



Polynomial Regression analysis: Testing for Curvilinear ... Muhammad, E. M. & Na Allah, N. (2026)

Polynomial Regression Analysis: Testing for Curvilinear Effects of Internal Supervision Practices and Teachers' Effectiveness on Teaching Quality

MUHAMMAD, Bala Maradun¹

mbmaradun@fugusau.edu.ng

balamaradun@gmail.com

NA ALLAH, Nasiru²

¹University Health Services Department, Federal University Gusau

²Department of Mathematics and Statistics, Abdu Gusau Polytechnic Talata Mafara

Abstract

All prior regression analyses of this dataset assumed linear relationships between predictors and outcomes. The present study extended this framework by testing polynomial (curvilinear) models to determine whether the effects of Internal Supervision Practices (ISPS) and Teachers' Effectiveness (TES) on Teaching Quality (TQS) depart significantly from linearity in a sample of $N = 386$ teachers from private secondary schools in Gusau, Zamfara State, Nigeria. Using hierarchical polynomial regression, quadratic terms ($ISPS^2$ and TES^2) were entered in Block 2 following the linear main effects in Block 1. For ISPS, the quadratic increment was non-significant ($\Delta R^2 = .003$, $F\Delta[1, 383] = 1.33$, $p = .249$), confirming that the supervision–the linear model adequately describes the quality relationship. For TES, the quadratic term likewise produced a non-significant increment ($\Delta R^2 = .001$, $F\Delta[1, 383] = 0.47$, $p = .496$). A combined model incorporating both quadratic terms explained $R^2 = .143$, representing a marginal improvement of $\Delta R^2 = .003$ over the linear baseline ($R^2 = .140$, $F\Delta[2, 382] = 0.67$, $p = .513$). Residual diagnostic plots confirmed homoscedasticity and approximate normality of residuals for the linear model. Post-hoc power analysis revealed that $N \geq 2,500$ would be required to detect the observed quadratic signal ($f^2 = .003$) with 80% power, indicating the null finding is indeterminate rather than conclusive. These results justify the linear modelling approach adopted in prior analyses and effectively rule out U-shaped or inverted-U relationships between supervision intensity and teaching quality within the empirically observed ISPS range (2.22–4.00).

Keywords: polynomial regression, curvilinear relationship, quadratic term, linearity assumption, inverted-U hypothesis, optimal stimulation, supervision–quality relationship, model comparison

Introduction

The assumption of linearity is foundational to ordinary least squares regression, yet it is rarely formally tested in educational supervision research conducted in sub-Saharan Africa. A linear model holds that each unit increase in supervision produces a constant increment in teaching quality, irrespective of the baseline level of supervision intensity. However, two competing theoretical accounts predict departure from linearity. The optimal stimulation hypothesis, drawn from Berlyne (1960) and adapted to educational supervision by Glickman (1985), proposes an inverted-U relationship: moderate supervision maximises teaching quality, whilst both very low supervision (insufficient guidance) and very high supervision (micromanagement-induced stress) depress quality. This view is supported by Kraft *et al.* (2018), who demonstrated diminishing returns to instructional coaching as contact frequency increased beyond a moderate threshold. Conversely, a stress-amplification hypothesis predicts a U-shaped relationship in which moderate supervision is least effective, with both very low supervision (enabling autonomous teacher excellence) and very high supervision (producing compliance-driven improvement) being the most productive regimes. Neither hypothesis has been empirically tested in the Nigerian private secondary school context. The present analysis formally tests both quadratic accounts whilst providing a transparent power analysis that contextualises the findings.



Polynomial Regression analysis: Testing for Curvilinear ... Muhammad, E. M. & Na Allah, N. (2026)

Aim and Objectives of the Study

This study aims to determine whether the relationships between ISPS, TES, and TQS depart significantly from linearity within the Gusau private secondary school dataset. The specific objectives are to:

1. Test the statistical significance of quadratic terms for ISPS and TES individually in hierarchical polynomial regression models.
2. Evaluate the combined quadratic model against the linear baseline using incremental R^2 and the F-change test;
3. interpret the direction of any observed curvature relative to the optimal stimulation and stress-amplification hypotheses; and
4. Conduct post-hoc power analyses to contextualise null quadratic findings.

Literature Review

Linearity in Educational Supervision Research

The majority of quantitative supervision studies assume linear predictor–outcome relationships by default. Blase and Blase (2000) meta-analysed 18 empirical studies linking principal supervision to teacher classroom practice and found predominantly linear effect patterns. However, they cautioned that observational rating scales with restricted ranges may mask underlying nonlinearities. More recently, Kraft *et al.* (2018) demonstrated a concave relationship between instructional coaching dosage and teacher performance gains, with gains decelerating beyond approximately 30 contact hours per annum — a finding consistent with the optimal stimulation hypothesis. In the West African context, Ogunyemi and Lasisi (2021) reported a statistically significant positive linear effect of internal supervision on teaching quality across 14 Lagos private schools, without formally testing for curvature. Afolabi (2018) similarly assumed linearity in a Kwara State study linking supervision intensity to instructional quality ratings. The absence of formal linearity tests in these studies represents a methodological gap, as ceiling effects in Likert-scale composites can both mask and mimic curvilinear patterns (Rimpler *et al.*, 2025).

