BAREKENG: Journal of Mathematics and Its Applications
June 2026 Volume 20 Issue 2 Page 1775–1790
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol20iss2pp1775-1790

QUANTILE BASED PLS-SEM WITH WILD BOOTSTRAP

Abdul Malik Balami [1], Bambang Widjanarko Otok [2*], Santi Wulan Purnami [3]

[1,2,3] *Department of Statistics, Faculty of Science and Data Analytics, Institut Teknologi Sepuluh Nopember Kampus ITS-Sukolilo, Surabaya 60111, Indonesia*

[1] *Department of Mathematics, Faculty of Science and Technology, Universitas Pattimura Jln. Ir. M. Putuhena, Kampus Unpatti-Poka, Ambon, 97233, Indonesia*

Corresponding author's e-mail: * ***Bambang_wo@its.ac.id***

Article Info

Article History:

Received: 30th August 2025
Revised: 30th September 2025
Accepted: 31st October 2025
Available Online: 26th January 2026

Keywords:

PLS-SEM;
Quantile Regression;
Structural Equation Modeling;
Wild Bootstrap.

ABSTRACT

Partial Least Squares SEM (PLS-SEM) is the recommended technique for structural equation modeling (SEM), which assesses correlations between latent components concurrently, particularly for small samples and non-normal data. But because traditional PLS-SEM only calculates average correlations between constructs, it runs the risk of overlooking variances in the quantile distribution. Consequently, the creation of the Quantile PLS-SEM approach, which incorporates quantile regression, provides a means to examine correlations across the entire data distribution. To improve estimation, wild bootstrap is used to address heteroscedasticity issues and produce more reliable inferences. The purpose of this study is to develop and apply Quantile based PLS-SEM with Wild Bootstrap to analyze the gizi data status of the Indonesian population based on the Survey Status Gizi Indonesia 2024. The analysis's findings indicate that specific and sensitive interventions have a significant impact on the gizi status of different quantities.

How to cite this article:

A. M. Balami, B. W. Otok and S. W. Purnami., "QUANTILE BASED PLS-SEM WITH WILD BOOTSTRAP", *BAREKENG: J. Math. & App.*, vol. 20, no. 2, pp. 1775-1790, Jun, 2026.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article · Open Access

1. INTRODUCTION

There is a pressing need for statistical analysis techniques that can capture variation and heterogeneity in the interactions between variables in an era of data that is becoming more complicated and diverse. One of the most popular methods for concurrently assessing connections between latent components is structural equation modeling, or SEM [1]. With a composite-based methodology distinct from covariance-based SEM, Partial Least Squares SEM (PLS-SEM) is one of the most versatile SEM methods, especially for non-normal data and small samples [2].

The average relationship (conditional mean) between constructs is typically the only estimate made by conventional PLS-SEM, which may overlook variations in effects on other data distribution components. When heteroscedasticity or diverse effect patterns exist across multiple quantiles of the distribution, as in complicated socioeconomic or health data, this condition becomes problematic [3]. A more thorough depiction of the variability in effects between variables is provided by the creation of Quantile PLS-SEM, which incorporates quantile regression into the estimation process and provides a way to analyze relationships across multiple quantiles rather than just the mean [4]. The bootstrap methodology is the primary approach for measuring estimation uncertainty in PLS-SEM for statistical inference [5]. Confidence intervals and hypothesis testing may be erroneous due to the constraints of the traditional bootstrap when handling heteroscedastic data [6]. By generating residuals calibrated to the real variance pattern in the original data, the wild bootstrap approach provides a more efficient substitute that yields more reliable and correct conclusions [7].

Significant methodological advancements in SEM analysis are enabled by combining Quantile PLS-SEM with the wild bootstrap, which allows researchers to obtain more trustworthy conclusions and robust estimates even in less-than-ideal data settings [8]. Applications in the social, economic, health, and environmental domains, where data frequently display heteroscedasticity and non-normal distributions, are especially pertinent to this approach [9]. Therefore, research that develops and applies Quantile-based PLS-SEM using a wild bootstrap approach is a crucial step in enhancing the validity and reliability of findings from structural relationship analysis while broadening the study's reach to encompass the full range of complex data distributions.

2. RESEARCH METHODS

2.1 Structural Equation Modeling (SEM)

SEM is a statistical analysis method that estimates multiple equations simultaneously by combining factor analysis and multiple regression [10]. SEM produces and performs two tasks: (i) concurrently estimating many interconnected equations with structural model output, as well as (ii) using measurement model output to represent latent variables (construct/latent/unobserved variables) based on indicator variables (manifest/observed). The following is the structural equation model:

$$\eta = \mathrm{B}\eta + \Gamma\xi + \zeta. \tag{1}$$

The endogenous latent variable η (eta) has a size of m × 1, the exogenous latent independent random variable vector ξ (xi) has a size of n × 1, and the coefficient matrix B shows how the endogenous latent variable affects other variables of size m × m, while Γ shows the relationship between ξ and η of size m × n. The random error vector ζ (zeta) is of size m × 1, using zero as the intended value. The structural model equation of latent variables makes the following assumptions: $(\mathrm{I}-\mathrm{B})^{-1})$ is a nonsingular matrix, ζ is not associated with ξ, and $E(\eta) = 0, E(\xi) = 0,$ and $E(\zeta) = 0$. Confirmatory Factor Analysis Model (CFA) or measurement model [11]. The following is one way to express the measuring model (CFA):

$$y = \Lambda_y\eta + \varepsilon\Sigma(\kappa), \tag{2}$$

$$x = \Lambda_x\xi + \delta. \tag{3}$$

Eqs. (2) and (3) show that the covariance matrix in SEM Σ(θ) looks like this:

$$\Sigma(\kappa) = \begin{bmatrix} \Sigma_{yy} & \Sigma_{yx} \\ \Sigma_{xy} & \Sigma_{xx} \end{bmatrix}. \tag{4}$$

Additionally, the structural parameter's function for $\Sigma_{yy}(\theta) = [\Lambda_y[(I - B)^{-1}(\Gamma\Phi\Gamma' + \Psi)(I - B')^{-1}]\Lambda'_y + \Theta_\varepsilon]$, $\Sigma_{yx}(\theta) = \Lambda_y(I - B)^{-1}\Gamma\Phi\Lambda_x'$ dan $\Sigma_{xx}(\theta) = \Lambda_x\Phi\Lambda_x' + \Theta_\delta$.

2.1.1 SEM Partial Least Square (SEM PLS)

Herman Wold first introduced partial least squares (PLS) in 1975. When the theory behind the model design is weak, this model was created as a backup. Because PLS has more flexible assumptions and can be applied to data at any scale, it is a robust analysis method. There are three steps involved in estimation using the PLS method. A straightforward regression or multiple regression iteration process that considers the structural model/inner model, measurement model/outer model, and weight estimation/weight relation is the initial stage in PLS estimation. The latent variable scores, which are linear combinations of indicator/manifest variables, are then determined from the estimated weights. Estimating the structural model coefficients (inner model) and the measurement model coefficients (outer model) are the second and third steps after acquiring the latent variable scores [12]. Ordinary least squares estimation is used in a sequence of simple and multiple regressions that make up the PLS algorithm [13].

