Method Article

EFFECT OF MULTICOLLINEARITY ON MULTILEVEL NON-LINEAR MODEL

(suggest an alternative title:
THE EFFECT OF INCREASING SAMPLE SIZE ON MULTICOLINEARITY IN NON-LINEAR STAGE LOGISTIC REGRESSION MODELS)

Abstract **(It is complete and appropriate)**
A multilevel logistic regression model demonstrating high correlations among predictor variables is susceptible to multicollinearity. This phenomenon contributes to an escalation in the variances of parameter estimates, thereby resulting in inaccurate inferences regarding the relationships between the response and explanatory factors. The primary objective of this study is to investigate the impact of multicollinearity on multilevel nonlinear models. The research aims to assess whether the quantity of independent variables influences the variance inflation factor and to explore the effect of altering the correlation degree at one level on multicollinearity within a multilevel nonlinear model. Additionally, the research seeks to determine how multicollinearity affects the standard errors of parameters in a multilevel nonlinear model. In a 2-level logistic regression, a binary variable served as the dependent variable, while pre-established normal variables functioned as regressors. The Monte Carlo analysis incorporated three distinct correlation strengths (0.2, 0.5, and 0.9) and sample sizes (500, 100, and 30). The Variance Inflation Factor was employed for multicollinearity diagnosis. The outcomes revealed that, within the logistic multilevel regression model, an increase in sample size correlated with a reduction in multicollinearity. Notably, the influence of multicollinearity on standard errors in a multilevel nonlinear model was more pronounced. It was observed that expanding the sample size remains an effective strategy to mitigate multicollinearity errors in a multilevel nonlinear model. This approach is particularly crucial due to the reliance on maximum likelihood estimation (MLE) in logistic regression, as opposed to ordinary least squares (OLS) regression, which contrasts with the methodology of OLS regression.
Key words: Multilevel, Logistic Regression, Multicollinearity, Standard Errors, Nonlinear model, Variance Inflation Factor

Introduction **(Suggests that you should use general rules for scientific writing, in the following order: Introduction, Materials and Methods, Results and Interpretation, Discussion, Conclusion, Acknowledgments, and References)**

Multicollinearity testing to determine whether there is a perfect linear relationship between the independent variables. The consequence of this multicollinearity is that the regression coefficient is not certain or the standard error is infinite. Multicollinearity can be seen with VIF (variance inflation factor). If the VIF value is <10 and the tolerance value is above 0.10, then there are no symptoms of multicollinearity and vice versa.

Multicollinearity values indicate a high level of linear intercorrelation between explanatory variables in multiple regression models and cause incorrect regression analysis results. Multicollinearity diagnostic tools include variance inflation factor (VIF), condition index and condition number, and variance decomposition proportion (VDP). Multicollinearity can be expressed by the coefficient of determination (Rh2) of a multiple regression model with one explanatory variable (Xh) as the model response variable and another (X i [i≠h] as the explanatory variable. variance (σh2) of the regression coefficients that form the regression model final is proportional to VIF (11 -RH2). Therefore, increasing Rh2 (strong multicollinearity) increases σh2 . Larger σh2 results in unreliable probability values and confidence intervals of regression coefficients. The square root of the ratio of the maximum eigenvalue to each eigenvalue of the correlation matrix of the standard explanatory variables is referred to as the condition index and the condition number is the maximum condition index. Multicollinearity occurs when the VIF is higher than 5 to 10 or the condition index is higher than 10 to 30. However, multicollinearity cannot indicate multicollinear explanatory variables. The VDP obtained from eigenvectors can identify multicollinear variables by showing the amount of inflation σh 2 according to the respective condition index. When two or more VDPs, corresponding to a higher general condition index of 10 to 30, are higher than 0.8 to 0.9, then the corresponding explanatory variables are multicollinear.

Pinheiro and Bates [1] advocate the use of nonlinear latent coefficient models due to their interpretability, parsimony, and, most importantly, their validity beyond the observed data range. Precision in approximating the genuine, typically nonlinear, regression function within the observed data range can be enhanced by increasing the order of a polynomial model. However, high-order polynomial models not only frequently lead to multicollinearity issues but also fail to provide theoretical insights into the underlying mechanism generating the data [2].
The Richards function, a generalization of non-linear functions [3], describes growth curves, with specific instances including the logistic, Gompertz, and monomolecular functions [4]. Presently, one can select curves of the form $y = f_1(x) + f_2(x) + e$ where the first component, $f_1(x)$ may have the form:

- Gompertz: $c_1 \exp(-c_2 \exp(-c_3 x))$
- logistic: $\frac{c_1}{(1+s\exp(c_2-c_3x))}$
- Monomolecular: $c_1(1+ s \exp(c_2 - c_3 x))$
- Exponential: $c_1 \exp(-c_2 x)$
- Power: $c_1 x^{bc2}$

The second component, $f_2(x)$, may have the form:

- Gompertz: $cd_1 \exp(d_2 \exp(-d_3 x))$

• logistic: $\frac{d_1}{(1+s\exp(d_2-d_3x))}$
• Monomolecular: $d_1(1+s\exp(d_2-d_3x))$
• Exponential: $d_1\exp(-d_2x)$
• Power: $d_1x^{d_2}$
The sign of the term $\exp(b_2 - b_3x)$ is denoted by 's' in the aforementioned curves, and it may assume values of either 1 or -1. Opting for a curve from this family is preferable over a polynomial curve due to the distinctive physical meanings associated with the parameters in the first three functions. These parameters can often be fitted to a set of responses with an equivalent level of precision. Parameter c_1 signifies the time asymptotic value of the observed characteristic, parameter c_2 represents the potential increase (or decrease) in the function's value during the interval t_1 to t_p, and parameter b_3 indicates the growth rate.

Similar to linear hierarchical models, the coefficients c_1, c_2, …, d_3 can be expressed as level-2 outcome variables [5].
$c_1= \delta_1+u_1$
$c_2= \delta_2+u_2$
$c_3= \delta_3+u_3$
$d_1= \delta_4+u_4$
$d_2= \delta_5+u_5$
$c_3= \delta_6+u_6$
Assuming zero means and a covariance matrix Φ, the normal distribution is attributed to the level-2 residuals u_1,u_2, … , u_6. Within the LISREL framework, it is also conceivable to posit that a level-2 covariate has an impact on the values of any random coefficient, thereby introducing a general association.
$c_1= \delta_1+ \gamma_1z_1+u_1$
$c_2= \delta_2+ \gamma_2z_2+u_2$
$d_3= \delta_3+ \gamma_3z_3+u_3$
$d_1= \delta_4+ \gamma_4z_4+u_4$
$d_2= \delta_5+ \gamma_5z_5+u_5$
$c_3= \delta_6+ \gamma_6z_6+u_6$
where z_i denotes the value of a covariate and γ_i the corresponding coefficient.
Generally, a single component ($f_1(x)$) is adequate for elucidating various monotone growing or decreasing growth patterns. However, for more complex patterns, two-component regression models prove useful. Within the LISREL framework, the user has the flexibility to select any of the five curve types for component 1 and combine it with a different curve type for component 2. Valid choices are, for example,
• Logistic
• Monomolecular + Gompertz
• Exponential + logistic
• Logistic + logistic
The model parameters encompass the vector of fixed coefficients (δ), the vector of covariate coefficients (γ), the variance (σ^2) of the level-1 measurement errors, and the covariance matrix (Π) of the level-2 residuals, all of which are unknown.

Multilevel Non-Linear Regression Models Estimation Procedure

Given that y represents a linear combination of the random coefficients, it follows that it exhibits a normal distribution within linear multilevel models. For example, the intercept-and-slopes-as-outcomes model is

$$y = c_1 + c_2 x + e$$

where $c_1 = u_1$, $c_2 = u_2$ and (u_1,u_2) is assumed to be normally distributed.

Nonlinear multilevel regression models preclude the normal distribution of y since they cannot be articulated as a linear combination of their coefficients [6]. A method for assessing the probability density function of y involves employing multiple integrals.

$$f(y) = \int_{b1} \dots \int_{c3} f(y, c_1, \dots, d_3) c g_1 \dots g d c_3$$

that, in general, cannot be solved in closed form.

To evaluate the likelihood function

$L = \prod_{i=1}^{n} f(y_i)$,

Application of numerical integration techniques is imperative in this context. Assumed distributions for e and the set (c_1,c_2 ,c_3,d_1,d_2, d_3) are $N(0,\Phi^2)$ and $N(0,\sigma^2)$, respectively.

In the multilevel procedure, a Gauss quadrature approach is employed for the numerical evaluation of integrals. To initiate the ML technique, it is crucial for the unknown parameters to possess appropriate initial values. A comprehensive account of the estimating process is provided by Cudeck and du Toit [7].

i. How nonlinear models is affected with Multicollinearity?

When the increment $\Delta b_{(i-1)}$ is sufficiently small, the solution for these algorithms can be derived from the first-order Taylor series approximation of the error function SSE [8].

$SSE(b_{(i)}) = SSE(b_{(i-1)}) + Z(\Delta b_{(i-1)})$ 1.12

where the elements of the matrix Z are the derivatives

$Z_{ij} = \{\partial e_i / \partial b_j\}$ 1.13

Therefore, at step "i" the algorithm search to minimize the updated error function 3) with respect to the new value of b: $b_{(i)}$

$b_{(i)} = b_{(i-1)} - (Z'Z)^{-1}.Z.SSE(b_{(i-1)})$ 1.14

This adaptive searching methodology is reminiscent of the pseudolinear regression approach [9, 10, 11], presenting a computationally less demanding alternative to the prediction error method [12]. The pseudolinear regression method is also referred to as approximate maximum likelihood or the extended least squares method.

This expectation arises from the fact that the linear approximation is obtained through the first-order Taylor expansion. In general, the elements of the Hessian matrix (H) for a sum-of-squares error function as in equation 2) are

$H_{ij} = \{\partial^2 SSE / \partial b_i \partial b_j\} = \{(\partial e / \partial b_i).(\partial e / \partial b_j) + e.(\partial^2 e / \partial b_i \partial b_j)\}$ 1.15

H can then be approximated by the matrix $Z'.Z$:

$(Z'.Z)_{ij} = \{(\partial e / \partial b_i).(\partial e / \partial b_j)\}$ 1.16

by discarding the term e. $(\partial^2 e / \partial b_i \partial b_j)$ of the equation 1.15.

This signifies that in the minimization process of equation 1.14, the Hessian matrix, when multiplied by the gradient of the error function ($\Delta SSE = Z$), yields the Newton step, also known as the Newton direction.