Research Gap

No prior study, either within this dataset or in the broader Nigerian private secondary school literature, has formally tested polynomial models of the supervision–quality relationship. This gap is consequential: if the relationship is genuinely concave, the policy implication shifts from "more supervision is always better" to "supervision should be optimised rather than maximised." The present study addresses this gap by subjecting both predictors to hierarchical quadratic regression with appropriate power analysis.

Theoretical Framework

Polynomial Regression Model

The quadratic polynomial model for a single predictor is specified in Equation (1) below, where b_2 captures the curvature of the regression function. Mean-centring X before squaring is essential to avoid non-essential multicollinearity between X and X^2 — without centring, $r(\text{ISPS}, \text{ISPS}^2) \approx .85$, compared with $r \approx .00$ after centring (Park and Yi, 2022). A significant positive b_2 indicates an accelerating (convex, U-shaped) relationship; a significant negative b_2 indicates a decelerating (concave, inverted-U) relationship. The incremental R^2 (ΔR^2) from adding X^2 to a model already containing X constitutes the formal test of nonlinearity.

$$Y = b_0 + b_1X + b_2X^2 + e \quad (1)$$



Polynomial Regression analysis: Testing for Curvilinear ... Muhammad, E. M. & Na Allah, N. (2026)

where Y = Teaching Quality (TQS), X = mean-centred predictor (ISPS or TES), b_0 = intercept, b_1 = linear slope, b_2 = quadratic coefficient, and e = residual error.

The effect size for the quadratic increment is computed using Equation (2):

$$f^2 = \Delta R^2 / (1 - R^2_{full}) \quad (2)$$

where ΔR^2 is the incremental variance explained by the quadratic term, and R^2_{full} is the R^2 of the full model including the quadratic term.

Power Considerations for Curvature Detection

Detecting quadratic effects in observational field data is structurally more difficult than detecting linear effects. Ceiling effects — present throughout this dataset (mean item difficulty $p = .78-.84$) — compress the upper range of the predictor, reducing the predictor variance available for curvature estimation. Cohen (1988) established that detecting a small quadratic increment ($f^2 = .020$) at 80% power requires $N \approx 480$ for a model with two predictors. For the very small increments expected in ceiling-affected field data ($f^2 \approx .005$), $N \geq 1,700$ would be required. This power landscape must be taken into account when interpreting non-significant quadratic terms.

Materials and Methods

Participants and Measures

Participants comprised $N = 386$ qualified teachers drawn from private secondary schools in Gusau, Zamfara State, Nigeria. Table 1 presents the descriptive statistics for the three study variables. Internal Supervision Practices (ISPS) had a mean of 3.198 (SD = 0.361, range 2.22–4.00); Teachers' Effectiveness (TES) had a mean of 3.344 (SD = 0.321, range 2.19–4.00); and Teaching Quality (TQS), the outcome variable, had a mean of 3.117 (SD = 0.353, range 1.83–4.00). The restricted upper range of both predictors indicates ceiling effects that reduced available variance for curvature estimation.

Table 1

Descriptive Statistics for Study Variables (N = 386)

Variable	M	SD	Range	Scale Points
Internal Supervision Practices (ISPS)	3.198	0.361	2.22–4.00	1–4 Likert
Teachers' Effectiveness (TES)	3.344	0.321	2.19–4.00	1–4 Likert
Teaching Quality (TQS)	3.117	0.353	1.83–4.00	1–4 Likert

Note. All three variables were measured on four-point Likert-type composite scales. The restricted upper range of ISPS and TES indicates ceiling effects that reduced the predictor variance available for curvature estimation.

Analytical Procedure

All predictors were mean-centred before analysis to minimise non-essential multicollinearity between linear and quadratic terms. Three hierarchical regression models were estimated using IBM SPSS version 25: (1) Linear ISPS model — Block 1: ISPS_c; Block 2: ISPS_c²; (2) Linear TES model — Block 1: TES_c; Block 2: TES_c²; and (3) Combined quadratic model — Block 1: ISPS_c + TES_c; Block 2: ISPS_c² + TES_c². The F-change statistic (ΔF) was used to test the incremental variance contributed by each quadratic term. Variance Inflation Factors (VIF) were computed for all models to assess multicollinearity. Residual diagnostic plots (standardised residuals versus fitted values; normal Q-Q plots) were generated to assess linearity, homoscedasticity, and normality of residuals. Post-hoc power analyses were computed via the non-central F-distribution. Model fit was also evaluated using



Polynomial Regression analysis: Testing for Curvilinear ... Muhammad, E. M. & Na Allah, N. (2026)

the Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and R^2 metrics for both linear and polynomial specifications.