In PLS, the relationships between latent variables are described by the structural model, often known as the inner model [14]. The following is a linear equation representation of the model equation:

$$\xi_j = \sum_i \beta_{ji}\xi_i + \zeta_j. \tag{5}$$

$E(\zeta_j) = 0, E(\xi_i\zeta_j) = 0$, and β_{ji} is the path coefficient, or coefficient of the link between latent variables i and j. Every indicator block in PLS can be associated with its latent variable to construct the measurement model or outer model. Using simple regression, the following equation may be expressed for a reflecting indicator block:

$$x_{jk} = \lambda_{jk}\xi_j + \varepsilon_{jk}, \tag{6}$$

$E(\varepsilon_{jk}) = E(\xi_j\varepsilon_{jk}) = 0$, ε_{jk} is the residual or error of each measurement variable, and λ_{jk} is the coefficient loading of the relationship between the latent variable j (ξ_j) and its indicator k (x_{jk}).

2.2 Quantile Regression

Koenker and Bassett in 1978 developed the regression analysis technique known as quantile regression. Several quantile functions of a distribution Y are estimated using this method as a function of X. When there is heterogeneity in the data distribution, quantile regression is highly helpful [15]. Consider the following data: $\{X_{1i}, X_{2i}, \dots, X_{ki}, Y_i\}$,, a collection of paired random variables with quantiles $\tau \in (0,1)$ that are independently and identically distributed.

The following is the general equation for linear quantile regression:

$$Y_i = \beta_0(\tau) + \beta_1(\tau)X_{1i} + \cdots + \beta_k(\tau)X_{ki} + \varepsilon_i(\tau). \tag{7}$$

Eq. (7) can therefore be expressed using the linear model form shown below:

$$Y = X\beta(\tau) + \varepsilon(\tau). \tag{8}$$

2.2.1 Calculated Quantile Regression for Parameters

A paired data set $\{x_{1i}, x_{2i}, x_{3i}, \cdots x_{ki}, y_i\}$ $i = 1,2, \cdots n$, $j = 1,2 \cdots k$ is a collection of paired random variables with quantile $\tau \in (0,1)$ that are independently and identically distributed [16]. The following is the definition of the conditional probability distribution function for the data:

$$F(Y|x_i) = P(Y \le y|x_i). \tag{9}$$

Considering the response variable y's τ-quantile [17], the inverse function $F^{-1}(\tau) = inf\{y: F(y) \ge \tau\}$. Accordingly, the following is the definition of the general equation of linear quantile regression for conditional quantiles:

$$Q_y(\tau|x) = F_y^{-1}(y|x_i) = inf\{y: F_y(y|x) \ge \tau\}.$$

Consequently, the following is the conditional quantile function model:

$$Q_y(\tau|x_i) = X^T\beta(\tau), \tau \in (0,1). \tag{10}$$

The response variable's general linear quantile regression equation is expressed as follows:

$$y_i = \beta_0(\tau) + \beta_1(\tau)x_{1i} + \cdots + \beta_k(\tau)x_{ki} + \varepsilon_i(\tau). \tag{11}$$

Eq. (11) can therefore be expressed in the matrix form shown below:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{k1} \\ 1 & x_{12} & x_{22} & \cdots & x_{k2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & x_{2n} & \cdots & x_{kn} \end{bmatrix} \begin{bmatrix} \beta_0(\tau) \\ \beta_1(\tau) \\ \vdots \\ \beta_k(\tau) \end{bmatrix} + \begin{bmatrix} \varepsilon_1(\tau) \\ \varepsilon_2(\tau) \\ \vdots \\ \varepsilon_n(\tau) \end{bmatrix},$$

in order to produce the linear model that follows:

$$\mathrm{y} = \mathrm{X}\beta(\tau) + \varepsilon(\tau). \tag{12}$$

Quantile regression has the principle of minimizing the sum of squared residuals to find estimators, as in the OLS method. The τth quantile regression of F_Y can be obtained from Eq. (13) as follows:

$$E\rho(Y-\hat{y}) = \int_{-\infty}^{\hat{y}} (1-\tau)(y-\hat{y})f(y)dy + \int_{\hat{y}}^{\infty} \tau(y-\hat{y})f(y)dy. \tag{13}$$

Eq. (14) is therefore reduced to zero:

$$\frac{\partial}{\partial \hat{y}} E[\rho_\tau(Y-\hat{y})] = 0 \tag{14}$$

$$\frac{\partial}{\partial \hat{y}} \int_{-\infty}^{\hat{y}} (1-\tau)(y-\hat{y})f(y)dy + \int_{\hat{y}}^{\infty} \tau(y-\hat{y})f(y)dy = 0$$

$$\frac{\partial}{\partial \hat{y}} \left[(1-\tau)\int_{-\infty}^{\hat{y}} (y-\hat{y})f(y)dy + \tau \int_{\hat{y}}^{\infty} (y-\hat{y})f(y)dy\right] = 0$$

$$(1-\tau)\left[(y-\hat{y})f(y)\Big|_{-\infty}^{\hat{y}} + \int_{-\infty}^{\hat{y}} \frac{\partial}{\partial \hat{y}}(y-\hat{y})f(y)dy\right] +$$

$$\tau\left[(y-\hat{y})f(y)\Big|_{\hat{y}}^{-\infty} + \int_{\hat{y}}^{\infty} \frac{\partial}{\partial \hat{y}}(y-\hat{y})f(y)dy\right] = 0$$

$$(1-\tau)\left[(y-\hat{y})f(y)\Big| y=\hat{y} + \int_{-\infty}^{\hat{y}} f(y)dy\right] + \tau\left[(y-\hat{y})f(y)\Big| y=\hat{y} + \int_{\hat{y}}^{\infty} f(y)dy\right] = 0$$

$$(1-\tau)[0 + F_Y(\hat{y})] + \tau\big[0 - \big(1 - F_Y(\hat{y})\big)\big] = 0$$

$$(1-\tau)F_Y(\hat{y}) - \tau\big(1 - F_Y(\hat{y})\big) = 0$$

$$(1-\tau)F_Y(\hat{y}) - \tau - \tau F_Y(\hat{y}) = 0$$

$$F_Y(\hat{y}) - \tau = 0,$$

then acquired

$$F_Y(\hat{y}) = \tau.$$

So that the solution of F_Y is the τ quantile. By minimizing the sum of the absolute values of the mistakes with weight (τ) for positive errors and weight $(1-\tau)$ for negative errors, the τ quantile regression is produced. The following solution is the result of this:

$$\hat{\beta}(\tau) = \min_{\beta \in R^{p+1}} \tau \sum y \geq x|y - X^T\beta| + (1-\tau)\sum y < x|y - X^T\beta|, \tag{15}$$

or can be expressed as follows in Eq. (16):

$$\hat{\beta}(\tau) = \min_{\beta \in R^{p+1}} \sum_{i=1}^{n} \rho_\tau(u_i), \tag{16}$$

where:

$$\rho_\tau(u_i) = \begin{cases} (\tau - 1)u_i \text{ ,with } u_i < 0 \\ \tau u_i, \text{ with } u_i \geq 0 \end{cases},$$

with

$\hat{\beta}(\tau)$: parameter estimator;

τ : quantile index with $\tau \in (0,1)$;

$\rho_\tau(u_i)$: loss function;

u_i : error of the parameter estimator.

