{2i} \end{aligned} \tag{4.2}$$

The equation model is as follows

$$\begin{aligned} \hat{f}_{1i} &= \frac{1}{2}a_{01} + b_{11}X_{1i} + \gamma_{11}\cos X_{1i} + \gamma_{21}\cos 2X_{1i} + b_{21}X_{2i} + \gamma_{31}\cos X_{2i} + \gamma_{41}\cos 2X_{2i} \\ \hat{f}_{2i} &= \frac{1}{2}a_{02} + b_{12}X_{1i} + \gamma_{12}\cos X_{1i} + \gamma_{22}\cos 2X_{1i} + b_{22}X_{2i} + \gamma_{32}\cos X_{2i} + \gamma_{42}\cos 2X_{2i} \\ &+ b_{32}Y_{1i} + \gamma_{52}\cos \mathrm{Y}_{1i} + \gamma_{62}\cos 2Y_{1i} \end{aligned} \tag{4.3}$$

With the form of the matrix X is as follows

$$\mathbf{X} = \begin{pmatrix} \frac{1}{2} & x_{11} & \cos x_{11} & \cdots & \cos 2x_{21} & 0 & 0 & 0 & \cdots & 0 \\ \frac{1}{2} & x_{12} & \cos x_{12} & \cdots & \cos 2x_{22} & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{2} & x_{1n} & \cos x_{1n} & \cdots & \cos 2x_{2n} & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 & \frac{1}{2} & x_{11} & \cos x_{11} & \cdots & \cos 2y_{11} \\ 0 & 0 & 0 & \cdots & 0 & \frac{1}{2} & x_{12} & \cos x_{12} & \cdots & \cos 2y_{12} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & \frac{1}{2} & x_{1n} & \cos x_{1n} & \cdots & \cos 2y_{1n} \end{pmatrix} \tag{4.4}$$

And $\underset{\sim}{a}$ are as follows

$$\underset{\sim}{a}^T = \begin{pmatrix} \frac{1}{2}a_{01} & b_{11} & \cdots & \gamma_{41} & \frac{1}{2}a_{02} & b_{12} & \cdots & \gamma_{62} \end{pmatrix} \tag{4.5}$$

So the dimension of the matrix for nonparametric path analysis of the series at the oscillation level = 2 is

$$\underset{\sim}{f}_{2nx1} = \boldsymbol{X}_{2nx17}\alpha_{17x1} \tag{4.6}$$

where:

$\underset{\sim}{f}\left(X_{ij}\right)$: Vector nonparametric regression function of the i-th observation nonparametric of the j-th exogenous variable

$\boldsymbol{X}_{ij}$: The j-th exogenous variable matrix on the i-th observation

$\underset{\sim}{\alpha}_{ij}$: Parameter vector of the j-th observation of the i-th exogenous variable

Proof

Before obtaining a model for the nonparametric path analysis of the Fourier series, the process was first obtained from (a) multiple linear regression analysis; (b) simple linear path analysis; and (c) nonparametric regression analysis as follows:

First part:

The simple linear regression model can be expressed as follows:

$$Y_i = \beta_0 + \beta_i X_i + \varepsilon_i \tag{4.7}$$

where:
Y_i : the value of the response variable on the i-th observation
β_0 : intercept parameter
β_1 : slope parameter
X_i : the value of the predictor variable on the i-th observation
ε_i : error on observation i

If there is more than one predictor variable, multiple linear regression analysis is used. Multiple linear regression model can be expressed in equation (4.8)

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \varepsilon_i \tag{4.8}$$

The general linear regression model can be expressed in equation (4.9)

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_p X_{ip-1} + \varepsilon_i \tag{4.9}$$

where:
Y_i : the value of the response variable on the i-th observation
β_0 : intercept parameter
$\beta_1, \beta_2, \dots, \beta_p$: slope parameters
$X_{i1}, \dots, X_{ip-1}$: the value of the predictor variable in the i-th observation.
i : 1, 2, ..., n.
n : number of observations
ε_i : error on observation i

Solving the problem of parameter estimation in multiple linear regression analysis that has more than two predictor variables can be solved by the matrix method. Equation 4.9 is a general equation of the population multiple linear regression model with the number of predictor variables as many as p-1 pieces. If there are n observations and p predictor variables, the regression equation can be written as follows:

$$\begin{aligned}
Y_1 &= \beta_0 + \beta_1 X_{11} + \beta_2 X_{12} + \cdots + \beta_p X_{1p-1} + \varepsilon_1 \\
Y_2 &= \beta_0 + \beta_1 X_{21} + \beta_2 X_{22} + \cdots + \beta_p X_{2p-1} + \varepsilon_2 \\
Y_3 &= \beta_0 + \beta_1 X_{31} + \beta_2 X_{32} + \cdots + \beta_p X_{3p-1} + \varepsilon_3 \\
&\vdots \quad \vdots \quad \vdots \qquad \vdots \qquad \vdots \quad \vdots \\
Y_n &= \beta_0 + \beta_1 X_{n1} + \beta_2 X_{n2} + \cdots + \beta_p X_{np-1} + \varepsilon_n
\end{aligned} \tag{4.10}$$

