

The impact of ignoring random features of predictor and moderator variables on sample size for precise interval estimation of interaction effects

Gwown Shieh

Published online: 7 May 2011
© Psychonomic Society, Inc. 2011

Abstract The influence of the joint distribution of predictor and moderator variables on the identification of interactions has been well described, but the impact on sample size determinations has received rather limited attention within the framework of moderated multiple regression (MMR). This article investigates the deficiency in sample size determinations for precise interval estimation of interaction effects that can result from ignoring the stochastic nature of continuous predictor and moderator variables in MMR. The primary finding of our examinations is that failure to accommodate the distributional properties of regressors can lead to underestimation of the necessary sample size and distortion of the desired interval precision. In order to take account of the randomness of regressor variables, two general and effective procedures for computing sample size estimates are presented. Moreover, corresponding programs are provided to facilitate use of the suggested approaches. This exposition helps to correct drawbacks in the existing techniques and to advance the practice of reporting confidence intervals in MMR analyses.

Keywords Moderation · Precision · Sample size

Moderated multiple regression (MMR) has been extensively employed to study the interaction effects between predictor

and moderator variables in management, psychology, education, and related disciplines. It follows from the comprehensive reviews of Stone-Romero, Alliger, and Aguinis (1994), Aguinis (1995), Aguinis and Stone-Romero (1997), Aguinis, Beaty, Boik, and Pierce (2005), and related work that most of the methodological research in MMR has been concerned with the statistical power of hypothesis testing for detecting moderating effects. Although null hypothesis significance testing is useful in various applications, the dichotomous accept–reject decision ignores other useful information in its analysis. As an alternative, the notion of interval estimation has been stressed in studies such as Hahn and Meeker (1991), Steiger and Fouladi (1997), and Smithson (2003). Accordingly, the inferential procedures of interval estimators are strongly recommended by Wilkinson and the American Psychological Association Task Force on Statistical Inference (1999), as well as the *Publication Manual of the American Psychological Association* (American Psychological Association, 2001). Since confidence intervals constructed with the desired reliability are more informative about the location of a targeted parameter, they should be the best reporting strategy in practical study. However, the methodological artifacts and statistical implications associated with interval estimation of moderating effects have received little attention within the framework of MMR.

The interactional formulation of MMR can be viewed as a special case of the statistical linear models, and so the inferential procedures of hypothesis testing and interval estimation of moderation can be conducted with standard methods and software packages for linear regression analysis. In this article, we consider the simple interaction models with criterion variable Y , predictor variable X , moderator variable Z , their cross-product term XZ , and a normal error term ϵ in the formulation of $Y = \beta_I + X\beta_X + Z\beta_Z + XZ\beta_{XZ} + \epsilon$, where both the predictor X and moderator Z are continuous

Electronic supplementary material The online version of this article (doi:10.3758/s13428-011-0103-y) contains supplementary material, which is available to authorized users.

G. Shieh (✉)
Department of Management Science,
National Chiao Tung University,
1001 Ta Hsueh Road,
Hsinchu, Taiwan 30050, Republic of China
e-mail: gwshieh@mail.nctu.edu.tw

variables. Naturally, the special consideration of continuous moderator and predictor variables incurs the important notion of two different regression formulations. Because of the inherent nature of continuous measurements of the two regressor variables, they are not typically fixed in advance and are available only after the data has been collected. In order to recognize this stochastic feature of regressor variables, the appropriate strategy is to consider a random regression or unconditional formation, rather than a fixed or conditional setting, in experimental designs where the factors are under the control of investigators. The intrinsic appropriateness and theoretical properties of fixed and random regression models have been discussed in Cramer and Appelbaum (1978) and Sampson (1974). Essentially, the inferential procedures of hypothesis testing and interval estimation are the same under both fixed and random formulations. The distinction between the two modeling approaches, however, becomes crucial when power, coverage probability, and corresponding sample size calculations are to be made.

In the context of MMR, the distinct formulations of fixed and random modeling of the simple interaction models were especially emphasized in Shieh (2009, 2010). Specifically, Shieh (2009) considered the power calculation and sample size determination for significance tests of moderating effects. The procedure takes account of the critical factors of strength of moderator effect, magnitude of error variation, and distributional property of predictor and moderator variables. On the other hand, Shieh (2010) incorporated the random nature of continuous moderator and predictor variables into two approaches to sample size computation for precise interval estimation of interaction effects. One approach provides the necessary sample size so that the designated interval for the least squares estimator of moderating effects attains the specified coverage probability. The other approach gives the sample size required to ensure, with a given tolerance probability, that a confidence interval of moderating effects will be within a specified range. The vital discrepancies between the conditional and unconditional settings in power and precision analyses are also closely evaluated in Shieh (2009, 2010). The results reveal substantial detrimental effects of failing to account for the randomness of predictor and moderator variables. Thus, the conventional fixed modeling formulation may not be applicable to the MMR with continuous regressor variables, because it tends to give insufficient sample sizes and inevitably leads to poor statistical performance.