The sum of squared errors is minimized in the OLS estimation of a linear model on y. In contrast, quantile regression estimation of a linear model on y is achieved by minimizing the predicted value of $\rho_\tau(u)$ [18], which is the value of the asymmetric loss function. The loss function's asymmetry will next be demonstrated using the following justification:

provided

$$\rho_\tau = [\tau I(u \geq 0) + (1 - \tau))I((u < 0))]|u| = [\tau - I(u < 0)]u,$$

with

$$I(u \geq 0) = \begin{cases} 1, u \geq 0 \\ 0, u < 0 \end{cases} \text{ and } |u| = \begin{cases} u, u \geq 0 \\ -u, u < 0 \end{cases}',$$

where

u : error from parameter estimator;

$I(u)$: indicator function that has been defined.

So that it can be proven

$$\rho_\tau = \begin{cases} \tau u, u \geq 0 \\ (\tau - 1)u, u < 0 \end{cases}.$$

Eq. (16) can only be solved numerically rather than analytically. The simplex algorithm is a numerical technique. Barrodale and Robert created the simplex algorithm approach in 1974. With the aid of computation, this algorithmic approach provides solutions to linear programming problems with numerous decision variables [19].

2.3 Bootstrap

In order to lessen the unreliability that comes with using the normal distribution incorrectly, Efron in 1979 created the bootstrap approach [20]. The standard deviation of B replications is used to compute the bootstrap standard error of $\hat{\theta}$.

$$\widehat{se}(\hat{\theta}_B) = \sqrt{\frac{\sum_{b=1}^{B}\left(\hat{\theta}^*_{(b)} - \hat{\theta}^*_{(.)}\right)^2}{B - 1}},$$

where B is the number of resampling sets of size n with replacement, and $\left(\hat{\theta}^*_{(.)}\right) = \frac{\sum_{b=1}^{B}\hat{\theta}^*_{(b)}}{B}$. The $\hat{\theta}$ statistic derived from the bth resampling $(b = 1, \ldots, B)$ is $\hat{\theta}^*_{(b)}$.

2.4 Wild Bootstrap

Liu introduced the wild bootstrap method as an innovative approach to address heteroskedasticity in regression models, particularly the instability of error variances [21]. Before this development, the standard bootstrap technique was already widely used to examine sample distributions and construct confidence intervals without making assumptions about the underlying distribution. However, the traditional bootstrap method tends to perform poorly when applied to data with non-constant error variances, a common issue in economic and financial analyses. To overcome this limitation, Liu proposed the wild bootstrap method, which modifies residuals in regression analysis by multiplying them by random Rademacher values, which take only two values, +1 and –1, each with probability 0.5. In this study, the wild bootstrap, a variant of the traditional bootstrap that incorporates randomization in the resampling process, is employed to provide a more robust inference framework for models affected by heteroskedasticity.

According to Feng, He, and Hu, the wild bootstrap model for quantile regression is a resampling technique intended to enhance inference in quantile regression estimates, especially when heteroscedasticity and fixed designs are present [22]. Because the error distribution cannot be presumed to be symmetric, typical bootstrap techniques like residual bootstrap are inappropriate when quantile regression minimizes the asymmetric loss function $\rho_\tau(u) = u(\tau - I(u < 0))$. It makes it impossible to assume that the error distribution is symmetric, rendering conventional bootstrap techniques like residual bootstrap incorrect. By modifying residual resampling using random weights (wild weights), wild bootstrap provides a solution; nevertheless, it is modified to account for the quantile loss function's asymmetry [23].

2.4.1 Methods for Quantile Regression Using Wild Bootstrap

1. Fit the data to the quantile regression model:

$$\underset{\beta \in \mathbb{R}p}{\text{argmin}} \sum_{i=1}^{n} \rho_\tau(y_i - x_i^T \beta(\tau)), \quad i = 1,2,\dots,n \quad \tau \in (0,1).$$

   Obtain residuals:

$$\hat{e}_i = y_i - x_i^T \beta(\tau).$$

2. Create a residual bootstrap:

$$e_i^* = w_i |\hat{e}_i|,$$

   where w_i represents a random weight drawn from a distribution that satisfies specific requirements.

3. Create a bootstrapped data set.

$$y_i^* = x_i^T \beta(\tau) + e_i^*.$$

4. Refit model using bootstrapped data:

   Utilizing quantile regression on $\{y_i^*, x_i\}$, compute $\hat{\beta}^*(\tau)$.

5. Steps 2–4 should be repeated multiple times. For inference tasks, such as computing standard errors and confidence intervals, use the distribution of the bootstrapped estimates ($\hat{\beta}^*$ distribution).

The 2024 *Survey Status Gizi Indonesia* (SSGI) publication served as the source of data for this investigation. The data gathered from this survey is essential for assessing the general health of the population in different parts of Indonesia. Table 1 provides an explanation of the variables used in this study using a number of distinct variables:

Table 1. Research Variable

Laten Variable	Manifest Variable
Nutritional status ($\boldsymbol{\eta}$)	Underweight ($y_{1.1}$) Stunting ($y_{1.2}$) Wasting ($y_{1.3}$) Overweight ($y_{1.4}$)
Specific Interventions ($\boldsymbol{\xi_1}$)	Percentage of pregnant women with chronic energy deficiency (CED) who receive additional nutritional intake ($x_{1.1}$) Percentage of pregnant women who consume at least 90 tablets of iron supplements during pregnancy ($x_{1.2}$) Percentage of adolescent girls who consume iron supplements ($x_{1.3}$) Percentage of infants under 6 months of age receiving exclusive breastfeeding ($x_{1.4}$) Percentage of children aged 6-23 months who receive complementary foods to breast milk ($x_{1.5}$)

