CAUCHY: Jurnal Matematika Murni dan Aplikasi
Volume 10(1) (2025), Pages 53-69
p-ISSN: 2086-0382; e-ISSN: 2477-3344

Development of Semiparametric Smoothing Spline Path Analysis on Cashless Society

Muhammad Rafi Hasan Nurdin[1], Muhammad Ohid Ullah[2*], Adji Achmad Rinaldo Fernandes[1], Eni Sumarminingsih[1], Solimun[1],

[1]Department of Statistics, Brawijaya University, Malang, Indonesia
[2]Departement of Statistics, Shahjalal University of Science and Technology, Bangladesh

Email: rafihasan@student.ub.ac.id, ohid-sta@sust.edu*

ABSTRACT

Path analysis has a linearity assumption that can be tested using the Ramsey Regression Specification Error Test (RESET). If the linearity results are partially linear and not linear, a semiparametric model will be used. One of the semiparametric methods is smoothing spline, which requires determining the spline polynomial order in estimating the semiparametric path function. Determining the spline polynomial order is challenging because there is no standard test for it. This study aims to develops a modified Ramsey RESET to identify the optimal spline smoothing order and conduct a simulation study on the Ramsey RESET algorithm. Ramsey RESET development algorithm was successfully applied to identify the optimal smoothing spline polynomial order. The modified Ramsey RESET algorithm is applied to cashless data, and the results are used to estimate a multi-group semiparametric smoothing spline function with a dummy variable approach. This estimation yields a goodness of fit of 94.14%, indicating that Product Quality and the Moderating Effect of Cashless Usage Frequency can explain Cashless User Satisfaction and Cashless User Loyalty by 94.14%, with the remaining 5.86% explained by variables outside the research model. Based on the research results, the research contribution is the Ramsey RESET algorithm to identify the best spline polynomial order.

Keywords: Semiparametric Path Analysis; Smoothing Spline; Ramsey RESET

INTRODUCTION

Path analysis is a statistical method used to test the direct and indirect effects of exogenous variables on endogenous variables [1]. In path analysis, variables are categorized differently compared to regression analysis. Regression analysis uses predictor and response variables while path analysis uses exogenous and endogenous variables. Path analysis involves at least one exogenous variable, intervening endogenous variable, and pure endogenous variable. Intervening endogenous variables are intermediate between exogenous and pure endogenous variables [2]. Path analysis has assumptions that must be met first, namely the assumption of linearity. Linearity is the nature of the relationship between variables is linear. The linearity assumption test uses the Ramsey Regression Specification Error Test (RESET) to determine whether the variable relationship is linear, quadratic, or cubic [3].

Submitted: November 14, 2024 Reviewed: December 28, 2024 Accepted: January 12, 2025
DOI: http://dx.doi.org/10.18860/ca.v10i1.29846

In this study, Ramsey RESET was used because of its broad ability to detect model misspecification without requiring an explicit form of non-linear relationship. In addition, this method is easy to implement on research data with a path analysis framework. However, in comparison, there are other methods such as polynomial regression diagnostics, which uses polynomial terms to check for non-linear relationships that may exist, and splines models, which offer great flexibility in capturing non-linear patterns without the need to define the shape of the relationship in advance [4]. While these alternative methods have their advantages, they are often more computationally complex and require additional interpretation that may be beyond the scope of this study. Therefore, we chose Ramsey RESET as a suitable method to ensure the model is free from misspecification.

The linearity test using Ramsey RESET is done by testing two regression equations. The first equation is an equation of exogenous to endogenous variables. In contrast, the second equation is the first equation added (similar to the concept of auxiliary regression) with additional predictor variables in the form of quadratic response variables and cubic response variables from the first equation. The relationship between the two variables is linear when hypothesis testing on the additional variables in the second equation is insignificant. This indicates that adding quadratic and cubic endogenous variables is unnecessary. The Ramsey RESET results will determine whether the relationship between variables is linear or non-linear. This non-linear meaning is possible in certain non-linear or non-parametric forms [5].

Parametric path analysis can be used when all research variable relationships are linear. If the entire relationship is non-linear then the parametric model cannot be used. An alternative model that can accommodate all non-linear research variable relationships is nonparametric and if some relationships are linear and some are not linear then semiparametric is used [6]. One approach to nonparametric and semiparametric is the spline approach. The spline approach can be used because it can follow the pattern of the relationship between exogenous variables and endogenous variables and is very flexible because spline is one of the approaches in nonparametric regression which has a special character that can adjust changes in data behavior [7].

Spline estimation has two methods: truncated spline and smoothing spline. Both methods have good flexibility in data adjustment so that various curve shapes can be obtained based on different smoothing coefficients [8]. The smoothing spline method requires the order of spline polynomials in the estimation of nonparametric or semiparametric path analysis functions [9]. Determining the order of the spline polynomial becomes an obstacle due to the different characteristics and diversity of the research data. Different characteristics or distribution of data in the relationship between two variables cause the order of the smoothing spline polynomial used to greatly affect the goodness of fit [10].

Fernandes et al. and Pratama et al. conducted research using nonparametric and semiparametric path analysis methods with smoothing spline estimation [11][12]. In the Smoothing Spline Path Analysis and Ramsey RESET conducted by previous researchers did not have testing in determining the order of the smoothing spline polynomial. The novelty of this research is to develop Ramsey RESET to determine the order of the smoothing spline polynomial. Ramsey RESET modification has been done by Solimun and Fernandes on quadratic or more than quadratic, cubic or more than cubic, and quartic or more than quartic models [5]. Solimun and Fernandes research is still limited to detecting non-linearity for quadratic, cubic, and quartic, non-linear models such as smoothing spline approach has not been done. So development of Ramsey RESET non-linear models in this study will use a nonparametric path function with a spline smoothing polynomial order of 2 to 5 in determining the best order.

The application of Ramsey RESET development can be applied to various fields, one of which is in the banking sector, namely the cashless society. Cash in English means cash and less means less, so cashless is less financial transactions or without using cash. The topic of cashless is interesting to study because the development of payment systems innovates from year to year which is the impact of advances in information technology. Indonesian people are starting to adapt to non-cash transactions directly or in using digital platforms. Cashless can lead to how adaptive people's behavior is to the development of cashless payment methods, where over time there is a possibility that the use of cash is decreasing [13]. Public loyalty to the use of cashless is caused by several factors, namely Product Quality and User Satisfaction.

Product quality is one of the determinants of consumer loyalty because good quality will make consumers satisfied with the products provided [14]. Product quality and consumer satisfaction that arise can make consumers loyal [15]. Loyal consumers tend to use products or services repeatedly on the same product. This statement is supported based on the results of research by [16] which show that product quality has a positive and significant effect on customer satisfaction and loyalty and customer satisfaction has a positive and significant effect on customer loyalty. The difference between this study and previous research is in the presence of moderating variables, namely Frequency of Use.

The moderation variable of Cashless Frequency of Use makes path analysis modeling divided into two groups, namely the low moderation group and the high moderation group. The low moderation group is people who still use the cash payment method while the high moderation group is people who often use the cashless payment method. Based on the explanation of the problems above, this research contributes to the field of science in the form of developing Ramsey RESET to identify the order of the smoothing spline polynomial which will be applied to the multi-group smoothing spline semiparametric path analysis on cashless society users and this research will develop the estimation of the smoothing spline semiparametric function with a dummy approach. This research also contributes to provide recommendations for banks on cashless applications such as mobile banking and internet banking.

METHODS

The statistical method used is semiparametric multi-group smoothing spline path analysis. In this study, it is limited to 2 groups, namely the low moderation group and the high moderation group. The purpose of this study is to develop a multi-group smoothing spline semiparametric function estimation with a dummy approach. The research model is shown in Figure 1.