Notably, there is a considerable recent literature pertaining to the illuminating applications of precise interval estimation in multiple linear regression (see Kelley 2008; Krishnamoorthy & Xia 2008; Kelley & Maxwell 2008; and the references therein). Instead of a direct accept-or-reject conclusion in a simple hypothesis test, it is arguable that

confidence intervals generally provide more information about the interested parameter value with a quantitative bound and assurance level. Accordingly, researchers should become methodologically conscious that the mere statistical significance of a targeted parameter is inadequate to warrant the conclusion that the effect is substantial and practically important. In view of the relatively scarce description of interval estimation in MMR, it is prudent to contribute to the documentation and examination of confidence intervals in different perspectives. The precision considerations in Shieh (2010) may not be the only criteria of practical importance. Just as in the instance of Kelley and Maxwell, two other useful principles related to the statistical properties of a confidence interval deal with the control of expected width and tolerance probability of interval width within a designated value. However, the explication of Kelley and Maxwell is confined to the case of a fixed regressor modeling framework, so that the corresponding sample size procedures may not be appropriate for the interval estimation of interaction effects between continuous predictor and moderator variables. Moreover, it is also noted in Kelley and Maxwell that the sampling distribution of an estimated regression coefficient depends on whether the regressors are of fixed or random nature. Thus, the difference should be properly recognized in sample size planning for the simple interaction models described above. The aim of this article is to contribute to the design of MMR study by illustrating how the specification of inherent features associated with continuous predictor and moderator variables in sample size calculations influences the resulting precision of confidence intervals for the prescribed two interval width appraisals in the inference of an interactive effect. Consequently, the presentation here complements the work of Shieh (2010) with distinct precision criteria in sample size determinations.

Due to the complexity within the random regression framework, there appears to be a lack of applicable sample size procedures in the literature that accommodate the considerations of expected confidence interval width and tolerance probability of interval width within a designated value. Therefore, it is essential to extend the development and exposition of sample size methodology for precise interval estimation of interaction effects in two distinct aspects. One method gives the minimum sample size, such that the expected confidence interval width is within the designated bound. The other provides the sample size needed to guarantee, with a given tolerance probability, that the width of a confidence interval will not exceed the planned range. It is important to realize that the simplicity of a fixed setup may be appealing for inducing computational shortcuts, but it does not involve all of the key factors in sample size calculation and, thus, is generally error prone. Accordingly, theoretical implications and numerical

examinations are presented to demonstrate that the sample size procedures for fixed regression, although they share many similarities with those for random regression, have some distinct disadvantages in MMR applications.

Confidence intervals of interaction effects

Consider the simple interaction model or MMR model within the fixed modeling framework:

$$Y_i = \beta_I + X_i\beta_X + Z_i\beta_Z + X_iZ_i\beta_{XZ} + \varepsilon_i, \quad (1)$$

where Y_i is the value of the response variable Y ; X_i and Z_i are the values of the continuous predictor X and moderator Z ; ε_i are $iidN(0, \sigma^2)$ random errors for $i = 1, \dots, N$; and $\beta_I, \beta_X, \beta_Z$, and β_{XZ} are unknown parameters. To examine the existence and magnitude of a moderating effect, we are concerned with the distributional property for the least squares estimator $\hat{\beta}_{XZ} \sim N(\beta, V(\hat{\beta}_{XZ}))$ of β_{XZ} , where $V(\hat{\beta}_{XZ})$ is the variance of $\hat{\beta}_{XZ}$ and one useful expression is $V(\hat{\beta}_{XZ}) = \sigma^2/SSE_{XZ}$ where $SSE_{XZ} = (1 - R_{XZ}^2)S_Y^2$ is the residual sum of squares for the regression of the product term XZ on the X and Z variables, R_{XZ}^2 is the corresponding coefficient of determination, $S_Y^2 = \sum_{i=1}^N (U_i - \bar{U})^2$, $U_i = X_iZ_i$, $i = 1, \dots, N$, and $\bar{U} = \sum_{i=1}^N U_i/N$. Moreover, the natural estimator $\hat{V}(\hat{\beta}_{XZ})$ of $V(\hat{\beta}_{XZ})$ is $\hat{V}(\hat{\beta}_{XZ}) = \hat{\sigma}^2/SSE_{XZ}$, where $\hat{\sigma}^2 = (1 - R^2)S_Y^2/(N - 4)$ is the usual unbiased estimator of σ^2 , R^2 is the coefficient of determination of the model in Eq. 1, $S_Y^2 = \sum_{i=1}^N (Y_i - \bar{Y})^2$, and $\bar{Y} = \sum_{i=1}^N Y_i/N$. For inferential purposes, we focus on the interval estimation procedure of β_{XZ} here. According to the standard results (Rencher, 2000, Section 8.6), a $100(1 - \alpha)\%$ confidence interval of β_{XZ} is

$$\left(\hat{\beta}_{XZ} - t_{N-4, \alpha/2} \left\{ \hat{\sigma}^2/SSE_{XZ} \right\}^{1/2}, \hat{\beta}_{XZ} + t_{N-4, \alpha/2} \left\{ \hat{\sigma}^2/SSE_{XZ} \right\}^{1/2} \right),$$

where $t_{N-4, \alpha/2}$ is the $100(1 - \alpha/2)$ th percentile of the t distribution with $N - 4$ degrees of freedom. The half-width of the $100(1 - \alpha)\%$ confidence interval is denoted by

$$H = t_{N-4, \alpha/2} \left\{ \hat{\sigma}^2/SSE_{XZ} \right\}^{1/2}. \quad (2)$$