From the above equation, it can be written in matrix form as follows:

$$\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \vdots \\ Y_n \end{pmatrix} = \begin{pmatrix} 1 & X_{11} & X_{12} & \cdots & X_{1p-1} \\ 1 & X_{21} & X_{22} & \cdots & X_{2p-1} \\ 1 & X_{31} & X_{32} & \cdots & X_{3p-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots & X_{np-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_n \end{pmatrix} \tag{4.11}$$

The second part:

It is known that the simple path analysis model is presented in equation (4.5) and the model in (4.6)

$$
\begin{aligned}
Y_{1i} &= f_1(X_{1i}, X_{2i}) + \varepsilon_{1i} \\
Y_{2i} &= f_2(X_{1i}, X_{2i}, Y_{1i}) + \varepsilon_{2i}
\end{aligned}
\tag{4.12}
$$

$$
\begin{aligned}
Y_{1i} &= \beta_{10} + \beta_{11}X_1 + \beta_{12}X_2 + \varepsilon_{1i} \\
Y_{2i} &= \beta_{20} + \beta_{21}X_1 + \beta_{22}X_2 + \beta_{23}Y_1 + \varepsilon_{2i}
\end{aligned}
\tag{4.13}
$$

With matrix form:

$$\underset{\sim}{Y}_{2nx1} = \boldsymbol{X}_{2nx7}\underset{\sim}{\beta}_{7x1} + \underset{\sim}{\varepsilon}_{2nx1} \tag{4.14}$$

$$
\begin{bmatrix} Y_{11} \\ Y_{12} \\ \vdots \\ Y_{1n} \\ Y_{21} \\ Y_{22} \\ \vdots \\ Y_{2n} \end{bmatrix}
=
\begin{bmatrix} X_X & \underset{\sim}{0}_{nx4} \\ \underset{\sim}{0}_{nx3} & X_{XY} \end{bmatrix}
\begin{bmatrix} \beta_{10} \\ \beta_{11} \\ \beta_{12} \\ \beta_{20} \\ \beta_{21} \\ \beta_{22} \\ \beta_{23} \end{bmatrix}
+
\begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \vdots \\ \varepsilon_{1n} \\ \varepsilon_{21} \\ \varepsilon_{22} \\ \vdots \\ \varepsilon_{2n} \end{bmatrix}
$$

where

$$
X_X = \begin{bmatrix} 1 & X_{11} & X_{21} \\ 1 & X_{12} & X_{22} \\ & \vdots & \\ 1 & X_{1n} & X_{2n} \end{bmatrix};\;
X_{XY} = \begin{bmatrix} 1 & X_{11} & X_{21} & Y_{11} \\ 1 & X_{12} & X_{22} & Y_{12} \\ & & \vdots & \\ 1 & X_{1n} & X_{2n} & Y_{1n} \end{bmatrix}
$$

With:

Y_{hi} : h-th endogenous variable and i-th observation

X_i : the value of the predictor variable on the i-th observation

β : predictor variable parameters

ε_{hi} : random error of the h-th endogenous variable, i-th observation

Third Part:

After knowing the equations and multiple linear regression models, nonparametric regression models can be made as presented in equations (4.15) and (4.16).

$$Y_{1i} = f_1(X_{1i}, X_{2i}) + \varepsilon_{1i} \tag{4.15}$$

$$\hat{f}_\lambda(x_i) = \hat{b}(\lambda)x_i + \frac{1}{2}\hat{a}_0(\lambda) + \sum_{k=1}^{2}\hat{a}_k(\lambda)\cos kx_i \tag{4.16}$$

With equations and matrix forms such as the following equation:

$$\underset{\sim}{f}_{2nx1} = \boldsymbol{X}_{2nx7}\alpha_{7x1} \tag{4.17}$$

$$
\begin{pmatrix} y_1 \\ y_2 \\ y_1 \\ \vdots \\ y_M \end{pmatrix}
=
\begin{pmatrix}
x_{11} & 1 & \cos x_{11} & \cos 2x_{11} & x_{21} & \cos x_{21} & \cos 2x_{21} \\
x_{12} & 1 & \cos x_{12} & \cos 2x_{12} & x_{22} & \cos x_{22} & \cos 2x_{22} \\
x_{13} & 1 & \cos x_{13} & \cos 2x_{13} & x_{23} & \cos x_{23} & \cos 2x_{23} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
x_{1n} & 1 & \cos x_{1n} & \cos 2x_{1n} & x_{2n} & \cos x_{2n} & \cos 2x_{2n}
\end{pmatrix}
\begin{pmatrix} b \\ \frac{1}{2}a_0 \\ a_1 \\ \vdots \\ a_7 \end{pmatrix}
\tag{4.18}
$$

where:

$\underset{\sim}{f}\left(X_{ij}\right)$: Vector nonparametric regression function of the i-th observation nonparametric of the j-th exogenous variable

$\boldsymbol{X}_{ij}$: The j-th exogenous variable matrix on the i-th observation

$\underset{\sim}{\alpha}_{ij}$: Parameter vector of the j-th observation of the i-th exogenous variable