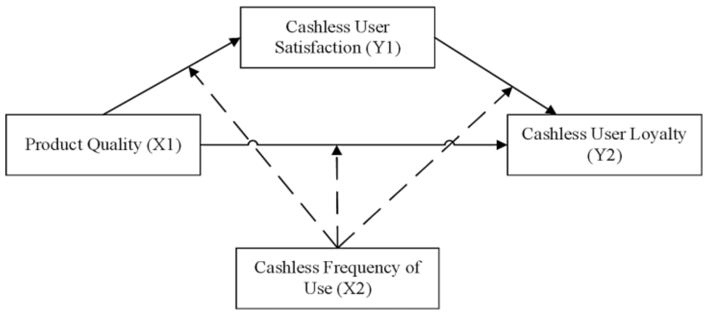

Figure 1. Research Model

Nonparametrik Smoothing Spline Estimation

The basis function estimation for nonparametric smoothing spline path analysis is as follows [11].

$$f = \mathbf{T}\beta + \mathbf{V}\alpha \tag{1}$$

With:

T is a matrix of size $n\times m$ is β a vector of size m.

$$\mathbf{T} = \begin{bmatrix} \langle\eta_1,\phi_1\rangle & \langle\eta_1,\phi_2\rangle & \cdots & \langle\eta_1,\phi_m\rangle \\ \langle\eta_2,\phi_1\rangle & \langle\eta_2,\phi_2\rangle & \cdots & \langle\eta_2,\phi_m\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle\eta_n,\phi_1\rangle & \langle\eta_n,\phi_2\rangle & \cdots & \langle\eta_n,\phi_m\rangle \end{bmatrix}_{(n\times m)} \text{dan } \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{bmatrix}_{(m\times 1)} \tag{2}$$

The element a $\langle\eta_i,\phi_\ell\rangle$ is obtained from the formula as in equation (3).

$$\langle\eta_i,\phi_\ell\rangle = \mathcal{L}_x\phi_\ell$$

$$\langle\eta_i,\phi_\ell\rangle = \frac{x_i^{\ell-1}}{(\ell-1)!} \tag{3}$$

Description:

i : The i-th observation data (i = 1, 2, ..., n)

ℓ : polynomial order a smoothing spline to-ℓ (ℓ = 1, 2, ..., m)

V is a matrix of size $n\times n$ and α is a vector of size $n\times 1$.

$$\mathbf{V} = \begin{bmatrix} \langle\xi_1,\xi_1\rangle & 0 & \cdots & 0 \\ 0 & \langle\xi_2,\xi_2\rangle & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \langle\xi_n,\xi_n\rangle \end{bmatrix}_{(n\times n)} \text{dan } \alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}_{(n\times 1)} \tag{4}$$

With:

$$\langle\xi_i,\xi_i\rangle = \mathcal{L}_x\xi_i$$

$$\langle\xi_i,\xi_i\rangle = R_1(x_i,x_i)$$

$$\langle\xi_i,\xi_i\rangle = \int_a^b \frac{(x_i-u)_+^{m-1}(x_i-u)_+^{m-1}}{\left((m-1)!\right)^2}du$$

$$\langle\xi_i,\xi_i\rangle = \int_a^b \frac{(x_i-u)_+^{2(m-1)}}{\left((m-1)!\right)^2}du, i = 1,2,\dots,n \tag{5}$$

If a $x_i \epsilon [0,1]$ in the spline polynomial order m=2, then equation (5) can be written as in equation (6) [17].

$$\langle\xi_i,\xi_i\rangle = \int_0^1 \frac{(x_i-u)_+^{2(2-1)}}{\left((2-1)!\right)^2}du$$

$$\langle\xi_i,\xi_i\rangle = \int_0^1 (x_i-u)^2\,du$$

$$\langle\xi_i,\xi_i\rangle = \left[ux_i^2 - u^2x_i + \frac{u^3}{3}\right]_0^1$$

$$\langle\xi_i,\xi_i\rangle = x_i^2 - x_i + \frac{1}{3} \tag{6}$$

Based on the equation that has been described, the function estimation for the simple path analysis model in Figure 1 (without moderating variables) is as follows [18].

X_1 toward Y_1

$$f_1(x_{1i}) = \sum_{\ell 1=1}^{2}\beta_{1\ell 1-1}x_{1i}^{\ell 1-1} + \sum_{i=1}^{n}\alpha_{1i}\left(x_{1i}^2 - x_{1i} + \frac{1}{3}\right) \tag{7}$$

X_1 toward Y_2

$$f_{2.1}(x_{1i}) = \sum_{\ell 2=1}^{2}\beta_{2\ell 2-1}x_{1i}^{\ell 2-1} + \sum_{i=1}^{n}\alpha_{2i}\left(x_{1i}^2 - x_{1i} + \frac{1}{3}\right) \tag{8}$$

Y_1 toward Y_2

$$f_{2.2}(y_{1i}) = \sum_{\ell 3=1}^{2}\beta_{3\ell 3-1}y_{1i}^{\ell 3-1} + \sum_{i=1}^{n}\alpha_{3i}\left(y_{1i}^2 - y_{1i} + \frac{1}{3}\right) \tag{9}$$

The estimation of the joint function of X_1 and Y_1 toward Y_2 is as follows.

$$f_2(x_{1i},y_{1i}) = \beta_{20} + \beta_{21}x_{1i} + \beta_{22}y_{1i} + \sum_{i=1}^{n}\alpha_{2i}\left(x_{1i}^2 - x_{1i} + y_{1i}^2 - y_{1i} + \frac{2}{3}\right) \tag{10}$$

Based on the estimation of the nonparametric smoothing spline function with order 2, the estimation of order 3 to 5 can be obtained. Estimation of nonparametric smoothing spline using the Penalized Weighted Least Square (PWLS) method [11]. The reason for choosing the PWLS method is because the results of Fernandes et al's research show that estimating the nonparametric smoothing spline curve with the PWLS approach is better at estimating data patterns compared to the PLS approach [19]. Smoothing spline requires optimal smoothing parameters through Generalized Cross Validation (GCV) [20]. The value of the smoothing parameter is chosen from the minimum GCV value [21].

If a $x_i \in [0,1]$ in the spline polynomial order m=3, then the result of equation (5) is shown in equation (11).

$$\langle \xi_i, \xi_i \rangle = \frac{1}{4}\left[x_{1i}^4 - 2x_{1i}^3 + 2x_{1i}^2 - x_{1i} + \frac{1}{5}\right] \quad (11)$$

If a $x_i \in [0,1]$ in the spline polynomial order m=4, then the result of equation (5) is shown in equation (12).

$$\langle \xi_i, \xi_i \rangle = \frac{1}{36}\left[x_{1i}^6 - 3x_{1i}^5 + 5x_{1i}^4 - 5x_{1i}^3 + 3x_{1i}^2 - x_{1i} + \frac{1}{7}\right] \quad (12)$$

If a $x_i \in [0,1]$ in the spline polynomial order m=5, then the result of equation (5) is shown in equation (13).

$$\langle \xi_i, \xi_i \rangle = \frac{1}{576}\left[x_{1i}^8 - 4x_{1i}^7 + \frac{28}{3}x_{1i}^6 - 14x_{1i}^5 + 14x_{1i}^4 - \frac{28}{3}x_{1i}^3 + 4x_{1i}^2 - x_{1i} + \frac{1}{9}\right] \quad (13)$$

Polynomial order 2 to 5 smoothing spline estimation is used to develop Ramsey RESET in identifying the best order and develop estimation of multi-group smoothing spline semiparametric function with dummy approach.

Ramsey RESET

Ramsey RESET was first developed by Ramsey in 1969 [24], this test is the most widely used method for testing linearity. The steps of RESET testing according to [25] are as follows.