Thus, the actual half-width H depends on the sample size N , confidence coefficient $1 - \alpha$, variance estimate $\hat{\sigma}^2$, and observed values of predictor and moderator variables through the quantity of SSE_{XZ} .

Due to the nature of continuous measurements encountered in practical research, the regressor variables typically cannot be controlled and are available only after observation. Hence,

in order to extend this concept and applicability to MMR, the continuous predictor and moderator variables $\{(X_i, Z_i), i = 1, \dots, N\}$ in Eq. 1 are assumed to have a joint probability function $g(X_i, Z_i)$ with finite moments. Moreover, the form of $g(X_i, Z_i)$ does not depend on any of the unknown parameters ($\beta_I, \beta_X, \beta_Z, \beta_{XZ}$) or σ^2 . Consequently, both $\hat{\sigma}^2$ and SSE_{XZ} are realized values of random variables. It readily follows from Eq. 2 that the statistical properties of $\hat{\sigma}^2$ and SSE_{XZ} jointly determine the underlying distributional feature for the half-width H of a confidence interval.

When planning a study, researchers wish to ensure that the confidence interval is narrow enough to produce meaningful findings despite the stochastic nature of an interval width. In order to obtain an informative interval, it is necessary to specify not only the required confidence level and desired precision, but also the appropriate sample size. The most common approach is to determine the required sample size such that the expected half-width of a $100(1 - \alpha)\%$ confidence interval is within the designated bound

$$E[H] = \delta, \quad (3)$$

where the expectation $E[H]$ is taken with respect to the joint distribution of $\hat{\sigma}^2$ and SSE_{XZ} , and $\delta (> 0)$ is a constant. Alternatively, one may compute the sample size needed to guarantee with a given tolerance probability that the width of a $100(1 - \alpha)\%$ confidence interval will not exceed the planned range

$$P\{H < \omega\} = 1 - \gamma, \quad (4)$$

where $1 - \gamma$ is the specified tolerance level and $\omega (> 0)$ is a constant. Note that these two concerns of expected magnitude and tolerance probability of half-width have been discussed in Kupper and Hafner (1989) for simpler situations of one- and two-sample problems and in Kelley and Maxwell (2008) for multiple linear regression. Although the notion of expected width is widely covered in standard texts for sample size determination, the assurance of tolerance probability approach is recommended by Kupper and Hafner. However, it is noteworthy that the two principles are closely related to the two distinct criteria of unbiasedness and consistency in statistical point estimation. Therefore, each measure is of theoretical importance and practical interest in its own right. Furthermore, although the results are not completely comparable, it typically requires a larger sample size to meet the necessary assurance of tolerance probability than the control of a designated expected width.

Sample size procedures for precise interval estimation

It was mentioned above that Kelley and Maxwell (2008, p. 173) focused on the situation in which regressor

variables are fixed. The particular conditional design assumes that the same set of regressor values would be examined for repeated studies. Since $\hat{\sigma}^2$ is an unbiased estimator of σ^2 , the use of population notation in the expression of $V(\hat{\beta}_{XZ})$ in Kelley and Maxwell (2008, Equation 34) implicitly assumes that $SSE_{XZ} \doteq (N-4)(1-\rho_{XZ}^2)\sigma_{XZ}^2$ or $E[SSE_{XZ}] \doteq (N-4)(1-\rho_{XZ}^2)\sigma_{XZ}^2$, where ρ_{XZ}^2 is the coefficient of determination for the regression of product term XZ on the X and Z variables and σ_{XZ}^2 is the variance of the product term XZ . Although the form of $V(\hat{\beta}_{XZ})$ involves some unstated clarifications, the formulation suggests a simplified approximation for $E[\hat{V}(\hat{\beta}_{XZ})] \doteq \sigma^2 / \{(N-4)(1-\rho_{XZ}^2)\sigma_{XZ}^2\}$ under a random regression setup. Hence, the corresponding approximate expected half-width is

$$E[H] \doteq t_{N-4, \alpha/2} [\sigma^2 / \{(N-4)(1-\rho_{XZ}^2)\sigma_{XZ}^2\}]^{1/2}. \quad (5)$$