From the equations in the simple linear regression analysis model, simple path analysis, and nonparametric regression analysis that have been described, it can be obtained a function that is formed as in equations (4.1) and (4.2), so that the following matrix is obtained:

$$\underset{\sim}{f}_{2nx1} = \boldsymbol{X}_{2nx17}\alpha_{17x1} \tag{4.19}$$

$$\mathbf{X} = \begin{pmatrix} \frac{1}{2} & x_{11} & \cos x_{11} & \cdots & \cos 2x_{21} & 0 & 0 & 0 & \cdots & 0 \\ \frac{1}{2} & x_{12} & \cos x_{12} & \cdots & \cos 2x_{22} & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{2} & x_{1n} & \cos x_{1n} & \cdots & \cos 2x_{2n} & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 & \frac{1}{2} & x_{11} & \cos x_{11} & \cdots & \cos 2y_{11} \\ 0 & 0 & 0 & \cdots & 0 & \frac{1}{2} & x_{12} & \cos x_{12} & \cdots & \cos 2y_{12} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & \frac{1}{2} & x_{1n} & \cos x_{1n} & \cdots & \cos 2y_{1n} \end{pmatrix} \tag{4.20}$$

where:

$\underset{\sim}{f}\left(X_{ij}\right)$: Vector nonparametric regression function of the i-th observation nonparametric of the j-th exogenous variable

$\boldsymbol{X}_{ij}$: The j-th exogenous variable matrix on the i-th observation

$\underset{\sim}{\alpha}_{ij}$: Parameter vector of the j-th observation of the i-th exogenous variable

Theorem 4.1

If the data is given following the nonparametric flow analysis model in the cross-section data as presented in Lemma 4.1, then the parameter estimation method that can minimize the number of squares of error is the ordinary least square method. Then the estimator of the Fourier series by minimizing $\underset{\sim}{\varepsilon}'\underset{\sim}{\varepsilon}$ is:

$$\underset{\beta \in R^{[(K+1)*(P_1+P_2)]+2}}{Min}(\underset{\sim}{\varepsilon}'\underset{\sim}{\varepsilon}) = \underset{\beta \in R^{[(K+1)*(P_1+P_2)]+2}}{Min}\left(\underset{\sim}{Y} - \mathbf{X}\underset{\sim}{\alpha}\right)'\left(\underset{\sim}{Y} - \mathbf{X}\underset{\sim}{\alpha}\right) \tag{4.21}$$

with $\hat{\underset{\sim}{f}} = \mathbf{X}\underset{\sim}{\alpha}$

$$\mathbf{X}=\begin{pmatrix} \frac{1}{2} & x_{11} & \cos x_{11} & \cdots & \cos 2x_{21} & 0 & 0 & 0 & \cdots & 0 \\ \frac{1}{2} & x_{12} & \cos x_{12} & \cdots & \cos 2x_{22} & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{2} & x_{1n} & \cos x_{1n} & \cdots & \cos 2x_{2n} & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 & \frac{1}{2} & x_{11} & \cos x_{11} & \cdots & \cos 2y_{11} \\ 0 & 0 & 0 & \cdots & 0 & \frac{1}{2} & x_{12} & \cos x_{12} & \cdots & \cos 2y_{12} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & \frac{1}{2} & x_{1n} & \cos x_{1n} & \cdots & \cos 2y_{1n} \end{pmatrix}$$

$$\hat{\alpha}(\lambda)=\left(n^{-1}X^{'}X+\lambda D\right)^{-1}n^{-1}X^{'}y$$

Proof:

Based on the Nonparametric Regression model, $f(x_i)$ is approached by the Fourier series as follows:

Minimize ε_i^2

$$Min\{\sum_{i=1}^{n}\varepsilon_i^2\}=Min\{\sum_{i=1}^{n}\left(y_i-f\left(x_i\right)\right)^2\} \tag{4.22}$$

In addition to minimizing equation (2.19), a penalty is also given for the size of the smoothness of the function f as follows:

$$\int_0^{\pi}\frac{2}{\pi}(\mathrm{f}^{(2)}(\mathrm{x}))^2\,dx \tag{4.23}$$

Thus the estimator for the regression curve f can be obtained from completing the optimization using Penalized Least Square (PLS)

$$Min\{n^{-1}\sum_{i=1}^{n}\left(y_i-f\left(x_i\right)\right)^2+\lambda\int_0^{\pi}\frac{2}{\pi}\left(f^{(2)}\left(x\right)\right)^2dx\} \tag{4.24}$$

To solve equation (2.24), first find the value of P(a) in the following way:

$$P(a)=\int_0^{\pi}\frac{2}{\pi}(\mathrm{f}^{(2)}(\mathrm{x}))^2\,dx$$

$$P(a)=\frac{2}{\pi}\int_0^{\pi}\left(\sum_{k=1}^{K}\left(k^2a_k\cos kx\right)^2+2\sum_{k<j}^{K}\sum_{k<j}^{K}\left(k^2a_k\cos kx\right)\left(j^2a_j\,\mathrm{cosj}\,x\right)\right)dx$$

$$P(a)=\sum_{k=1}^{K}k^4a_k^2 \tag{4.25}$$

$$P(a)=\int_0^{\pi}\frac{2}{\pi}\left(\mathrm{f}^{(2)}(\mathrm{x})\right)^2dx$$

$$=\int_0^{\pi}\frac{2}{\pi}\left(\frac{d^2}{dx^2}\left(bx+\frac{1}{2}a_0+\sum_{k=1}^{K}a_k\cos kx\right)\right)^2dx$$