Laten Variable	Manifest Variable
	Percentage of children under five years of age (toddlers) with malnutrition who receive malnutrition treatment services ($x_{1.6}$)
	Percentage of children under five years of age (toddlers) whose growth and development are monitored ($x_{1.7}$)
	Percentage of children under five years of age (toddlers) who are malnourished and receive additional nutritional intake ($x_{1.8}$)
	Percentage of children under five years of age (toddlers) who receive complete basic immunizations ($x_{1.9}$)
Sensitive Intervention (ξ_2)	Percentage of family planning (FP) services after childbirth ($x_{2.1}$)
	Percentage of unwanted pregnancies ($x_{2.2}$)
	Coverage of prospective couples of childbearing age (PUS) who receive health checks as part of marriage services ($x_{2.3}$)
	Percentage of households with access to safe drinking water in priority districts/cities ($x_{2.4}$)
	Percentage of households with access to proper sanitation (domestic wastewater) in priority districts/cities ($x_{2.5}$)
	Coverage of National Health Insurance Contribution Assistance Recipients ($x_{2.6}$)
	Coverage of families at risk of stunting who receive assistance ($x_{2.7}$)
	Number of poor and vulnerable families receiving conditional cash transfers ($x_{2.8}$)
	Percentage of target beneficiaries who have a good understanding of stunting in priority locations ($x_{2.9}$)
	Number of poor and vulnerable families receiving social assistance in the form of food ($x_{2.10}$)
	Percentage of villages/subdistricts that have stopped open defecation free (ODF) ($x_{2.11}$)

***Data source**: SSGI 2024*

2.5 Steps in Research

The methods and techniques of analysis that will be used to achieve the research goal are as follows:

1. Presenting a conceptual model based on theory.

   The theoretical framework on the context, causes, and effects of toddler stunting served as the foundation for the development of the Status conceptual framework, which also took into account the policies and measures in place to hasten the reduction of stunting in Indonesia. A number of factors, including Specific Intervention Indicators and Sensitive Intervention Indicators, affect toddlers' nutritional status, which is measured as underweight, stunted, wasted, or overweight [24].

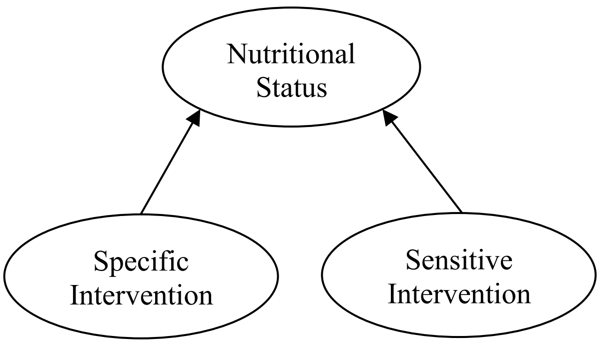

Figure 1. Conceptual model based on theory

2. Creating a path diagram.
3. Validity testing in the outer model and the structural model/inner model is done by PLS-SEM modeling.
4. For every hidden variable, get the weighted factor scores. The Quantile Regression modeling analysis using PLS-SEM structural equations will make use of these weighted factor scores.

5. Finding the estimated values of the SEM quantile regression parameters.
6. Using bootstrap and wild bootstrap for hypothesis testing.
7. Interpreting and summarizing the results.

3. RESULTS AND DISCUSSION

3.1 Estimating Model Parameters

This work employed a route approach to obtain the model parameter coefficients using PLS. The following are the λ coefficients for exogenous and endogenous variables:

$$\lambda_{x1.1} = 0.697\lambda_{x1.2} = 0.821\lambda_{x1.3} = 0.673\lambda_{x1.4} = 0.764\lambda_{x1.5} = 0.744\lambda_{x1.6} = 0.505$$
$$\lambda_{x1.7} = 0.751\lambda_{x1.8} = 0.717\lambda_{x1.9} = 0.754\lambda_{x2.1} = 0.605\lambda_{x2.2} = 0.077\lambda_{x2.3} = 0.723$$
$$\lambda_{x2.4} = 0.652\lambda_{x2.5} = 0.656\lambda_{x2.6} = 0.409\lambda_{x2.7} = 0.665\lambda_{x2.8} = 0.496\lambda_{x2.9} = 0.745$$
$$\lambda_{x2.10} = 0.499\lambda_{x2.11} = 0.618\lambda_{y1.1} = 0.961\lambda_{y1.2} = 0.797\lambda_{y1.3} = 0.824\lambda_{y1.4} = -0.481.$$

Coefficient γ:

$\gamma_{1.1} = -0.137$ and $\gamma_{2.1} = -0.214$.

3.2 Assessment of Outer Model Measurement Models

The measurement model (outer model) was evaluated for each PLS scheme employed, namely the path scheme. Assessing each indicator's validity and reliability in relation to its latent variable was part of the measurement model evaluation process.

1. Validity

A metric used to characterize the relationship between indicator scores and latent variables is called validity. First, the validity indicators displayed by the factor loading values (λ) are examined. The indicator is deemed legitimate if the loading value (λ) is more than or equal to 0.5. If λ is less than or equal to 0.5, the indicator is deemed invalid and needs to be eliminated from the analysis because it shows that it is not trustworthy enough to measure the hidden variable. There are still factor loading values (λ) <0.5 in the λ coefficients, namely in indicators $X_{2.2}$, $X_{2.6}$, $X_{2.8}$, and $X_{2.10}$, which represent the specific intervention service variable, and indicator $Y_{1.4}$, which represents the nutritional status variable. A loading factor value (λ) < 0.5 indicates that the indicator is invalid and should be removed from the analysis because it indicates that the indicator is not good enough to measure the latent variable. Therefore, indicators $Y_{1.4}$, $X_{2.2}$, $X_{2.6}$, $X_{2.8}$, and $X_{2.10}$ are not used in the PLS analysis.

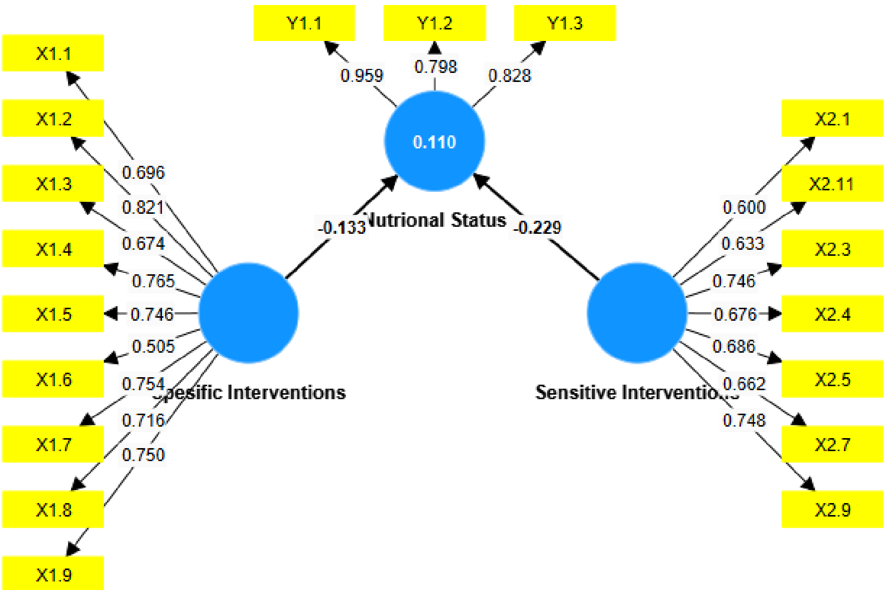

Figure 2. Path Diagram for Structural Equations using Path Scheme
***Source**: Processed results from smartPLS 4*

According to Fig. 2, each indicator of the latent variables Nutritional Status, Sensitive Intervention Services, and Specific Intervention Services has a loading factor (λ) values greater than 0.5. Therefore, it can be concluded that every indicator employed to measure hidden variables is excellent and reliable.

