PARAMETER
JURNAL MATEMATIKA, STATISTIKA DAN TERAPANNYA

e-ISSN. 2829-6303
Vol. 05 No.1, April 2026
Pages 207-218

Check for updates

Parameter Estimation of Partial Spline Regression Using Weighted Least Squares with Moving Average Approach

Jiran Julita [1*], **Jerry Dwi Trijoyo Purnomo** [2], **I Nyoman Budiantara** [3]

[1,2,3]Department of Statistics, Faculty of Science and Data Analytics, Institut Teknologi Sepuluh Nopember

Jl. Mechanical Engineering No. 175, Keputih, Sukolilo District, Surabaya, East Java 60111, Indonesia

E-mail Correspondence Author: *jiranjulita57@mail.com

Abstract

Semiparametric regression combines parametric and nonparametric components within a single framework. The partial spline model adopts a linear parametric component and a truncated spline nonparametric component. When heteroscedasticity is present, Ordinary Least Squares (OLS) loses efficiency and Weighted Least Squares (WLS) is required. When the variance structure is unknown, the weight matrix must be estimated from the data. Silverman (1985) proposed a moving average of squared generalized residuals for nonparametric regression, but this procedure cannot be directly extended to the partial spline setting due to the simultaneous presence of both components. This study derives closed-form WLS estimators $\widehat{\boldsymbol{\beta}} = \boldsymbol{E}(K)\boldsymbol{W}\boldsymbol{y}$ *and* $\widehat{\boldsymbol{\gamma}} = \boldsymbol{F}(K)\boldsymbol{W}\boldsymbol{y}$ *via partial differentiation and substitution, yielding the regression curve estimator* $\widehat{\mathbf{Y}} = \mathbf{T}(X, Z, K)\mathbf{W}\mathbf{y}$ *with optimal knots selected by minimizing the Generalized Cross-Validation (GCV) criterion. The proposed estimators provide an efficient and theoretically grounded solution for partial spline regression under heteroscedastic error conditions whose functional form is unknown.*

Keywords: *Generalized cross validation, moving average, semiparametric regression, truncated spline, weighted least squares.*

: https://doi.org/10.30598/parameterv5i1pp207-218

Research Article ***Open Acces***
Email: jurnalparameter@gmail.com
Homepage: https://ojs3.unpatti.ac.id/index.php/parameter

Submitted: February 2026
Accepted: April 2026

1. INTRODUCTION

Regression analysis is one of the fundamental statistical tools used to examine and estimate the functional relationship between a response variable and one or more predictor variables [1]. Based on the knowledge about the shape of the regression curve, the estimation approach is classified into three categories: parametric, nonparametric, and semiparametric [2]. The parametric approach is used when the shape of the regression curve is entirely known, such as linear or quadratic forms. The nonparametric approach is employed when the curve shape is completely unknown and the data are allowed to speak for themselves [3]. In many real-world situations, however, neither approach alone is adequate, as part of the relationship between variables follows a known structure while another part does not. This motivates the semiparametric regression model, which combines both parametric and nonparametric components within a single framework [4], [5].

One of the most widely studied forms of semiparametric regression is the partial spline model, in which the parametric component assumes a linear pattern and the nonparametric component is approximated using spline functions [6]. The partial spline model was first formally introduced by Engle et al. [7] in the context of modeling the relationship between weather conditions and electricity sales, and has since been applied across a wide range of disciplines. Among available spline estimators, the truncated spline has received considerable attention due to its excellent statistical interpretability and its flexibility in accommodating changes in data behavior across different sub-intervals of the predictor domain [8], [9], [10]. The truncated spline estimator belongs to the class of piecewise polynomial functions joined at points called knots, and the accuracy of the resulting model depends heavily on the appropriate selection of these knot points [11]. Optimal knot selection is typically achieved by minimizing the Generalized Cross-Validation (GCV) criterion, which possesses asymptotically optimal properties, is invariant under linear transformation, and does not require prior knowledge of the error variance [12], [13], [14].

Parameter estimation in partial spline regression is commonly performed using either the Ordinary Least Squares (OLS) or the Weighted Least Squares (WLS) method [15]. The OLS estimator produces the Best Linear Unbiased Estimator (BLUE) only when the classical assumptions hold, in particular the homoscedasticity assumption-that is, the error variance is constant across all observations [16]. When this assumption is violated and heteroscedasticity is present, the OLS estimator remains unbiased but loses its minimum variance property, making it no longer the most efficient estimator [17]. Furthermore, parameter inference based on OLS under heteroscedasticity-including hypothesis tests and confidence intervals-may produce misleading results. The WLS method addresses this problem by assigning each observation a weight inversely proportional to its error variance, thereby down-weighting high-variability observations and restoring the efficiency of the estimator [18].

A fundamental challenge in implementing WLS is that the error variances are rarely known in advance and must be estimated from the data. When the weight matrix is misspecified-for instance, assumed to be the identity matrix despite the presence of heteroscedasticity-the resulting estimator may remain inefficient [19]. Silverman [20] proposed an iterative reweighting procedure to address this in the nonparametric regression context, in which the weights are estimated via a local moving average of squared generalized residuals. In this procedure, an initial unweighted fit is obtained, from which generalized residuals r_i^* are computed. These squared residuals are then smoothed over a local window of k neighboring observations to produce an adaptive variance estimate $\hat{w}_i^{-1}$ for each data point. This approach captures the local variance structure of the data without requiring a parametric specification of the heteroscedasticity function, making it especially suitable when the form of variance non-constancy is unknown.

However, a critical research gap exists: the local moving average weight estimation procedure of Silverman [20] was developed exclusively for the nonparametric regression setting and cannot be directly applied to the partial spline model. This limitation arises from three fundamental complications. First, computing the generalized residuals in partial spline regression requires an initial simultaneous estimate from both the parametric component $g(x_i) = \beta_0 + \beta_1 x_i$ and the nonparametric component $f(z_i)$, whereas Silverman's [20] procedure assumes a single nonparametric component only. Second, the hat matrix in partial spline regression takes the combined form $\mathbf{T}(\mathbf{X}, \mathbf{Z}, K) = \mathbf{XE}(K) + \mathbf{ZF}(K)$, which differs structurally from the simpler hat matrix in Silverman's [20] nonparametric framework, requiring a different derivation of $\hat{\sigma}^2$ and r_i^*. Third, minimizing the WLS criterion in partial spline regression yields two mutually dependent estimating equations for $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\gamma}}$, requiring a substitution procedure to obtain independent closed-form estimators-a step entirely absent in the purely nonparametric setting. Despite the broad development of semiparametric regression with spline truncated estimators [8], [9], [10], [15], none of the existing studies has derived WLS estimators for partial spline regression with a weight matrix constructed adaptively from a local moving average of generalized residuals, leaving an open methodological problem in efficient estimation under unknown heteroscedastic error conditions.