Accordingly, the sample size N_{EHS} needed for the expected half-width of a $100(1-\alpha)\%$ confidence interval to fall within the designated bound δ is the minimum integer N , such that

$$t_{N-4, \alpha/2} [\sigma^2 / \{(N-4)(1-\rho_{XZ}^2)\sigma_{XZ}^2\}]^{1/2} \leq \delta. \quad (6)$$

A similar expression is presented in Kelley and Maxwell (Equation 35). On the other hand, an analogous argument applies for the tolerance probability consideration given in Eq. 4. Thus, $P\{H < \omega\} \doteq P\{t_{N-4, \alpha/2} [\hat{\sigma}^2 / \{(N-4)(1-\rho_{XZ}^2)\sigma_{XZ}^2\}]^{1/2} < \omega\}$, or equivalently,

$$P\{H < \omega\} \doteq P\{K < (\omega/t_{N-4, \alpha/2})^2 (N-4)^2 (1-\rho_{XZ}^2)\sigma_{XZ}^2 / \sigma^2\}, \quad (7)$$

where $K = (N-4)\hat{\sigma}^2 / \sigma^2 \sim \chi^2(N-4)$ and $\chi^2(\nu)$ is a chi-square distribution with ν degrees of freedom. The sample size N_{PHS} required to guarantee with a given tolerance probability $(1-\gamma)$ that the width of a $100(1-\alpha)\%$ confidence interval will not exceed the planned range ω is the smallest integer N such that

$$\chi^2(N-4, \gamma) \leq (\omega/t_{N-4, \alpha/2})^2 (N-4)^2 (1-\rho_{XZ}^2)\sigma_{XZ}^2 / \sigma^2 \quad (8)$$

where $\chi^2(N-4, \gamma)$ is the $100(1-\gamma)$ th percentile of a $\chi^2(N-4)$ distribution. The formula in Kelley and Maxwell (Equation 35) basically provides the same inequality, although the degrees of freedom of the chi-square distribution there was incorrectly expressed as $N-1$ instead of $N-4$. To compute the necessary sample sizes, a standard iterative search can be conducted to find N_{EHS} and N_{PHS} with the inequalities given in Eqs. 6 and 8, respectively. As noted above, these two simplified procedures rely primarily on the straightforward approximation of $E[SSE_{XZ}] \doteq (N-4)(1-\rho_{XZ}^2)\sigma_{XZ}^2$, and they can alternatively be viewed as the exact methods under

the fixed regressor setting as in Kelley and Maxwell. However, their usage in the random regression scenario raises natural concerns about the resulting accuracy, for the obvious reason that the underlying variability of the residual sum of squares SSE_{XZ} is ignored. The potential deficiency of the sample sizes N_{EHS} and N_{PHS} for precise interval estimation of interaction effects will be examined later in the numerical investigations.

The prescribed discussion emphasizes that the underlying statistical properties of a confidence interval half-width are uniquely defined by the joint distribution of the regressors. Accordingly, the underlying property of predictor and moderator variables should be incorporated into the sample size procedures as much as possible. In order to provide a unified and feasible technique, we adopt the large-sample approach to avoid reliance on any specific distributional assumption of the predictor and moderator variables. To this end, the theoretical properties of the suggested asymptotic approximation to the distribution of $W^* = SSE_{XZ} / (N-1)$ is presented in Eq. 11 in the Appendix. The results facilitate the proposed sample size procedures for constructing precise confidence intervals under the two precision criteria described in Eqs. 3 and 4.

With the approximate expected half-width presented in Eq. 12, the sample size N_{EHP} needed for the expected half-width of a $100(1-\alpha)\%$ confidence interval to fall within the designated bound δ is the minimum integer N , such that

$$t_{N-4, \alpha/2} (N-1)^{-1/2} E[\hat{\sigma}] \cdot E[W^{*-1/2}] \leq \delta. \quad (9)$$

In contrast, the approximate probability in Eq. 13 suggests another useful approach to sample size determination. Consequently, the sample size N_{PHP} required to guarantee, with a given tolerance probability $(1-\gamma)$, that the width of a $100(1-\alpha)\%$ confidence interval will not exceed the planned range ω is the smallest integer N , such that

$$E[\Phi(\omega^* W^*)] \geq 1-\gamma. \quad (10)$$

Clearly, the two inequalities in Eqs. 9 and 10 are more involved than the corresponding formulas in Eqs. 6 and 8. Instead of substituting the term SSE_{XZ} with a constant estimate, as in the simplified procedures, the extra complexity in the suggested formulations is employed to reflect the embedded stochastic characteristic of SSE_{XZ} . The usefulness of the sophisticated inequalities is shown next in the numerical example.