$b_{(i)} = b_{(i-1)} - H^{-1}.\Delta SSE$ 1.17

Exact solutions determined by the Hessian matrix are feasible only for linear problems with finite datasets [13]. In contrast, for nonlinear structures, such as those encountered in infinite datasets, the Newton direction asymptotically converges towards the global minimum. Consequently, it becomes imperative to estimate and update the Hessian for finite samples at each stage.
The challenge lies in the fact that, irrespective of the sample size, the pseudolinear regression procedure necessitates the inversion of the Hessian matrix, specifically in its approximated form Z'.Z. For $(Z'.Z)^{-1}$ to be invertible, it is essential that the columns of $(Z'\times Z)^{-1}$ exhibit linear independence. Consequently, the issue of multicollinearity arises in the iterative optimization processes of nonlinear models [14].

ii. Effects of Multicollinearity

Multicollinearity within the design matrix has a notable impact on the least squares estimation of parameters [15]. The variances of the parameters are given

$$\mathrm{Var}(\hat{\beta}_j) = \sigma^2 \frac{1}{1-\mathrm{R}_j^2} \qquad (\mathrm{R}_j^2 \leq 1)\ j=1, 2, \ldots, p \qquad 1.18$$

Where $\hat{\beta}$ is the least squares estimator of δ. $\frac{1}{1-\mathrm{R}_j^2}$ is the diagonal elements of $(X'X)^{-1}$ matrix in equation (1.18). Also, R_j^2 is the coefficient of multiple determination from the regression of x_j on the remaining $p-1$ regressor variables. If there is strong multicollinearity between x_j and any other regressors, the value of R^2_j will be close to unity ($\mathrm{R}_j^2 \to 1$), diagonal elements of variance-covariance matrix will be grow considerable ($\frac{1}{1-\mathrm{R}_j^2} \to \infty$)
In the assessment of different parameter estimators or the proximity of an estimator to the true parameter, mean square error is commonly employed [16, 17].
In the presence of multicollinearity in the model, the mean square error can be altered using the formula,

$$\begin{aligned} E(L_1^2) &= E[(\hat{\beta}_j - \delta)'(\hat{\beta}_j - \delta)] = \mathrm{MSE}(\hat{\beta}) \\ &= \sum_{j=i}^{p} E(\hat{\beta}_j - \beta_j)^2 \\ &= \sum_{j=1}^{p} Var(\hat{\beta}_j) \qquad\qquad 1.19 \\ &= \sigma^2 \mathrm{Tr}(X'X)^{-1} \\ &= \sigma^2 \sum_{j=1}^{p} \frac{1}{\lambda_j}, \qquad \lambda_j > 0,\ j = 1, 2, \ldots p \end{aligned}$$

can be used. The values of λj represent the eigenvalues of the $X'X$ correlation matrix. In equation (1.19), $L_1^2 = (\hat{\beta}j - \delta)'(\hat{\beta}j - \delta)$ denotes the square of the distance from the estimator to the true parameter. In cases where multicollinearity in the design matrix renders the $X'X$ correlation matrix ill-conditioned, at least one λ_j will be considerably small, leading to the exclusion of estimation values from the real parameter [50].
To identify multicollinearity, the literature commonly refers to the examination of the correlation matrix, variance inflation factor, and eigenvalue eigenvector analysis of the correlation matrix [18, 16]. Multicollinearity typically arises from model selection; for example, incorporating explanatory variables with a high degree of correlation can result in an undesirable multicollinearity issue. The specific objectives of this study encompass:

1. To explore the impact of a change in the degree of correlation at one level on the multicollinearity value in a multilevel nonlinear model.
2. To assess whether the number of independent variables influences the value of the variance inflation factor.

3. To examine the influence of multicollinearity on the parameter standard errors in the multilevel nonlinear model.

THEORETICAL FRAMEWORK

Introduction

In multiple regression analysis, multicollinearity is a major difficulty that frequently results from the assumption of linear independence among the explanatory variables being broken. A statistical problem known as multicollinearity occurs when there is a strong correlation between two or more of the independent variables in a regression model [19]. This issue is commonly discussed in terms of its degree rather than mere presence or absence.

Belsley [20] notes the absence of a clear definition of multicollinearity in the literature. When data vectors representing two variates lie on the same line, they are termed collinear, forming a subspace of dimension one. In a broader sense, *k* variates are considered collinear if their vectors are linear combinations of others and lie in a subspace of dimension fewer than *k*. While "exact multicollinearity" is rare in practice, a more comprehensive understanding is required to address the issue in statistical estimation. By accounting for small angles between vectors, multicollinearity is extended to more than two variates, suggesting a high degree of correlation. Additionally, multicollinearity can make survival analysis variable selection more difficult [21].

In regression analysis, multicollinearity becomes problematic when independent variables are correlated or exhibit intercorrelations [22]. This issue can lead to dependencies between the estimated regression coefficients and associated predictor variables. Greene [23] identifies multicollinearity by observing wide swings in parameter estimates with minor changes in data, coefficients with inconsistent sign magnitudes and very large standard errors, even when statistically significant and with a high R^2.

The negative influence of Multicollinearity

Regression based on multicollinearity data has two negative effects:

Computational problem

The application of a condition number signifies its role in providing an estimate of the sensitivity of a linear equation's solution to variations in the constituent parts of the system [24]. In computational terms, this implies that solutions to a system of linear equations or a set of least-squares normal equations will inherently possess a limited interpretative accuracy, determined by the conditioning of the data concerning the condition number. Essentially, the condition number serves as a multiplicative factor, allowing imprecision in the data to propagate into imprecision in the solution of the linear system of equations [25].

To illustrate, if data are trusted to four digits and the condition number of (X^TX) is 10^3, a change in the fifth place of data could impact the least-squares solution in its second (5-3) significant digit. However, a small change in the data's last place might affect the solution $Z = A^{-1}C$ in the $(d\text{-}r)^{th}$ place. Consequently, only the first digit can be considered reliable, with all other digits potentially rendered useless and subject to arbitrary changes in X that do not influence the accuracy of the information. It is evident that none of the significant numbers in C could be trusted if the condition number of X^TX were 10^4 or 10^5 in this scenario. While computational adjustments can mitigate this issue when calculating least-squares estimates, complete elimination remains unattainable.

Statistical problem

In statistical analysis, the presence of multicollinearity within a data matrix is widely acknowledged to result in diminished precision of statistical estimates, manifesting as elevated conditional variances [26]. Poorly conditioned data, characterized by nearly linearly dependent series, contribute little independent information, adversely affecting regression-based estimation, forecasting, and hypothesis testing by inflating variances.

Users of linear regression often encounter situations where confidence or forecast intervals appear excessively large, signaling inadequately conditioned data. Additionally, tests of significance may yield inconclusive results due to elevated error variance induced by collinearity [27]. While obtaining fresh, well-conditioned data is the most immediate remedy, practical constraints often limit this option. Thus, diagnostic tools become invaluable for identifying and isolating multicollinearity [28]. These tools enable researchers to assess whether corrective measures, such as gathering new data or employing Bayesian techniques, are warranted.

The statistical treatment of multicollinearity as a problem traces back to Farar and Glauber's [29] work in 1967, where they proposed a method based on an orthogonal data matrix null hypothesis. Their approach involves a statistical test against the null hypothesis, indicating the absence of multicollinearity if accepted. Critics like Haitovsky (1969) [30] challenged this orthogonality null hypothesis, advocating for tests against a perfect singularity null hypothesis. Kumar [31] also critiqued the assumption of stochasticity in Farar and Glauber's method, emphasizing that it does not align with the fixed X assumption in the standard regression model.

While multicollinearity introduces computational challenges and diminishes statistical precision, it is not a statistically testable scenario within the linear regression model [32]. Instead, it requires a combination of statistical and numerical methods. Silvey [33] notes the similarity of this approach to the singular value decomposition method in numerical analysis, evident through the presence of small eigenvalues in the $X'X$ matrix.

The singular-value decomposition (SVD)

Any matrix X, with n observations on p variates and dimensions n x p, can be decomposed as

$X = UDV'$ 2.1

where U is n x p, D is p x p, and V is p x p. An alternative formulation,

${}_nX_p = {}_nU_n\,{}_nD_n\,{}_pV_p^T$ 2.2

In this formulation, D always maintains full rank, denoted as r = rank(X).

The singular value of X is defined as $U'U = V'V = I$ and D, where μ_k, k = 1,…, p, represents nonnegative elements on the diagonal. The centrality or scaling of X does not alter this definition. However, for the subsequent multicollinearity diagnostic, it is recommended to scale X to have equal (unit) column length. If the data are relevant to a model with a constant term, X should include both uncentered data and a column of ones. Using the centered data matrix X in this context is discouraged, as it may obscure the constant's role in any underlying dependencies and yield inaccurate diagnostic outcomes.

X is a matrix of predefined variables, which may include lagged values of the response variable y. While the singular-value decomposition (SVD) differs from eigenvalues and eigenvectors, it is closely connected to these concepts. The relationship that $X'X = VD^2V'$ signifies that V is an orthogonal matrix diagonalizing X'X. Thus, the squares of the singular values in D^2, the diagonal components of D^2, correspond to the eigenvalues of the real symmetric matrix X'X. The columns of U are the p eigenvectors of X'X related to its p nonzero eigenvalues, and the orthogonal columns

of V are the eigenvectors of X'X. The information provided by the eigensystem of X'X is encompassed by the SVD of the matrix X. However, practical considerations justify the preference for the SVD method. Firstly, it directly relates to the data matrix X, rather than the cross-product matrix X'X. Second, the concept of the condition number of X is more accurately described in terms of the singular values of X (spectral norm) rather than the square roots of the eigenvalues of X'X. Third, introducing the idea of the singular-value decomposition to a broader audience is beneficial due to its extensive practical and analytical applications in matrix algebra. Finally, though computationally comparable technically, the eigensystem and the SVD of a matrix differ. In instances where X is poorly conditioned, algorithms for computing the SVD of X offer higher numerical stability than those for the eigensystem of X'X.
In practical terms, it is advisable to perform the multicollinearity diagnostic using the stable SVD approach for X rather than employing an algorithm to determine the eigenvalues and eigenvectors of X'X.

Detection of Multicollinearity

Test of the determinant of the correlation matrices:

The determinant of the matrix, constrained within the correlation domain, spans a range from zero to one. Specifically, it attains a value of one in instances of orthogonal independent variables and decreases to zero in scenarios of complete linear dependency. The proximity of the determinant to zero signifies an escalation in multicollinearity. However, this method does not offer the ability to identify the specific variables contributing to multicollinearity [34] (Peck & Montgomery, 1981).

Analysis of the correlation matrix:

In the course of this procedure, scrutiny is applied to the non-diagonal elements (r_{ij}) of the correlation matrix. The correlation coefficient (r_{ij}) becomes particularly relevant when examining the variables X_i and X_j, which are subjects of study, as it approaches unity in absolute values when these variables exhibit a significant linear dependency. High correlation coefficients, indicative of multicollinearity, emerge when variables demonstrate strong linear interdependence. However, it is imperative to note that the significance of this requirement is heightened when dealing with more than two independent variables, as the absence of robust correlation between two variables does not guarantee the absence of multicollinearity [37].