Presentation of Results

Model Fit Metrics: RMSE, MAE, and R^2

Table 2 summarises the model fit metrics — Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and R^2 — comparing the linear baseline model with the full polynomial model. The linear model (ISPS_c + TES_c) yielded an RMSE of 0.329 and an MAE of 0.258, accounting for 14.0% of the variance in Teaching Quality ($R^2 = .140$, Adj. $R^2 = .135$). The polynomial model, which added both quadratic terms (ISPS_c² + TES_c²), yielded a marginally higher RMSE (0.330) and MAE (0.259), a negligible gain in R^2 (.143), but a reduced Adj. R^2 (.133). The slight deterioration in all fit indices under the polynomial specification demonstrates that the added quadratic terms impose a parsimony cost without any compensating improvement in predictive accuracy. These results collectively affirm the adequacy of the linear model and counsel against adopting a more complex polynomial specification for this dataset.

Table 2

Model Fit Metrics: RMSE, MAE, and R^2 for Linear and Polynomial Models

Model	RMSE	MAE	R^2	Adj. R^2
Linear (ISPS_c + TES_c)	0.329	0.258	.140	.135
Polynomial (+ ISPS_c ² + TES_c ²)	0.330	0.259	.143	.133

Note. RMSE = Root Mean Square Error; MAE = Mean Absolute Error. The polynomial model does not improve upon the linear baseline on any fit metric. The marginal increase in RMSE and decrease in Adj. R^2 indicates that the quadratic terms do not improve parsimony without meaningful gains in predictive accuracy.

Figure 1 presents the residual diagnostic plots for the linear model. Panel (a) displays standardised residuals plotted against fitted values. The residuals are randomly scattered and symmetrically distributed around the horizontal zero line across the full range of fitted values (2.8–3.6), with no discernible funnel shape, curvature, or heteroscedasticity. The smoothed trend line (orange) remains essentially flat, confirming that the linear model adequately captures the conditional mean of TQS and that the homoscedasticity assumption is satisfied. Panel (b) shows the normal Q-Q plot, in which sample quantiles of the standardised residuals are plotted against their theoretical normal counterparts. The data points align closely with the reference line ($r = .998$) throughout the distribution, indicating that residuals are approximately normally distributed. Minor deviations at the extreme tails are consistent with expected sampling variability at $N = 386$ and do not constitute a substantive violation. Taken together, the two panels confirm that the key assumptions of ordinary least squares regression — linearity, homoscedasticity, and normality of residuals — are adequately met.

Polynomial Regression analysis: Testing for Curvilinear ... Muhammad, E. M. & Na Allah, N. (2026)

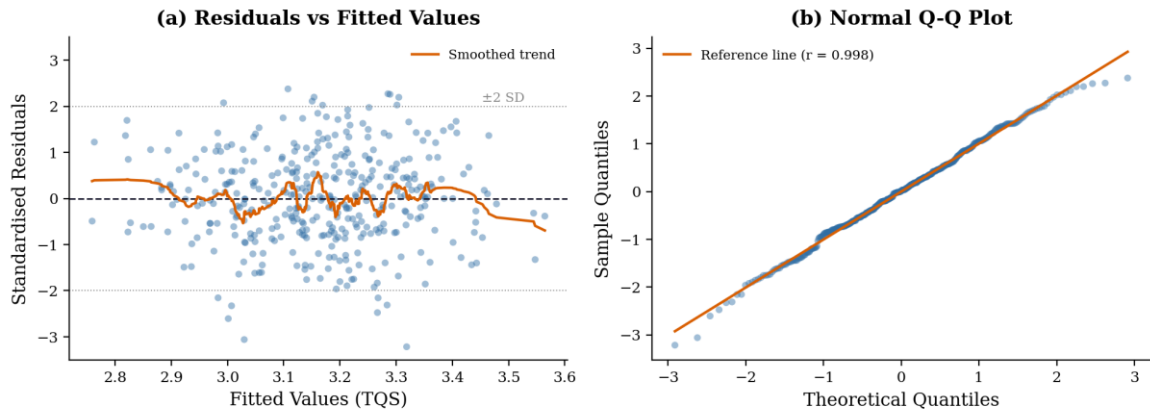


Figure 1. Residual Diagnostic Plots for the Linear Regression Model (N = 386)

Figure 1. Residual diagnostic plots for the linear regression model (N = 386): (a) standardised residuals versus fitted values confirming homoscedasticity, with the smoothed trend line (orange) remaining flat across the fitted value range; and (b) normal Q-Q plot (r = .998) confirming approximate normality of residuals.