Numerical example

To illustrate a typical sample size problem most frequently encountered in the planning stage of an MMR study, we

consider the hypothetical research framework presented in Shieh (2010) for assessing interaction effects between the length of time in the position (X) and managerial ability (Z) on the self-assurance of managers (Y). Due to the prospective nature of advanced research planning, the general guideline suggests that a successful pilot study can offer plausible and reasonable planning values of the vital characteristics for calculating the necessary sample size. On the basis of the pilot data in Table 5 in Shieh (2010), the empirical distribution of the 60 observed configurations of predictor and moderator variables may be utilized to reconstruct or approximate the actual distributional feature of the two variables. However, the prescribed sample size procedures differ in their use of the pilot information and lead to substantially distinct results. The key difference is that the simplified method treats the regressor variables as constants, but it is more reasonable to regard them as random variables, and so the suggested procedures assume they were drawn from a multivariate population.

For the 60 observed values of $\mathbf{X}_i = (X_i, Z_i, X_i Z_i)^T$ and empirical probability $1/60$ for $i = 1, \dots, 60$, the moment matrices for the quantities in the Appendix can be obtained by

$$\mu \doteq \sum_{i=1}^{60} \mathbf{X}_i / 60, \Sigma \doteq \sum_{i=1}^{60} (\mathbf{X}_i - \mu)(\mathbf{X}_i - \mu)^T / 60,$$

and

$$\psi \doteq \sum_{i=1}^{60} [(\mathbf{X}_i - \mu)(\mathbf{X}_i - \mu)^T \otimes (\mathbf{X}_i - \mu)(\mathbf{X}_i - \mu)^T] / 60.$$

Moreover, it follows that $\mu_{W^*} = (1 - \rho_{XZ}^2)\sigma_{XZ}^2 \doteq 1.2348$ and $\sigma_{W^*}^2 \doteq 22.6511 / (N - 1)$, as shown in Shieh (2010). In planning a research study with these prior inputs, the minimum sample size needed to control the expected half-width of a 95% confidence interval of interaction effects within the designated bound 0.15 can be computed with the inequalities in Eqs. 6 and 9. The resulting sample sizes are

$N_{EHS} = 145$ and $N_{EHP} = 156$, and the corresponding approximate expected half-widths given in Eq. 12 are 0.1560 and 0.1497, respectively. Hence, the sample size N_{EHS} determined by the simplified method is too small to satisfy the desired precision bound. This phenomenon should continue to exist in other settings. Moreover, the smallest sample sizes required to guarantee with a given tolerance probability of .90 that the width of a 95% confidence interval will not exceed the planned range 0.15 are $N_{PHS} = 165$ and $N_{PHP} = 216$, based on the inequalities in Eqs. 8 and 10, respectively. Notably, the approximate tolerance probabilities obtained with Eq. 13 are .6762 and .9019. The sizable difference between $N_{PHP} - N_{PHS} = 216 - 165 = 51$ yields the major deficit of $.9019 - .6762 = .2257$ in tolerance probability.

This numerical investigation exemplifies the fundamental deficiency that overlooking the stochastic nature in regressor variables may lead to a serious underestimation of the sample size required to obtain a designated expected half-width or to ensure the adequate probability of achieving the desired half-width for the confidence intervals. To facilitate the application of the proposed approaches, the SAS/IML (SAS Institute, 2008) programs employed to perform the sample size calculations are available as [supplemental materials](#) from brm.psychonomic-journals.org/content/supplemental. Users can easily identify the statements containing the key values in this exposition and then modify the program to accommodate their own specifications.

Simulation study

To further evaluate the performance of the sample size formulas with respect to the prescribed precision criteria in Eqs. 3 and 4 under various parameter specifications, the MMR model defined in Eq. 1 with bivariate normal predictor, and moderator variables is used as the base for a

Table 1 Computed sample size, approximate expected half-width, and simulated expected half-width for the half-width of a 90% two-sided interval of β_{XZ} at specified half-width $\delta = 0.15$ with bivariate

normal predictor and moderator variables ($\beta_{XZ} = 1, \sigma^2 = 1, \mu_X = \mu_Z = 0, \sigma_X^2 = \sigma_Z^2 = 1$, correlation ρ)

ρ	Simplified Method				Proposed Method			
	N_{EHS}	Approximate Expected Half-Width	Simulated Expected Half-Width	Error	N_{EHP}	Approximate Expected Half-Width	Simulated Expected Half-Width	Error
.1	125	0.1499	0.1551	-0.0052	129	0.1497	0.1525	-0.0028
.3	117	0.1494	0.1555	-0.0061	122	0.1497	0.1520	-0.0023
.5	103	0.1493	0.1570	-0.0078	111	0.1492	0.1508	-0.0015
.7	87	0.1496	0.1598	-0.0102	97	0.1498	0.1496	0.0002
.9	73	0.1492	0.1623	-0.0131	83	0.1492	0.1494	-0.0002

Table 2 Computed sample size, approximate expected half-width, and simulated expected half-width for the half-width of a 90% two-sided interval of β_{XZ} at specified half-width $\delta = 0.20$ with bivariate

normal predictor and moderator variables ($\beta_{XZ} = 1, \sigma^2 = 1, \mu_X = \mu_Z = 0, \sigma_X^2 = \sigma_Z^2 = 1, \text{correlation } \rho$)