Eigenvectors and eigenvalues analysis in the correlation matrix:

The matrix condition number (CN), evaluated based on the symmetry of the matrix, indicates the relationship between the largest and smallest eigenvalues. A small CN suggests the presence of one or more close linear dependencies among the variables [36], reflected in one or more small eigenvalues ($\lambda_1,\lambda_2,\ldots, \lambda_p$) of the correlation matrix.
According to Montgomery and Peck [34], multicollinearity is deemed insignificant if CN < 100. The severity of multicollinearity is classified as moderate to strong if 100 < CN < 1000 and severe if CN > 1000.
To assess the approximate nature of linear dependencies among variables, an analysis of eigenvalues can be conducted [37]. In these analyses, $R = V\Lambda V'$, where Λ is an orthogonal matrix with dimensions p x p, and its columns (v_1, v_2, …, v_p) represent the normalized eigenvectors of R. Λ is a diagonal matrix with dimensions p x p (where p is the number of variables used for the R correlation matrix), and its elements are the eigenvalues λ_j (j = 1,2,…,p) of R.

A near-zero eigenvalue (λ_j) indicates a linear correlation between observations, with the characteristics of this dependency expressed by the components of the eigenvector (v_j) associated with that specific eigenvalue.

Inflation factors of the variance:

Obtaining the matrix $X'X$ in correlation form, [38] provides the diagonal elements of the matrix $R_{xx}^{-1} = (X'X)^{-1}$, representing the variance inflation factors (VIF) [34]. Multicollinearity can be identified using these VIF criteria. In cases of linear dependency between variables, the variance of the j^{th} least squares regression coefficient is denoted by $v_{jj}\sigma^2$, where v_{jj} serves as the factor affecting the variance of $\hat{\beta}$. The impact of multicollinearity on the least squares regression coefficients becomes significant when VIF values exceed 10.

According to Freund and Wilson [39], multicollinearity does not substantially affect the coefficient estimators as long as the VIFs are less than ($\frac{1}{1-R^2}$), where R^2 is the coefficient of determination for the model Y considering all X variables. In this scenario, the correlation between the independent variables and the dependent variable Y is stronger than their intercorrelation.

Splitting into singular values

Expressed as $X = UDV'$, where X is an n x p matrix with n observations and p variables, the decomposition of $X'X$ is given by $(UDV')'(UDV') = VD^2V'$ [37]. The matrix U is of dimensions n x p, with columns representing eigenvectors corresponding to the eigenvalues of $X'X$. The normalised eigenvectors of $X'X$ constitute matrix V, which is p x p. The matrices $U'U = V'V = I_p$ ensure orthogonality. The diagonal matrix D, of dimensions p x p, contains non-negative diagonal components μ_j (j = 1, 2, p), representing the unique values of X. The expression X = UDV' thus signifies the decomposition of X into its singular values.

The size of singular values is affected by the degree of poor conditioning in matrix X. As the singular value for each approximate linear dependency increases, so does the size of the singular values. This relationship is captured by the matrix's condition index (η_k), defined as

$$\eta_k = \frac{\mu_{max}}{\mu_k}, \ k = 1, 2, \ldots, p$$

thus, $\eta_k \geq 1$, for all k

the variance of the least squares estimator of δ for each k is denoted by

$$V(\hat{\beta}) = \sigma^2 (X'X)^{-1} = \sigma^2 V\Lambda^{-1}V$$

Or for the k^{th} component of $\hat{\beta}$,

$$V(\hat{\beta}_k) = \sigma^2 \sum_{j=1}^{p} \frac{v_{kj}^2}{\mu_j^2} = \sigma^2 \sum_{j=1}^{p} \frac{v_{kj}^2}{\lambda_j}$$

Except for σ^2, the k^{th} element of $V\Lambda^{-1}V$ diagonal is the k^{th} inflation factor of the variance, thus

$$(VIF)_k = = \sum_{j=1}^{p} \frac{v_{kj}^2}{\mu_j^2} = \sum_{j=1}^{p} \frac{v_{kj}^2}{\lambda_j}$$

The presence of one or more small singular values, or eigenvalues, will cause inflation in the $\hat{\beta}_j$ variance [35].

The variance splitting process can also be done to measure the degree of multicollinearity by

$$\Pi_{jk} = \frac{\frac{v_{kj}^2}{\mu_j^2}}{(VIF)_k} \ j = 1, 2, \ldots, p$$

The elements Π_{jk} are ordered in a $\Pi_{p \times p}$ dimension matrix.

The elements of each column of matrix Π are constituted by the variance of each $\hat{\beta}_k$, denoted as the variance inflation factor $(VIF)_k$ associated with the i^{th} singular value. Low values of μ_j and elevated proportions of Π_{jk} indicate an association between multicollinearity and the specific singular value, resulting in the inflation of $\hat{\beta}$ variances

METHODOLOGY

Monte-Carlo Experiment

The Monte-Carlo experiment, in its probabilistic approach, involves observing and selecting random numbers to simulate the original problem's physical random process [40]. Conversely, the analytical method of the Monte-Carlo experiment leverages theoretical mathematics to address issues that are currently solved numerically but remain unsolvable theoretically [40].

The Monte-Carlo studies follow a concise procedure:

1. For Levels 1 and 2, the experimenter creates correlated normal random variables using Cholesky decomposition, a matrix operation.
2. For logistic regression, a binary dependent variable is created.
3. The researcher defines a model, assigning specified numerical values as parameters believed to represent the actual parameter values.

These three stages are iteratively executed for R-replications. By treating the generated data as real-world data, the experimenter obtains parameter estimates for each replication, including estimates for various estimation strategies and sample sizes.

Multilevel Model Structure

The nonlinear MLMs consisted of Level 1 (X_1, X_2, X_3 and X_4) and Level 2 (Z_1, Z_2, Z_3 and Z_4) predictors of the Level 1 outcome (Y). The simulated level 1 model is:

$$Logit(Y_{ij}) = \beta_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + \beta_{3j}X_{3ij} + \beta_{4j}X_{4ij} + r_{ij} \qquad 3.1$$

The simulated level 2 model is

$$\beta_{0j} = \gamma_{00} + \gamma_{01}Z_{1j} + \gamma_{02}Z_{2j} + \gamma_{03}Z_{3j} + \beta_{4j}X_{4ij} + U_{0j} \qquad 3.2$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}Z_{1j} + \gamma_{12}Z_{2j} + \gamma_{13}Z_{3j} + \gamma_{14}Z_{4j} + U_{1j} \qquad 3.3$$

$$\beta_{2j} = \gamma_{20} + \gamma_{21}Z_{1j} + \gamma_{22}Z_{2j} + \gamma_{23}Z_{3j} + \gamma_{24}Z_{4j} + U_{2j} \qquad 3.4$$

$$\beta_{3j} = \gamma_{30} + \gamma_{31}Z_{1j} + \gamma_{32}Z_{2j} + \gamma_{33}Z_{3j} + \gamma_{42}Z_{4j} + U_{4j} \qquad 3.5$$

The estimated coefficients encompassed the Level 1 error variance component, $\sigma^2 r$, and all fixed-effect coefficients (γ's). Each predictor in the models was standardized to have a variance of one and a mean of zero. Additionally, the mean for Y was set to zero. The slope variances (τ_{11}, τ_{22}, τ_{33}) were consistently maintained at 0.05. Furthermore, the intraclass correlation coefficient (ICC) for Y was uniformly fixed at 0.20 across all conditions.

Sample size

The study manipulated the number of groups and level sizes to examine the influence of sample size. Three different group sizes were selected, consisting of 30, 100, and 1000, reflecting principles commonly observed in organizational literature and primarily derived from those articulated by Shieh and Fouladi [41].

Multilevel Variance Inflation Factor (MVIF)

A multilevel version of the MVIF will be employed, calculated in a manner analogous to Ordinary Least Squares (OLS) models. The Multilevel VIFs (MVIFs) are akin to the VIFs in OLS models and are obtained as the diagonals of the inverse of the predictor correlation matrix. In the case of two Level 1 predictors (X_1 and X_2) and two Level 2 predictors (Z_1 and Z_2), forming a four-by-four correlation matrix (X_1, X_2, Z_1, Z_2), the MVIFs are represented by the four values on the diagonal of the inverse. When cross-level interactions are present, a product term is introduced and incorporated into the correlation matrix.

Parameter estimates will be calculated and standard errors for a single sample drawn from a population with a multilevel data structure. The modeling approach employed resemble that of the previous OLS regression example.

Generating the Explanatory Variables

This study employed the formula for generating explanatory variables as established by McDonald and Galarneau [42], Gibbons [43], Wichern and Churchill [44], and Kibria (2003) [45]: The equation is given as:

$X_{ij} = (1-\rho^2)^{\frac{1}{2}} Z_{ij} + \rho Z_{ip},$ i= 1, 2, 3, …, n. j=1, 2, …p. 3.6

The correlation between any two explanatory variables is denoted by ρ, where p represents the number of explanatory variables, and Z_{ij} is an independent standard normal distribution with a mean of zero and unit variance. Values of ρ, ranging from 0.9 to 0.1, were considered in this investigation. Each level in the study involves a p-value of 4. The regression coefficients for the Monte-Carlo simulation research were specified as follows: β_1=4, β_2=2.5, β_3=1.8 and β_4=3.1. It is noteworthy that, as is often the case in simulation studies, the parameter values were chosen to satisfy the condition $\beta'\beta$=1 [46].

Generating the Dependent Variable

This study focuses on a dichotomous dependent variable, and logistic regression is employed as the nonlinear model. In logistic regression analysis, outcomes are commonly coded as either 0 or 1, with 0 indicating the absence of the desired outcome and 1 indicating its presence. The logistic regression model is formulated as follows, defining p as the likelihood of the outcome being 1:

$$\hat{P} = \frac{\exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p)}{1 + \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p)}$$

$\hat{p}$ is the expected probability that the outcome is present; X_1 through X_p are distinct independent variables; and δ_0 through δ_p are the regression coefficients.

The iterative equations for estimating the parameters of δ of a logistic regression model are

$\beta^{t+1} = \beta^t + (X'WX)^{-1}X'S$

where both $W = diag\{\hat{p}_i(1-\hat{P}_i)$ and $S = Y - \hat{Y}$ are evaluated at δ^t. The variance-covariance matrix for the estimated parameters is

$$Var(\hat{\beta}) = (\hat{X}WX)^{-1}$$

To estimate the parameters and their variances, the information matrix $(X'WX)$ need to be inverted. Since W is a non-singular diagonal matrix, $X'WX = X'W^{1/2}W^{1/2}X = L'L$, where $L\, W^{1/2}X$. It can be shown that Rank(X) = Rank(L).