1. Calculating the estimated value of the first regression parameter using Ordinary Least Square (OLS) in the equation (14).
   $$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i \quad (14)$$
   Calculating the coefficient of determination (R_{first}^2).
   $$R_{first}^2 = 1 - \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} \quad (15)$$
2. Calculating the estimated values of the second regression parameter.
   $$\hat{Y}_i^* = \hat{\beta}_0^* + \hat{\beta}_1^* X_i + \hat{\beta}_2 \hat{Y}_i^2 + \hat{\beta}_3 \hat{Y}_i^3 \quad (16)$$
   Calculating the second coefficient of determination (R_{second}^2).
   $$R_{second}^2 = 1 - \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i^*)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} \quad (17)$$
3. The hypotheses tested are as follows.
   $H_0: \beta_2 = \beta_3 = 0$
   H_1: there is at least one $\beta_j \neq 0; j = 2,3$
4. Calculate the value of the F test statistic using the equation formula (18).
   $$F_{count} = \frac{(R_{second}^2 - R_{first}^2)/2}{(1 - R_{second}^2)/(n-k)} \sim F_{2,n-k} \quad (18)$$
   Description:
   n : number of observations; i = 1, 2, 3, …, n
   k : number of parameters in the second regression

   Determining the critical value approximates the F distribution at α = 5%. The test criterion is if the test statistic $F_{count} \geq F_{table}$ or p-value < α then H_0 rejected, meaning the model is not linear.

RESULTS AND DISCUSSION

Ramsey RESET Modification

The Ramsey RESET modification uses the foundation of the Classic Ramsey which was first developed in 1969. In Classic Ramsey, the first equation is a form to detect linear and the second equation is a non-linear form in the form of quadratic and quartic which cannot detect non-linear models with a smoothing spline approach. Based on the weaknesses in the classic Ramsey RESET, it is necessary to develop to detect the order of the smoothing spline polynomial. Modifications are made to the second equation with a nonparametric smoothing spline path equation with polynomial orders 2 to 5 and for the first equation no changes are made. The steps of the Ramsey RESET modification to identify the smoothing spline polynomial order are as follows.

1. Calculating the estimated value of the first regression parameter using OLS using the equation (14) and calculate the first coefficient of determination (R^2_{first}) using the equation (15). The first step is a linear form that will be used as a basis as in the classic Ramsey RESET.
2. Forms the second regression equation shown in equation (19).
   $$Y^*_{mi} = f_m(x_i) + \epsilon_i \qquad (19)$$
   Calculating the estimated value of the second regression parameter.
   $$Y^*_{mi} = \hat{f}_m(x_i) \qquad (20)$$
3. Where m indicates the order of the spline smoothing polynomial so that there will be 4 function estimates shown in the equation (21) to (24). Equations (21) to (24) are used for the non-linear model with spline smoothing approach which will be compared with the linear model. This concept is the same as the classic Ramsey RESET except that the second regression equation is modified. Equation (21) is a regression equation with polynomial smoothing spline order 2 (m=2).
   $$Y^*_{2i} = \hat{f}_2(x_i)$$
   $$\hat{Y}^*_{2i} = \hat{\beta}_{20} + \hat{\beta}_{21}x_i + \sum_{i=1}^{n} \alpha_{2i}\left(x_i^2 - x_i + \frac{1}{3}\right) \qquad (21)$$
   The regression equation with spline smoothing polynomial order 3 (m=3) is shown in equation (22).
   $$Y^*_{3i} = \hat{f}_3(x_i)$$
   $$\hat{Y}^*_{3i} = \hat{\beta}_{30} + \hat{\beta}_{31}x_i + \hat{\beta}_{32}x_i^2 + \sum_{i=1}^{n} \alpha_{3i}\left(\frac{1}{4}\left(x_{1i}^4 - 2x_{1i}^3 + 2x_{1i}^2 - x_{1i} + \frac{1}{5}\right)\right) \qquad (22)$$
   The regression equation with spline smoothing polynomial order 4 (m=4) is shown in equation (23).
   $$Y^*_{4i} = \hat{f}_4(x_i)$$
   $$\hat{Y}^*_{4i} = \hat{\beta}_{40} + \hat{\beta}_{41}x_i + \hat{\beta}_{42}x_i^2 + \hat{\beta}_{43}x_i^3 + \sum_{i=1}^{n} \alpha_{4i}\left(\frac{1}{36}\left(x_{1i}^6 - 3x_{1i}^5 + 5x_{1i}^4 - 5x_{1i}^3 + 3x_{1i}^2 - x_{1i} + \frac{1}{7}\right)\right) \qquad (23)$$
   The regression equation with spline smoothing polynomial order 5 (m=5) is shown in equation (24).
   $$Y^*_{5i} = \hat{f}_5(x_i)$$
   $$\hat{Y}^*_{5i} = \hat{\beta}_{50} + \hat{\beta}_{51}x_i + \hat{\beta}_{52}x_i^2 + \hat{\beta}_{53}x_i^3 + \hat{\beta}_{54}x_i^4 + \sum_{i=1}^{n} \alpha_{5i}\left(\frac{1}{576}\left(x_{1i}^8 - 4x_{1i}^7 + \frac{28}{3}x_{1i}^6 - 14x_{1i}^5 + 14x_{1i}^4 - \frac{28}{3}x_{1i}^3 + 4x_{1i}^2 - x_{1i} + \frac{1}{9}\right)\right) \qquad (24)$$
   Calculating the coefficient of determination ($R^2_{m,second}$) for equations (21) to (24).
   $$R^2_{m,second} = 1 - \frac{\sum_{i=1}^{n}(Y_i - \hat{f}_m)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} \qquad (25)$$
4. Calculate the value of the F test statistic using the equation formula (26).

$$F_{m,count} = \frac{(R^2_{m,second} - R^2_{first})/(m-1)}{(1-R^2_{second})/(n-k)} \sim F_{m-1,n-k} \quad (26)$$

Description:

n : number of observations; $i = 1,2,3,\dots,n$

m : order of polynomial smoothing spline

k : number of parameters in the second regression

Determining the critical value approximates the F distribution at α = 5%. The test criterion is if the test statistic $F_{count} \geq F_{table}$ or p-value < α then H_0 is rejected, meaning that the model is not linear with polynomial smoothing spline orders 2, 3, 4, and 5. This Ramsey RESET modification is then used in the case of cashless society data.

Linearity Test

This study has a moderating variable "Frequency of Cashless Use" so that multi-groups are formed, namely people who rarely use cashless payment methods (Group 1) and people who often use cashless payments (Group 2). The division of Group 1 and Group 2 is based on the average score on the "Cashless Frequency of Use" variable. Classical Ramsey RESET and Ramsey RESET development were used to test linearity in both groups, the results of the linearity test for group 1 are shown in Table 1.

Table 1. Linearity Test Results Group 1

Relationships	p-value of Ramsey RESET	p-value of Ramsey RESET Modification			
		m=2	m=3	m=4	m=5
X_1 to Y_1	0.0254	**0.0285**	0.1392	0.2115	0.3899
X_1 to Y_2	0.0443	**0.0015**	0.0146	0.0349	0.0115
Y_1 to Y_2	**0.2220**	-	-	-	-

Based on Table 1, the p-value is less than 0.05 for the relationship between variables X_1 to Y_1 and X_1 to Y_2, which indicates that it is not linear, while Y_1 to Y_2 obtained a p-value of more than 0.05, indicating that the relationship is linear. The non-linear variable relationship is continued to the Ramsey RESET modification to identify the best spline smoothing polynomial order. The results of the Ramsey RESET modification obtained the smallest p-value at m = 2 which indicates that the best order for X_1 to Y_1 and X_1 to Y_2 is the order of the polynomial smoothing spline of 2. The results of the linearity test in group 2, namely the group of people who often use cashless payment methods, are shown in Table 2.