ρ	Simplified Method				Proposed Method				
	N_{EHS}	Approximate Expected Half-width	Simulated expected Half-width	Error	N_{EHP}	Approximate Expected Half-Width	Simulated Expected Half-Width	Error	
.1	73	0.1997	0.2111	-0.0114	78	0.1991	0.2029	-0.0038	
.3	68	0.1998	0.2128	-0.0130	75	0.1986	0.2011	-0.0025	
.5	60	0.1999	0.2174	-0.0175	68	0.1997	0.2019	-0.0022	
.7	52	0.1983	0.2203	-0.0220	60	0.1982	0.2014	-0.0031	
.9	44	0.1979	0.2258	-0.0279	52	0.1954	0.2009	-0.0055	

Monte Carlo exposition. A similar bivariate normal assumption was made for related MMR treatments in McClelland and Judd (1993), O'Connor (2006), and Shieh (2009, 2010). Specifically, the coefficient parameters and variance of the simple interaction model are set as $\beta_I = \beta_X = \beta_Z = \beta_{XZ} = 1$ and $\sigma^2 = 1$, respectively. Moreover, the predictor and moderator (X, Z) variables are jointly normally distributed with mean $(0, 0)$, variance $(1, 1)$, and correlation ρ . The correlation parameter ρ between the predictor variable and moderator variable is set at the five levels of .1–.9 in increments of .2.

With the specifications described above, the simulation study is conducted in two steps. First, under the selected values of coefficient parameters, error, distribution configurations of the bivariate predictor, and moderator distribution, the estimates of sample sizes required for precise interval estimation are computed with confidence levels $(1 - \alpha) = .90$ and $.95, \delta = \omega = 0.15$ and 0.20 , and tolerance probability $(1 - \gamma) = .90$. These levels were selected to represent reasonably the range of specifications and sample sizes used in typical research settings. The sample sizes N_{EHS} and N_{EHP} required for controlling the expected half-width are presented in Tables 1, 2, 3 and 4, while those sample sizes, N_{PHS} and N_{PHP} , ensuring the probability of a bounded confidence

interval are summarized in Tables 5, 6, 7 and 8. It can be readily seen that the computed sample sizes N_{EHP} are generally larger than N_{EHS} for controlling the expected half-width in Tables 1, 2, 3 and 4, although the difference is more pronounced between the sample size estimates N_{PHS} and N_{PHP} for the assurance of tolerance probability in Tables 5, 6, 7 and 8. With the reported sample sizes N_{EHS} and N_{PHS} , the actual values of the approximate expected half-width and approximate tolerance probability given in Eqs. 5 and 7, respectively, are also calculated. Similarly, the approximated expected half-width and approximated tolerance probability associated with computed sample sizes N_{EHP} and N_{PHP} are calculated on the basis of Eqs. 12 and 13, respectively. As was expected, the resulting approximate expected half-width is slightly less than the designated expected half-width, whereas the approximate tolerance probability is marginally greater than the chosen tolerance probability.

In the second step, the accuracy of the sample size procedures is examined through a Monte Carlo simulation study. Estimates of the true expected half-width and tolerance probability associated with a given sample size N and various parameter configurations are computed through a Monte Carlo simulation of 10,000 independent

Table 3 Computed sample size, approximate expected half-width, and simulated expected half-width for the half-width of a 95% two-sided interval of β_{XZ} at specified half-width $\delta = 0.15$ with bivariate

normal predictor and moderator variables ($\beta_{XZ} = 1, \sigma^2 = 1, \mu_X = \mu_Z = 0, \sigma_X^2 = \sigma_Z^2 = 1, \text{correlation } \rho$)

ρ	Simplified Method				Proposed Method				
	N_{EHS}	Approximate Expected Half-Width	Simulated Expected Half-Width	Error	N_{EHP}	Approximate Expected Half-Width	Simulated Expected Half-Width	Error	
.1	176	0.1498	0.1531	-0.0033	179	0.1499	0.1518	-0.0019	
.3	164	0.1495	0.1536	-0.0041	168	0.1499	0.1520	-0.0020	
.5	143	0.1500	0.1560	-0.0060	151	0.1496	0.1514	-0.0019	
.7	121	0.1500	0.1575	-0.0075	131	0.1495	0.1504	-0.0008	
.9	101	0.1498	0.1593	-0.0095	111	0.1500	0.1506	-0.0006	

Table 4 Computed sample size, approximate expected half-width, and simulated expected half-width for the half-width of a 95% two-sided interval of β_{XZ} at specified half-width $\delta = 0.20$ with bivariate

normal predictor and moderator variables ($\beta_{XZ} = 1, \sigma^2 = 1, \mu_X = \mu_Z = 0, \sigma_X^2 = \sigma_Z^2 = 1, \text{correlation } \rho$)