Hence, a non-full rank data matrix X implies the matrix $L'L$ to be ill-conditioned. As a consequence, the estimates and variances of the parameters of logistic regression model will be affected by the presence of dependencies among the columns of X.

Random variables simulation with a defined correlation

Cholesky decomposition was used, a matrix operation effective for generating correlated data, akin to taking a square root with scalars. This method is employed to mimic multivariate data that adheres to a specific covariance matrix. The Cholesky decomposition of a covariance matrix M involves a lower triangular matrix L, satisfying M = L L'.

By generating a vector z with random, normally distributed values having a mean of zero and a variance of one (with a length equal to the dimension of M), the product Lz was employed to create a realization of our multivariate distribution. The variance of Lz equals L times the variance of z. Since L is essentially a constant, the variance of z is represented by the identity matrix I, as the random integers have variance one and are independently distributed. Thus, Variance(Lz) = LIL' = LL' = M. Consequently, we generate random data conforming to the intended covariance matrix.

Results of Findings

To investigate the impact of the number of independent variables on the VIF and understand how alterations in the degree of correlation at one level influence multicollinearity in a multilevel nonlinear model, the analysis focuses on VIF, a prominent method for detecting multicollinearity issues. VIF assesses the rate at which variances and covariances escalate.

Neter et al. [47] recommended focusing on the highest VIF value, often indicative of multicollinearity problems when exceeding 10. The simulation results, summarized in Tables 1 to 6, were derived from experiments with a sample size of thirty. In each experiment, four predictors were present at Level One, and one Level Two predictor was sequentially eliminated. The correlation between Level One predictors remained constant at 0.9, while the correlation between Level Two predictors varied between 0.2 and 0.5.

Table 1: Level One with $\rho = 0.9$, 4 predictors. Level Two with ρ=0.2, 4 Predictors. n=30

	Estimate	Standard Error	z-Score	P-value	VIF
Constant	-2.18664242	1.05910778	-2.06460802	0.0389601	
Z	0.882953	0.81938968	1.07757398	0.28122392	12.8300071
Y	2.12038655	1.26826768	1.67187621	0.09454874	18.0996
X	-0.92168803	1.65733031	-0.55612814	0.57812326	12.8938724
W	-1.57278369	1.33550473	-1.17766988	0.23892824	13.3038761
D	-16.9697456	8.9436142	-1.89741476	0.05777322	15.110131
C	24.7572123	24.4237117	1.01365479	0.3107475	6.02391233
B	9.59300308	4.87902311	1.96617291	0.04927864	7.26918455
A	-3.59890154	2.7958542	-1.2872279	0.19801486	8.98650973

Table 2: Level One with $\rho = 0.9$, 4 predictors. Level Two with ρ=0.2, 3 Predictors. n=30

	Estimate	Standard Error	z-Score	P-value	VIF
(Intercept)	-2.18664242	1.05910778	-2.06460802	0.0389601	
Z	0.882953	0.81938968	1.07757398	0.28122392	12.9211517

Y	2.12038655	1.26826768	1.67187621	0.09454874	6.57826422
X	-0.92168803	1.65733031	-0.55612814	0.57812326	8.02851961
W	-1.57278369	1.33550473	-1.17766988	0.23892824	10.9308619
D	-16.9697456	8.9436142	-1.89741476	0.05777322	7.34734441
C	24.7572123	24.4237117	1.01365479	0.3107475	9.62721296
B	9.59300308	4.87902311	1.96617291	0.04927864	7.74892755

Table 3: Level One with $\rho = 0.9$, 4 predictors. Level Two with $\rho=0.2$, 2 Predictors. n=30

	Estimate	Standard Error	z-Score	P-value	VIF
(Intercept)	-2.18664242	1.05910778	-2.06460802	0.0389601	
Z	0.882953	0.81938968	1.07757398	0.28122392	12.2867826
Y	2.12038655	1.26826768	1.67187621	0.09454874	5.76218159
X	-0.92168803	1.65733031	-0.55612814	0.57812326	6.94143214
W	-1.57278369	1.33550473	-1.17766988	0.23892824	11.851558
D	-16.9697456	8.9436142	-1.89741476	0.05777322	1.02860598
C	24.7572123	24.4237117	1.01365479	0.3107475	1.22481972

Tables 1, 2, and 3 illustrate VIF values for each predictor, all of which are below 10. This suggests that multicollinearity does not exert an influence on the coefficients of the multilevel nonlinear regression. The observed lack of impact could be attributed to the relatively low correlation ($\rho=0.2$) among the predictors. Conversely, a noteworthy observation was the discernible influence of the diminishing number of predictors on the VIF.

Table 4: Level One with $\rho = 0.9$, 4 predictors. Level Two with $\rho=0.5$, 4 Predictors. n=30

	Estimate	Standard Error	z-Score	P-value	VIF
(Intercept)	0.428623	0.490868	0.873195	0.382557	
Z	-0.56783	0.355441	-1.59754	0.110146	12.648678
Y	0.219655	0.476092	0.46137	0.644533	8.426218
X	0.455316	0.670519	0.679051	0.497106	9.856104
W	0.429266	0.49845	0.861203	0.389126	14.104909
D	-1.45637	1.57146	-0.92676	0.354051	2.528198
C	2.581878	5.494042	0.469941	0.638397	2.149806
B	0.634162	1.12954	0.561434	0.574502	2.395000
A	0.402727	0.657304	0.612695	0.540078	1.313592

Table 5: Level One with $\rho = 0.9$, 4 predictors. Level Two with $\rho=0.5$, 3 Predictors. n=30

	Estimate	Standard Error	z-Score	P-value	VIF
(Intercept)	0.3767538	0.5928932	0.6354497	0.52513517	
Z	-0.1284161	0.4425487	-0.2901739	0.77168320	14.263429
Y	-2.0389196	0.9564087	-2.1318496	0.03301921	19.571273
X	2.7764063	1.3673011	2.0305742	0.04229820	40.952278
W	-0.4401214	0.6274875	-0.7014027	0.48305176	19.001700

D	-5.1052587	4.1820805	-1.2207462	0.22218214	6.305815
C	14.6140259	13.6115925	1.0736456	0.28298157	4.526346
B	3.8850649	2.6270555	1.4788667	0.13917595	5.881204

Table 6: Level One with $\rho = 0.9$, 4 predictors. Level Two with ρ=0.5, 2 Predictors. n=30

	Estimate	Standard Error	z-Score	P-value	VIF
(Intercept)	0.088133739	0.6181332	0.14258049	0.88662150	
Z	0.004593639	0.3691077	0.01244525	0.99007038	12.626259
Y	-1.119023849	0.6742833	-1.65957525	0.09699993	13.136033
X	1.563868244	0.9636338	1.62288649	0.10461367	27.090726
W	-0.406214510	0.4802038	-0.84592114	0.39759674	14.096024
D	-7.282646642	10.5536686	-0.69005830	0.49015753	6.149020
C	1.785462556	1.7946057	0.99490523	0.31978236	6.368545

Most variables in Level One exhibit VIF values exceeding 10, as evident in Tables 4, 5, and 6. This suggests that multicollinearity is impacting these variables. However, it is noteworthy that all VIF values for Level Two are below 10, signifying the absence of multicollinearity in this level.

Table 7: Level One with $\rho = 0.9$, 4 predictors. Level Two with ρ=0.9, 4 Predictors. n=30

	Estimate	Standard Error	z-Score	P-value	VIF
(Intercept)	0.05575316	0.42749182	0.13041925	0.89623474	
Z	-0.24105328	0.30194211	-0.79834268	0.42467166	0.05575316
Y	0.33587676	0.39045871	0.86021069	0.38967291	-0.24105328
X	-0.5079478	0.68433589	-0.74224926	0.45793633	0.33587676
W	0.10725637	0.31900783	0.33621862	0.73670601	-0.5079478
D	2.22087375	2.10723251	1.05392914	0.29191536	0.10725637
C	-2.12851738	4.69272023	-0.45357858	0.65013217	2.22087375
B	0.31758025	0.89441038	0.35507219	0.72253551	-2.12851738
A	-1.02038004	0.81079811	-1.25848843	0.20821517	0.31758025

Table 8: Level One with $\rho = 0.9$, 4 predictors. Level Two with ρ=0.9, 3 Predictors. n=30

	Estimate	Standard Error	z-Score	P-value	VIF
(Intercept)	-0.46645983	0.4787782	-0.97427124	0.32992188	
Z	-0.26059648	0.33237176	-0.78405121	0.43301006	7.17228065
Y	0.46587606	0.43935808	1.0603562	0.28898258	7.98717126
X	-0.72425721	0.7534759	-0.96122147	0.33644082	6.55989843
W	0.25435822	0.37682285	0.67500743	0.49967105	7.14014901

D	-3.29101814	2.23095603	-1.47516047	0.14016941	4.0443979
C	1.87421353	7.18554594	0.26083105	0.7942228	2.69182657
B	3.32273862	1.56694174	2.12052467	0.03396182	3.31452198

Table 9: Level One with $\rho = 0.9$, 4 predictors. Level Two with ρ=0.9, 2 Predictors. n=30,

	Estimate	Standard Error	z-Score	P-value	VIF
(Intercept)	0.02060845	0.42775935	0.04817768	0.96157464	
Z	-0.27003472	0.2712651	-0.99546431	0.31951049	6.90205684
Y	0.23148145	0.35470534	0.6526021	0.51401283	6.8460401
X	-0.48068617	0.65318235	-0.7359142	0.46178291	7.39516561
W	0.27706381	0.317092	0.87376477	0.38224637	6.97829045
D	1.36558072	4.64143071	0.29421547	0.76859326	1.45833467
C	-0.67137335	0.81586726	-0.82289533	0.41056752	1.5481649

All Level One variable exhibit VIF values exceeding 10, as depicted in Tables 7, 8, and 9. This illustrates the pervasive impact of multicollinearity on each Level One variable and underscores how the correlation degree between variables influences the VIF value. While Tables 8 and 9 show no impact of multicollinearity on other Level Two variables, the VIF for variables X and Y (Level Two variables) in Table 7 also surpasses 10. This suggests that the number of variables influences the VIF value at Level Two.

Tables 10 through 18 present the outcomes for a sample size of 100, maintaining four Level One predictors across experiments, eliminating one Level Two predictor after each experiment, maintaining a consistent correlation of 0.9 among Level One predictors, and varying the correlations of 0.2, 0.5, and 0.9 between Level Two predictors.