Table 2. Linearity Test Results Group 2

Relationships	p-value of Ramsey RESET	p-value of Ramsey RESET Modification			
		m=2	m=3	m=4	m=5
X_1 to Y_1	0.0338	3.2×10^{-11}	$\mathbf{2.7\times10^{-11}}$	4.9×10^{-9}	2.9×10^{-8}
X_1 to Y_2	0.0407	$\mathbf{9.4\times10^{-8}}$	1.3×10^{-6}	1.1×10^{-5}	4.6×10^{-5}
Y_1 to Y_2	**0.6059**	-	-	-	-

Based on Table 2, the nature of the variable relationship in group 2 is obtained. The results in group 1 and group 2 obtained the same results, namely non-linear in X_1 to Y_1 and X_1 to Y_2 and linear in Y_1 to Y_2. The variable relationship X_1 to Y_1 is non-linear obtaining the smallest p-value when the order of the smoothing spline polynomial is 3 and X_1 to Y_2 is non-linear obtaining the smallest p-value with the smoothing spline polynomial order of 2. The linearity results in Tables 1 and 2 will be used to estimate the multi-group smoothing spline semiparametric path function with a dummy approach.

Development of Multi-Group Semiparametric Smoothing Spline Path Function Estimator

The estimation of the semiparametric smoothing spline path function developed is the result of the linearity test. The development of the multi-group smoothing spline semiparametric path function estimation using the dummy approach, the dummy used is 0 and 1, where the value 0 indicates that Group 1 (low moderation) is a community that rarely uses cashless payment methods and 1 for Group 2 (high moderation) is a community that often uses cashless payments shown in equation (27).

$$\begin{aligned}\hat{f}_1(x_{1i}) &= \hat{\beta}_{10} + \hat{\beta}_{11}D_i + \hat{\beta}_{12}x_{1i} + \hat{\beta}_{13}x_{1i}D_i + \hat{\beta}_{14}x_{1i}^2D_i \\ &\quad + \sum_{i=1}^{n_1}\hat{\alpha}_{1i}\left(x_{1i}^2 - x_{1i} + \tfrac{1}{3}\right)(1-D_i) \\ &\quad + \sum_{i=n_1+1}^{n}\hat{\alpha}_{1i}\left(\tfrac{1}{4}\left(x_{1i}^4 - 2x_{1i}^3 + 2x_{1i}^2 - x_{1i} + \tfrac{1}{5}\right)\right)(D_i) \qquad (27)\\ \hat{f}_2(x_{1i}, y_{1i}) &= \hat{\beta}_{20} + \hat{\beta}_{21}D_i + \hat{\beta}_{22}x_{1i} + \hat{\beta}_{23}x_{1i}D_i + \hat{\beta}_{24}y_{1i} + \hat{\beta}_{25}y_{1i}D_i \\ &\quad + \sum_{i=1}^{n_1}\hat{\alpha}_{2i}\left(x_{1i}^2 - x_{1i} + \tfrac{1}{3}\right)(1-D_i) + \sum_{i=n_1+1}^{n}\hat{\alpha}_{2i}\left(x_{1i}^2 - x_{1i} + \tfrac{1}{3}\right)(D_i)\end{aligned}$$

Equation (27) can be written in matrix form as follows.

$$\begin{bmatrix} y_{1(n\times1)} \\ y_{2(n\times1)} \end{bmatrix}_{(2n\times1)} = \begin{bmatrix} \mathbf{T}_{1(n\times5)} & \mathbf{0}_{(n\times6)} \\ \mathbf{0}_{(n\times5)} & \mathbf{T}_{2(n\times6)} \end{bmatrix}_{(2n\times11)} + \begin{bmatrix} \beta_{1(5\times1)} \\ \beta_{2(6\times1)} \end{bmatrix}_{(11\times1)} + \begin{bmatrix} \mathbf{V}_{1(n\times n)} & \mathbf{0}_{(n\times n)} \\ \mathbf{0}_{(n\times n)} & \mathbf{V}_{2(n\times n)} \end{bmatrix}_{(2n\times2n)} + \begin{bmatrix} \alpha_{1(n\times1)} \\ \alpha_{2(n\times1)} \end{bmatrix}_{(2n\times1)} + \begin{bmatrix} \epsilon_{1(n\times1)} \\ \epsilon_{2(n\times1)} \end{bmatrix}_{(2n\times1)} \qquad (28)$$

Description:

$$y_1 = \begin{bmatrix} y_{11} \\ y_{12} \\ \vdots \\ y_{1n} \end{bmatrix}_{(n\times1)} ;\ y_2 = \begin{bmatrix} y_{21} \\ y_{22} \\ \vdots \\ y_{2n} \end{bmatrix}_{(n\times1)}$$

$$\mathbf{T}_1 = \begin{bmatrix} 1 & D_1 & x_{11} & x_{11}D_1 & x_{11}^2D_1 \\ 1 & D_2 & x_{12} & x_{12}D_2 & x_{12}^2D_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & D_n & x_{1n} & x_{1n}D_n & x_{1n}^2D_n \end{bmatrix}_{(n\times5)} ;\ \mathbf{T}_1 = \begin{bmatrix} 1 & D_1 & x_{11} & x_{11}D_1 & y_{11} & y_{11}D_1 \\ 1 & D_2 & x_{12} & x_{12}D_2 & y_{12} & y_{12}D_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & D_n & x_{1n} & x_{1n}D_n & y_{1n} & y_{1n}D_n \end{bmatrix}_{(n\times6)}$$

$$\beta_1 = \begin{bmatrix} \beta_{10} \\ \beta_{11} \\ \beta_{12} \\ \beta_{13} \\ \beta_{14} \end{bmatrix}_{(5\times1)} ;\ \beta_2 = \begin{bmatrix} \beta_{20} \\ \beta_{21} \\ \beta_{22} \\ \beta_{23} \\ \beta_{24} \\ \beta_{25} \end{bmatrix}_{(6\times1)}$$

$$\mathbf{V}_1 = \begin{bmatrix} \mathbf{V}_{11(n_1\times n_1)} & \mathbf{0}_{(n_2\times n_2)} \\ \mathbf{0}_{(n_1\times n_1)} & \mathbf{V}_{12(n_2\times n_2)} \end{bmatrix}_{(n\times n)} ;\ \mathbf{V}_2 = \begin{bmatrix} \mathbf{V}_{21(n_1\times n_1)} & \mathbf{0}_{(n_2\times n_2)} \\ \mathbf{0}_{(n_1\times n_1)} & \mathbf{V}_{22(n_2\times n_2)} \end{bmatrix}_{(n\times n)}$$

$$\mathbf{V}_{11} = \begin{bmatrix} \left(x_{11}^2 - x_{11} + \frac{1}{3}\right)(1-D_1) & 0 & \cdots & 0 \\ 0 & \left(x_{12}^2 - x_{12} + \frac{1}{3}\right)(1-D_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \left(x_{1n_1}^2 - x_{1n_1} + \frac{1}{3}\right)(1-D_{n_1}) \end{bmatrix}_{(n_1\times n_1)}$$

$$\mathbf{V}_{12} = \begin{bmatrix} \frac{1}{4}\left(x_{1n_1+1}^4 - 2x_{1n_1+1}^3 + 2x_{1n_1+1}^2 - x_{1n_1+1} + \frac{1}{5}\right)(D_{n_1+1}) & 0 & \cdots & 0 \\ 0 & \frac{1}{4}\left(x_{1n_1+2}^4 - 2x_{1n_1+2}^3 + 2x_{1n_1+2}^2 - x_{1n_1+2} + \frac{1}{5}\right)(D_{n_1+2}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{4}\left(x_{1n}^4 - 2x_{1n}^3 + 2x_{1n}^2 - x_{1n} + \frac{1}{5}\right)(D_n) \end{bmatrix}_{(n_2\times n_2)}$$

$$\mathbf{V}_{21} = \begin{bmatrix} \left(x_{11}^2 - x_{11} + \frac{1}{3}\right)(1-D_1) & 0 & \cdots & 0 \\ 0 & \left(x_{12}^2 - x_{12} + \frac{1}{3}\right)(1-D_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \left(x_{1n_1}^2 - x_{1n_1} + \frac{1}{3}\right)(1-D_{n_1}) \end{bmatrix}_{(n_1 \times n_1)}$$