$$=\int_0^{\pi}\frac{2}{\pi}\left(\frac{d}{dx}\left(\frac{d}{dx}\left(bx+\frac{1}{2}a_0+\sum_{k=1}^{K}a_k\cos kx\right)\right)\right)^2dx$$

$$=\int_0^{\pi}\frac{2}{\pi}\left(\frac{d}{dx}\left(b+0-\sum_{k=1}^{K}a_k\,\mathrm{k}\sin kx\right)\right)^2 dx$$

$$=\int_0^{\pi}\frac{2}{\pi}\left(\frac{d}{dx}\left(b-\sum_{k=1}^{K}a_k\,\mathrm{k}\sin kx\right)\right)^2 dx$$

$$=\int_0^{\pi}\frac{2}{\pi}\left(0-\sum_{k=1}^{K}ka_k\,\mathrm{k}\cos kx\right)^2 dx$$

$$=\int_0^{\pi}\frac{2}{\pi}\left(\sum_{k=1}^{K}ka_k\,\mathrm{k}\cos kx\right)^2 dx$$

$$=\frac{2}{\pi}\int_0^{\pi}\left(\sum_{k=1}^{K}\left(ka_k\,\mathrm{k}\cos kx\right)^2+2\sum_{k<j}^{K}\sum_{k<j}^{K}\left(k^2a_k\cos kx\right)\left(j^2a_j\,\mathrm{cosj}\,x\right)\right)dx$$

$$=\frac{2}{\pi}\int_0^{\pi}\sum_{k=1}^{K}\left(k^2a_k\cos kx\right)^2 dx+\frac{2}{\pi}\int_0^{\pi}2\sum_{k<j}^{K}\left(k^2a_k\cos kx\right)\left(j^2a_j\,\mathrm{cosj}\,x\right)dx \tag{4.26}$$

Suppose $A=\frac{2}{\pi}\int_0^{\pi}\sum_{k=1}^{K}\left(k^2a_k\cos kx\right)^2 dx$

$B=\frac{2}{\pi}\int_0^{\pi}2\sum_{k<j}^{K}\left(k^2a_k\cos kx\right)\left(j^2a_j\,\mathrm{cosj}\,x\right)dx$

a) Calculating the integral of the equation A

$$A=\frac{2}{\pi}\int_0^{\pi}\sum_{k=1}^{K}\left(k^2a_k\cos kx\right)^2 dx$$

$$=\frac{2}{\pi}\sum_{k=1}^{K}\int_0^{\pi}k^4a_k^2\cos^2 kx dx$$

$$=\frac{2}{\pi}\sum_{k=1}^{K}k^4a_k^2\int_0^{\pi}\cos^2 kx dx$$

$$=\frac{2}{\pi}\sum_{k=1}^{K}k^4a_k^2\int_0^{\pi}\frac{1+2\cos kx}{2}dx$$

$$=\frac{2}{\pi}\sum_{k=1}^{K}k^4a_k^2\left(\frac{1}{2}\left[t+\frac{2}{k}\sin kx\right]_0^{\pi}\right)$$

$$=\frac{2}{\pi}\sum_{k=1}^{K}k^4a_k^2\left(\frac{1}{2}\pi+\left(\frac{1}{2}\right)\left(\frac{1}{2k}\right)\sin\left(2k\pi\right)-\sin(0)\right)$$

$$=\sum_{k=1}^{K}k^4a_k^2 \tag{4.27}$$

b) Calculating the integral of the equation B

$$B=\frac{2}{\pi}\int_0^{\pi}2\sum_{k<j}^{K}\left(k^2a_k\cos kx\right)\left(j^2a_j\,\mathrm{cosj}\,x\right)dx$$

$$=\frac{2}{\pi}2\sum_{k<j}^{K}\int_0^{\pi}\left(k^2a_k\,\mathrm{cosk}\,x\right)\left(j^2a_j\,\mathrm{cosj}\,x\right)dx$$

$$= \frac{4}{\pi}\sum_{k<j}^{K}\int_{0}^{\pi} k^2 j^2 a_k a_j \cos kx \cos jx dx$$

$$= \frac{4}{\pi}\sum_{k<j}^{K}(kj)^2 a_k a_j \int_{0}^{\pi} \cos kx \cos jx dx$$

$$= \frac{4}{\pi}\sum_{k<j}^{K}(kj)^2 a_k a_j \int_{0}^{\pi} \frac{\cos(k+j)x+\cos(k-j)x}{2}dx$$

$$= \frac{4}{\pi}\sum_{k<j}^{K}(kj)^2 a_k a_j \int_{0}^{\pi} \frac{1}{2}\cos(k+j)x\, dx + \int_{0}^{\pi}\frac{1}{2}\cos(k-j)x dx$$

$$= \frac{4}{\pi}\sum_{k<j}^{K}(kj)^2 a_k a_j \left(\frac{1}{2}\frac{1}{(k+j)}\left(\sin(k+j)\pi - \sin(0)\right)\right)$$