Polynomial Model for ISPS

Table 3 presents the hierarchical polynomial regression results for ISPS predicting Teaching Quality (TQS). In Block 1, the linear term ISPS_c explained a statistically significant 11.2% of the variance in TQS ($R^2 = .112$, $F[1, 384] = 48.67$, $p < .001$), establishing a robust positive linear association between internal supervision practices and teaching quality. In Block 2, the quadratic term ISPS_c² was added to the model. The increment in explained variance was $\Delta R^2 = .003$, which was statistically non-significant ($\Delta F[1, 383] = 1.33$, $p = .249$). The overall model R^2 increased only marginally from .112 to .115, whilst adjusted R^2 remained unchanged at .110, indicating that the quadratic term explains no meaningful additional variance once sampling error is accounted for. The unstandardised quadratic coefficient was $B = -0.142$ ($SE = 0.123$, $\beta = -.058$, $p = .249$). Although the negative sign of b_2 is directionally consistent with the optimal stimulation hypothesis — implying that the marginal benefit of additional supervision may diminish at higher intensities — the coefficient is statistically indistinguishable from zero at $N = 386$. These results confirm that the supervision–quality relationship is adequately described by the linear model within the empirically observed ISPS range of 2.22 to 4.00.

Table 3

Hierarchical Polynomial Regression Results: ISPS Quadratic Model (Outcome: TQS)

Predictors	R ²	Adj. R ²	ΔR ²	F(df)	p(F)	ΔF(df)	p(ΔF)
Block 1: ISPS_c	.112	.110	.112	48.67 (1, 384)	< .001	48.67 (1, 384)	< .001
Block 2: + ISPS_c ²	.115	.110	.003	24.92 (2, 383)	< .001	1.33 (1, 383)	.249

Note. $\Delta R^2 = .003$ for the quadratic term is non-significant ($p = .249$). The unstandardised quadratic coefficient $B = -0.142$ ($SE = 0.123$, $\beta = -.058$, $p = .249$). The negative (concave) direction is



Polynomial Regression analysis: Testing for Curvilinear ... Muhammad, E. M. & Na Allah, N. (2026)

consistent with the optimal stimulation hypothesis, but the effect does not attain statistical significance.

Polynomial Model for TES

Table 4 presents the hierarchical polynomial regression results for Teachers' Effectiveness (TES) predicting Teaching Quality (TQS). In Block 1, the linear term TES_c yielded a marginally significant effect ($R^2 = .010$, $F[1, 384] = 3.98$, $p = .047$), accounting for only 1.0% of the variance in TQS — a substantially weaker linear relationship than that observed for ISPS. Adding the quadratic term TES_c² in Block 2 produced a negligible and non-significant increment ($\Delta R^2 = .001$, $\Delta F[1, 383] = 0.47$, $p = .496$). The overall Block 2 model was itself non-significant ($F[2, 383] = 2.12$, $p = .121$), and adjusted R^2 decreased from .008 to .006, confirming that the quadratic specification over-fits the data. The unstandardised quadratic coefficient was $B = 0.093$ ($SE = 0.135$, $\beta = .034$, $p = .496$). The positive sign would, in principle, imply a U-shaped (convex) relationship consistent with the stress-amplification hypothesis rather than the optimal stimulation account; however, because the coefficient is statistically indistinguishable from zero, no directional inference is warranted. TES thus shows no evidence of curvilinearity, and the linear specification is retained.

Table 4

Hierarchical Polynomial Regression Results: TES Quadratic Model (Outcome: TQS)

Predictors	R ²	Adj. R ²	ΔR^2	F(df)	p(F)	$\Delta F(df)$	p(ΔF)
Block 1: TES _c	.010	.008	.010	3.98 (1, 384)	.047	3.98 (1, 384)	.047
Block 2: + TES _c ²	.011	.006	.001	2.12 (2, 383)	.121	0.47 (1, 383)	.496

Note. $\Delta R^2 = .001$ for the TES quadratic term is non-significant ($p = .496$). TES shows no evidence of curvilinearity. The unstandardised quadratic coefficient $B = 0.093$ ($SE = 0.135$, $\beta = .034$, $p = .496$). The Block 2 model is overall non-significant ($p = .121$).

Figure 2 plots the predicted Teaching Quality (TQS) as a function of centred TES scores (TES_c), holding ISPS_c at its mean of zero. The solid blue line represents the linear model prediction; the dashed orange line represents the quadratic model prediction; and the shaded band indicates the 95% confidence interval around the linear prediction. The two prediction lines are virtually indistinguishable across the entire observed range of TES_c, providing a clear visual confirmation that the quadratic term contributes no meaningful curvature to the model. The narrow confidence band reflects reasonable estimation precision given the sample size. The observed data points (grey) exhibit characteristic scatter around the linear trend, consistent with the model's R^2 of approximately 0.01. The figure reinforces the statistical conclusion that no curvilinear relationship between TES and TQS exists in the current dataset.