To address this gap, this study proposes a novel approach: the derivation of closed-form WLS parameter estimators for a partial spline regression model with a linear parametric component and a spline truncated nonparametric component, where the weight matrix is constructed adaptively from a local moving average of squared generalized residuals following Silverman [20]. The novelty of this study lies in three contributions: (1) the extension of Silverman's [20] local moving average reweighting procedure to the partial spline semiparametric framework, (2) the derivation of independent closed-form estimators $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\gamma}}$ via partial differentiation and substitution under the adapted weight matrix [21], and (3) the integration of GCV-based optimal knot selection into the WLS estimation procedure for partial spline regression. The proposed approach contributes a methodologically rigorous and efficient estimation procedure for semiparametric regression under heteroscedastic error conditions whose form is unknown.

2. THEORETICAL REVIEW

2.1 Truncated Spline Semiparametric Regression

Truncated splines constitute a widely employed nonparametric regression framework. They are smooth, segmented polynomial functions capable of adapting to varying data patterns. Model estimation in spline regression depends critically on knot points-intersection points that demarcate transitions in the pattern and functional behavior across sub-intervals [22]. Given paired data (y_i, x_i, z_i), $i = 1, 2, \dots, n$, where y_i is the response variable while x_i is the predictor variable that follows a parametric pattern and z_i is the predictor variable that follows a nonparametric pattern. The relationship pattern of y_i, x_i and z_i can be expressed in a regression model as follows:

$$y_i = \mathbf{x}_i' \boldsymbol{\beta} + f(z_i) + \varepsilon_i, \qquad i = 1, 2, \dots, n \tag{1}$$

Next, if the regression curve $f(z_i)$ in **Equation (1)** is approximated by a truncated spline function with knots $K_1, K_2, \dots, K_r$ then:

$$f(z_i) = \gamma_1 z_i + \gamma_2 z_i^2 + \cdots + \gamma_m z_i^m + \gamma_{1+m}(z_1 - K_1)_+^m + \cdots + \gamma_{r+m}(z_i - K_r)_+^m \quad (2)$$

$$f(z_i) = \sum_{j=1}^{m} \gamma_j z_i^j + \sum_{k=1}^{r} \gamma_{k+m}(z_i - K_k)_+^m \quad (3)$$

With

$$(z_i - K_k)_+^m = \begin{cases} (z_i - K_k)^m & , z_i \geq K_k \\ 0 & , z_i < K_k \end{cases} \quad (4)$$

γ is a parameter and $(z_i - K_k)_+^m$ is a truncated function. Degree m is the degree of the polynomial equation. Typically used m = 1, 2, and 3. A first-degree polynomial regression curve is usually called a linear regression curve, a second-degree polynomial regression curve is usually called a quadratic regression curve, while a third-degree polynomial regression curve is usually called a cubic regression curve. The knot points K_1, $K_2, \dots, K_r$ are knot points that indicate the curve's behavior pattern at different subintervals.

As a simple illustration, suppose a truncated linear spline with $m = 1$, $r = 3$, three knot points at $z_{1i} = K_1$, $z_{2i} = K_2$, $z_{3i} = K_3$ is given. Then, it can be presented in the form:

$$f(z_i) = \gamma_1 z_i + \gamma_2(z_i - K_1)_+^1 + \gamma_3(z_i - K_2)_+^1 + \gamma_4(z_i - K_3)_+^1 \quad (5)$$

The spline function $f(z_i)$ can also be presented in the form:

$$f(z_i) = \begin{cases} \gamma_1 z_i & , z_i < K_1 \\ \gamma_1 z_i + \gamma_2(z_i - K_1)_+^1 & , K_1 \leq z_i < K_2 \\ \gamma_1 z_i + \gamma_2(z_i - K_1)_+^1 + \gamma_3(z_i - K_2)_+^1 & , K_1 \leq z_i < K_2 \\ \gamma_1 z_i + \gamma_2(z_i - K_1)_+^1 + \gamma_3(z_i - K_2)_+^1 + \gamma_4(z_i - K_3)_+^1 & , z_i \geq K_3 \end{cases} \quad (6)$$

The semiparametric truncated spline regression equation in **Equation (1)** can be expressed as follows:

$$y_i = \mathbf{x}_\mathbf{i}'\boldsymbol{\beta} + \sum_{j=1}^{m} \gamma_j z_i^j + \sum_{k=1}^{r} \gamma_{k+m}(z_i - K_k)_+^m + \varepsilon_i \quad (7)$$

where y_i is the i-th response variable, $\mathbf{x}_\mathbf{i}' = [1 \quad x_{i1} \quad x_{i2} \quad \cdots \quad x_{ip}]$ is the predictor variable for the parametric component, z_i is the predictor variable for the nonparametric component, $\mathbf{x}_\mathbf{i}'\boldsymbol{\beta}$ is the parametric component, $\boldsymbol{\beta} = [\beta_0 \quad \beta_1 \quad \cdots \quad \beta_h \quad \cdots \quad \beta_p]_{(p+1)\times 1}^T$ unknown parameters. $f(z_i)$ is a nonparametric component function whose pattern is unknown and ε_i is a random error, where $\varepsilon_i \sim \mathrm{N}(0, \sigma^2)$.

2.2 Moving Average (MA)

In the context of nonparametric and semiparametric regression, the local MA approach is used to estimate weights based on residual variance patterns. Smoothing with MA can reduce the influence of extreme random variations and highlight local trends, so that the resulting weights can more accurately reflect heteroscedasticity characteristics [23]. The weight values are obtained from the local moving average using the following formula [20]:

$$\widehat{w}_i^{-1} = (n_i - m_i + 1)^{-1} \sum_{t=m_i}^{n_i} r_{it}^{*2} \tag{8}$$

with $m_i = \max(1, i-k)$ and $n_i = \min(n, i+k)$, where k is the window size and r_i^* is generalized residuals as follows:

$$r_i^* = \frac{w_i^{\frac{1}{2}}\{y_i - \hat{y}(x_i, z_i)\}}{\hat{\sigma}\{1 - n^{-1} tr\mathbf{T}(\mathbf{X}, \mathbf{Z}, K)\mathbf{W}\}^{\frac{1}{2}}} \tag{9}$$

where $\hat{y}(x_i, z_i)$ is semiparametric model estimator, $\hat{\sigma}$ is the square root of the residual variance estimate $\hat{\sigma}^2$ which is calculated as follows:

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n} w_i\{y_i - \hat{y}(x_i, z_i)\}^2}{n - tr\mathbf{T}(\mathbf{X}, \mathbf{Z}, K)\mathbf{W}} \tag{10}$$

2.3 Weighted Least Squares

Two requirements must be met for parametric or semiparametric regression analysis: the variance of the random error in the model is assumed to be homogenous, and the variance-covariance error matrix is known [8]. Without sacrificing its bias and consistency, the WLS approach preserves the estimator's efficiency. With the exception of adding a new variable called $\boldsymbol{w_i}$ as a weighting factor, the WLS approach is substantially the same as the least squares method [24]. WLS (Weighted Least Square) is a technique that can be utilized to satisfy the first premise. By reducing the number of squares error between the observation and the model, the WLS approach is utilized to estimate the parameter.

$$\mathbf{W} = \begin{bmatrix} \widehat{w}_1^{-1} & 0 & \cdots & 0 \\ 0 & \widehat{w}_2^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \widehat{w}_n^{-1} \end{bmatrix} \tag{11}$$

$$\min_{\boldsymbol{\beta},\boldsymbol{\gamma}}\{(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\boldsymbol{\gamma})'\mathbf{W}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\boldsymbol{\gamma})\} \tag{12}$$

2.4 Generalized Cross Validation

GCV is a statistical optimization technique employed to identify the most suitable parameters in a regression model. The GCV-based approach to knot selection is considered optimal owing to its asymptotically optimal properties, which hold even for large samples, making it closely related to sample size behavior. GCV is also invariant in the sense that its value remains stable even when the dataset undergoes modifications [25].

$$\begin{aligned} GCV(K) &= \frac{n^{-1}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\left(n^{-1} trace(\mathbf{I} - \mathbf{T}(\mathbf{X}, \mathbf{Z}, K)\mathbf{W})\right)^2} \\ &= \frac{MSE(K)}{\left(n^{-1} trace(\mathbf{I} - \mathbf{T}(\mathbf{X}, \mathbf{Z}, K)\mathbf{W})\right)^2} \end{aligned} \tag{13}$$

where $\mathbf{T}(X, Z, K)\mathbf{W}$ is the hat matrix of the partial spline model.

2.5 Analysis Steps

The analysis steps to obtain weight estimates in the partial spline model using the moving average approach are as follows:

1) Assume paired data $(x_{i1}, x_{i2}, \dots, x_{ip}, z_{i1}, z_{i2}, \dots, z_{iq}, y_i)$ follows a semiparametric regression model

$$y_i = g(x_{i1}, x_{i2}, \dots, x_{ip}) + f(z_{i1}, z_{i2}, \dots, z_{iq}) + \varepsilon_i \tag{14}$$

2) The regression curve $g(x_i)$ is a parametric component, approximated by a linear function

$$g(x_{i1}, x_{i2}, \dots, x_{ip}) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} \tag{15}$$

3) The regression curve $f(z_i)$ is a nonparametric component, approximated by a truncated spline function

$$f(z_{i1}, z_{i2}, \dots, z_{iq}) = \sum_{l=1}^{q} \left(\sum_{j=1}^{m} \gamma_{jl} z_{il}^{j} + \sum_{k=1}^{r} \gamma_{(m+k)l} (z_{il} - K_{kl})_{+}^{m} \right) \tag{16}$$

4) Given a truncated spline semiparametric regression model

$$y_i = \sum_{h=0}^{p} \beta_h x_{ih} + \sum_{l=1}^{q} \left(\sum_{j=1}^{m} \gamma_{jl} z_{il}^{j} + \sum_{k=1}^{r} \gamma_{(m+k)l} (z_{il} - K_{kl})_{+}^{m} \right) + \varepsilon_i \tag{17}$$

It can also be presented in the following matrix form:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon} \tag{18}$$

5) Obtaining parameter estimates $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ by completing Weighted Least Square optimization, determining the value of the weight matrix $\mathbf{W}$ using the moving average approach in **Equation (8)**

$$\min_{\boldsymbol{\beta},\boldsymbol{\gamma}} \{ (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\boldsymbol{\gamma})' \mathbf{W} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\boldsymbol{\gamma}) \} \tag{19}$$

6) Obtain the weighted truncated spline semiparametric regression model estimator.

$$\hat{y}_i = \sum_{h=0}^{p} \hat{\beta}_h x_{ih} + \sum_{l=1}^{q} \left(\sum_{j=1}^{m} \hat{\gamma}_{jl} z_{il}^{j} + \sum_{k=1}^{r} \hat{\gamma}_{(m+k)l} (z_{il} - K_{kl})_{+}^{m} \right) \tag{20}$$

3. RESULTS AND DISCUSSION

3.1. Truncated Spline Semiparametric Regression Model

Suppose a set of observational data is available in the form of paired data, namely $(x_{i1}, x_{i2}, \dots, x_{ip}, z_{i1}, z_{i2}, \dots, z_{iq}, y_i)$ The data is assumed to have a functional relationship between the response variable (y) and the predictor variable consisting of two components, namely the parametric component predictor variable and the nonparametric component predictor variable. The relationship between these variables is modeled using semiparametric regression (partial splines), which can be generally expressed in **Equation (14)** The model states that the response variable (y) is the sum of two principal components: the parametric component $g(x_{i1}, x_{i2}, \dots, x_{ip})$ and the nonparametric component $f(z_{i1}, z_{i2}, \dots, z_{iq})$, and the random ε_i assumed to be normally distributed with a mean of zero and a variance of σ^2.

The parametric component $g(x_{i1}, x_{i2}, \dots, x_{ip})$ represents the linear relationship between the response variable and the predictor variables. In this study, these components are approximated using a linear function expressed as follows:

$$g(x_{i1}, x_{i2}, \dots, x_{ip}) = \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip} \tag{22}$$

where β_0 is the intercept and $\beta_1, \beta_2, \dots, \beta_p$ are regression parameters indicating the magnitude of the influence of each parametric variable on the response variable. Meanwhile, the nonparametric component $f(z_{i1}, z_{i2}, \dots, z_{iq})$ is used to model nonlinear relationships whose functional form is not explicitly known. In this study, this nonparametric component is approximated using a truncated spline function of degree m with a number of knots $(K_1, K_2, \dots, K_r)$. This approach allows the model flexibility in capturing changes in data patterns at certain intervals. Mathematically, this nonparametric function can be written as the following equation:

$$f(z_{i1}, z_{i2}, \dots, z_{iq}) = f(z_{i1}) + f(z_{i2}) + \cdots + f(z_{iq}) = \sum_{l=1}^{q} f(z_{il}) \tag{23}$$