Table 10: Level One with $\rho = 0.9$, 4 predictors. Level Two with ρ=0.2, 4 Predictors. n=100

	Estimate	Standard Error	z-Score	P-value	VIF
(Intercept)	0.11447783	0.21888917	0.52299448	0.60097809	
Z	0.23905193	0.15350267	1.55731441	0.11939586	7.01912802
Y	-0.22864157	0.25227978	-0.90630159	0.36477624	7.70756776
X	-0.23882671	0.35542198	-0.67195256	0.5016139	8.51005711
W	0.14009423	0.20791368	0.67380957	0.50043239	7.66687306
D	-1.08234147	0.61024817	-1.77360872	0.07612789	1.28687401
C	-3.16793907	2.59338911	-1.22154406	0.22188009	1.31494942
B	0.24967593	0.42095542	0.59311726	0.55310268	1.34647078
A	0.03081048	0.24218899	0.12721667	0.89876891	1.17113971

Table 11: Level One with $\rho = 0.9$, 4 predictors. Level Two with ρ=0.2, 3 Predictors. n=100

	Estimate	Standard Error	z-Score	P-value	VIF
(Intercept)	0.07058262	0.21316468	0.33111778	0.74055552	
Z	0.15746509	0.14220671	1.10729717	0.26816547	6.43790192
Y	-0.17375202	0.2443976	-0.71093995	0.47712145	7.8092688

X	-0.0471682	0.33469392	-0.14092937	0.88792574	8.09114877
W	0.08366117	0.19775491	0.42305481	0.67225527	7.56432478
D	-0.48542689	0.58699628	-0.82696757	0.40825544	1.15055654
C	-0.11010659	2.19589761	-0.05014195	0.96000927	1.09789838
B	0.41144701	0.55765931	0.73781071	0.46062948	1.23396306

Table 12: Level One with $\rho = 0.9$, 4 predictors. Level Two with ρ=0.2, 2 Predictors. n=100

	Estimate	Standard Error	z-Score	P-value	VIF
(Intercept)	0.0982901	0.21296992	0.46152103	0.64442484	
Z	0.2023663	0.14679553	1.37855894	0.16803078	6.78325026
Y	-0.2074628	0.24978331	-0.8305711	0.40621597	8.12874616
X	-0.06517078	0.33302156	-0.19569538	0.8448486	7.90283864
W	0.06556562	0.19502138	0.33619709	0.73672225	7.36121007
D	-0.18967388	0.37812201	-0.50162084	0.61593426	1.1392319
C	-1.68371637	2.28715341	-0.73616241	0.46163186	1.18259837

Tables 10, 11, and 12 illustrate that as the sample size increased, the impact of multicollinearity diminished, affecting only variable A in Table 10. The VIF values for all other variables in both Levels One and Two are below 10. This suggests that multicollinearity exerts no influence on the coefficients in the multilevel nonlinear regression when the VIF is less than 10. This observation may be attributed to the modest correlation (ρ=0.2) among the predictors. However, it was observed that the reduction in the number of variables corresponded with a decrease in VIF, indicating a relationship between the number of predictors and VIF.

Table 13: Level One with $\rho = 0.9$, 4 predictors. Level Two with ρ=0.5, 4 Predictors. n=100

	Estimate	Standard Error	z-Score	P-value	VIF
(Intercept)	0.11447783	0.21888917	0.52299448	0.60097809	
Z	0.23905193	0.15350267	1.55731441	0.11939586	7.01912802
Y	-0.22864157	0.25227978	-0.90630159	0.36477624	7.70756776
X	-0.23882671	0.35542198	-0.67195256	0.5016139	8.51005711
W	0.14009423	0.20791368	0.67380957	0.50043239	7.66687306
D	-0.97383305	0.70154552	-1.38812525	0.16509891	1.70072762
C	-3.70555483	3.00309843	-1.23391055	0.21723624	1.78873712
B	0.28035709	0.48381297	0.5794741	0.56226932	1.82256026
A	0.03553238	0.27930596	0.12721667	0.89876891	1.49851064

Table 14: Level One with $\rho = 0.9$, 4 predictors. Level Two with ρ=0.5, 3 Predictors. n=100

	Estimate	Standard Error	z-Score	P-value	VIF

(Intercept)	0.07058262	0.21316468	0.33111778	0.74055552	
Z	0.15746509	0.14220671	1.10729717	0.26816547	6.43790192
Y	-0.17375202	0.2443976	-0.71093995	0.47712145	7.8092688
X	-0.0471682	0.33469392	-0.14092937	0.88792574	8.09114877
W	0.08366117	0.19775491	0.42305481	0.67225527	7.56432478
D	-0.59413987	0.67233297	-0.88369885	0.3768588	1.50940607
C	-0.46334617	2.64224168	-0.17536101	0.86079598	1.46031377
B	0.32455378	0.43988759	0.73781071	0.46062948	1.77927114

Table 15: Level One with $\rho = 0.9$, 4 predictors. Level Two with ρ=0.5, 2 Predictors. n=100

	Estimate	Standard Error	z-Score	P-value	VIF
(Intercept)	0.09652227	0.21333527	0.45244406	0.65094911	
Z	0.19153423	0.14560605	1.31542776	0.18836618	6.70837228
Y	-0.1676876	0.24390466	-0.6875129	0.49175958	7.77540985
X	-0.09020903	0.32968536	-0.27362159	0.78437544	7.79556246
W	0.0688218	0.19494405	0.35303362	0.72406323	7.35865816
D	-0.22213642	2.19202652	-0.10133838	0.91928185	1.00992568
C	-0.26939258	0.36268027	-0.74278256	0.45761334	1.04916854

Tables 13, 14, and 15 illustrate that, with a sample size increased to 100, the VIF values for all variables in both Level One and Level Two of the multilevel nonlinear regression are below 10. This observation suggests that multicollinearity exerts no influence on these variables under the expanded sample size condition.

Table 16: Level One with $\rho = 0.9$, 4 predictors. Level Two with ρ=0.9, 4 Predictors. n=100

	Estimate	Standard Error	z-Score	P-value	VIF
(Intercept)	0.33926658	0.21830352	1.55410492	0.12015939	
Z	0.10039055	0.16130088	0.62238065	0.5336916	8.75370487
Y	-0.3916471	0.27394233	-1.42966989	0.15281179	10.446227
X	0.80375112	0.33734005	2.38261399	0.01719021	8.0426806
W	-0.20082483	0.18827789	-1.06664052	0.2861342	8.52330422
D	-2.16532376	1.5236384	-1.42115331	0.15527219	7.47566895
C	3.54029076	5.77036641	0.61352963	0.5395262	6.24065519
B	0.54730946	0.97757927	0.55986197	0.57557359	6.4869721
A	-0.0547511	0.6821492	-0.08026265	0.93602837	6.9341913

Table 17: Level One with $\rho = 0.9$, 4 predictors. Level Two with ρ=0.9, 3 Predictors. n=100

	Estimate	Standard Error	z-Score	P-value	VIF

(Intercept)	0.34479734	0.22526998	1.53059601	0.12586927	
Z	0.12700942	0.15843753	0.80163724	0.42276283	8.54247298
Y	-0.34859232	0.27520963	-1.26664287	0.20528302	10.6561337
X	0.65942723	0.32056715	2.05706427	0.03968004	7.30351812
W	-0.21186878	0.18870592	-1.12274583	0.26154546	8.67045011
D	1.88760565	1.61356772	1.16983355	0.24206796	7.99515882
C	-4.74538595	6.44073802	-0.73677674	0.46125813	7.95587878
B	-0.47425259	1.04737094	-0.4528029	0.65069068	8.21785844

Table 18: Level One with $\rho = 0.9$, 4 predictors. Level Two with ρ=0.9, 2 Predictors. n=100

	Estimate	Standard Error	z-Score	P-value	VIF
(Intercept)	0.3473448	0.21334828	1.62806472	0.10351116	
Z	0.11455148	0.15985121	0.71661314	0.47361283	8.76989206
Y	-0.36378986	0.2711663	-1.34157474	0.17973392	10.4365753
X	0.69309477	0.31892529	2.17321985	0.02976377	7.32552956
W	-0.19058474	0.18112177	-1.05224646	0.29268649	8.11258847
D	-1.0109329	4.85884619	-0.20806028	0.8351819	5.34199484
C	0.15454423	0.76062085	0.2031817	0.83899302	5.27793249

The impact of multicollinearity in the multilevel nonlinear model diminishes with an expanded sample size of 100, despite the high correlation among variables. Tables 16, 17, and 18 reveal that the VIF values for Level One variables C and D exceed 10, highlighting the effect of multicollinearity on these variables. However, no other Level One or Level Two variables are affected by multicollinearity.

Tables 19 through 21 present the outcomes obtained with a sample size of 500, maintaining four predictors for Level One throughout the experiments, removing one Level Two predictor after each experiment, maintaining a fixed correlation of 0.9 between Level One predictors, and varying correlations of 0.2, 0.5, and 0.9 between Level Two predictors.

Table 19: Level One with ρ=0, 4 predictors. Level Two with ρ=0.2, 4 Predictors. n=500

	Estimate	Standard Error	z-Score	P-value	VIF
(Intercept)	0.33826566	0.0918432	3.68307803	0.00023043	
Z	-0.039689	0.06245896	-0.63544118	0.52514072	7.214048
Y	-0.00721659	0.09855764	-0.07322205	0.94162942	7.2705297
X	0.03550608	0.14149146	0.25094149	0.80185935	7.57851303
W	0.01573136	0.08555138	0.18388201	0.85410602	7.86112653
D	0.14388598	0.2439545	0.58980661	0.55532031	1.22899878
C	-1.76445041	1.03507237	-1.70466381	0.08825714	1.26573036
B	0.16442824	0.17959778	0.91553608	0.35991027	1.21413886
A	-0.07813571	0.10366443	-0.75373699	0.45100716	1.18694254

Table 20: Level One with ρ=0.9, 4 predictors. Level Two with ρ=0.2, 3 Predictors. n=500.

	Estimate	Standard Error	z-Score	P-value	VIF
(Intercept)	0.33130813	0.09205867	3.59888016	0.00031959	
Z	-0.04403348	0.06251729	-0.7043408	0.48122056	7.24618485
Y	-0.01217774	0.09865858	-0.12343316	0.9017641	7.2835254
X	0.02484308	0.14005023	0.17738691	0.85920449	7.44061267
W	0.03128379	0.08450917	0.37018215	0.71124678	7.67808379
D	-0.30563059	0.23056273	-1.32558541	0.18497706	1.08281566
C	-0.32978375	0.93864075	-0.35134182	0.72533193	1.05144368
B	0.21933835	0.23936058	0.91635121	0.35948272	1.08000666

Table 21: Level One with ρ=0.9, 4 predictors. Level Two with ρ=0.2, 2 Predictors. n=500

	Estimate	Standard Error	z-Score	P-value	VIF
(Intercept)	0.34320625	0.09141209	3.75449523	0.00017369	
Z	-0.05018856	0.06251756	-0.8027915	0.42209526	7.24659316
Y	-0.0166593	0.09808902	-0.16983856	0.8651371	7.21863168
X	0.02655251	0.13980141	0.18993017	0.84936385	7.41483639
W	0.04205829	0.08552008	0.49179427	0.62286479	7.86778745
D	-0.13163358	0.15432097	-0.85298571	0.39366723	1.13328639
C	0.64787372	0.95740068	0.67670071	0.49859587	1.10802556

Tables 19, 20, and 21 illustrate that, with an increased sample size (n = 500), the VIF values for each variable in both Level One and Level Two of the multilevel nonlinear regression are below 10. This observation indicates the absence of multicollinearity impact. The lower correlation among variables (ρ=0.5) is likely a contributing factor to this outcome.