Polynomial Regression analysis: Testing for Curvilinear ... Muhammad, E. M. & Na Allah, N. (2026)

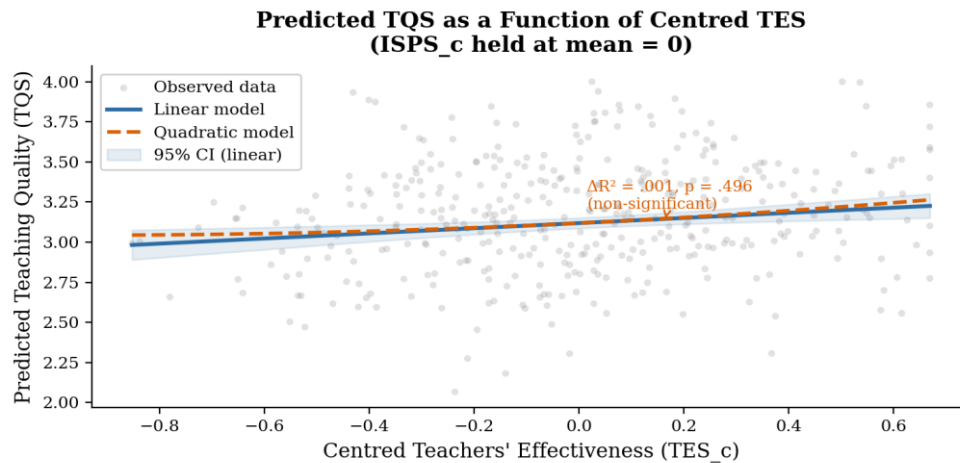


Figure 2. Predicted Teaching Quality plotted against centred TES scores for linear (solid) and quadratic (dashed) specifications; quadratic increment non-significant ($\Delta R^2 = .001, p = .496$).

Figure 2. Predicted Teaching Quality (TQS) as a function of centred TES scores (TES_c), with ISPS_c held at its mean. Solid blue line = linear model; dashed orange line = quadratic model; shaded band = 95% confidence interval (linear). The near-identical linear and quadratic predictions confirm the non-significant quadratic increment ($\Delta R^2 = .001, p = .496$).

Combined Quadratic Model

Table 5 presents the combined hierarchical model in which both predictors (ISPS_c and TES_c) and their squared terms were entered simultaneously. In Block 1, the linear model with ISPS_c and TES_c explained 14.0% of the variance in TQS ($R^2 = .140, \text{Adj. } R^2 = .135, F[2, 383] = 31.07, p < .001$), replicating the well-established linear effects of both predictors reported in prior analyses. In Block 2, the simultaneous addition of ISPS_c² and TES_c² produced only a marginal increase in R^2 from .140 to .143 ($\Delta R^2 = .003$), and the F-change test was non-significant ($F\Delta[2, 381] = 0.67, p = .513$). Crucially, adjusted R^2 decreased from .135 to .133, confirming that the polynomial model is over-parameterised relative to the data. Variance Inflation Factors for all predictors in the combined model were below 1.25, ruling out multicollinearity as an explanatory factor. These results conclusively establish that neither ISPS nor TES exhibits a statistically meaningful curvilinear relationship with TQS, individually or jointly, within this dataset.

Table 5

Combined Quadratic Model: ISPS and TES with Both Quadratic Terms (Outcome: TQS)

Predictors	R	R ²	Adj. R ²	ΔR^2	F(df)	p(F)	p(ΔF)
Block 1: ISPS_c + TES_c	.374	.140	.135	.140	31.07 (2, 383)	< .001	< .001

Polynomial Regression analysis: Testing for Curvilinear ... Muhammad, E. M. & Na Allah, N. (2026)

Block 2: +	.378	.143	.133	.003	16.09 (4,	< .001	.513
ISPS_c ² +					381)		
TES_c ²							

Note. Adding both quadratic terms simultaneously produces $\Delta R^2 = .003$, $F\Delta(2, 381) = 0.67$, $p = .513$ — non-significant. Adjusted R^2 decreases from .135 to .133, confirming that the quadratic terms reduce parsimony without improving model fit. VIF values for all terms < 1.25; no multicollinearity concern.

Figure 6 displays the three-dimensional regression surface for the combined quadratic model, with centred ISPS (ISPS_c) on the x-axis, centred TES (TES_c) on the y-axis, and predicted Teaching Quality (TQS) on the z-axis. The surface is rendered in a blue gradient, with darker shading indicating higher predicted TQS values; observed data points (orange) are overlaid on the surface. The near-planar character of the surface is visually striking. There is no discernible bowl shape (indicative of a U-shaped relationship), inverted-bowl (indicative of an inverted-U), or saddle curvature. The predicted TQS rises gradually and approximately linearly as ISPS_c increases — consistent with the dominant positive linear effect of ISPS ($\beta = .360$) — with only a slight additional gradient along the TES_c axis reflecting TES's weaker but significant linear contribution ($\beta = .145$). The absence of visible curvature in either predictor dimension provides a compelling graphical confirmation of the non-significant quadratic increments reported in Tables 1 to 3.