where,

$$\sum_{l=1}^{q} f(z_{il}) = \sum_{l=1}^{q} \left(\sum_{j=1}^{m} \gamma_{jl} z_{il}^{j} + \sum_{k=1}^{r} \gamma_{(m+k)l} (z_{il} - K_{kl})_+^m \right), \quad i = 1,2, \dots, n \tag{24}$$

where γ_{jl} and $\gamma_{(k+m)l}$ are the spline parameters to be estimated, while $(z_{il} - K_{kl})_+$ is a truncated spline function that is zero for $z_{il} < K_{kl}$ and $(z_{il} - K_{kl})$ for $z_{il} \geq K_{kl}$.
By substituting the parametric and nonparametric components into the semiparametric regression model (partial spline), the complete semiparametric spline regression model equation can be written as follows:

$$y_i = \sum_{h=0}^{p} \beta_h x_{ih} + \sum_{l=1}^{q} \left(\sum_{j=1}^{m} \gamma_{jl} z_{il}^{j} + \sum_{k=1}^{r} \gamma_{(m+k)l} (z_{il} - K_{kl})_+^m \right) + \varepsilon_i \quad , i = 1,2, \dots, n \tag{25}$$

Based on Equation (25) it can be explained as follows:

$$\begin{aligned} y_i = {} & \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + \gamma_{11} z_{i1}^1 + \gamma_{21} z_{i1}^2 + \cdots + \gamma_{m1} z_{i1}^m + \\ & \gamma_{(m+1)1} (z_{i1} - K_{11})_+^m + \gamma_{(m+2)1} (z_{i1} - K_{21})_+^m + \cdots + \gamma_{(m+r)1} (z_{i1} - K_{r1})_+^m \\ & + \cdots + \gamma_{1q} z_{iq}^1 + \gamma_{2q} z_{iq}^2 + \cdots + \gamma_{mq} z_{iq}^m + \gamma_{(m+1)q} (z_{iq} - K_{1q})_+^m + \gamma_{(m+2)q} \\ & (z_{iq} - K_{2q})_+^m + \cdots + \gamma_{(m+r)q} (z_{iq} - K_{rq})_+^m + \varepsilon_i \end{aligned} \tag{26}$$

Equation (26) can also be presented in the following matrix form like Equation (18) With

$$\mathbf{Y} = [y_1 \quad y_2 \quad \cdots \quad y_n]_{(n\times1)}^T \tag{28}$$

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{p1} \\ 1 & x_{12} & x_{22} & \cdots & x_{p2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & x_{2n} & \cdots & x_{pn} \end{bmatrix}_{n\times(p+1)} \tag{29}$$

$$\boldsymbol{\beta} = [\beta_0 \quad \beta_1 \quad \cdots \quad \beta_p]_{(p+1)\times1}^T \tag{30}$$

$$\boldsymbol{\varepsilon} = [\varepsilon_1 \quad \varepsilon_2 \quad \cdots \quad \varepsilon_n]_{(n\times1)}^T \tag{31}$$

$$\boldsymbol{\gamma} = [\gamma_{11} \quad \cdots \quad \gamma_{m1} \quad \gamma_{(m+1)1} \quad \cdots \quad \gamma_{(m+r)1} \quad \cdots \quad \gamma_{1q} \quad \cdots \quad \gamma_{mq} \quad \gamma_{(m+1)q} \quad \cdots \quad \gamma_{(m+r)q}]_{1\times(m+r)q}^T \tag{32}$$

$$\mathbf{Z} = \begin{bmatrix} z_{11} & \cdots & z_{11}^m & (z_{11} - K_{11})_+^m & \cdots & (z_{11} - K_{r1})_+^m & \cdots & z_{1q} & \cdots & z_{1q}^m & (z_{1q} - K_{1q})_+^m & \cdots & (z_{1q} - K_{rq})_+^m \\ z_{21} & \cdots & z_{21}^m & (z_{21} - K_{11})_+^m & \cdots & (z_{21} - K_{r1})_+^m & \cdots & z_{2q} & \cdots & z_{2q}^m & (z_{2q} - K_{1q})_+^m & \cdots & (z_{2q} - K_{rq})_+^m \\ \vdots & \ddots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \ddots \\ z_{n1} & \cdots & z_{n1}^m & (z_{n1} - K_{11})_+^m & \cdots & (z_{n1} - K_{r1})_+^m & \cdots & z_{nq} & \cdots & z_{nq}^m & (z_{nq} - K_{1q})_+^m & \cdots & (z_{nq} - K_{rq})_+^m \end{bmatrix}_{n\times(m+r)q} \tag{33}$$

3.2 Estimation of Semiparametric Regression Model in Weighted Least Squares Method with Moving Average

To overcome the problem of heteroscedasticity, an estimation is carried out using the Weighted Least Squares (WLS) method. Furthermore, by using the weighting matrix W, the weight values are obtained in **Equation (35)** through the moving average approach with the W matrix, namely

$$\mathbf{W} = \begin{bmatrix} \widehat{w}_1^{-1} & 0 & \cdots & 0 \\ 0 & \widehat{w}_2^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \widehat{w}_n^{-1} \end{bmatrix} \quad (34)$$

With

$$\widehat{w}_i^{-1} = (n_i - m_i + 1)^{-1} \sum_{t=m_i}^{n_i} r_{it}^{*2} \quad (35)$$

The estimates of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ in **Equation (27)** can be obtained by solving the optimization using Weighted Least Square (WLS).

$$\boldsymbol{\varepsilon} = \mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\boldsymbol{\gamma} \quad (36)$$

$$\min_{\boldsymbol{\beta},\boldsymbol{\gamma}}\{(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\boldsymbol{\gamma})'\mathbf{W}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\boldsymbol{\gamma})\} \quad (37)$$

The optimization solution above is carried out with the following explanation.