Table 22: Level One with ρ=0.9, 4 predictors. Level Two with ρ=0.5, 4 Predictors. n=500

	Estimate	Standard Error	z-Score	P-value	VIF
(Intercept)	0.33826566	0.0918432	3.68307803	0.00023043	
Z	-0.039689	0.06245896	-0.63544118	0.52514072	7.214048
Y	-0.00721659	0.09855764	-0.07322205	0.94162942	7.2705297
X	0.03550608	0.14149146	0.25094149	0.80185935	7.57851303
W	0.01573136	0.08555138	0.18388201	0.85410602	7.86112653
D	0.23646327	0.28503675	0.8295887	0.40677137	1.6777819
C	-1.99267194	1.19752683	-1.66398938	0.09611455	1.74851713
B	0.19573112	0.20612512	0.94957431	0.3423286	1.6661471
A	-0.0901105	0.11955165	-0.75373699	0.45100716	1.53781214

Table 23: Level One with ρ=0.9, 4 predictors. Level Two with ρ=0.5, 3 Predictors. n=500

	Estimate	Standard Error	z-Score	P-value	VIF
(Intercept)	0.33130813	0.09205867	3.59888016	0.00031959	
Z	-0.04403348	0.06251729	-0.7043408	0.48122056	7.24618485

Y	-0.01217774	0.09865858	-0.12343316	0.9017641	7.2835254
X	0.02484308	0.14005023	0.17738691	0.85920449	7.44061267
W	0.03128379	0.08450917	0.37018215	0.71124678	7.67808379
D	-0.33880127	0.27700575	-1.22308389	0.22129801	1.56298171
C	-0.55370523	1.0999995	-0.50336862	0.61470513	1.43859363
B	0.17301643	0.18881017	0.91635121	0.35948272	1.53402323

Table 24: Level One with ρ=0.9, 4 predictors. Level Two with ρ=0.5, 2 Predictors. n=500

	Estimate	Standard Error	z-Score	P-value	VIF
(Intercept)	0.33522297	0.09186881	3.64893118	0.00026333	
Z	-0.04379517	0.06259783	-0.69962766	0.48415987	7.27194424
Y	-0.02001281	0.09816139	-0.20387657	0.83844996	7.23162773
X	0.02153173	0.13996204	0.15383979	0.87773606	7.43557061
W	0.04011871	0.08535521	0.47002063	0.63834028	7.83879271
D	-0.78230369	0.92081159	-0.84958062	0.39555829	1.00946712
C	-0.10346734	0.14682484	-0.70469919	0.48099746	1.0253948

Owing to the substantial sample size (n=500), the VIF values for all variables in both Level One and Level Two of the multilevel nonlinear regression are below 10, as indicated in Tables 22, 23, and 24. This observation suggests the absence of any multicollinearity impact on the variables. The diminished correlation among variables (ρ=0.5) is a plausible explanation for this outcome.

Table 25: Level One with ρ=0.9, 4 predictors. Level Two with ρ=0.9, 4 Predictors. n=500

	Estimate	Standard Error	z-Score	P-value	VIF
(Intercept)	0.24199385	0.09225768	2.62302128	0.00871538	
Z	-0.04900362	0.06521755	-0.75138703	0.45241977	7.63743343
Y	0.24639825	0.10212527	2.41270592	0.01583459	7.51330147
X	-0.10925029	0.13886671	-0.78672768	0.43144129	7.06695187
W	-0.10899118	0.08378364	-1.30086472	0.19330476	8.06666065
D	-0.22252405	0.63704586	-0.34930616	0.72685947	7.97934325
C	-0.4685628	2.39541404	-0.19560827	0.84491679	7.18207975
B	0.82097658	0.45356065	1.81007014	0.07028491	9.27337523
A	-0.41969979	0.27760149	-1.51187874	0.13056471	7.6089514

Table 26: Level One with ρ=0.9, 4 predictors. Level Two with ρ=0.9, 3 Predictors. n=500

	Estimate	Standard Error	z-Score	P-value	VIF
(Intercept)	0.2460905	0.09169124	2.68390421	0.0072768	
Z	-0.04683302	0.06507374	-0.71969151	0.47171496	7.65006701
Y	0.22596181	0.10200686	2.21516286	0.02674889	7.52751065
X	-0.09879229	0.13853729	-0.71310975	0.47577785	7.07465256
W	-0.10337404	0.0835248	-1.23764487	0.21584777	8.05950185

D	-0.01221831	0.62556002	-0.0195318	0.98441687	7.35890212
C	1.24586239	2.40896027	0.51717847	0.60503158	7.00000788
B	-0.08191298	0.39944722	-0.20506585	0.83752067	7.03337185

Table 27: Level One with ρ=0.9, 4 predictors. Level Two with ρ=0.9, 2 Predictors. n=500

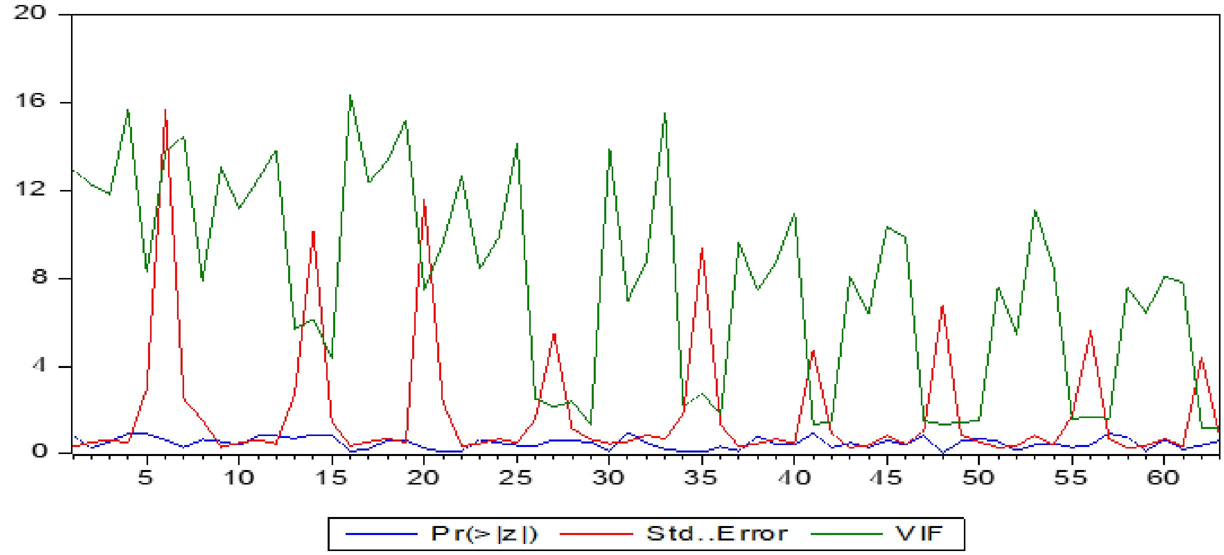
	Estimate	Standard Error	z-Score	P-value	VIF
(Intercept)	0.33522297	0.09186881	3.64893118	0.00026333	
Z	-0.04379517	0.06259783	-0.69962766	0.48415987	7.62426773
Y	-0.02001281	0.09816139	-0.20387657	0.83844996	7.40197289
X	0.02153173	0.13996204	0.15383979	0.87773606	7.01363592
W	0.04011871	0.08535521	0.47002063	0.63834028	8.04524435
D	-0.78230369	0.92081159	-0.84958062	0.39555829	5.46603009
C	-0.10346734	0.14682484	-0.70469919	0.48099746	5.46163797

As depicted in Tables 25, 26, and 27, the expansive sample size (n=500) mitigates the impact of the elevated correlation among variables. Notwithstanding, the VIF values for all variables in both Level One and Level Two of the multilevel nonlinear regression remain below 10. This observation underscores the negligible influence of multicollinearity on the variables.

EFFECT MULTICOLLINEARITY ON STANDARD ERRORS OF THE PARAMETERS IN MULTILEVEL NONLINEAR MODEL

An examination of the interplay among P-values, Standard Errors, and Variance Inflation Factor is presented across diverse sample sizes, correlation magnitudes, and diminishing variable quantities. Graphical depictions of these outcomes are illustrated in Figures 1, 2, and 3 for sample sizes of 30, 100, and 500, respectively.

Figure 1: When n=30, ρ=0.2, 0.5 and 0.9 (1 – 21 = 0.9; 22 – 42 = 0.5; 43 - 63 = 0.2)

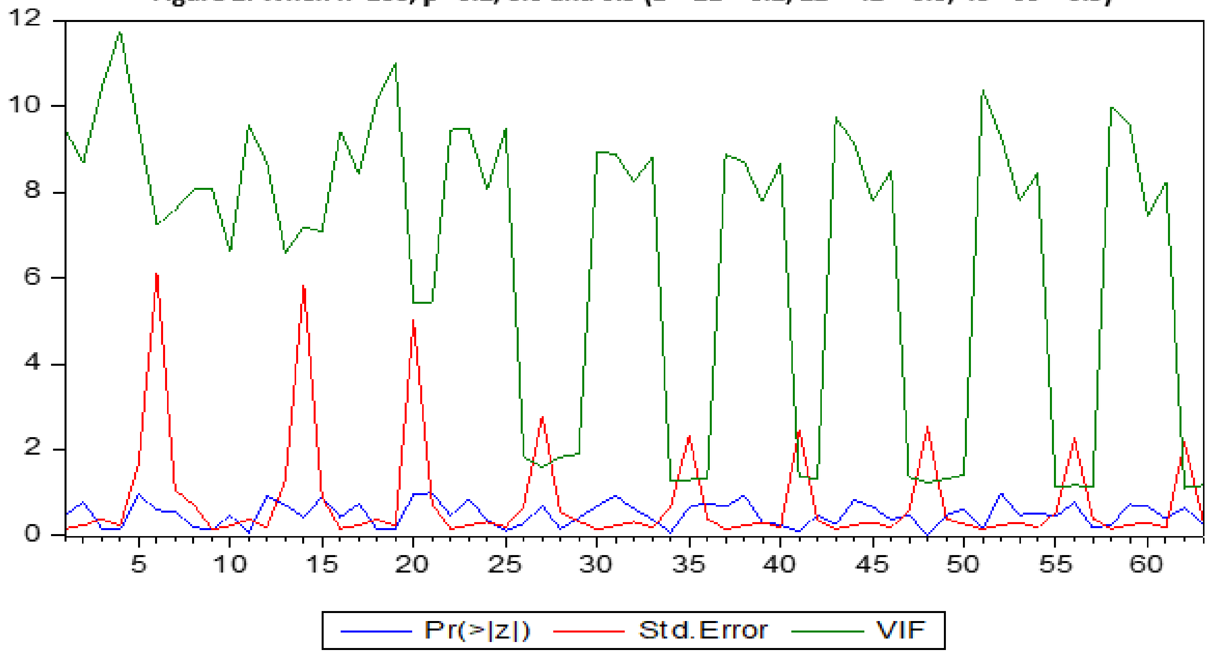


The X-axis represents the number of observations, totaling sixty-three across all experiments (twenty-one observations in each experiment with four variables in Level 1 and four in Level 2, four variables in Level 1 and three in Level 2, and four variables in Level 1 and two in Level 2). The experiment was conducted three times, with sample sizes of 30, 100, and 500, each denoted on the scale of five per unit.