2. Reliability

A coefficient called reliability indicates how consistent the data are. If the data at several points in time are similar, the study is deemed credible. For a variable to be considered trustworthy, its composite reliability value must be better than 0.7. Table 2 displays the output results, which include each variable's composite reliability (CR) values.

Table 2. Reliability Value

Latent Variable	Composite reliability
Specific Interventions	0.809
Nutritional Status	0.836
Sensitive Intervention	0.911

Data source*: Processed Results from smartPLS 4*

From Table 2, every indicator employed to measure latent variables is dependable since each latent variable has a composite reliability (CR) value higher than 0.7. These factors lead to the conclusion that the measurement model is good since it satisfies the criterion for validity and reliability.

3.3 Evaluation of Structural Models (Inner Model)

The parameter coefficient estimates and their significance levels were examined to evaluate the structural model and investigate the connections among the previously proposed latent components. R-square of 0.110 is one metric that may be used to assess the structural model (inner model).

1. Hypothesis Testing

The parameters λ and γ are tested as part of hypothesis testing. The t-test is the statistical test that smartPLS uses.

a. Testing the measurement model hypothesis (outer model)

The following hypothesis is applied in order to determine the importance of the outer model parameters:

$H_0: \lambda_i = 0,$
$H_1: \lambda_i \neq 0.$

At a significance level of α equal to 5 percent, reject H_0 if $t_{stat} > t_{table}$ or $p-value < \alpha$. The t-table value is 1.960. Table 3 displays the findings of the measurement model's t-statistic test.

Table 3. Bootstrap t-Statistic Significance Test

Latent Variable	Manifest Variable	Loading	Standard Error	t -statistics	p- values
Specific Interventions	$x_{1.1}$	0.696	0.055	12.606	0.000
	$x_{1.2}$	0.821	0.033	24.912	0.000
	$x_{1.3}$	0.674	0.043	15.816	0.000
	$x_{1.4}$	0.765	0.034	22.711	0.000
	$x_{1.5}$	0.746	0.037	19.978	0.000
	$x_{1.6}$	0.505	0.059	8.577	0.000
	$x_{1.7}$	0.754	0.040	18.866	0.000
	$x_{1.8}$	0.716	0.049	14.564	0.000
	$x_{1.9}$	0.750	0.031	24.213	0.000
Sensitive Interventions	$x_{2.1}$	0.600	0.048	12.517	0.000
	$x_{2.3}$	0.746	0.039	19.324	0.000

Latent Variable	Manifest Variable	Loading	Standard Error	t -statistics	p- values
	$x_{2.4}$	0.676	0.051	13.277	0.000
	$x_{2.5}$	0.686	0.048	14.389	0.000
	$x_{2.7}$	0.662	0.048	13.642	0.000
	$x_{2.9}$	0.748	0.039	19.292	0.000
	$x_{2.11}$	0.633	0.052	12.060	0.000
Nutritional Status	$y_{1.1}$	0.959	0.006	172.507	0.000
	$y_{1.2}$	0.798	0.032	25.260	0.000
	$y_{1.3}$	0.828	0.025	33.038	0.000

Data source*: Processed results from smartPLS 4*

Table 3 shows that the loading factor values link the Specific Intervention construct to the indicator variables $x_{1.1}$, $x_{1.2}$, $x_{1.3}$, $x_{1.4}$, $x_{1.5}$, $x_{1.6}$, $x_{1.7}$, $x_{1.8}$, and $x_{1.9}$. The Sensitive Intervention construct to the indicator variables $x_{2.1}$, $x_{2.3}$, $x_{2.4}$, $x_{2.5}$, $x_{2.7}$, $x_{2.9}$ and $x_{2.11}$. The association between the nutritional status construct and the indicator variables $y_{1.1}$, $y_{1.2}$ and $y_{1.3}$ has a t-statistic value > 1.960 at a significance threshold of α = 0.05, and each has a loading value (λ) ≥ 0.5.

The following equation was created using the loading factor and standard error values found in Table 3.

i. Exogenous latent variable 1 (Specific Intervention)

$$x_{1.1} = 0.696\xi_1 + 0.055$$
$$x_{1.2} = 0.821\xi_1 + 0.033$$
$$x_{1.3} = 0.674\xi_1 + 0.043$$
$$x_{1.4} = 0.765\xi_1 + 0.034$$
$$x_{1.5} = 0.746\xi_1 + 0.037$$
$$x_{1.6} = 0.505\xi_1 + 0.059$$
$$x_{1.7} = 0.754\xi_1 + 0.040$$
$$x_{1.8} = 0.716\xi_1 + 0.049$$
$$x_{1.9} = 0.750\xi_1 + 0.031$$

ii. Exogenous latent variable 2 (Sensitive Intervention)

$$x_{2.1} = 0.600\xi_2 + 0.048$$
$$x_{2.3} = 0.746\xi_2 + 0.039$$
$$x_{2.4} = 0.676\xi_2 + 0.051$$
$$x_{2.5} = 0.686\xi_2 + 0.048$$
$$x_{2.7} = 0.662\xi_2 + 0.048$$
$$x_{2.9} = 0.748\xi_2 + 0.039$$
$$x_{2.11} = 0.828\xi_2 + 0.025$$

iii. Latent variable endogenous (Nutritional Status)

$$y_{1.1} = 0.978\eta + 0.010$$
$$y_{1.2} = 0.917\eta + 0.034$$
$$y_{1.3} = 0.953\eta + 0.032.$$

b. Examining the inner model's structural model hypothesis

The following hypothesis was applied in order to test the inner model parameters.

i. Sensitive intervention (ξ_1) on Nutritional Status (η)

$$H_0: \gamma_{11} = 0,$$
$$H_1: \gamma_{11} \neq 0.$$

ii. Specific intervention (ξ_2) on Nutritional Status (η)

$$H_0: \gamma_{12} = 0,$$
$$H_1: \gamma_{12} \neq 0.$$

The t-table value, with a significance level of α = 5 percent, is 1.960. $t_{stat} > t_{table}$ or $p-value < \alpha$, reject H_0. Table 4 displays the findings of the path coefficient estimation.