$$+\left(\frac{1}{2}\frac{1}{(k0j)}\left(\sin(k-j)\pi - \sin(0)\right)\right)$$

$$= \frac{4}{\pi}\sum_{k<j}^{K}(kj)^2 a_k a_j \left(\frac{1}{2(k+j)}(0-0) + \frac{1}{2(k-j)}(0-0)\right)$$

$$= \frac{4}{\pi}\sum_{k<j}^{K}(kj)^2 a_k a_j (0)$$

$$= 0 \qquad (4.28)$$

So that:

$$P(a) = \frac{2}{\pi}\int_{0}^{\pi}\sum_{k=1}^{K}\left(k^2 a_k \cos kx\right)^2 dx + \frac{2}{\pi}\int_{0}^{\pi} 2\sum_{k<j}^{K}\left(k^2 a_k \cos kx\right)\left(j^2 a_j \cos jx\right)dx$$

$$= \sum_{k=1}^{K} k^4 a_k^2 + 0$$

$$= \sum_{k=1}^{K} k^4 a_k^2 \qquad (4.29)$$

Since f is a continuous function, f can be approximated by the function x, with

$$f(x) = bx + \frac{1}{2}a_0 + \sum_{k=1}^{K} a_k \cos kx \qquad (4.30)$$

Based on equation (4,29) it can be written

$$\mathrm{Min}\left\{n^{-1}\sum_{i=1}^{n}\left(y_i - f(x_i)\right)^2 + \lambda\int_0^p \frac{2}{p}\left(f^{(2)}(x)\right)^2 dt\right\}$$

$$= \mathrm{Min}\left\{n^{-1}\sum_{i=1}^{n}\left[y_i - bx - \frac{1}{2}a_0 - \sum_{k=1}^{K} a_K \cos kx\right]^2 + \lambda\sum_{k=1}^{K} k^4 a_k^2\right\}$$

$$= \mathrm{Min}\{n^{-1}(y - \mathbf{Xa})'(y - \mathbf{Xa}) + \lambda \mathbf{a'Da}\}$$

$$= Min\{n^{-1}y' y - n^{-1}a' X' y - n^{-1}(a' X' y)' + a'(n^{-1}X' X + \lambda D)a\} \qquad (4.31)$$

where:

$\mathbf{D} = diag\left(0, 0, 1^4, 2^4, \ldots, K^4\right)$

Suppose that equation (4.31) is called , by subtracting $Q\left(a\right)$ partially with respect to a and equating to zero we get:

$$\frac{\partial Q\left(a\right)}{\partial a} = 0 - 2n^{-1}\boldsymbol{X}'y + 2\left(n^{-1}\boldsymbol{X}'\boldsymbol{X} + \lambda\boldsymbol{D}\right)\boldsymbol{a}$$

$$\hat{\alpha}\left(\lambda\right) = \left(n^{-1}X'X + \lambda D\right)^{-1} n^{-1}X'y \tag{4.32}$$

Based on the properties of the Fourier series estimator, it can be written in matrix form

$$\underset{\sim}{f} = \mathbf{X}\underset{\sim}{a} + \underset{\sim}{e} \tag{4.33}$$

$$\underset{\sim}{\hat{f}} = \mathbf{X}\underset{\sim}{a} \tag{4.34}$$

dimana:

$$\underset{\sim}{a} = \left(b, \frac{1}{2}a_0, a_1, \ldots, a_K\right) \tag{4.35}$$

$$\mathbf{X} = \begin{pmatrix} x_1 & 1 & \cos x_1 & \cos 2x_1 & \cdots & \cos Kx_1 \\ x_2 & 1 & \cos x_2 & \cos 2x_2 & \cdots & \cos Kx_2 \\ x_3 & 1 & \cos x_3 & \cos 2x_3 & \cdots & \cos Kx_3 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ x_n & 1 & \cos x_n & \cos 2x_n & \cdots & \cos Kx_n \end{pmatrix} \tag{4.36}$$

4.2. Assumption of linearity

Path analysis modeling will be carried out on the relationship between the variables Character, Capacity, Capital, Collateral, and Condition on time to pay through the willingness to pay. The path analysis modeling steps begin with a linearity test to determine the form of the relationship between variables. The Ramsey Reset test is used to determine whether the relationship between variables has a linear or non-linear relationship. Table 1 presents the results of the Ramsey Reset test between the variables used in the study.

Table 1. Linearity Assumption Test Results

Research variable			p-value	Conclusion
Character (X1)	→	Willingness to Pay (Y1)	0.909	Linear
Capacity (X2)	→	Willingness to Pay (Y1)	0.001	Not Linear
Capital (X3)	→	Willingness to Pay (Y1)	0.234	Linear
Collateral (X4)	→	Willingness to Pay (Y1)	0.129	Linear
Condition (X5)	→	Willingness to Pay (Y1)	0.029	Not Linear
Character (X1)	→	Obedient Paying Behavior (Y2)	0.130	Linear
Capacity (X2)	→	Obedient Paying Behavior (Y2)	0.092	Linear
Capital (X3)	→	Obedient Paying Behavior (Y2)	0.159	Linear
Collateral (X4)	→	Obedient Paying Behavior (Y2)	0.517	Linear
Condition (X5)	→	Obedient Paying Behavior (Y2)	0.178	Linear
Willingness to Pay (Y1)	→	Obedient Paying Behavior (Y2)	0.033	Not Linear

From the results of the linearity test above, it is known that there are three non-linear relationships between variables, namely the relationship between variables X2 and Y1, X5 and Y1, and Y1 and Y2. Thus, modeling can be done by nonparametric path analysis.