$$\mathbf{V}_{22} = \begin{bmatrix} \left(x_{1n_1+1}^2 - x_{1n_1+1} + \frac{1}{3}\right)(D_{n_1+1}) & 0 & \cdots & 0 \\ 0 & \left(x_{1n_1+2}^2 - x_{1n_1+2} + \frac{1}{3}\right)(D_{n_1+2}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \left(x_{1n}^2 - x_{1n} + \frac{1}{3}\right)(D_n) \end{bmatrix}_{(n_2 \times n_2)}$$

$$\alpha_1 = \begin{bmatrix} \alpha_{11} \\ \alpha_{12} \\ \vdots \\ \alpha_{1n_1} \\ \alpha_{1n_1+1} \\ \alpha_{1n_1+2} \\ \vdots \\ \alpha_{1n} \end{bmatrix}_{(n\times1)} ;\ \alpha_2 = \begin{bmatrix} \alpha_{21} \\ \alpha_{22} \\ \vdots \\ \alpha_{2n_1} \\ \alpha_{2n_1+1} \\ \alpha_{2n_1+2} \\ \vdots \\ \alpha_{2n} \end{bmatrix}_{(n\times1)} ;\ \epsilon_1 = \begin{bmatrix} \epsilon_{11} \\ \epsilon_{12} \\ \vdots \\ \epsilon_{1n} \end{bmatrix}_{(n\times1)} ;\ \epsilon_2 = \begin{bmatrix} \epsilon_{21} \\ \epsilon_{22} \\ \vdots \\ \epsilon_{2n} \end{bmatrix}_{(n\times1)}$$

Description:

n_1 : Number of observations in Group 1

n_2 : Number of observations in Group 2 ; $n_2 = n_1 + 1, n_1 + 2, \ldots, n$

n : Number of observations in Group 1 dan 2

If the dummy value (D_i) for Group 1 is given a value of 0, the estimation of the semiparametric smoothing spline path function of the low moderation group is shown in equation (29).

$$\begin{aligned} \hat{f}_1(x_{1i}) &= \hat{\beta}_{10} + \hat{\beta}_{12}x_{1i} + \textstyle\sum_{i=1}^{n_1} \hat{\alpha}_{1i}\left(x_{1i}^2 - x_{1i} + \frac{1}{3}\right) \\ \hat{f}_2(x_{1i}, y_{1i}) &= \hat{\beta}_{20} + \hat{\beta}_{22}x_{1i} + \hat{\beta}_{24}y_{1i} + \textstyle\sum_{i=1}^{n_1} \hat{\alpha}_{2i}\left(x_{1i}^2 - x_{1i} + \frac{1}{3}\right) \end{aligned} \tag{29}$$

Estimating the semiparametric smoothing spline multi-group path function for Group 2 when Di is equal to 1, equation (30) is obtained.

$$\begin{aligned} \hat{f}_1(x_{1i}) &= (\hat{\beta}_{10} + \hat{\beta}_{11}) + (\hat{\beta}_{12} + \hat{\beta}_{13})x_{1i} + \hat{\beta}_{14}x_{1i}^2 + \\ &\quad \textstyle\sum_{i=n_1+1}^{n} \hat{\alpha}_{1i}\left(\frac{1}{4}\left(x_{1i}^4 - 2x_{1i}^3 + 2x_{1i}^2 - x_{1i} + \frac{1}{5}\right)\right) \\ \hat{f}_2(x_{1i}, y_{1i}) &= (\hat{\beta}_{20} + \hat{\beta}_{21}) + (\hat{\beta}_{22} + \hat{\beta}_{23})x_{1i} + (\hat{\beta}_{24} + \hat{\beta}_{25})y_{1i} + \\ &\quad \textstyle\sum_{i=n_1+1}^{n} \hat{\alpha}_{2i}\left(x_{1i}^2 - x_{1i} + \frac{1}{3}\right) \end{aligned} \tag{30}$$

The estimation of the semiparametric smoothing spline path function in equation (30) has additional parameters, namely $\hat{\beta}_{11}, \hat{\beta}_{13}, \hat{\beta}_{14}, \hat{\beta}_{21}, \hat{\beta}_{23}$, and $\hat{\beta}_{25}$ from equation (29). Additional parameters $\hat{\beta}_{11}, \hat{\beta}_{13}, \hat{\beta}_{21}, \hat{\beta}_{23}$, and $\hat{\beta}_{25}$ is the difference between groups 1 and 2, if it is positive then the influence between variables in group 2 is higher than group 1 and vice versa if it is negative then the influence between variables in group 2 is lower than group 1. In group 2 there is a $\hat{\beta}_{14}$ parameter because the relationship between X_1 and Y_1 is not linear with a smoothing spline polynomial order of 3 so that there is a form of X_1^2.

Estimation of Semiparametric Path Function of Multi-Group Spline Smoothing Based on Modified RRT on Cashless User Loyalty Data

The first step is selecting the smoothing coefficient using the Generalized Cross Validation (GCV) method. GCV selection is done with a maximum of 100 iterations and iteration stops when the difference between the i-th and i-1 GCV is less than 10^{-4}, The results of iteration with the help of RStudio software are shown in Table 3.

Table 3. GCV and Lambda Result

Iterations	GCV	λ_1	λ_2
1	0.37798	0.99992	0.01295
2	0.32522	0.99989	0.13795
3	0.32500	0.99989	0.14115
4	0.32321	0.99988	0.17190
5	0.32049	0.99987	0.24804
6	0.31871	0.99985	0.33570
7	0.31710	0.99982	0.46969
8	0.31577	0.99978	0.65663
9	0.31465	0.99972	0.92803
10	0.31376	0.99963	1.30165
11	0.31315	0.99953	1.77245
12	0.31281	0.99941	2.28465
13	0.31267	0.99931	2.74768
14	0.31264	0.99923	3.07020

Based on Table 3, the optimum GCV is 0.31264 with $\lambda_1 = 0.99923$ and $\lambda_2 = 3.06020$. GCV is the goodness of fit in the model to determine the optimal smoothing parameter [22]. The smoothing parameter controls the balance between the fit of the curve to the data and the smoothness of the curve [23]. The GCV value and λ value form a multi-group smoothing spline semiparametric path analysis function. The parameters in the estimation of the function formed are equation (27) applied to Cashless Data shown in Table 4.

Table 4. Value of $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\alpha}}$

	Y_1	Y_2
$\hat{\beta}_{k0}$	3.335	2.126
$\hat{\beta}_{k1}$	-1.379	0.345
$\hat{\beta}_{k2}$	0.075	0.340
$\hat{\beta}_{k3}$	0.427	0.226
$\hat{\beta}_{k4}$	0.068	0.065
$\hat{\beta}_{k5}$	-	0.364
$\hat{\alpha}_{k1}$	-0.006	-0.001
$\hat{\alpha}_{k2}$	0.001	0.005
⋮	⋮	⋮
$\hat{\alpha}_{k90}$	0.009	-0.011
$\hat{\alpha}_{k91}$	-0.006	-0.001
$\hat{\alpha}_{k92}$	0.012	-0.008
⋮	⋮	⋮
$\hat{\alpha}_{k200}$	-0.017	0.007

The values in Table 4 can form two functions of multi-group smoothing spline semiparametric path analysis. The first function obtained is the function of Product Quality on Cashless User Satisfaction and the Second function is the function of Product Quality and Cashless User Satisfaction on Cashless User Loyalty. Both functions are shown in equation (31).