ρ	Simplified Method				Proposed Method				
	N_{EHS}	Approximate Half-Width	Expected Simulated Half-Width	Error	N_{EHP}	Approximate Half-Width	Expected Simulated Half-Width	Error	
.1	102	0.1995	0.2068	-0.0073	106	0.1994	0.2034	-0.0040	
.3	95	0.1994	0.2091	-0.0096	101	0.1992	0.2027	-0.0034	
.5	84	0.1990	0.2116	-0.0125	92	0.1996	0.2003	-0.0008	
.7	71	0.1998	0.2164	-0.0166	81	0.1994	0.2000	-0.0007	
.9	60	0.1990	0.2188	-0.0198	69	0.1989	0.2007	-0.0018	

data sets. For each replicate, N sets of predictor and moderator values are generated from the designated bivariate normal distribution. These values of predictor and moderator, in turn, determine the mean responses for generating N normal outcomes with the simple interaction model. Next, the half-width estimate H is computed, and the simulated expected half-width is the mean of the 10,000 replicates of H . Alternatively, the simulated tolerance probability is the proportion of the 10,000 replicates whose values of H are less than or equal to the specified bound ω . The adequacy of the sample size procedure for precise interval estimation is determined by the following formula: error = approximated expected half - width – simulated expected half - width, or error = approximated tolerance probability – simulated tolerance probability. The simulated expected half-width, simulated tolerance probability, and associated error are also summarized in Tables 1, 2, 3, 4, 5, 6, 7 and 8.

Examination of the sample sizes in these tables reveals the general pattern that when all other factors remain constant, the sample size increases with increasing confidence level $(1 - \alpha)$, with decreasing half-width bound δ or ω and with decreasing correlation ρ . Therefore, for both simplified and proposed methods, the largest sample size

$N_{EHS} = 176$ and $N_{EHP} = 179$ for $(1 - \alpha) = .95, \delta = 0.15$ and $\rho = .1$ in Table 3, whereas the smallest sample sizes $N_{EHS} = 44$ and $N_{EHP} = 52$ for $(1 - \alpha) = .90, \delta = 0.20$ and $\rho = .9$ in Table 2. Accordingly, the largest and smallest sample sizes N_{PHS} are 198 and 54, and N_{PHP} are 232 and 84 in Tables 7 and 6, respectively.

Furthermore, as can be seen from the errors in Tables 1, 2, 3 and 4 concerning the precision of expected half-width, the inequality in Eq. 6 for computing the sample size N_{EHS} produces accurate expected half-width for the cases of $\delta = 0.15$ in Tables 1 and 3. Most of the resulting absolute errors are less than 0.01, with only two exceptions (0.0102 and 0.0131) in Table 1. In contrast, the performance associated with the situations of $\delta = 0.20$ in Tables 2 and 4 degrade slightly but are still reasonable, with the absolute errors between 0.0073 and 0.0279. In short, the simplified approach tends to give reliable yet marginally smaller than required sample sizes. However, the accuracy improves with larger sample sizes, because large errors occur with smaller sample sizes. Comparatively, the errors of the expected half-width associated with sample sizes N_{EHP} in Tables 1, 2, 3 and 4 clearly show that the inequality of Eq. 9 performs extremely well because all absolute errors are less than or equal to 0.0055 for the 20 cases examined

Table 5 Computed sample size, approximate tolerance probability, and simulated tolerance probability for the half-width of a 90% two-sided interval of β_{XZ} at specified half-width $\omega = 0.15$ and tolerance

probability 0.90 with bivariate normal predictor and moderator variables ($\beta_{XZ} = 1, \sigma^2 = 1, \mu_X = \mu_Z = 0, \sigma_X^2 = \sigma_Z^2 = 1, \text{correlation } \rho$)

ρ	Simplified Method				Proposed Method			
	N_{PHS}	Approximate Tolerance Probability	Simulated Tolerance Probability	Error	N_{PHP}	Approximate Tolerance Probability	Simulated Tolerance Probability	Error
.1	144	.9059	.6650	0.2409	173	.9023	.8990	0.0033
.3	135	.9116	.6456	0.2660	167	.9012	.9099	-0.0087
.5	119	.9047	.6120	0.2927	155	.9001	.9115	-0.0114
.7	102	.9049	.6075	0.2974	140	.9025	.9198	-0.0173
.9	86	.9019	.5884	0.3135	123	.9020	.9232	-0.0212

Table 6 Computed sample size, approximate tolerance probability, and simulated tolerance probability for the half-width of a 90% two-sided interval of $\hat{\beta}_{XZ}$ at specified half-width $\omega = 0.20$ and tolerance

probability .90 with bivariate normal predictor and moderator variables ($\beta_{XZ} = 1, \sigma^2 = 1, \mu_X = \mu_Z = 0, \sigma_X^2 = \sigma_Z^2 = 1, \text{correlation } \rho$)

ρ	Simplified Method				Proposed Method			
	N_{PHS}	Approximate Tolerance Probability	Simulated Tolerance Probability	Error	N_{PHP}	Approximate Tolerance Probability	Simulated Tolerance Probability	Error
.1	87	.9082	.6508	0.2574	110	.9033	.9050	-0.0017
.3	82	.9157	.6348	0.2809	108	.9040	.9144	-0.0104
.5	73	.9150	.5991	0.3159	102	.9019	.9208	-0.0189
.7	63	.9134	.5930	0.3204	94	.9035	.9293	-0.0258
.9	54	.9178	.5857	0.3321	84	.9024	.9311	-0.0287

here. Although the simplified inequality of Eq. 6 is accurate enough for practical use, it is consistently outperformed by the improved formula in Eq. 9.