$$\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\boldsymbol{\gamma})'\mathbf{W}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\boldsymbol{\gamma}) \quad (38)$$

$$= (\mathbf{y}' - \boldsymbol{\beta}'\mathbf{X}' - \boldsymbol{\gamma}'\mathbf{Z}')\mathbf{W}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\boldsymbol{\gamma}) \quad (39)$$

$$= (\mathbf{y}' - \boldsymbol{\beta}'\mathbf{X}' - \boldsymbol{\gamma}'\mathbf{Z}')(\mathbf{W}\mathbf{y} - \mathbf{W}\mathbf{X}\boldsymbol{\beta} - \mathbf{W}\mathbf{Z}\boldsymbol{\gamma}) \quad (40)$$

$$= (\mathbf{y}'\mathbf{W}\mathbf{y} - \mathbf{y}'\mathbf{W}\mathbf{X}\boldsymbol{\beta} - \mathbf{y}'\mathbf{W}\mathbf{Z}\boldsymbol{\gamma} - \boldsymbol{\beta}'\mathbf{X}'\mathbf{W}\mathbf{y} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{W}\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{W}\mathbf{Z}\boldsymbol{\gamma} - \boldsymbol{\gamma}'\mathbf{Z}'\mathbf{W}\mathbf{y} + \boldsymbol{\gamma}'\mathbf{Z}'\mathbf{W}\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\gamma}'\mathbf{Z}'\mathbf{W}\mathbf{Z}\boldsymbol{\gamma}) \quad (41)$$

$$= (\mathbf{y}'\mathbf{W}\mathbf{y} - 2\boldsymbol{\beta}'\mathbf{X}'\mathbf{W}\mathbf{y} - 2\boldsymbol{\gamma}'\mathbf{Z}'\mathbf{W}\mathbf{y} + 2\boldsymbol{\beta}'\mathbf{X}'\mathbf{W}\mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{W}\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\gamma}'\mathbf{Z}'\mathbf{W}\mathbf{Z}\boldsymbol{\gamma}) \quad (42)$$

$$= Q(\boldsymbol{\beta},\boldsymbol{\gamma}) \quad (43)$$

To obtain an estimate of the parameters, partial derivatives are performed. The partial derivative of $\boldsymbol{Q(\beta,\gamma)}$ is given by:

$$\frac{\partial(Q(\boldsymbol{\beta},\boldsymbol{\gamma}))}{\partial(\boldsymbol{\beta})} = \frac{\partial(\mathbf{y}'\mathbf{W}\mathbf{y} - 2\boldsymbol{\beta}'\mathbf{X}'\mathbf{W}\mathbf{y} - 2\boldsymbol{\gamma}'\mathbf{Z}'\mathbf{W}\mathbf{y} + 2\boldsymbol{\beta}'\mathbf{X}'\mathbf{W}\mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{W}\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\gamma}'\mathbf{Z}'\mathbf{W}\mathbf{Z}\boldsymbol{\gamma})}{\partial(\boldsymbol{\beta})} \quad (44)$$

$$\frac{\partial(Q(\boldsymbol{\beta},\boldsymbol{\gamma}))}{\partial(\boldsymbol{\beta})} = -2\mathbf{X}'\mathbf{W}\mathbf{y} + 2\mathbf{X}'\mathbf{W}\mathbf{Z}\boldsymbol{\gamma} + \mathbf{2X}'\mathbf{W}\mathbf{X}\boldsymbol{\beta} \quad (45)$$

if the partial derivative above is equated to zero, namely:

$$\frac{\partial(Q(\boldsymbol{\beta},\boldsymbol{\gamma}))}{\partial(\boldsymbol{\beta})} = \mathbf{0} \quad (46)$$

obtained:

$$-2\mathbf{X}'\mathbf{W}\mathbf{y} + 2\mathbf{X}'\mathbf{W}\mathbf{Z}\boldsymbol{\gamma} + \mathbf{2X}'\mathbf{W}\mathbf{X}\widehat{\boldsymbol{\beta}} = \mathbf{0} \quad (47)$$

$$\mathbf{2X}'\mathbf{W}\mathbf{X}\widehat{\boldsymbol{\beta}} = 2\mathbf{X}'\mathbf{W}\mathbf{y} - 2\mathbf{X}'\mathbf{W}\mathbf{Z}\boldsymbol{\gamma} \quad (48)$$

So the parameter estimator $\widehat{\boldsymbol{\beta}}$ is as follows:

$$\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}(\mathbf{X}'\mathbf{W}\mathbf{y} - \mathbf{X}'\mathbf{W}\mathbf{Z}\boldsymbol{\gamma}) \quad (49)$$

Next, to obtain the estimator $\boldsymbol{\gamma}$, we perform partial derivatives.

$$\frac{\partial(Q(\boldsymbol{\beta},\boldsymbol{\gamma}))}{\partial(\boldsymbol{\gamma})} = \frac{\partial(\mathbf{y'Wy} - 2\boldsymbol{\beta}'\mathbf{X'Wy} - 2\boldsymbol{\gamma}'\mathbf{Z'Wy} + 2\boldsymbol{\beta}'\mathbf{X'WZ}\boldsymbol{\gamma} + \boldsymbol{\beta}'\mathbf{X'WX}\boldsymbol{\beta} + \boldsymbol{\gamma}'\mathbf{Z'WZ}\boldsymbol{\gamma})}{\partial(\boldsymbol{\gamma})} \tag{50}$$

$$\frac{\partial(Q(\boldsymbol{\beta},\boldsymbol{\gamma}))}{\partial(\boldsymbol{\gamma})} = -2\mathbf{Z'Wy} + 2\mathbf{Z'WX}\boldsymbol{\beta} + 2\mathbf{Z'WZ}\boldsymbol{\gamma} \tag{51}$$

Then **Equation (51)** is equated to zero, namely:

$$\frac{\partial(Q(\boldsymbol{\beta},\boldsymbol{\gamma}))}{\partial(\boldsymbol{\gamma})} = \mathbf{0} \tag{52}$$

obtained:

$$-2\mathbf{Z'Wy} + 2\mathbf{Z'WX}\boldsymbol{\beta} + 2\mathbf{Z'WZ}\hat{\boldsymbol{\gamma}} = \mathbf{0} \tag{53}$$

$$2\mathbf{Z'WZ}\hat{\boldsymbol{\gamma}} = 2\mathbf{Z'Wy} - 2\mathbf{Z'WX}\boldsymbol{\beta} \tag{54}$$

So that the parameter estimator $\hat{\boldsymbol{\gamma}}$ is obtained

$$\hat{\boldsymbol{\gamma}} = (\mathbf{Z'WZ})^{-1}(\mathbf{Z'Wy} - \mathbf{Z'WX}\boldsymbol{\beta}) \tag{55}$$

Because the estimator $\hat{\boldsymbol{\beta}}$ still contains $\boldsymbol{\gamma}$ and the estimator $\hat{\boldsymbol{\gamma}}$ still contains $\boldsymbol{\beta}$, the following substitution is made between **Equation (49)** and **Equation (55)**.