$$\hat{f}_1(x_{1i}) = 3{,}335 - 1{,}379D_i + 0{,}075x_{1i} + 0{,}427x_{1i}D_i + 0{,}068x_{1i}^2D_i - 0{,}006\left(x_{11}^2 - x_{11} + \frac{1}{3}\right)(1 - D_1) + 0{,}001\left(x_{12}^2 - x_{12} + \frac{1}{3}\right)(1 - D_2) + \cdots + 0{,}009\left(x_{190}^2 - x_{190} + \frac{1}{3}\right)(1 - D_{90}) - 0{,}006\left(\frac{1}{4}\left(x_{191}^4 - 2x_{191}^3 + 2x_{191}^2 - x_{191} + \frac{1}{5}\right)\right)(D_{91}) + 0{,}012\left(\frac{1}{4}\left(x_{192}^4 - 2x_{192}^3 + 2x_{192}^2 - x_{192} + \frac{1}{5}\right)\right)(D_{92}) + \cdots - 0{,}017\left(\frac{1}{4}\left(x_{1200}^4 - 2x_{1200}^3 + 2x_{1200}^2 - x_{1200} + \frac{1}{5}\right)\right)(D_{200})$$

$$\hat{f}_2(x_{1i}, y_{1i}) = 2{,}126 + 0{,}345D_i + 0{,}340x_{1i} + 0{,}226x_{1i}D_i + 0{,}065y_{1i} + 0{,}364y_{1i}D_i - 0{,}001\left(x_{11}^2 - x_{11} + \frac{1}{3}\right)(1 - D_1) + 0{,}005\left(x_{12}^2 - x_{12} + \frac{1}{3}\right)(1 - D_2) + \cdots - 0{,}011\left(x_{190}^2 - x_{190} + \frac{1}{3}\right)(1 - D_{90}) - 0{,}001\left(x_{191}^2 - x_{191} + \frac{1}{3}\right)(D_{91}) - 0{,}008\left(x_{192}^2 - x_{192} + \frac{1}{3}\right)(D_{92}) + \cdots + 0{,}007\left(x_{1200}^2 - x_{1200} + \frac{1}{3}\right)(D_{200}) \quad (31)$$

The estimation of the semiparametric smoothing spline path function in equation (31) can be divided into two, namely the estimation of Group 1 and Group 2 functions. The estimation of the Group 1 semiparametric smoothing spline path function is shown in equation (32).

$$\hat{f}_1(x_{1i}) = 3{,}335 + 0{,}075x_{1i} - 0{,}006\left(x_{11}^2 - x_{11} + \frac{1}{3}\right) + 0{,}001\left(x_{12}^2 - x_{12} + \frac{1}{3}\right) + \cdots + 0{,}009\left(x_{190}^2 - x_{190} + \frac{1}{3}\right)$$

$$\hat{f}_2(x_{1i}, y_{1i}) = 2{,}126 + 0{,}340x_{1i} + 0{,}065y_{1i} - 0{,}001\left(x_{11}^2 - x_{11} + \frac{1}{3}\right) + 0{,}005\left(x_{12}^2 - x_{12} + \frac{1}{3}\right) + \cdots - 0{,}011\left(x_{190}^2 - x_{190} + \frac{1}{3}\right) \quad (32)$$

Figures 2 to 3 visualize the estimation of the semiparametric smoothing spline path function of the community group with low frequency of cashless use (Group 1).

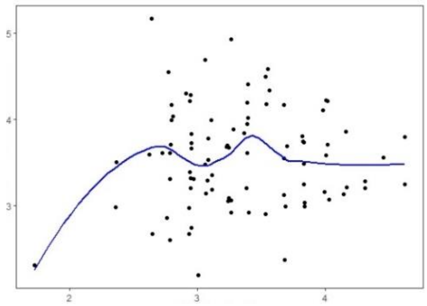

Figure 2. Product Quality to Cashless User Satisfaction Group 1

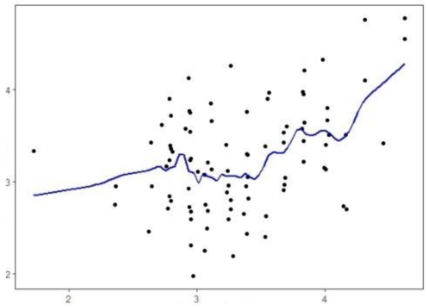

Figure 3. Product Quality to Cashless User Loyalty Group 1

Figures 2 and 3 show the visualization of the nonparametric smoothing spline approach to the relationship of the Product Quality variable to the Cashless User

Satisfaction and Cashless User Loyalty variables in Group 1. The blue estimator curve visually follows the pattern of the data because there is a penalty for each observation.

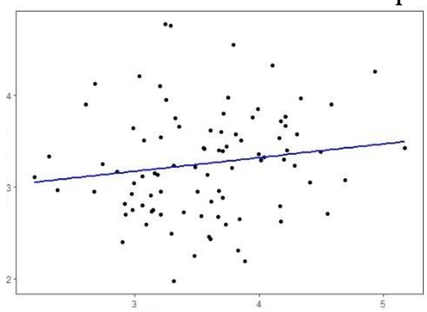

Figure 4. Cashless User Satisfaction to Cashless User Loyalty Group 1

The diagram in Figure 4 shows the relationship between Cashless User Satisfaction and Cashless User Loyalty. The path curve shows an increasing pattern. In other words, the higher the level of user satisfaction in using cashless payment methods will increase user loyalty. The estimation of the Group 2 semiparametric smoothing spline path function is shown in equation (33).

$$\hat{f}_1(x_{1i}) = 1{,}956 + 0{,}502x_{1i} + 0{,}068x_{1i}^2 - 0{,}006\left(\frac{1}{4}\left(x_{191}^4 - 2x_{191}^3 + 2x_{191}^2 - x_{191} + \frac{1}{5}\right)\right) + 0{,}012\left(\frac{1}{4}\left(x_{192}^4 - 2x_{192}^3 + 2x_{192}^2 - x_{192} + \frac{1}{5}\right)\right) + \cdots - 0{,}017\left(\frac{1}{4}\left(x_{1200}^4 - 2x_{1200}^3 + 2x_{1200}^2 - x_{1200} + \frac{1}{5}\right)\right) \quad (33)$$

$$\hat{f}_2(x_{1i}, y_{1i}) = 2{,}471 + 0{,}566x_{1i} + 0{,}429y_{1i} - 0{,}001\left(x_{191}^2 - x_{191} + \frac{1}{3}\right) - 0{,}008\left(x_{192}^2 - x_{192} + \frac{1}{3}\right) + \cdots + 0{,}007\left(x_{1200}^2 - x_{1200} + \frac{1}{3}\right)$$

Figures 5 to 7 visualize the estimation of the semiparametric function of the community group with a high frequency of cashless use.

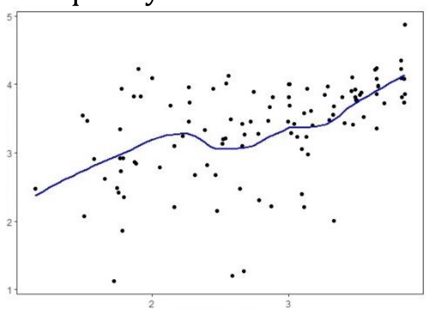

Figure 5. Product Quality to Cashless User Satisfaction Group 2

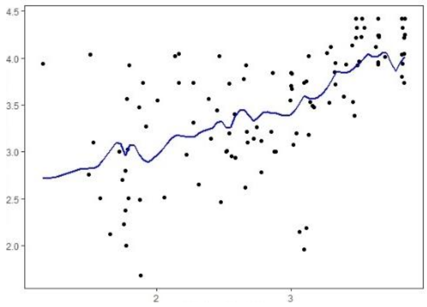

Figure 6. Product Quality to Cashless User Loyalty Group 2

Figure 5 and 6, the blue estimator curve visually follows the pattern of the data because there is a penalty for each observation. This can be interpreted that the function estimation with a nonparametric approach is more flexible and more able to follow the data pattern.