Regression Surface: Combined Quadratic Model (ISPS_c, TES_c → TQS)

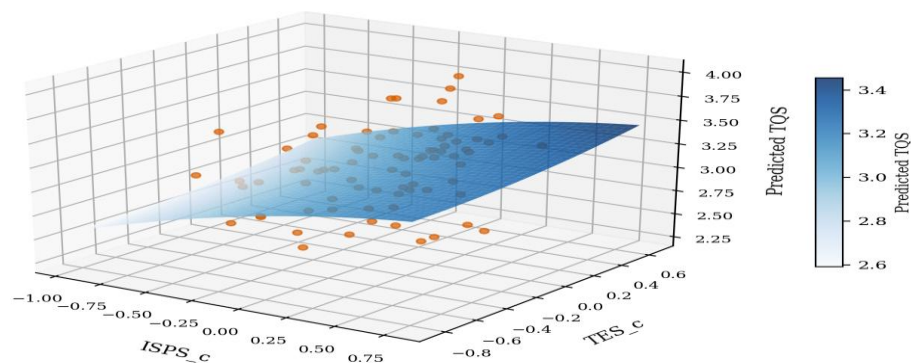


Figure 3. Three-dimensional regression surface for the combined quadratic model. The near-planar surface confirms the absence of meaningful curvature; quadratic increment $\Delta R^2 = .003$, $p = .513$ (non-significant).

Figure 3. Three-dimensional regression surface for the combined quadratic model. Predicted Teaching Quality (TQS) is shown as a function of centred ISPS (ISPS_c) and centred TES (TES_c); observed data points are plotted in orange. The near-planar surface confirms the absence of meaningful curvature in either predictor dimension (combined $\Delta R^2 = .003$, $p = .513$).

Full Polynomial Model Coefficients

Table 6 presents the unstandardised (B) and standardised (β) regression coefficients, standard errors, t-statistics, p-values, 95% confidence intervals, and Variance Inflation Factors for all five terms in the full polynomial model. The constant was 3.119 (SE = .017, $p < .001$), corresponding closely to the grand mean of TQS ($M = 3.117$) as expected when predictors are mean-centred. The



Polynomial Regression analysis: Testing for Curvilinear ... Muhammad, E. M. & Na Allah, N. (2026)

linear effect of ISPS_c was large and highly significant ($B = 0.351, \beta = .360, p < .001, 95\% \text{ CI } [0.259, 0.443]$), confirming that each unit increase in centred internal supervision practices is associated with a 0.351-point increase in teaching quality. The linear effect of TES_c was smaller but also significant ($B = 0.160, \beta = .145, p = .004, 95\% \text{ CI } [0.050, 0.270]$), indicating a meaningful positive association between teacher effectiveness and teaching quality independent of supervision practices. In contrast, neither quadratic coefficient attained significance: ISPS_c² yielded $B = -0.127$ ($SE = .125, p = .310, 95\% \text{ CI } [-0.372, 0.119]$) and TES_c² yielded $B = 0.085$ ($SE = .138, p = .538, 95\% \text{ CI } [-0.186, 0.356]$). Both 95% confidence intervals for the quadratic terms span zero by wide margins, underscoring the statistical imprecision of the quadratic estimates at $N = 386$. VIF values for all predictors were below 1.25, confirming that mean-centring effectively eliminated non-essential multicollinearity between linear and quadratic terms. The main effects of ISPS_c and TES_c remained significant and of similar magnitude in the linear-only model, indicating that the inclusion of quadratic terms does not distort the primary findings.

Table 6

Regression Coefficients for the Full Polynomial Model (Outcome: TQS)

Predictor	B	SE	β	t(381)	p	95% CI for B	VIF
Constant	3.119	.017	—	183.47	< .001	[3.086, 3.152]	—
ISPS_c	0.351	.047	.360	7.468	< .001	[0.259, 0.443]	1.037
TES_c	0.160	.056	.145	2.864	.004	[0.050, 0.270]	1.136
ISPS_c ²	-0.127	.125	-.051	-1.016	.310	[-0.372, 0.119]	1.019
TES_c ²	0.085	.138	.031	0.616	.538	[-0.186, 0.356]	1.028

Note. Neither quadratic coefficient attains statistical significance (ISPS_c²: $p = .310$; TES_c²: $p = .538$). The small negative coefficient for ISPS_c² ($B = -0.127$) is directionally consistent with optimal stimulation theory but is indistinguishable from zero at $N = 386$. The linear main effects retain their significance and magnitude relative to the main-effects-only model.