4.4. Comparison of GCV of each lambda on each oscillation

4.4.1. Selection of the Best Model for Each Oscillation

Estimation of nonparametric path function when K=2,3,4,5 is performed on a value of 0.2-0.9. First, λ the optimal value will be selected by looking at the GCV value. The minimum/smallest GCV value will result λ in an optimal value. Table 2 shows the GCV value of each value of λ.

Table 2. Selection of the Best Smoothing Parameters

OSCILLATION	LAMBDA	GCV
2	**0,2**	**89937.560**
	0,3	1560106.000
	0,4	4598515.000
	0,5	664767.600
	0,6	332864.500
	0,7	227785.800
	0,8	178524.700
	0,9	150476.400
3	0,2	94060.860
	0,3	73789.970
	0,4	66028.350
	0,5	61945.170
	0,6	59429.910
	0,7	57725.930
	0,8	56495.590
	0,9	**55565.640**
4	0,2	56588.770
	0,3	53643.340
	0,4	52256.130
	0,5	51449.470
	0,6	50922.030
	0,7	50550.240
	0,8	50274.070
	0,9	**50060.820**
5	0,2	56588.770
	0,3	53643.340

OSCILLATION	LAMBDA	GCV
	0,4	52256.130
	0,5	51449.470
	0,6	50922.030
	0,7	50550.240
	0,8	50274.070
	0,9	**50060.820**

In table 2, it gives the smallest lambda value at each oscillation level. Furthermore, the calculation of the coefficient of determination is carried out to determine which path analysis model is the best at oscillation levels 2, 3, 4, and 5. The coefficient of determination for each oscillation is shown in table 3.

Table 3. Coefficient of Determination Value for Each Oscillation

Oscillation	**Coefficient Of Determination**
2	0.762
3	0.775
4	0.780
5	0.770

Based on Table 3 shows that K = 4 has the largest coefficient of determination, which is 78% so that the best model is at the time of oscillation 4 with a lambda of 0.9 with many parameters that must be estimated as many as 31 and 37. R^2 of 78% indicates that the variance is correct. Time to Pay can be explained by character, customer capacity, capital, collateral, and condition of economy by 78% while the rest is explained by other variables. The following is a nonparametric path analysis function with the best Fourier series approach:

$$
\begin{aligned}
\hat{f}_1 = {} & 2.737 + 0.348x_{1i} + 0.001cosx_{1i} + 0.001cos2x_{1i} - 0.001cos3x_{1i} + 0.001cos4x_{1i} + 0.001cos5x_{1i} \\
& + 0.328x_{2i} + 0.373cosx_{2i} - 0.001cos2x_{2i} + 0.001cos3x_{2i} + 0.001cos4x_{2i} + 0.001cos5x_{2i} + 0.001x_{3i} \\
& - 0.269cosx_{3i} - 0.194cos2x_{3i} + 0.001cos3x_{3i} + 0.001cos4x_{3i} + 0.001cos5x_{3i} + 0.001x_{4i} \\
& + 0.001cosx_{4i} - 0.460cos2x_{4i} - 0.252cos3x_{4i} + 0.001cos4x_{4i} + 0.001cos5x_{4i} + 0.001x_{5i} \\
& + 0.001cosx_{5i} + 0.001cos2x_{5i} + 0.166cos3x_{5i} + 0.015cos4x_{5i} - 0.008cos5x_{5i}
\end{aligned}
$$

$$
\begin{aligned}
\hat{f}_2 = {} & 0.001 + 0.001x_{1i} + 0.001cosx_{1i} + 0.001cos2x_{1i} + 0.159cos3x_{1i} + 0.125cos4x_{1i} + 0.011cos5x_{1i} \\
& + 0.002x_{2i} + 0.001cosx_{2i} + 0.001cos2x_{2i} + 0.001cos3x_{2i} - 0.158cos4x_{2i} - 0.157cos5x_{2i} + 0.030x_{3i} \\
& + 0.001\cos x_{3i} + 0.001cos2x_{3i} + 0.001cos3x_{3i} + 0.001cos4x_{3i} - 0.004cos5x_{3i} + 0.080x_{4i} + 0.001\cos x_{4i} \\
& + 0.001cos2x_{4i} + 0.001cos3x_{4i} + 0.001cos4x_{4i} + 0.001cos5x_{4i} + 0.424x_{5i} - 0.403\cos x_{5i} - 0.014os2x_{5i} \\
& + 0.003cos3x_{5i} - 0.001cos4x_{5i} + 0.001cos5x_{5i} + 0.001y_{1i} - 2.141\cos y_{1i} - 1.259cos2y_{1i} - 0.010cos3y_{1i} \\
& + 0.002cos4y_{1i} - 0.001cos5y_{1i}
\end{aligned}
$$

The actual and predicted data plots on the nonparametric path function of the Fourier Series when the oscillation is equal to 4 can be seen in the figure below:

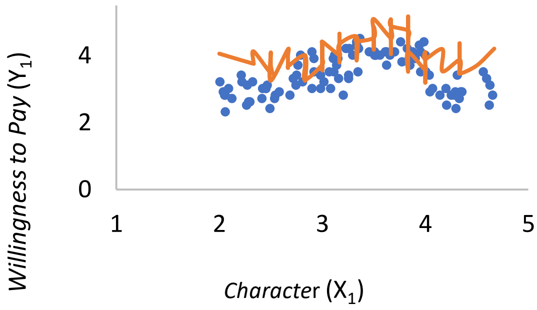

Figure 1. The plot of actual and predicted data between variables X_1 and Y_1

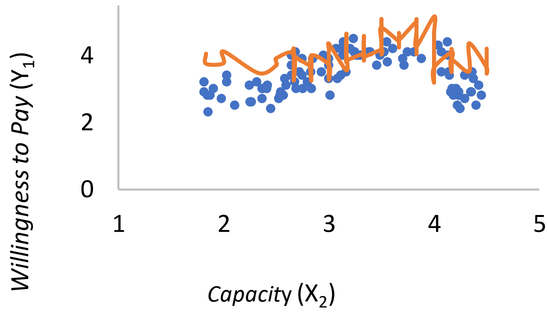

Figure 2. The plot of actual and predicted data between variables X_2 and Y_1

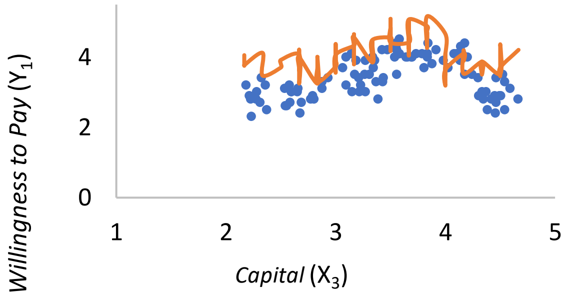

Figure 3. The plot of actual and predicted data between variables X_3 and Y_1

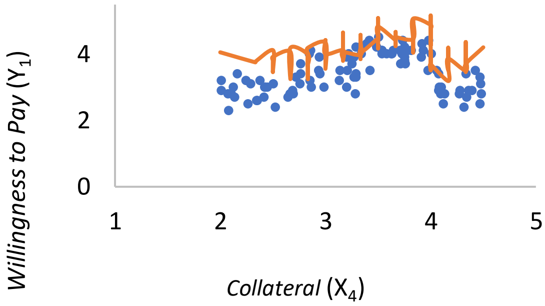

Figure 4. The plot of actual and predicted data between variables X_4 and Y_1

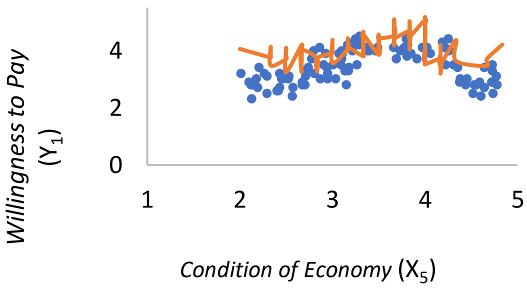

Figure 5. The plot of actual and predicted data between variables X_5 and Y_1

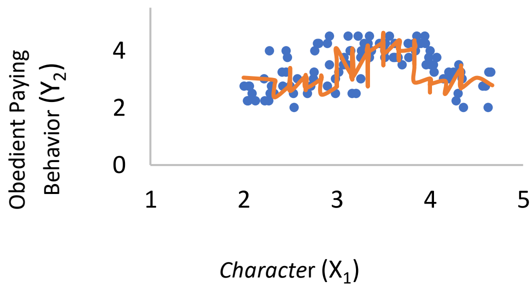

Figure 6. The plot of actual and predicted data between variables X_1 and Y_2

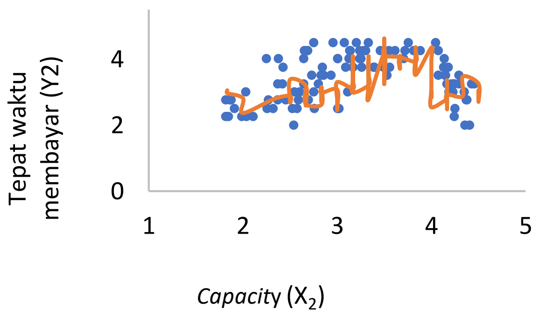

Figure 7. The plot of actual and predicted data between variables X_2 and Y_2

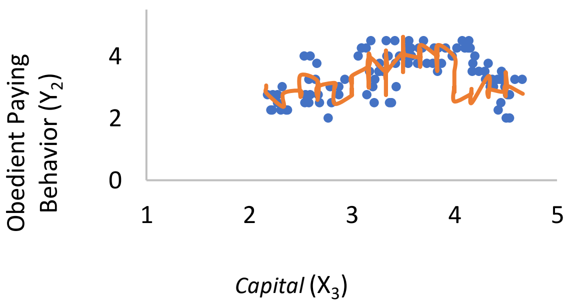

Figure 8. The plot of actual and predicted data between variables X_3 and Y_2

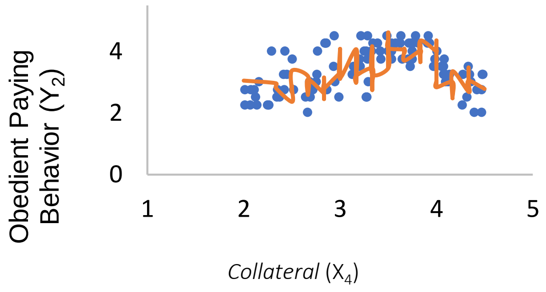

Figure 9. The plot of actual and predicted data between variables X_4 and Y_2

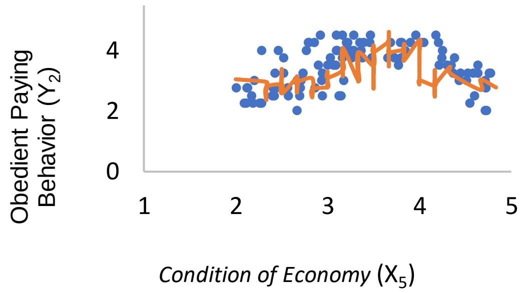

Figure 10. The plot of actual and predicted data between variables X_5 and Y_2

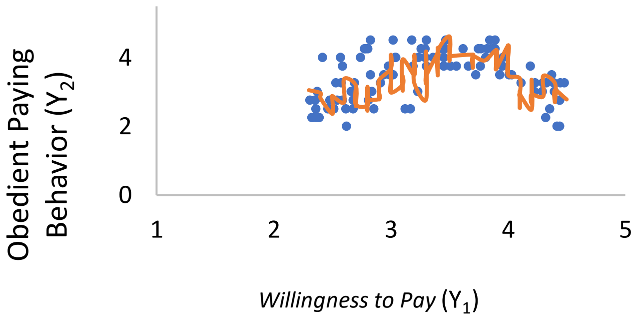

Figure 11. The plot of actual and predicted data between variables Y_1 and Y_2

Based on the data plot formed, it can be seen that the path function estimation using the Fourier Series approach can describe the model well. This can be shown by the actual data that is around the predicted line of the path function of the Fourier Series approach, resulting in a small residual.

Conclusions

1) The function estimation in nonparametric path analysis with Fourier series approach is as follows:
   $\hat{\alpha}(\lambda)=\left(n^{-1}X^{'}X+\lambda D\right)^{-1}n^{-1}X^{'}y$
2) The nonparametric path function that is formed on each oscillation comes from the lambda which has the optimal value, which is when the GCV is the smallest. The best model for each oscillation is:
   a) K=2 while $\lambda=0.2$