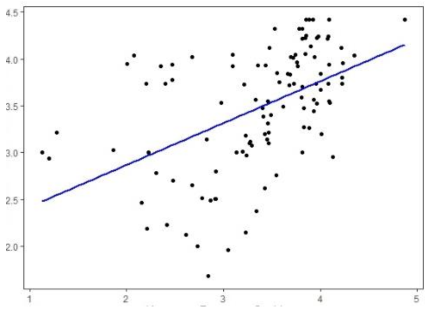

Figure 7. Cashless User Satisfaction to Cashless User Loyalty Group 2

The diagram in Figure 7 shows the relationship between the Cashless User Satisfaction variable and the Cashless User Loyalty variable. The path curve shows an increasing pattern, where the higher the level of user satisfaction in using cashless payment methods will increase user loyalty. The effect on Group 2 is higher than Group 1 because users with high frequency of cashless use (Group 2) already feel accustomed to using and have a higher attachment than users with low frequency so that Group 2 user satisfaction has a greater influence in strengthening loyalty. In addition, people in Group 2 have experienced various situations in using cashless methods, both positive and negative. The satisfaction felt from positive experiences can significantly increase loyalty as people are better able to see practical benefits in daily use.

Hypothesis Test

The next step is to test the significance of function estimator. The hypothesis tested is as follows.
Group 1
X_1 to Y_1
H_0: $\beta_{12}X_1 = 0$ (Product Quality has no affect Cashless User Satisfaction)
H_1: $\beta_{12}X_1 \neq 0$ (Product Quality affect Cashless User Satisfaction)
X_1 to Y_2
H_0: $\beta_{22}X_1 = 0$ (Product Quality has no affect Cashless User Loyalty)
H_1: $\beta_{22}X_1 \neq 0$ (Product Quality affect Cashless User Loyalty)
Y_1 to Y_2
H_0: $\beta_{24}Y_1 = 0$ (Cashless User Satisfaction has no affect Cashless User Loyalty)
H_1: $\beta_{24}Y_1 \neq 0$ (Cashless User Satisfaction affect Cashless User Loyalty)
Group 2
X_1 to Y_1
H_0: $(\beta_{12} + \beta_{13})X_1 = 0$ (Product Quality has no affect Cashless User Satisfaction)
H_1: $(\beta_{12} + \beta_{13})X_1 \neq 0$ (Product Quality affect Cashless User Satisfaction)
X_1 to Y_2
H_0: $(\beta_{22} + \beta_{23})X_1 = 0$ (Product Quality has no affect Cashless User Loyalty)
H_1: $(\beta_{22} + \beta_{23})X_1 \neq 0$ (Product Quality affect Cashless User Loyalty)
Y_1 to Y_2
H_0: $(\beta_{24} + \beta_{25})Y_1 = 0$ (Cashless User Satisfaction has no affect Cashless User Loyalty)
H_1: $(\beta_{24} + \beta_{25})Y_1 \neq 0$ (Cashless User Satisfaction affect Cashless User Loyalty)

The test criteria used are when alpha is less than 0.05 then reject H_0 and when more than 0.05 then accept H_1. The results of the hypothesis test using jackknife resampling conducted 500 times times are as follows.

Table 5. Hypothesis Testing Results

Group	Relationships	Parameter in Function	Coefficient	Test Statistic t	p-value
1	$X_1 \rightarrow Y_1$	$\beta_{12}X_1$	0.032	15.03	<0.001*
	$X_1 \rightarrow Y_2$	$\beta_{22}X_1$	0.056	7.981	<0.001*
	$Y_1 \rightarrow Y_2$	$\beta_{24}Y_1$	0.033	1.733	0.083 ns
2	$X_1 \rightarrow Y_1$	$(\beta_{12} + \beta_{13})X_1$	0.001	5.541	<0.001*
	$X_1 \rightarrow Y_2$	$(\beta_{22} + \beta_{23})X_1$	0.419	-3.753	0.002*
	$Y_1 \rightarrow Y_2$	$(\beta_{24} + \beta_{25})Y_1$	-0.073	-5.628	<0.001*

Descriptions:
* : Significant at $\alpha = 0{,}05$
ns : Not significant at $\alpha = 0{,}05$

Based on Table 5, a significant or insignificant decision is obtained between exogenous variables on endogenous variables. Exogenous variables on endogenous variables are said to be significant if the p-value is less than 0.05 (marked with * on the p-value). In Group 1, the results show that Product Quality has a significant effect on Cashless User Satisfaction and Cashless User Loyalty. The higher the product quality will tend to increase cashless user satisfaction and loyalty. In some conditions, the higher the product quality can reduce the satisfaction and loyalty of cashless users. The decline is because user satisfaction and loyalty to the product are subjective and can change over time. Changes in trends or user preferences can make features that were once considered quality enhancements less relevant. As a result, despite product quality, users may not feel satisfied and loyal if the product does not meet people's current needs.

The Cashless User Satisfaction variable does not have a significant effect on Cashless User Loyalty in Group 1. Users with low frequency tend to have less attachment or strong habits with cashless payment methods. For people with low frequency of cashless use, satisfaction is not enough to form stable habits or preferences to continue using cashless methods loyally because people are not yet dependent on cashless methods in their daily lives. People who rarely use cashless may be more comfortable or loyal to traditional payment methods such as using cash. As a result, satisfaction in using cashless does not have a significant influence because people already have a strong preference for other methods.

In Group 2, the Product Quality variable has a significant effect on Cashless User Satisfaction and Cashless User Loyalty. Product Quality as a major factor in user experience. For users who frequently use cashless payment methods, product quality plays a major role in shaping user satisfaction. Quality here includes aspects such as product durability, conformity to specifications, and product durability. Product quality that meets users' expectations will create a satisfying experience and support users' preference to continue using cashless methods.

For frequent cashless users, product quality is not only a momentary factor but also a long-term factor that shapes the loyalty of cashless users. Efforts to maintain and improve product quality will be very effective in maintaining the loyalty of high-frequency user groups, as they rely more on these products in their daily lives. In addition, improving product quality will encourage users who rarely use cashless to adapt to advances in payment methods. The expectation of high product quality is that the use of cashless methods has increased.

Cashless User Satisfaction has a significant effect on Cashless User Loyalty. Cashless user satisfaction reflects the extent to which this product meets needs in terms of presence, responsiveness, and timeliness of service. Extensive service presence facilitates user access in various situations, while service responsiveness helps solve problems quickly and effectively, giving users a sense of security. Timeliness in transaction processing also increases users' confidence in the speed and reliability of cashless products. When these three aspects are consistently fulfilled, users feel satisfied and encouraged to continue using cashless services, ultimately forming strong loyalty.

The coefficient of determination of the estimated function of Product Quality on Cashless User Satisfaction is 0.6553 or 65.53%. The coefficient of determination of 65.53% indicates that Product Quality can explain Cashless User Satisfaction by 65.53% and the remaining 34.47% is explained by variables other than Product Quality. The coefficient of determination in estimating the function of Product Quality and Cashless User Satisfaction on Cashless User Loyalty is 0.8301 or 83.01%. The figure 83.01 indicates that Product Quality and Cashless User Satisfaction can explain Cashless User Loyalty by 83.01% and the remaining 16.99% is explained by variables other than Product Quality and Cahless User Satisfaction.

The two coefficients of determination of the estimated function can be seen that the additional Cashless User Satisfaction variable increases the influence by 17.48% on Cashless User Loyalty. Based on the results of the calculation, the total coefficient of determination is 94.14%, which indicates that Product Quality and Cashless Usage Frequency Moderation can explain Cashless User Satisfaction and Cashless User Loyalty by 94.14% and variables outside the research model explain the remaining 5.86%. Hasil koefisien determinasi ini tergolong sangat baik. In cashless data, researchers analyzed using parametric path analysis to obtain a coefficient of determination of 0.6313. These results are very small when compared to 0.9414. So based on the coefficient of determination, it can be concluded that for diffuse and non-patterned data patterns it is very suitable to use semiparametric path analysis.