The Y-axis displays serial numbers with a scale of four per unit.

The depicted results reveal a discernible pattern in the performance of Standard Errors and P-values for each level of multicollinearity (refer to Figure 1). It was observed that Standard Errors and multicollinearity intersected at a sample size of 30, and an increase in Standard Errors was noted with higher VIF values. The findings indicated that multicollinearity led to elevated Standard Errors in the parameter estimates, as illustrated in Figure 1, where both Standard Errors and VIF values escalated with the degree of correlation.

Figure 2: When n=100, ρ=0.2, 0.5 and 0.9 (1 – 21 = 0.2; 22 – 42 = 0.5; 43 - 63 = 0.9)

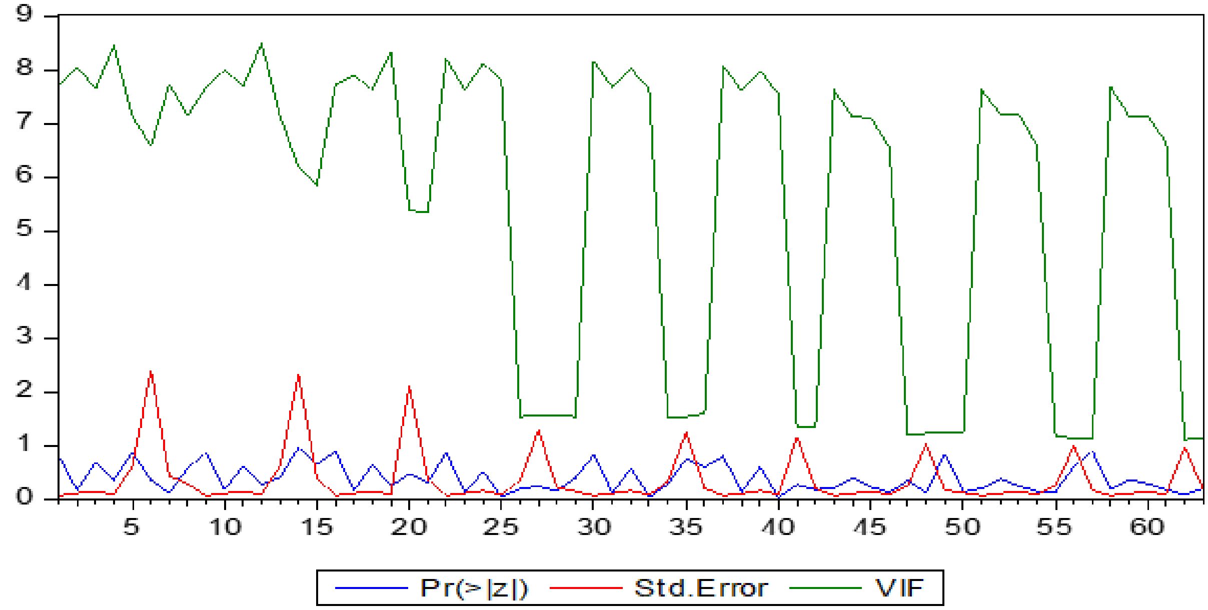


The X-axis represents the number of observations, totaling sixty-three across all experiments (twenty-one observations in each experiment with four variables in Level 1 and four in Level 2, four variables in Level 1 and three in Level 2, and four variables in Level 1 and two in Level 2). The experiment was conducted three times, with sample sizes of 30, 100, and 500, each denoted on the scale of five per unit.

The Y-axis displays serial numbers with a scale of four per unit.

Referring to Figure 2, the trend suggests that for a correlation degree of 0.9, Standard Errors disassociate from multicollinearity at a sample size of 100. The convergence of multicollinearity and standard errors was observed as the degree of correlation between variables decreased. Notably, an increase in VIF values corresponded to an elevation in Standard Errors. The findings indicated that multicollinearity led to heightened Standard Errors in the parameter estimates, as illustrated in Figure 2, where both Standard Errors and VIF values escalated with the degree of correlation.

Figure 3: When n=100, ρ=0.2, 0.5 and 0.9 (1 – 21 = 0.2; 22 – 42 = 0.5; 43 - 63 = 0.9)

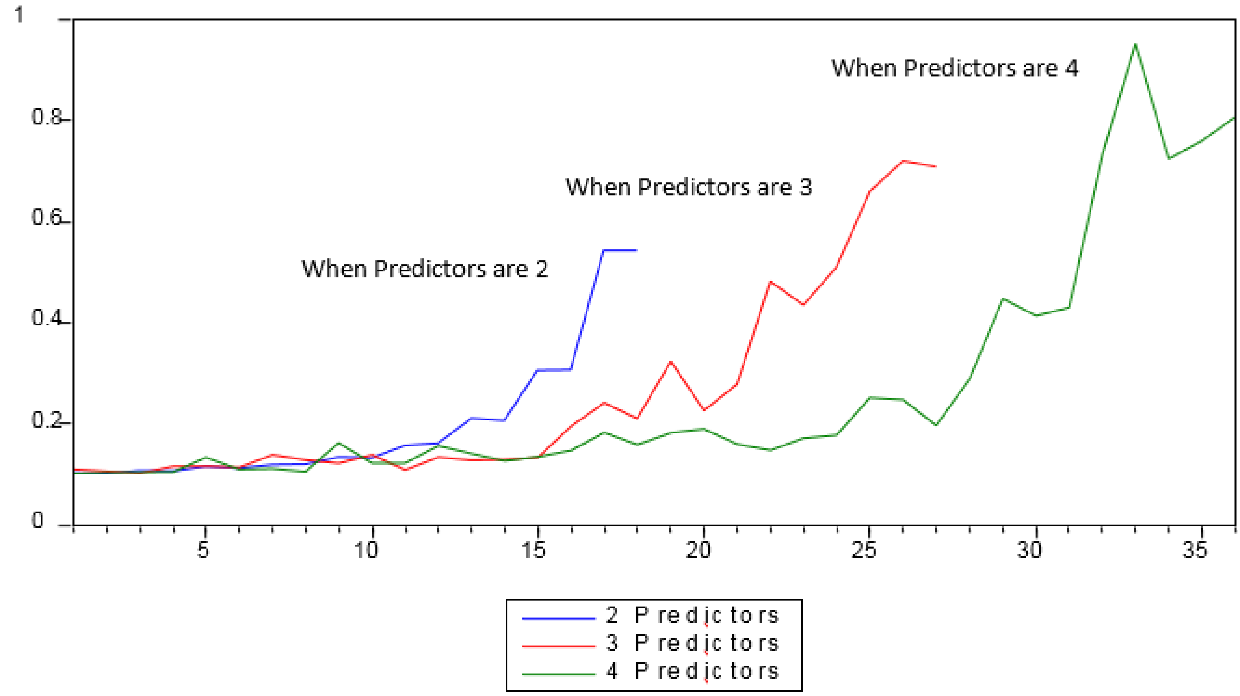


The X-axis denotes the number of observations, totaling sixty-three across all experiments: twenty-one observations in each experiment (with four variables in Level 1 and four in Level 2, four variables in Level 1 and three in Level 2, and four variables in Level 1 and two in Level 2). The experiment was conducted three times, with sample sizes of 30, 100, and 500, each represented on the scale of five per unit.

On the Y-axis, serial numbers are indicated, with a scale of four per unit.

In Figure 3, the trend of VIF, Standard Errors, and P-values is displayed for a sample size of 500. The pattern reveals that, owing to varying degrees of predictor correlation, the Standard Errors deviate from the VIF. It is noteworthy that Standard Errors increase with rising VIF values. The findings underscore that multicollinearity contributes to heightened Standard Errors in the parameter estimates.

Figure 4: VIF of variables of Level 2 against degree of correlation



The Variance Inflation Factor (VIF) for Level 2 variables in the multilevel model exhibits an upward trend as the degree of correlation increases, as illustrated in Figure 4. This pattern indicates that heightened multicollinearity corresponds to an increased correlation among the variables. Multicollinearity is observed when a model includes multiple factors interrelated with each other and the response variable.

DISCUSSION, CONCLUSIONS AND RECOMMENDATION

Discussion of Findings

In the Logistic Multilevel model, a binary variable served as the dependent variable, and predetermined correlated predictors were employed as the independent variables. The standard errors of the logit coefficients increased with a higher correlation among the explanatory factors. The results indicated that an escalation in multicollinearity corresponded to an augmentation in the standard errors of the parameter estimates. This outcome aligns with the observations of Adeboye, Fagoyinbo, and Olatayo [48], who noted a significant impact of multicollinearity on the standard errors of regression coefficients. The presence of multicollinearity poses notable challenges to regression analysis. The precision of estimating the influence of one variable on the outcome, after adjusting for others, tends to be less accurate when multicollinearity is prevalent, compared to scenarios where predictors are uncorrelated. Multicollinear predictors share information about the outcome variable, leading to potential overfitting in logistic regression analysis. Hence, optimal logistic regression models exhibit a high correlation between predictor variables and the outcome while maintaining, at most, moderate correlations among predictor variables themselves.

Conclusion

This study investigates the impact of multicollinearity on multilevel non-linear models, utilizing redefined correlated predictors for both Level One and Level Two generated through simulated data. The detection of multicollinearity is demonstrated by a high Variance Inflation Factor (VIF), with a VIF ranging from five to ten indicating potentially problematic levels of association. The study reveals that the VIF value tends to increase with the degree of association.