$$\hat{\boldsymbol{\beta}} = (\mathbf{X'WX})^{-1}\left(\mathbf{X'Wy} - \mathbf{X'WZ}\left((\mathbf{Z'WZ})^{-1}(\mathbf{Z'Wy} - \mathbf{Z'WX}\boldsymbol{\beta})\right)\right) \tag{56}$$

$$= (\mathbf{X'WX})^{-1}\left(\mathbf{X'Wy} - \mathbf{X'WZ}\left((\mathbf{Z'WZ})^{-1}\mathbf{Z'Wy} - (\mathbf{Z'WZ})^{-1}\mathbf{Z'WX}\boldsymbol{\beta}\right)\right) \tag{57}$$

$$= (\mathbf{X'WX})^{-1}\left(\mathbf{X'Wy} - \mathbf{X'WZ}(\mathbf{Z'WZ})^{-1}\mathbf{Z'Wy} + \mathbf{X'WZ}(\mathbf{Z'WZ})^{-1}\mathbf{Z'WX}\boldsymbol{\beta}\right) \tag{58}$$

$$= (\mathbf{X'WX})^{-1}\mathbf{X'Wy} - (\mathbf{X'WX})^{-1}\mathbf{X'WZ}(\mathbf{Z'WZ})^{-1}\mathbf{Z'Wy} + (\mathbf{X'WX})^{-1}\mathbf{X'WZ}(\mathbf{Z'WZ})^{-1}\mathbf{Z'WX}\boldsymbol{\beta} \tag{59}$$

Then the components containing $\hat{\boldsymbol{\beta}}$ are combined as follows:

$$\hat{\boldsymbol{\beta}} - \left[(\mathbf{X'WX})^{-1}\mathbf{X'WZ}(\mathbf{Z'WZ})^{-1}\mathbf{Z'WX}\right]\hat{\boldsymbol{\beta}} = (\mathbf{X'WX})^{-1}\mathbf{X'Wy} - (\mathbf{X'WX})^{-1}\mathbf{X'WZ}(\mathbf{Z'WZ})^{-1}\mathbf{Z'Wy} \tag{60}$$

$$\hat{\boldsymbol{\beta}}\left[\mathbf{I} - (\mathbf{X'WX})^{-1}\mathbf{X'WZ}(\mathbf{Z'WZ})^{-1}\mathbf{Z'WX}\right] = (\mathbf{X'WX})^{-1}\mathbf{X'Wy} - (\mathbf{X'WX})^{-1}\mathbf{X'WZ}(\mathbf{Z'WZ})^{-1}\mathbf{Z'Wy} \tag{61}$$

So the estimator $\hat{\boldsymbol{\beta}}$ is obtained by the following equation:

$$\hat{\boldsymbol{\beta}} = \left[\mathbf{I} - (\mathbf{X'WX})^{-1}\mathbf{X'WZ}(\mathbf{Z'WZ})^{-1}\mathbf{Z'WX}\right]^{-1}\left[(\mathbf{X'WX})^{-1}\mathbf{X'Wy} - (\mathbf{X'WX})^{-1}\mathbf{X'WZ}(\mathbf{Z'WZ})^{-1}\mathbf{Z'Wy}\right] \tag{62}$$

$$= \left[\mathbf{I} - (\mathbf{X'WX})^{-1}\mathbf{X'WZ}(\mathbf{Z'WZ})^{-1}\mathbf{Z'WX}\right]^{-1}(\mathbf{X'WX})^{-1}\left[\mathbf{X'Wy} - \mathbf{X'WZ}(\mathbf{Z'WZ})^{-1}\mathbf{Z'Wy}\right] \tag{63}$$

$$= \left[\mathbf{I} - (\mathbf{X'WX})^{-1}\mathbf{X'WZ}(\mathbf{Z'WZ})^{-1}\mathbf{Z'WX}\right]^{-1}(\mathbf{X'WX})^{-1}\left[\mathbf{X'} - \mathbf{X'WZ}(\mathbf{Z'WZ})^{-1}\mathbf{Z'}\right]\mathbf{Wy} \tag{64}$$

$$= \mathbf{E}(K)\,\mathbf{Wy} \tag{65}$$

where $\mathbf{I}$ is the identity matrix, $\boldsymbol{K}$ is the knot point of $\mathbf{1}, \mathbf{2}, \ldots, \boldsymbol{r}$ and the matrix

$$\mathbf{E}(K) = \left[\mathbf{I} - (\mathbf{X'WX})^{-1}\mathbf{X'WZ}(\mathbf{Z'WZ})^{-1}\mathbf{Z'WX}\right]^{-1}(\mathbf{X'WX})^{-1}\left[\mathbf{X'} - \mathbf{X'WZ}(\mathbf{Z'WZ})^{-1}\mathbf{Z'}\right] \tag{66}$$

Likewise for the parameter estimator $\hat{\boldsymbol{\gamma}}$, so that we obtain:

$$\hat{\boldsymbol{\gamma}} = (\mathbf{Z'WZ})^{-1}\left(\mathbf{Z'Wy} - \mathbf{Z'WX}\left((\mathbf{X'WX})^{-1}(\mathbf{X'Wy} - \mathbf{X'WZ}\boldsymbol{\gamma})\right)\right) \tag{67}$$

$$= (\mathbf{Z'WZ})^{-1}\left(\mathbf{Z'Wy} - \mathbf{Z'WX}\left((\mathbf{X'WX})^{-1}\mathbf{X'Wy} - (\mathbf{X'WX})^{-1}\mathbf{X'WZ}\boldsymbol{\gamma}\right)\right) \quad (68)$$

$$= (\mathbf{Z'WZ})^{-1}\left(\mathbf{Z'Wy} - \mathbf{Z'WX}(\mathbf{X'WX})^{-1}\mathbf{X'Wy} + \mathbf{Z'WX}(\mathbf{X'WX})^{-1}\mathbf{X'WZ}\boldsymbol{\gamma}\right) \quad (69)$$

$$= (\mathbf{Z'WZ})^{-1}\mathbf{Z'Wy} - (\mathbf{Z'WZ})^{-1}\mathbf{Z'WX}(\mathbf{X'WX})^{-1}\mathbf{X'Wy} + (\mathbf{Z'WZ})^{-1}\mathbf{Z'WX}(\mathbf{X'WX})^{-1}\mathbf{X'WZ}\boldsymbol{\gamma} \quad (70)$$

The components containing $\hat{\boldsymbol{\gamma}}$ are combined into one left-hand side as follows:

$$\hat{\boldsymbol{\gamma}} - \left[(\mathbf{Z'WZ})^{-1}\mathbf{Z'WX}(\mathbf{X'WX})^{-1}\mathbf{X'WZ}\right]\hat{\boldsymbol{\gamma}} = (\mathbf{Z'WZ})^{-1}\mathbf{Z'Wy} - (\mathbf{Z'WZ})^{-1}\mathbf{Z'WX}(\mathbf{X'WX})^{-1}\mathbf{X'Wy} \quad (71)$$

$$\hat{\boldsymbol{\gamma}}\left[\mathbf{I} - (\mathbf{Z'WZ})^{-1}\mathbf{Z'WX}(\mathbf{X'WX})^{-1}\mathbf{X'WZ}\right] = (\mathbf{Z'WZ})^{-1}\mathbf{Z'Wy} - (\mathbf{Z'WZ})^{-1}\mathbf{Z'WX}(\mathbf{X'WX})^{-1}\mathbf{X'Wy} \quad (72)$$