Table 4. Estimation of Path Coefficient Values

Latent Variable	Parameter Coefficient	Standard deviation	t-statistics	p-value
Specific Interventions-> Nutritional Status	-0.133	0.055	2.412	0.016
Sensitive Intervention-> Nutritional Status	-0.229	0.057	4.040	0.000

***Data source**: Processed results from smartPLS*

The following explanation explains the influence of the link between factors in Table 4. Specific interventions affect nutritional status, The path parameter coefficient obtained from the relationship between the Sensitive Intervention variable and nutritional status is -0.133, with a t-statistic value of 2.412 > 1.960 (t_{table}), at a significance level of $\alpha = 5\%$, indicating that there is a significant effect between specific interventions and nutritional status. Sensitive interventions influence nutritional status. The path parameter coefficient obtained from the relationship between the specific intervention variable and nutritional status is -0.133 with a $t_{statistic}$ value of 4.040 > 1.960 (t_{table}) at a significance level of α = 5%, indicating a significant effect of sensitive interventions on nutritional status.

3.4 Quantile Regression in SEM for Nutritional Status

To determine the coefficient for each predictor variable, this study used the Quantile Regression approach to fit a model for each quantile that explains the degree to which the predictor variables affect nutritional status at that quantile. Furthermore, the degree to which the predictor factors influence nutritional status at each quantile is shown. Table 5 displays the results of the parameter estimation.

Table 5. Estimation of Path Quantile Regression

Parameter	Quantile (τ)		
	0.05	**0.5**	**0.95**
$\beta_0(\tau)$	-1.55578	-0.08226	1.66386
$\beta_1(\tau)$	-0.26339	-0.23091	0.01856
$\beta_2(\tau)$	0.22905	-0.19032	-0.57280

***Data source**: Processed results from RStudio*

The quantiles indicate that several predictor variables have both positive and negative effects on nutritional status. The factors with a major impact on nutritional status at each quantile will be examined next. Table 6 presents the results of the parameter significance test.

Table 6. p-Value from The Results of Parameter Estimation

Parameter	Quantile (τ)		
	0.05	**0.5**	**0.95**
$\beta_0(\tau)$	0.00000	0.08237	0.00000
$\beta_1(\tau)$	0.26947	0.00029	0.90690
$\beta_2(\tau)$	0.33682	0.00276	0.00034

***Data source**: Processed results from RStudio*

There is substantial variance in the quantile regression parameters across quantiles, as shown in Table 6, which shows the p-values from the estimation of the parameters at the three quantile levels (0.05, 0.5, and 0.95). With p-values of 0.00029 and 0.00276, respectively, the parameters Specific Interventions ($\beta_1(\tau)$) and Sensitive Interventions ($\beta_2(\tau)$) are significant at the median quantile but not at the lower or upper quantiles. The quantile regression approach is beneficial for capturing heterogeneity in covariate effects across the data distribution, as these results show that the influence of each parameter on the response variable varies across different portions of the distribution.

3.5 Bootstrap for each quantile in SEM

A quantile regression analysis was carried out using the bootstrap method to gain a more thorough understanding of each predictor's impact on the distribution of the response variable. This technique enables identification of shifts in the predictors' influence on the mean and on specific distributional segments, including the lower, middle, and upper quantiles. Three quantile levels, 0.05, 0.5, and 0.95, were employed in this instance to capture the dynamics of the predictor effects at different distributional positions. Additionally, as shown in Fig. 3 below, the predicted regression coefficients are displayed with 95% confidence intervals.

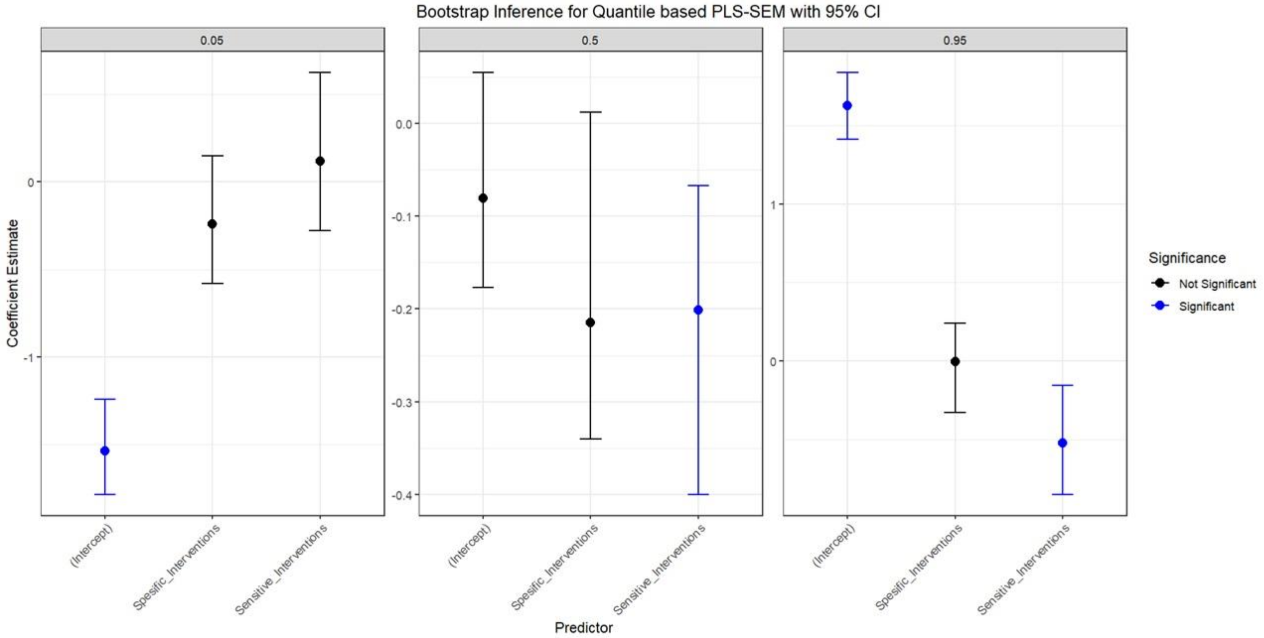

Figure 3. Bootstrap Quantile in SEM
***Source**: Processed results from RStudio*

Fig. 3 shows the results of Bootstrap Quantile Regression analysis with a 95% confidence interval at three quantiles (0.05, 0.5, and 0.95), which illustrates the effect of two predictors (Specific Intervention, and Sensitive Intervention) on the distribution of the response variable. At the 0.05 quantile, none of the predictors are significant. At the 0.5 quantile (median), the sensitive intervention predictor shows a significant negative effect on the response, while the other predictors are not significant. At the 0.95 quantile, sensitive intervention again shows a significant negative effect, while the intercept shows a significant positive effect. These results indicate that sensitivity has a significant effect that varies depending on the part of the response distribution being analyzed, with a tendency to decrease values at the median and upper quantiles, while specific intervention does not have a significant effect at all quantiles.