Furthermore, the research concludes that the multicollinearity within the logistic multilevel regression model decreases as the sample size increases. Consequently, increasing the sample size serves as a conventional approach to mitigate the risks associated with multicollinearity. This consistency is attributed to the use of maximum likelihood estimation (MLE) in logistic regression for parameter determination, in contrast to ordinary least squares (OLS). Since MLE relies on large-sample asymptotic normality, estimates become more reliable with a larger sample size, minimizing issues related to multicollinearity errors.

Recommendation

Researchers must recognize that while multicollinearity cannot be entirely eradicated, efforts can be made to mitigate its impact on the data. It is imperative for researchers to routinely assess the presence of multicollinearity before applying a multilevel nonlinear model. In conducting research, researchers can consider another statistic or means of conducting research if it can fulfil their research objectives. According to Atoyebi et al., [49], by using meta-analysis, the results will be immune to statistical issues such as multicollinearity.

Multicollinearity, instead of biasing results, elevates the standard errors of correlated explanatory variables. It is essential to note that multicollinearity primarily affects explanatory variables, not the dependent variable. Therefore, this study recommends estimating an equivalent linear regression model and specifying collinearity diagnostics if multicollinearity is detected in a logit model. In most cases, the methods employed to address multicollinearity are found to be satisfactory. The maximum likelihood technique's weight matrix should be utilized to adjust linear combinations by eliminating variables, ensuring the construction of robust predictive logistic regression models through thorough inspections and corrective actions to address multicollinearity.

Future simulation studies on this subject should explore the impact of additional levels with varying sample sizes. Researchers should investigate whether the findings remain applicable for larger sample sizes, considering instances of larger groups in the literature. Future investigations should manipulate the degree of correlation between the model's levels to encompass diverse scenarios. Furthermore, research on cross-level interactions should involve cross-level correlations among predictors and other modifications to the cross-level interaction coefficient.

A deeper understanding of the significance of multicollinearity in nonlinear multilevel models requires more real-world data research. This will provide the field with a more comprehensive comprehension of the expected values from the available data. Researchers should adjust their models based on empirical evidence and assess the meaningfulness of the modifications. This approach will enhance the reliability of nonlinear multilevel models when presenting results in practical contexts.

References

1. Pinheiro, J.C., and Bates, D.M. (2000). *Mixed-Effects Models in S and S-Plus*. Springer-Verlag:New York.
2. Mokhtari, S., Navidi, W. and Mooney, M., 2020. White-box regression (elastic net) modeling of earth pressure balance shield machine advance rate. *Automation in Construction*, *115*, p.103208.
3. Koops, W.J. (1988). Multiphasic Analysis of Growth Curves in Chickens. *Poultry Science*, 67, 33-42.
4. Richards, F.J., 1959. A flexible growth function for empirical use. *Journal of experimental Botany*, *10*(2), pp.290-301.
5. Maeda, Y., 2007. *Monte Carlo evidence regarding the effects of violating assumed conditions of two-level hierarchical models for cross-sectional data*. University of Minnesota.
6. Dedrick, R.F., Ferron, J.M., Hess, M.R., Hogarty, K.Y., Kromrey, J.D., Lang, T.R., Niles, J.D. and Lee, R.S., 2009. Multilevel modeling: A review of methodological issues and applications. *Review of educational research*, *79*(1), pp.69-102.
7. Cudeck, R., Jöreskog, K.G., Sörbom, D. and Du Toit, S., 2001. *Structural equation modeling: Present and future: A Festschrift in honor of Karl Jöreskog*. Scientific Software International.
8. Greene, W.H. (2011). *Econometric Analysis*. Prentice Hall; 7 edition (February 13, 2011).
9. Ljung L, Söderström T (1983). *Theory and practice of recursive Identification*. Cambridge, MA: MIT Press.
10. Stoica, P., SÖDERSTRÖM, T., Ahlén, A. and Solbrand, G., 1985. On the convergence of pseudo-linear regression algorithms. *International Journal of Control*, *41*(6), pp.1429-1444.
11. Janzack A (2004). *Identification of nonlinear systems using neural networks and polynomial models*. Springer, Berlin, D.
12. Ljung L (2002). Prediction error estimation method. *Circ Syst Signal Pr* 21(**1**): 11-21.
13. Thacker, W.C., 1989. The role of the Hessian matrix in fitting models to measurements. *Journal of Geophysical Research: Oceans*, *94*(C5), pp.6177-6196.
14. Roozbeh, M., Babaie–Kafaki, S. and Aminifard, Z., 2021. Two penalized mixed–integer nonlinear programming approaches to tackle multicollinearity and outliers effects in linear regression models. *Journal of Industrial & Management Optimization*, *17*(6).
15. Roozbeh, M., Babaie–Kafaki, S. and Aminifard, Z., 2021. Two penalized mixed–integer nonlinear programming approaches to tackle multicollinearity and outliers effects in linear regression models. *Journal of Industrial & Management Optimization*, *17*(6).
16. Belsley (1991), "Multicollinearity" diagnosing its presence and assessing the potential damage it causes working paper 154 national bureau of economic research Cambridge
17. Zhang, J. and Ibrahim, M., 2005. A simulation study on SPSS ridge regression and ordinary least squares regression procedures for multicollinearity data. *Journal of Applied Statistics*, *32*(6), pp.571-588.
18. Montgomery D. C., Peck E. A., Vining G. G.,(2001) Introduction to Linear Regression Analysis, John Wiley & Sons, Inc., New York.
19. Atoyebi, O.M., Atoyebi, S.B. and Ajao, A.F. 2023. The Impact of Mathematics Anxiety on the Mathematical Value of Secondary School Students in Nigeria. *Asian Journal of Advanced Research and Reports* 17(11), pp. 236–254.

20. Belsley (1976), "Multicollinearity" diagnosing its presence and assessing the potential damage it causes working paper 154 national bureau of economic research Cambridge
21. Atoyebi , S. B., Awodutire , P. O., Igbalajobi , M. M., and Anifowose , O. O., 2023. Survival Times of Breast Cancer Patients in Nigeria. *Asian Journal of Immunology*, 6(1), pp. 209–223.
22. Bowerman, B.L. and O'connell, R.T., 1990. Linear statistical models: An applied approach. *(No Title)*.
23. Greene (2000), "Econometric Analysis" (Fourth edition). Upper Saddle River, NJ: Prentice- Hall.
24. Gwelo, A.S., 2019. Principal components to overcome multicollinearity problem. *Oradea Journal of Business and Economics*, *4*(1), pp.79-91.
25. Reddy, T.A. and Henze, G.P., 2023. Parametric and Non-Parametric Regression Methods. In *Applied Data Analysis and Modeling for Energy Engineers and Scientists* (pp. 355-407). Cham: Springer International Publishing.
26. Deng, W.Q., Craiu, R.V. and Sun, L., 2022. Measuring the severity of multi-collinearity in high dimensions. *arXiv preprint arXiv:2203.10360*.
27. Memon, M.A., Cheah, J.H., Ramayah, T., Ting, H., Chuah, F. and Cham, T.H., 2019. Moderation analysis: issues and guidelines. *Journal of Applied Structural Equation Modeling*, *3*(1), pp.1-11.
28. Kalnins, A., 2018. Multicollinearity: How common factors cause Type 1 errors in multivariate regression. *Strategic Management Journal*, *39*(8), pp.2362-2385.
29. Farrar, D.E. and Glauber, R.R., 1967. Multicollinearity in regression analysis: the problem revisited. *The Review of Economic and Statistics*, pp.92-107.
30. Haitovsky, Y., 1969. Multicollinearity in regression analysis: Comment. *The Review of economics and statistics*, pp.486-489.
31. Kmenta (1970) "Elements of econometrics" Macmillan publishing co. New York.
32. Chen, H., 2023. A novel wind model downscaling with statistical regression and forecast: A case study for the Arctic cleaner energy. *Journal of Cleaner Production*, p.140217.
33. Silvey, S. D. (1969). Multicollinearity and Imprecise Estimation. Journal of the Royal Statistical Society, 31, 539–552
34. Montgomery, D. C. and Peck, E. A. (1981). Introduction to linear regression analysis. New York: John Wiley.
35. Belsley, D.A., Kuh, E. and Welsch, R.E.(1980), "Regression Diagnostics: Identifying influential Data and Sources of Collinearity". John Wiley & Sons., New York.
36. Kim, J.H., 2019. Multicollinearity and misleading statistical results. *Korean journal of anesthesiology*, *72*(6), pp.558-569.
37. Jin, S.S. and Jung, H.J., 2018. Vibration-based damage detection using online learning algorithm for output-only structural health monitoring. *Structural Health Monitoring*, *17*(4), pp.727-746.
38. Marquaridt, D.W., 1970. Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation. *Technometrics*, *12*(3), pp.591-612.
39. Freund, R.J., Wilson, W.J. and Sa, P., 1998. Statistical modeling of a response variable. *Regression Analysis; Academic Press: St. Louis, MI, USA*.
40. Rubinstein, R.Y. and Kroese, D.P., 2016. *Simulation and the Monte Carlo method*. John Wiley & Sons.

41. Shieh, Y-Y. and Fouladi, R. T. (2003). The effect of multicollinearity on multilevel modeling parameter estimates and standard errors. *Educational and Psychological Measurement, 63*(6), 951-985.
42. McDonald, G.C. and Galarneau, D.I., 1975. A Monte Carlo evaluation of some ridge-type estimators. *Journal of the American Statistical Association*, *70*(350), pp.407-416.
43. Gibbons, D.G., 1981. A simulation study of some ridge estimators. *Journal of the American Statistical Association*, *76*(373), pp.131-139.
44. Wichern, D.W. and Churchill, G.A., 1978. A comparison of ridge stimators. *Technometrics*, *20*(3), pp.301-311.
45. Kibria, B.G., 2003. Performance of some new ridge regression estimators. *Communications in Statistics-Simulation and Computation*, *32*(2), pp.419-435.
46. Muniz Jr, J., Batista Jr, E.D. and Loureiro, G., 2010. Knowledge-based integrated Production Management Model applied to automotive companies. *International Journal of Knowledge Management Studies*, *4*(3), pp.301-318.
47. Neter, J., Wasserman, W. and Kutner, M.H., 1983. *Applied linear regression models*. Richard D. Irwin.
48. Adeboye, N.O., Fagoyinbo, I.S. and Olatayo, T.O., 2014. Estimation of the effect of multicollinearity on the standard error for regression coefficients. *Journal of Mathematics*, *10*(4), pp.16-20.
49. Atoyebi, O.M. and Atoyebi, S.B., 2022. Do technology–based approaches reduce mathematics anxiety? A Systematic literature. *International Journal of Research and Innovation in Social Science (IJRISS)*, *6*(10), pp.502-509.
50. Marquardt, D.W. and Snee, R.D., 1975. Ridge regression in practice. *The American Statistician*, *29*(1), pp.3-20.