So that the estimator $\hat{\boldsymbol{\gamma}}$ obtains the following equation:

$$\hat{\boldsymbol{\gamma}} = \left[\mathbf{I} - (\mathbf{Z'WZ})^{-1}\mathbf{Z'WX}(\mathbf{X'WX})^{-1}\mathbf{X'WZ}\right]^{-1}\left[(\mathbf{Z'WZ})^{-1}\mathbf{Z'Wy} - (\mathbf{Z'WZ})^{-1}\mathbf{Z'WX}(\mathbf{X'WX})^{-1}\mathbf{X'Wy}\right] \quad (73)$$

$$= \left[\mathbf{I} - (\mathbf{Z'WZ})^{-1}\mathbf{Z'WX}(\mathbf{X'WX})^{-1}\mathbf{X'WZ}\right]^{-1}(\mathbf{Z'WZ})^{-1}\left[\mathbf{Z'Wy} - \mathbf{Z'WX}(\mathbf{X'WX})^{-1}\mathbf{X'Wy}\right] \quad (74)$$

$$= \left[\mathbf{I} - (\mathbf{Z'WZ})^{-1}\mathbf{Z'WX}(\mathbf{X'WX})^{-1}\mathbf{X'WZ}\right]^{-1}(\mathbf{Z'WZ})^{-1}\left[\mathbf{Z'} - \mathbf{Z'WX}(\mathbf{X'WX})^{-1}\mathbf{X'}\right]\mathbf{Wy} \quad (75)$$

$$= \mathbf{F}(K)\,\mathbf{Wy} \quad (76)$$

Matrix $\mathbf{I}$ is the identity matrix, $\boldsymbol{K}$ is the knot point with values of $\mathbf{1}, \mathbf{2}, \dots, \boldsymbol{r}$ and matrix

$$\mathbf{F}(K) = \left[\mathbf{I} - (\mathbf{Z'WZ})^{-1}\mathbf{Z'WX}(\mathbf{X'WX})^{-1}\mathbf{X'WZ}\right]^{-1}(\mathbf{Z'WZ})^{-1}\left[\mathbf{Z'} - \mathbf{Z'WX}(\mathbf{X'WX})^{-1}\mathbf{X'}\right] \quad (77)$$

After obtaining the estimators for the parametric and nonparametric components, the next step is to determine the estimator for the semiparametric truncated spline regression model using the Weighted Least Square estimation method using the moving average approach as follows.

$$\hat{\mathbf{Y}} = \hat{f}(\mathbf{X}, \mathbf{Z}) \quad (78)$$

$$= \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{Z}\hat{\boldsymbol{\gamma}} \quad (79)$$

$$= \mathbf{XE}(K)\,\mathbf{Wy} + \mathbf{ZF}(K)\,\mathbf{Wy} \quad (80)$$

$$= \left(\mathbf{XE}(K) + \mathbf{ZF}(K)\right)\mathbf{Wy} \quad (81)$$

$$= \mathbf{T}(\mathbf{X}, \mathbf{Z}, K)\mathbf{Wy} \quad (82)$$

Where $\mathbf{T}(\mathbf{X}, \mathbf{Z}, K) = \mathbf{XE}(K) + \mathbf{ZF}(K)$

4. CONCLUSION

Based on the theoretical derivation presented in this study, several conclusions can be drawn. First, the partial spline semiparametric regression model with a linear parametric component and a truncated spline nonparametric component can be expressed in matrix form as $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}$, where $\mathbf{X}$ is the parametric design matrix, $\mathbf{Z}$ is the truncated spline design matrix, and $\boldsymbol{\varepsilon}$ is the random error vector. Second, the weight matrix $\mathbf{W} = \mathrm{diag}(\hat{w}_1^{-1}, \hat{w}_2^{-1}, \dots, \hat{w}_n^{-1})$ is constructed adaptively from a local moving average of squared generalized residuals following Silverman (1985), extending this procedure from the nonparametric to the partial spline framework. Third, the WLS criterion $\min_{\boldsymbol{\beta},\boldsymbol{\gamma}}\{(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}-\mathbf{Z}\boldsymbol{\gamma})^{\top}\mathbf{W}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}-\mathbf{Z}\boldsymbol{\gamma})\}$ yields two mutually dependent estimating equations. Through partial differentiation and a substitution procedure, independent closed-form estimators are obtained:

$$\hat{\boldsymbol{\beta}} = \mathbf{E}(K)\mathbf{W}\mathbf{y}, \quad \hat{\boldsymbol{\gamma}} = \mathbf{F}(K)\mathbf{W}\mathbf{y}$$

Fourth, the regression curve estimator for the partial spline model is:

$$\hat{\mathbf{Y}} = \mathbf{T}(X, Z, K)\mathbf{W}\mathbf{y}$$

where $\mathbf{T}(X, Z, K) = \mathbf{XE}(K) + \mathbf{ZF}(K)$, and optimal knot points are selected by minimizing the GCV criterion.

This study contributes a methodologically rigorous extension of Silverman's [20] local moving average reweighting framework to the partial spline semiparametric regression setting, providing efficient parameter estimation under heteroscedastic error conditions whose functional form is unknown.

Funding Information

The authors declare that no external funding was received for this research. All activities related to this research were fully supported and funded by the authors themselves.

Author Contributions Statement

First Author: Conceptualization, methodology, formal analysis, validation, writing-original draft. Second Author: Methodology, formal analysis, writing-original draft, writing-review & editing. Third Author: Supervision, validation, writing-review & editing. All authors participated in the discussion of the research results and approved the full version of the paper.

Conflict of Interest Statement

No conflicts of interest were disclosed by the authors.

Data Availability

Data availability is not applicable to this paper as no new data were created or analyzed in this study.

REFERENCES

[1] D. N. Gujarati and D. C. Porter, *Basic Econometrics*, 5th ed. New York, USA: McGraw-Hill, 2013.

[2] A. P. Anisar, A. T. R. Dani, and R. Dani, "Estimation of a bi-response truncated spline nonparametric regression model on life expectancy and prevalence of underweight children in Indonesia," *BAREKENG: J. Ilmu Mat. Terap.*, vol. 17, no. 4, pp. 2011–2022, 2023, doi: 10.30598/barekengvol17iss4pp2011-2022.