3.6 Wild Bootstrap for each quantile in SEM

Based on the results from the standard bootstrap, further analysis can be performed using the wild bootstrap to address potential heteroscedasticity and structural error dependence in the quantile regression model. Wild bootstrap provides more robust confidence intervals by accounting for more realistic variations in random errors, thereby providing more accurate and reliable estimates for testing the significance of coefficients at various quantiles. By using wild bootstrap, we can reduce bias arising from error distributions that do not meet classical assumptions and thereby improve statistical inference in this quantile regression model.

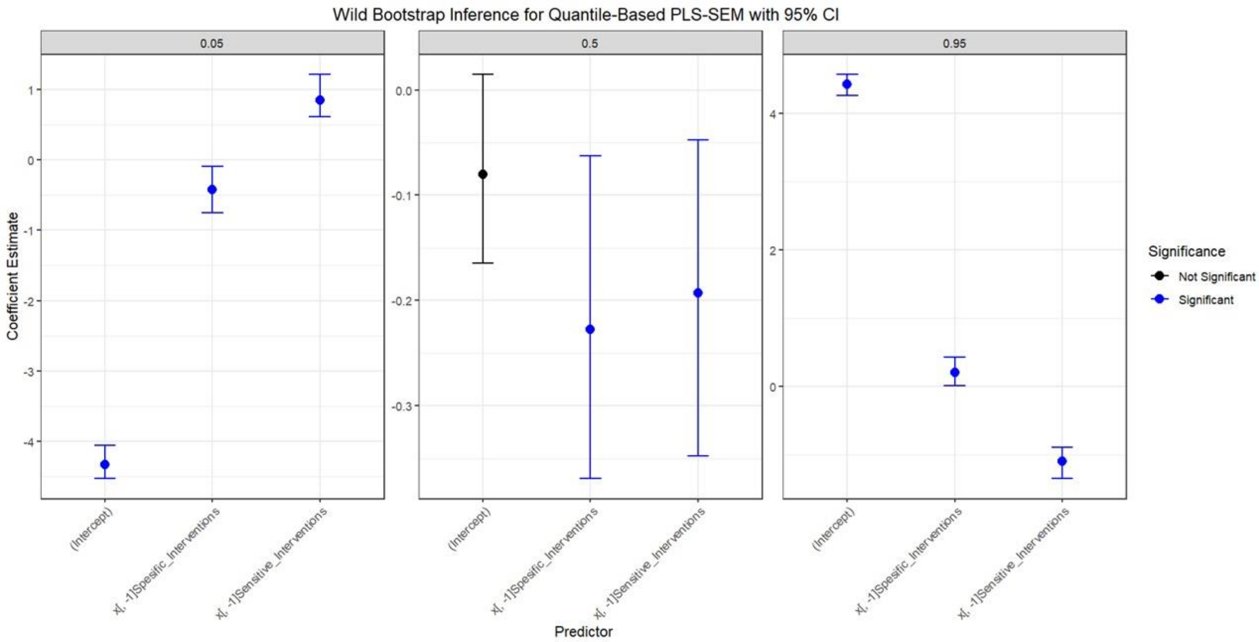

Figure 4. Wild Bootstrap Quantile in SEM
***Source**: Processed results from RStudio*

With negative coefficient estimates, Fig. 4 demonstrates that all predictors significantly impact the lower values of the response distribution at the 0.05 quantile. Only the sensitive intervention variable and the particular intervention variable are significantly negative at the 0.5 quantile; the intercept, which shows how predictors affect the median response value, is not significant. All three predictors, however, are significant at the 0.95 quantile, with the sensitive intervention variable having a negative influence and the intercept and specific intervention variable exhibiting large positive effects. Given that the direction and significance of effects can vary substantially across different portions of the distribution, these results highlight the need to account for the position within the distribution (quantile) when assessing the impact of predictors.

4. CONCLUSION

Based on the results of two approaches, Bootstrap Quantile in SEM and Wild Bootstrap Quantile in SEM, it can be concluded that the predictor consistently shows a significant negative effect on the median (quantile 0.5) and the upper quantile (0.95) of the response distribution, while at the lower quantile (0.05), the effect varies depending on the method used. In the conventional bootstrap analysis, the predictor does not exhibit a significant effect at any quantile, whereas in the wild bootstrap approach, it shows significance at the 0.05 and 0.95 quantiles. This indicates that the wild bootstrap method is more sensitive in detecting the significance of regression coefficients, especially under heteroskedasticity or unstable error distributions. Its main advantage lies in its ability to handle the heterogeneity of variance commonly found in real-world data, thereby producing more robust and accurate confidence interval estimates. Consequently, the use of the wild bootstrap in quantile regression enhances the reliability of statistical inference and provides a more comprehensive understanding of how predictor effects vary across the entire response distribution.

Author Contributions

Abdul Malik Balami: Conceptualization, Methodology, Writing-Original Draft, Software, Validation. Bambang Widjanarko Otok: Data Curation, Resources, Draft Preparation. Santi Wulan Purnami: Formal Analysis, Validation. All authors discussed the results and contributed to the final manuscript.

Funding Statement

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Acknowledgement

The authors would like to thank all individuals and organizations who provided support and assistance throughout this research. Special thanks to the participants who contributed valuable data, and to the teams who helped with the data processing and analysis. We also appreciate the technical and logistical support provided by various contributors throughout the study.

Declarations

The author declares that there are no conflicts of interest to report in this study.

Declaration of Generative AI and AI-assisted technologies

ChatGPT was used only as a language-editing aid. The authors confirm that no AI-generated content was incorporated without substantial author revision, and all statements were verified against the study's data and objectives.

REFERENCES

[1] G. Dash and J. Paul, "CB-SEM VS PLS-SEM METHODS FOR RESEARCH IN SOCIAL SCIENCES AND TECHNOLOGY FORECASTING," *Technol Forecast Soc Change*, vol. 173, Dec. 2021, doi: https://doi.org/10.1016/j.techfore.2021.121092.

[2] M. Sarstedt, J. F. Hair, C. M. Ringle, K. O. Thiele, and S. P. Gudergan, "ESTIMATION ISSUES WITH PLS AND CBSEM: WHERE THE BIAS LIES!," *J Bus Res*, vol. 69, no. 10, pp. 3998–4010, Oct. 2016, doi: https://doi.org/10.1016/j.jbusres.2016.06.007.

[3] C. Davino and V. E. Vinzi, "QUANTILE COMPOSITE-BASED PATH MODELING," *Adv Data Anal Classif*, vol. 10, no. 4, pp. 491–520, Dec. 2016, doi: https://doi.org/10.1007/s11634-015-0231-9.