CONCLUSIONS

Ramsey RESET development algorithm was successfully applied to identify the optimal smoothing spline polynomial order. The weakness of the Ramsey RESET development algorithm is the high computational complexity to determine the optimal spline smoothing polynomial order. The Ramsey RESET development is carried out on cashless society data, where the results of the cashless data linearity test are then used in developing the estimation of the semiparametric smoothing spline path analysis function in the condition of 2 groups using a dummy approach. Through the dummy approach, it can summarize the function estimation process and determine the difference in influence between Group 1 and Group 2.

Based on the results of hypothesis testing, banks can improve the quality of their mobile banking and internet banking. The effect of product quality on satisfaction is higher in people who rarely use cashless than people who often use cashless. Improving product quality can also affect the loyalty of using cashless products. Therefore, it is necessary to improve product quality in order to increase the frequency of using cashless for people who rarely use cashless methods and keep people who often use cashless methods from switching.

ACKNOWLEDGMENTS

I am deeply grateful to all who supported me in this research. Special thanks go to my advisor for invaluable guidance, the faculty and staff for their encouragement and resources, and my colleagues and friends for their insights. I also thank my family for their unwavering support and the research participants for their cooperation. Lastly, I acknowledge any funding bodies or organizations that contributed to this study (SK: 32/IV/KS/06/2022 and 1612.1/UN10.C10/TU/2022). Thank you all for your dedication and encouragement.

REFERENCES

[1] Marwan, Konadi, W., Kamaruddin, Sufi, I., & Akmal, Y. (2023). *Analisis Jalur dan Aplikasi SPSS Versi 25 Edisi Kedua.* Merdeka Kreasi.

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

[3] Hamid, M., Sufi, I., Konadi, W., & Yusrizal, A. (2017). *Analisis Jalur Dan Aplikasi Spss Versi 25 Edisi Pertama.* Kopelma Darussalam.

[4] Montgomery, D. C., Peck, E. A., & Vining, G. G. (2021). Introduction to Linear Regression Analysis. John Wiley & Sons.

[5] Solimun, & Fernandes, A. A. R. (2023). Innovation-Based Research Using Structural Flexibility and Acceptance Model (SFAM). *Cogent Business and Management. 10*(1), 2128255. https://doi.org/10.1080/23311975.2022.2128255

[6] Purnama, D. I. (2020). A Comparison between Nonparametric Approach: Smoothing Spline and B-Spline to Analyze The Total of Train Passangers in Sumatra Island. *EKSAKTA: Journal of Sciences and Data Analysis*. *1*(1), 73–80. https://doi.org/10.20885/EKSAKTA.vol1.iss1.art11

[7] Eubank, R. L. (2020). Nonparametric Regression and Spline Smoothing 2nd Edition. CRC Press.

[8] Kvam, P. H. & Vidakovic, B. (2019). Nonparametric Statistics with Applications to Science and Engineering 2nd Edition. John Wiley & Sons.

[9] Wahyuningsih, T. D., Handajani, S. S., & Indriati, D. (2018). *Penerapan Generalized Cross Validation dalam Model Regresi Smoothing Spline pada Produksi Ubi Jalar di Jawa Tengah. Indonesian Journal of Applied Statistics. 1*(2), 117-125. https://doi.org/10.13057/ijas.v1i2.26250

[10] Salam, N., Sukmawaty, Y., & Halida, A. (2022). *Estimasi Model Regresi Nonparametrik Dengan Metode B-Spline*. *Jurnal Sistem Media Bina Ilmiah. 16*(10), 7631–7638. https://doi.org/10.33578/mbi.v16i10.15

[11] Fernandes, A. A. R., Hutahayan, B., Solimun, Arisoesilaningsih, E., Yanti, I., Astuti, A. B., Nurjannah, & Amaliana, L. (2019). Comparison of Curve Estimation of the Smoothing Spline Nonparametric Function Path Based on PLS and PWLS In Various Levels of Heteroscedasticity. *IOP Conference Series: Materials Science and Engineering. 546*(5), 052024. https://doi.org/10.1088/1757-899X/546/5/052024

[12] Pratama, Y. M., Fernandes, A. A. R., Wardhani, N. W. S., & Hamdan, R. (2024). Nonparametric Smoothing Spline Approach in Examining Investor Interest Factors, *JTAM (Jurnal Teori Dan Aplikasi Matematika). 8*(2), 425-440.

https://doi.org/10.31764/jtam.v8i2.20192

[13] Massi, M., Sullivan, G., Strauß, M., & Khan, M. (2019). How Cashless Payments Help Economies Grow. Boston Consulting Group.

[14] Kotler, P., Keller, L. K., & Chernev, A. (2021). Marketing Management (16th ed.). Pearson Education.

[15] Rahi, S., Yasin, N. M., & Alnaser, F. M. (2017). Measuring the Role of Website Design, Assurance, Customer Service, and Brand Image Towards Customer Loyalty and Intention to Adopt Internet Banking. *Journal of Internet Banking and Commerce*. *22*(8), 1–18. https://www.icommercecentral.com/open-access/measuring-the-role-of-website-design-assurance-customer-service-and-brand-image-towards-customer-loyalty-and-intention-to-adopt-internet-banking.pdf

[16] Palilati, A. P., Umar, Z. A., & Niode, I. Y. (2022). *Pengaruh Kualitas Produk Dan Kepercayaan Terhadap Loyalitas Konsumen Dengan Kepuasan Konsumen Sebagai Variabel Intervening*. *Jurnal Ilmiah Manajemen dan Bisnis*. *5*(2), 534–542. https://doi.org/10.37479/jimb.v5i2.16492

[17] Fernandes, A. A. R., Solimun, & Arisoesilaningsih, E. (2017). Estimation of spline function in nonparametric path analysis based on penalized weighted least square (PWLS). *AIP Conference Proceedings*. *1913*(1). https://doi.org/10.1063/1.5016664

[18] Fernandes, A. A. R., Budiantara, I. N., Otok, B. W., & Suhartono. (2014). Spline estimator for bi-responses nonparametric regression model for longitudinal data, *Applied Mathematical Sciences*. *8*(114), 5-5665. https://doi.org/10.12988/ams.2014.47566

[19] Fernandes, A. A. R., Astuti, A. B., Amaliana, L., Yanti, I., Arisoesilaningsih, E., & Isaskar, R. (2019). *S*moothing Spline Nonparametric Path: Application for Green Product and Green Marketing Strategy towards Green Product Purchasing Intention. *In IOP Conference Series: Earth and Environmental Science. 239*(1), 012018. https://doi.org/10.1088/1755-1315/239/1/012018

[20] Zebua, H. I. (2021). *Pemodelan Kemiskinan di Sumatera Utara Menggunakan Regresi Nonparametrik Kernel dan Splines. Seminar Nasional Official Statistics. 2021*(1), 899–907. https://doi.org/10.34123/semnasoffstat.v2021i1.1087

[21] Lukas, M. A., Hoog, F. R. D., & Anderssen R. S. (2016). Practical Use of Robust GCV and Modified GCV for Spline Smoothing. *Computational Statistics*, *31*(1), 269-289. https://doi.org/10.1007/s00180-015-0577-7

[22] Ramsey, J. B. (1969). Tests for Specification Errors in Classical Linear Least-Squares Regression Analysis. *Journal of the Royal Statistical Society Series B: Statistical Methodology*. *31*(2), 350–371. https://www.jstor.org/stable/2984219

[23] Gujarati, D. N., & Porter, D. C. (2020). Essentials of Econometrics (5th ed.), McGraw-Hill.

[24] Lestari, B., Budiantara, I. N., & Chamidah, N. (2019). Smoothing Parameter Selection Method for Multiresponse Nonparametric Regression Model Using Smoothing Spline and Kernel Estimators Approaches. *In Journal of Physics: Conference Series*. 1397(1), 012064. https://doi.org/10.1088/1742-6596/1397/1/012064

[25] Peng, M., Xiang, L., & Wang, S. (2018). Semiparametric regression analysis of clustered survival data with semi-competing risks. *Computational Statistics & Data Analysis. 124*, 53-70. https://doi.org/10.1016/j.csda.2018.02.003