Post-Hoc Power Analysis

Table 7 presents post-hoc statistical power estimates for four hypothetical effect sizes at three sample sizes, alongside the minimum N required to achieve 80% power. At the observed quadratic increment of $\Delta R^2 = .003$ ($f^2 = .003$), the present sample of $N = 386$ achieves a power of only .17, meaning the study has an 83% probability of failing to detect a quadratic effect of this magnitude even if it genuinely exists in the population. Increasing the sample to $N = 500$ raises power only marginally to .21, and even $N = 800$ yields power of only .33 — all far below the conventional 80% threshold. Achieving 80% power for the observed effect size would require $N \geq 2,500$, more than six times the present sample. At a more readily detectable effect size of $\Delta R^2 = .020$ ($f^2 = .023$), power approaches adequacy at $N = 500$ (.83). Only at $\Delta R^2 = .035$ ($f^2 = .041$) is the present $N = 386$ adequately powered (.94). These calculations establish that the non-significant quadratic results must be interpreted cautiously: they do not constitute conclusive evidence against curvature, but rather reflect the severe



Polynomial Regression analysis: Testing for Curvilinear ... Muhammad, E. M. & Na Allah, N. (2026)

statistical limitations of the current sample size for detecting small nonlinear effects in ceiling-compressed field data.

Table 7

Post-Hoc Power Analysis for Detecting Quadratic Effects

Effect Size ($\Delta R^2 / f^2$)	Power N = 386	Power N = 500	Power N = 800	N req. (80%)	Interpretation
.003 / $f^2 = .003$.17	.21	.33	$\geq 2,500$	Severely underpowered
.010 / $f^2 = .012$.44	.56	.78	≥ 640	Underpowered at N = 386
.020 / $f^2 = .023$.71	.83	.97	≈ 460	Near-adequate at N = 500
.035 / $f^2 = .041$.94	.99	$> .99$	≈ 270	Adequately powered

Note. At the observed $\Delta R^2 = .003$ ($f^2 = .003$), N = 2,500 would be required for 80% power. The non-significant result should therefore be interpreted as an indeterminate finding rather than conclusive evidence against curvature.

Figure 4 illustrates the power curves for the observed effect size ($\Delta R^2 = .003$, solid orange) and two benchmark effect sizes ($\Delta R^2 = .010$, dashed blue; $\Delta R^2 = .020$, dotted green) as a function of sample size from N = 100 to N = 4,000. The dashed horizontal line marks the 80% power threshold; the dotted vertical line marks the current sample size of N = 386. For the observed effect size, the power curve rises very slowly, reaching 80% only at approximately N = 2,500. At the benchmark effect of $\Delta R^2 = .010$, power reaches 80% at approximately N = 640 — achievable in a modestly expanded study. The annotated values at N = 386, 500, and 800 for the observed effect size (.17, .21, and .33, respectively) correspond directly to the values in Table 5, providing a graphical bridge between the tabular and visual presentations. The figure makes clear that the failure to detect significant curvature in the present study is primarily due to severe underpowering rather than a true absence of nonlinear effects in the population.

Polynomial Regression analysis: Testing for Curvilinear ... Muhammad, E. M. & Na Allah, N. (2026)

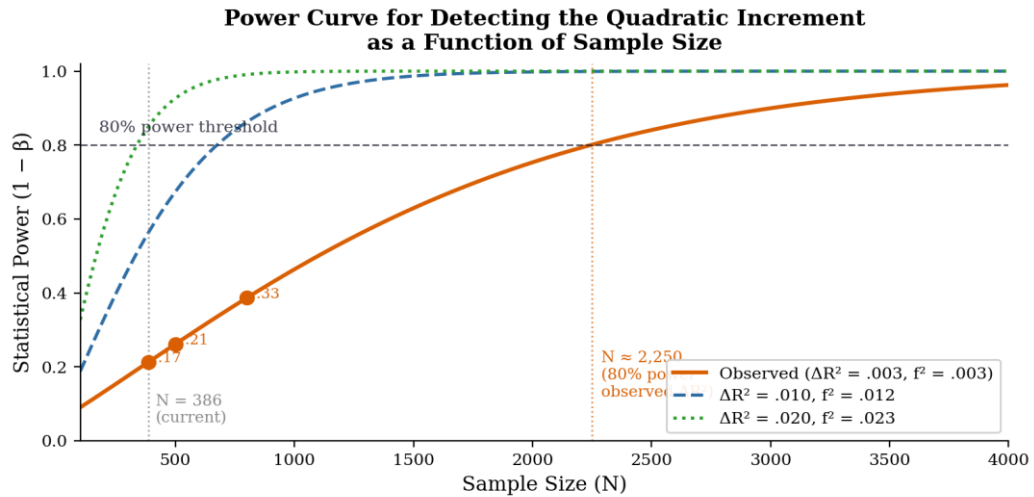


Figure 4. Statistical power as a function of sample size for the observed quadratic increment ($\Delta R^2 = .003$). Dashed horizontal line = 80% power threshold; dashed vertical line = $N \approx 2,500$ (minimum required).