[4] P. Dolce, C. Davino, and D. Vistocco, "QUANTILE COMPOSITE-BASED PATH MODELING: ALGORITHMS, PROPERTIES AND APPLICATIONS," *Adv Data Anal Classif*, vol. 16, no. 4, pp. 909–949, Dec. 2022, doi: https://doi.org/10.1007/s11634-021-00469-0.

[5] D. J. Lee, "BOOTSTRAP TESTS FOR STRUCTURAL BREAKS WHEN THE REGRESSORS AND THE SERIALLY CORRELATED ERROR TERM ARE UNSTABLE," *Bull Econ Res*, vol. 73, no. 2, pp. 212–229, 2021, doi: https://doi.org/10.1111/boer.12243.

[6] T. Bluhmki *et al.*, "A WILD BOOTSTRAP APPROACH FOR THE AALEN–JOHANSEN ESTIMATOR," *Biometrics*, vol. 74, no. 3, pp. 977–985, 2018, doi: https://doi.org/10.1111/biom.12861.

[7] M. A. Martin, "WILD BOOTSTRAP," *Wiley StatsRef: Statistics Reference Online*, vol. 2, pp. 1–6, 2017, doi: https://doi.org/10.1002/9781118445112.stat08015.

[8] C. Davino, P. Dolce, S. Taralli, and D. Vistocco, "COMPOSITE-BASED PATH MODELING FOR CONDITIONAL QUANTILES PREDICTION. AN APPLICATION TO ASSESS HEALTH DIFFERENCES AT LOCAL LEVEL IN A WELL-BEING PERSPECTIVE," *Soc Indic Res*, vol. 161, no. 2–3, pp. 907–936, Jun. 2022, doi: https://doi.org/10.1007/s11205-020-02425-5.

[9] H. Cheng, "QUANTILE-BASED PLS-SEM WITH BAG OF LITTLE BOOTSTRAPS," *Commun Stat Theory Methods*, vol. 53, no. 18, pp. 6561–6579, 2024, doi: https://doi.org/10.1080/03610926.2023.2248324.

[10] J. F. Hair, J. J. Risher, M. Sarstedt, and C. M. Ringle, "WHEN TO USE AND HOW TO REPORT THE RESULTS OF PLS-SEM," Jan. 14, 2019, *Emerald Group Publishing Ltd.* doi: https://doi.org/10.1108/EBR-11-2018-0203.

[11] H. Latan and R. Noonan, *PARTIAL LEAST SQUARES PATH MODELING: BASIC CONCEPTS, METHODOLOGICAL ISSUES AND APPLICATIONS*. Springer International Publishing, 2017. doi: https://doi.org/10.1007/978-3-319-64069-3.

[12] P. Tarka, "AN OVERVIEW OF STRUCTURAL EQUATION MODELING: ITS BEGINNINGS, HISTORICAL DEVELOPMENT, USEFULNESS AND CONTROVERSIES IN THE SOCIAL SCIENCES," *Qual Quant*, vol. 52, no. 1, pp. 313–354, Jan. 2018, doi: https://doi.org/10.1007/s11135-017-0469-8.

[13] M. A. Ashraf and H. Ahmed, "APPROACHES TO QUALITY EDUCATION IN TERTIARY SECTOR: AN EMPIRICAL STUDY USING PLS-SEM," *Educ Res Int*, vol. 2022, 2022, doi: https://doi.org/10.1155/2022/5491496.

[14] W. Zhang, B. Wang, and L. Zhou, "EVALUATION AND IMPROVEMENT OF ENGLISH-SPEAKING INSTRUCTION BASED ON PLS-SEM AND INTELLIGENT SPEECH SYSTEM," *Mobile Information Systems*, vol. 2022, 2022, doi: https://doi.org/10.1155/2022/1081327.

[15] J. Llorens-Terrazas, "AN ORACLE INEQUALITY FOR MULTIVARIATE DYNAMIC QUANTILE FORECASTING," *Journal of Business and Economic Statistics*, vol. 43, no. 3, pp. 603–614, 2025, doi: https://doi.org/10.1080/07350015.2024.2415107.

[16] T. Yoshida, "QUANTILE FUNCTION REGRESSION AND VARIABLE SELECTION FOR SPARSE MODELS," *Canadian Journal of Statistics*, vol. 49, no. 4, pp. 1196–1221, 2021, doi: https://doi.org/10.1002/cjs.11616.

[17] G. Ciuperca, "ADAPTIVE LASSO MODEL SELECTION IN A MULTIPHASE QUANTILE REGRESSION," *Statistics (Ber)*, vol. 50, no. 5, pp. 1100–1131, 2016, doi: https://doi.org/10.1080/02331888.2016.1151427.

[18] S. P. Kumah and J. O. Mensah, "ARE CRYPTOCURRENCIES CONNECTED TO GOLD? A WAVELET-BASED QUANTILE-IN-QUANTILE APPROACH," *International Journal of Finance and Economics*, vol. 27, no. 3, pp. 3640–3659, 2022, doi: https://doi.org/10.1002/ijfe.2342

[19] J. Chen and Y. Liu, "Small Area Quantile Estimation," *International Statistical Review*, vol. 87, no. S1, pp. S219–S238, 2019, doi: https://doi.org/10.1111/insr.12293.

[20] A. S. Blum and J. Salomon, "NOTHING COMES FROM NOTHING BUT THE RHO MESON: THE ORIGIN OF THE BOOTSTRAP CONCEPT IN PARTICLE PHYSICS," *Stud Hist Philos Sci*, vol. 114, p. 102082, Dec. 2025, doi: https://doi.org/10.1016/j.shpsa.2025.102082.

[21] C. Lamarche and T. Parker, "WILD BOOTSTRAP INFERENCE FOR PENALIZED QUANTILE REGRESSION FOR LONGITUDINAL DATA," *J Econom*, vol. 235, no. 2, pp. 1799–1826, 2023, doi: https://doi.org/10.1016/j.jeconom.2022.11.011.

[22] L. Amro, F. Konietschke, and M. Pauly, "INCOMPLETELY OBSERVED NONPARAMETRIC FACTORIAL DESIGNS WITH REPEATED MEASUREMENTS: A WILD BOOTSTRAP APPROACH," *Biometrical Journal*, vol. 66, no. 8, 2024, doi: https://doi.org/10.1002/bimj.70008.

[23] L. Merlo, L. Petrella, N. Salvati, and N. Tzavidis, "UNIFIED UNCONDITIONAL REGRESSION FOR MULTIVARIATE QUANTILES, M-QUANTILES, AND EXPECTILES," *J Am Stat Assoc*, vol. 119, no. 547, pp. 2154–2165, 2024, doi: https://doi.org/10.1080/01621459.2023.2250512.

[24] Kementerian Kesehatan Republik Indonesia, *SURVEI STATUS GIZI INDONESIA: SSGI 2024 DALAM ANGKA*. Jakarta, 2024