[3] A. T. R. Dani, R. Dani, and M. Fauziyah, "Mixed estimators of truncated spline-Epanechnikov kernel on nonparametric regression and its applications," *BAREKENG: J. Ilmu Mat. Terap.*, vol. 17, no. 4, pp. 2023–2032, 2023, doi: 10.30598/barekengvol17iss4pp2023-2032.

[4] N. Chamidah, B. Lestari, I. N. Budiantara, T. Saifudin, R. Rulaningtyas, A. Aryati, P. Wardani, and D. Aydin, "Consistency and asymptotic normality of estimator for parameters in multiresponse multipredictor semiparametric regression model," *Symmetry*, vol. 14, no. 2, p. 336, 2022, doi: 10.3390/sym14020336.

[5] L. Fu, Z. Lin, and Z. Su, "Semiparametric efficient estimation in high-dimensional partial linear regression models," *Scand. J. Stat.*, 2024, doi: 10.1111/sjos.12716.

[6] N. Y. Adrianingsih, "Estimator spline truncated of semiparametric regression," *J. Ilm. Wahana Pendidik.*, vol. 8, no. 2, pp. 208–212, 2022, doi: 10.5281/zenodo.6224937.

[7] R. F. Engle, C. W. J. Granger, J. Rice, and A. Weiss, "Semiparametric estimates of the relation between weather and electricity sales," *J. Am. Stat. Assoc.*, vol. 81, no. 394, pp. 310–320, 1986,

doi: 10.1080/01621459.1986.10478274.

[8] A. A. Khalil, I. N. Budiantara, and I. Zain, "Comparison of linear and quadratic bi-response semiparametric regression models using spline truncated," *J. Phys.: Conf. Ser.*, vol. 1511, no. 1, p. 012046, 2020, doi: 10.1088/1742-6596/1511/1/012046.

[9] H. Husain, I. N. Budiantara, and I. Zain, "Mixed estimator of spline truncated, Fourier series, and kernel in biresponse semiparametric regression model," *IOP Conf. Ser.: Earth Environ. Sci.*, vol. 880, no. 1, p. 012046, 2021, doi: 10.1088/1755-1315/880/1/012046.

[10] I. Sriliana, I. N. Budiantara, and V. Ratnasari, "A truncated spline and local linear mixed estimator in nonparametric regression for longitudinal data and its application," *Symmetry*, vol. 14, no. 12, p. 2687, 2022, doi: 10.3390/sym14122687.

[11] D. A. Maharani and D. R. S. Saputro, "Generalized cross validation (GCV) in smoothing spline nonparametric regression models," *J. Phys.: Conf. Ser.*, vol. 1808, no. 1, p. 012053, 2021, doi: 10.1088/1742-6596/1808/1/012053.

[12] R. Putra, M. G. Fadhlurrahman, and Gunardi, "Determination of the best knot and bandwidth in geographically weighted truncated spline nonparametric regression using generalized cross validation," *MethodsX*, vol. 10, p. 101994, 2023, doi: 10.1016/j.mex.2022.101994.

[13] Sifriyani, I. N. Budiantara, K. P. Candra, and M. Putri, "Selection of optimal knot point and best geographic weighted on geographically weighted spline nonparametric regression model," *MethodsX*, vol. 13, p. 102802, 2024, doi: 10.1016/j.mex.2024.102802.

[14] T. Handayani, Sifriyani, and A. T. R. Dani, "Nonparametric spline truncated regression with knot point selection method generalized cross validation and unbiased risk," *JTAM (J. Teori Apl. Mat.)*, vol. 7, no. 3, 2023, doi: 10.31764/jtam.v7i3.14034.

[15] B. Lestari, N. Chamidah, I. N. Budiantara, and D. Aydin, "Determining confidence interval and asymptotic distribution for parameters of multiresponse semiparametric regression model using smoothing spline estimator," *J. King Saud Univ. Sci.*, vol. 35, no. 5, p. 102664, 2023, doi: 10.1016/j.jksus.2023.102664.

[16] F. Virgantari, M. Widyastiti, and N. Ir Seno, "Comparison of weights in weighted least square method for handling heteroscedasticity on multiple regression model," *Int. J. Math. Stat. Comput.*, vol. 2, no. 2, pp. 60–67, 2024, doi: 10.46336/ijmsc.v2i2.93.

[17] W. Fransiska, S. Nugroho, and R. Rachmawati, "A comparison of weighted least square and quantile regression for solving heteroscedasticity in simple linear regression," *J. Stat. Data Sci.*, vol. 1, no. 1, 2022.

[18] M. Setyawati, N. Chamidah, A. Kurniawan, and D. Aydin, "Confidence interval for semiparametric regression model parameters based on truncated spline with application to COVID-19 dataset in Indonesia," *Data Metadata*, vol. 3, 2024, doi: 10.56294/dm2024.609.

[19] J. P. Romano and M. Wolf, "Resurrecting weighted least squares," *J. Econom.*, vol. 197, no. 1, pp. 1–19, 2017, doi: 10.1016/j.jeconom.2016.10.003.

[20] B. W. Silverman, "Some aspects of the spline smoothing approach to non-parametric regression curve fitting," *J. R. Stat. Soc. Ser. B*, vol. 47, no. 1, pp. 1–52, 1985.

[21] N. Chamidah, B. Lestari, I. N. Budiantara, and D. Aydin, "Estimation of multiresponse multipredictor nonparametric regression model using mixed estimator," *Symmetry*, vol. 16, no. 4, p. 386, 2024, doi: 10.3390/sym16040386.

[22] W. Alwi, M. Irwan, and Musfirah, "Penerapan regresi nonparametrik spline dalam memodelkan faktor-faktor yang mempengaruhi indeks pembangunan manusia (IPM) di Indonesia tahun 2018," *J. Mat.*, vol. 9, no. 2, 2021.

[23] M. P. Wand and M. C. Jones, *Kernel Smoothing*. London, UK: Chapman & Hall/CRC, 1995.

[24] H. Nisa, D. Kusnandar, and S. Martha, "Estimasi parameter metode weighted least square dalam mengatasi masalah heteroskedastisitas," *Bul. Ilm. Mat. Stat. Terapannya (Bimaster)*, vol. 09, no. 1, pp. 65–70, 2020.

[25] R. D. Fadlirhohim, A. T. R. Dani, and R. Dani, "Modeling stunting prevalence in Indonesia using spline truncated semiparametric regression," *BAREKENG: J. Ilmu Mat. Terap.*, vol. 18, no. 3, pp. 2015–2028, 2024.