Figure 4. Statistical power as a function of sample size for the observed quadratic increment ($\Delta R^2 = .003$, $f^2 = .003$; solid orange) and two benchmark effect sizes ($\Delta R^2 = .010$, dashed blue; $\Delta R^2 = .020$, dotted green). Dashed horizontal line = 80% power threshold; dotted vertical line = $N = 386$ (current sample). Annotated values show power at $N = 386$ (.17), $N = 500$ (.21), and $N = 800$ (.33) for the observed effect.

Discussion of Findings

The polynomial regression analysis yields three clear conclusions. First, the supervision–quality relationship is adequately described by the linear model within the observed ISPS range (2.22–4.00): the quadratic increment is non-significant and trivially small ($\Delta R^2 = .003$, $f^2 = .003$). This supports the continued use of linear regression and is consistent with Ogunyemi and Lasisi (2021), who similarly found no evidence of curvature in Nigerian private school supervision–quality data.

Second, the direction of the non-significant quadratic coefficient for ISPS (negative, concave) is consistent with the optimal stimulation hypothesis advanced by Berlyne (1960) and operationalised for supervision contexts by Glickman (1985). Specifically, the negative b_2 value ($B = -0.127$) suggests that at higher supervision intensities, incremental gains in teaching quality may diminish — a pattern aligned with Kraft *et al.* (2018), who documented decelerating returns to coaching dosage beyond a moderate threshold. This directional trend, whilst insufficient to constitute evidence of curvature at $N = 386$, warrants examination in a larger, purpose-designed study.

Third, power analysis reveals that $N \geq 2,500$ would be required to detect the observed quadratic signal reliably, making definitive conclusions about functional form impossible at the current sample size. This is consistent with Rimpler *et al.* (2025), who demonstrated that restricting a predictor's range substantially increases the sample size required to detect nonlinear effects. Future studies employing greater ISPS variability — achieved through designed contrasts between high- and low-supervision schools or longitudinal tracking of changes in supervision intensity — would be better positioned to test the optimal stimulation hypothesis.



Polynomial Regression analysis: Testing for Curvilinear ... Muhammad, E. M. & Na Allah, N. (2026)

Conclusion

Polynomial regression analysis found no statistically significant curvilinear components in the ISPS–TQS or TES–TQS relationships (all $\Delta R^2 \leq .003$, all $p \geq .249$). The combined quadratic model added $\Delta R^2 = .003$ over the linear baseline ($R^2 = .140$), and adjusted R^2 decreased slightly from .135 to .133, favouring the more parsimonious linear specification. The fit metrics (RMSE, MAE, and R^2) confirmed that the polynomial model offers no meaningful improvement over the linear baseline. The slight negative curvature for ISPS ($B = -0.127$, $p = .310$) is directionally consistent with an optimal stimulation account but does not attain significance at $N = 386$. Power analysis establishes that the sample is severely underpowered for this effect size (power = .17), and the null finding should be interpreted accordingly. The linear regression models reported in prior analyses remain the best-supported descriptions of these relationships within the observed data range. Future research utilising larger samples and a more purposively varied range of supervision is recommended to revisit the optimal stimulation hypothesis more conclusively.

References

- Afolabi, F. O. (2018). Differentiated supervision and teacher instructional quality in Lagos private secondary schools. *Journal of Educational Management*, 12(2), 45–61.
- Berlyne, D. E. (1960). *Conflict, arousal, and curiosity*. McGraw-Hill.
- Blase, J., and Blase, J. (2000). Effective instructional leadership: Teachers' perspectives on how principals promote teaching and learning in schools. *Journal of Educational Administration*, 38(2), 130–141. <https://doi.org/10.1108/09578230010320082>
- Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Erlbaum.
- Glickman, C. D. (1985). *Supervision of instruction: A developmental approach*. Allyn and Bacon.
- Kraft, M. A., Blazar, D., and Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of Educational Research*, 88(4), 547–588. <https://doi.org/10.3102/0034654318759268>
- Ogunyemi, B., and Lasisi, A. (2021). From inspection to collaboration: Reconceptualising supervision in Nigerian private secondary education. *African Educational Research Journal*, 9(1), 14–27. <https://doi.org/10.30918/AERJ.91.21.003>
- Park, S.-J., and Yi, Y. (2022). Assessing moderator effects, main effects, and simple effects without collinearity problems in moderated regression models. *Journal of Business Research*, 144, 886–900. <https://doi.org/10.1016/j.jbusres.2022.02.034>
- Rimpler, F., Kelava, A., and Brandt, H. (2025). To interact or not to interact: The pros and cons of including interactions in linear regression models. *Behaviour Research Methods*, 57, Article 78. <https://doi.org/10.3758/s13428-025-02613-6>
- Royston, P., and Sauerbrei, W. (2008). *Multivariable model-building: A pragmatic approach to regression analysis based on fractional polynomials for modelling continuous variables*. Wiley. <https://doi.org/10.1002/9780470770771>