   b) K=3 while $\lambda=0.9$
   c) K=4 while $\lambda=0.9$
   d) K=5 while $\lambda=0.9$
3) The best nonparametric path model that can describe the 5C variable on Time to Pay through Willingness to Pay is when the oscillation K=4 with R^2 is 78%.

References

Eubank, R. L. (1999). *Nonparametric regression and spline smoothing*. CRC press.

Fernandes, A. A. R., Solimun, & Arisoesilaningsih, E. (2017). Estimation of spline function in nonparametric path analysis based on penalized weighted least square (PWLS). In *AIP Conference Proceedings*, 1913(1), 020030.

Gujarati, D. N., & Porter, D. C. (2012). *Dasar-Dasar Ekonometrika*, Edisi 5. Jakarta: Salemba Empat.

Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). *Applied linear statistical models* (Vol. 5). New York: McGraw-Hill Irwin.

Nurjanah, F., Utami, T. W., & Nur, I. M. (2015). MODEL REGRESI NONPARAMETRIK DENGAN PENDEKATAN DERET FOURIER PADA POLA DATA CURAH HUJAN DI KOTA SEMARANG. *Jurnal Statistika Universitas Muhammadiyah Semarang*, *3*(2).

Prahutama, A. (2013). Model Regresi Nonparametrik dengan Pendekatan Deret Fourier pada Kasus Tingkat Pengangguran Terbuka di Jawa Timur. In *Prosiding Seminar Nasional Statistika UNDIP*.

Sholiha, A., Kuzairi, K., & Madianto, M. F. F. (2018). Estimator Deret Fourier Dalam Regresi Nonparametrik dengan Pembobot Untuk Perencanaan Penjualan Camilan Khas Madura. *Zeta-Math Journal*, *4*(1), 18-23.

Solimun, M. S. (2002). *Multivariate Analysis Structural Equation Modelling (SEM) Lisrel dan Amos*. *Malang: Fakultas MIPA Universitas Brawijaya*.

Solimun. (2010). *Metode Partial Least Square-PLS*. Malang: CV Citra Malang.

Tripena, A. (2009). Pemilihan parameter penghalus pada estimator deret Fourier dalam regresi nonparametrik. *Jurnal Ilmiah Matematika dan Pendidikan Matematika Universitas Jenderal Soedirman Purwokerto*, *2*, 2-16.

Wisisono, I.R.N, Nurwahidah, A.I., & Andriyana, Y. (2018). Regresi Nonparametrik dengan Pendekatan Deret Fourier pada Data Debit Air Sungai Citarum. *Jurnal Matematika "MANTIK&quot*, *4*(2), 75-82.