On the other hand, the performance of the simplified method is extremely poor and disturbing for the consideration of tolerance probability of bounded interval half-width. The discrepancies between the approximate tolerance probabilities and simulated tolerance probabilities in Tables 5, 6, 7, and 8 range from 0.2409 to 0.3321. Hence, the inequality in Eq. 8 severely overestimates the attained tolerance probability, and thus it underestimates the necessary sample size to meet the selected criterion. Similar findings were reported in Kupper and Hafner (1989) for interval estimation of one- and two-sample problems. In the present case, the parameter value of $(N - 4)(1 - \rho_{XZ}^2)\sigma_{XZ}^2$ is used in place of SSE_{XZ} , so the variability in SSE_{XZ} has been neglected in sample size calculations. The repercussions of ignoring the random feature of predictor and moderator variables on sample size calculation are detrimental and substantial. Hence, the simplified procedure should not be used in such a random regression setting, because it can lead to under allocation of sample size or overconfidence in the interval precision. Regarding the behavior of the proposed method, an

inspection of Tables 5, 6, 7 and 8 shows that the corresponding differences between the approximated and simulated tolerance probabilities are fairly small. Since the considered approach uses large sample approximation, the accuracy is affected, to some extent, for those situations with small sample sizes. The largest two deviations of -0.0258 and -0.0287 occur with the sample sizes 94 and 84 in Table 6 for $\rho = .7$ and $.9$, respectively. Obviously, the advantage of the proposed procedure over the simplified method persists in the case of ensuring the required tolerance level. Due to the behavior of the simplified method, the accurate performance of the suggested inequality outweighs the extra computational requirement. In light of these detailed empirical comparisons, the proposed methods are clearly superior to the simplified procedures in sample size calculations for precise interval estimation.

Concluding remarks

The purpose of the present article was to discuss the sample size issues surrounding the use of confidence intervals for the inference of interaction effects in MMR

Table 7 Computed sample size, approximate tolerance probability, and simulated tolerance probability for the half-width of a 95% two-sided interval of $\hat{\beta}_{XZ}$ at specified half-width $\omega = 0.15$ and tolerance

probability .90 with bivariate normal predictor and moderator variables ($\beta_{XZ} = 1, \sigma^2 = 1, \mu_X = \mu_Z = 0, \sigma_X^2 = \sigma_Z^2 = 1, \text{correlation } \rho$)

ρ	Simplified Method				Proposed Method			
	N_{PHS}	Approximate Tolerance Probability	Simulated Tolerance Probability	Error	N_{PHP}	Approximate Tolerance Probability	Simulated Tolerance Probability	Error
.1	198	.9020	.6598	0.2422	232	.9021	.8984	0.0037
.3	185	.9050	.6453	0.2597	223	.9025	.9022	0.0003
.5	163	.9004	.6271	0.2733	205	.9011	.9040	-0.0029
.7	140	.9097	.6096	0.3001	182	.9000	.9163	-0.0163
.9	118	.9114	.6074	0.3040	159	.9016	.9179	-0.0163

Table 8 Computed sample size, approximate tolerance probability, and simulated tolerance probability for the half-width of a 95% two-sided interval of $\hat{\beta}_{XZ}$ at specified half-width $\omega = 0.20$ and tolerance

probability .90 with bivariate normal predictor and moderator variables ($\beta_{XZ} = 1, \sigma^2 = 1, \mu_X = \mu_Z = 0, \sigma_X^2 = \sigma_Z^2 = 1, \text{correlation } \rho$)

ρ	Simplified Method				Proposed Method			
	N_{PHS}	Approximate Tolerance Probability	Simulated Tolerance Probability	Error	N_{PHP}	Approximate Tolerance Probability	Simulated Tolerance Probability	Error
.1	118	.9006	.6522	0.2484	145	.9033	.8973	0.0060
.3	111	.9093	.6320	0.2773	141	.9036	.9160	-0.0124
.5	98	.9014	.6049	0.2965	132	.9029	.9218	-0.0189
.7	85	.9127	.6043	0.3084	119	.9008	.9212	-0.0204
.9	72	.9097	.5940	0.3157	106	.9030	.9294	-0.0264

studies. The focus is on the information contained in the joint probability function of continuous predictor and moderator variables in a simple interaction model. During the planning stage for MMR research with limited resources, it is important to consider all possible features. We demonstrate that sample size estimates for precise interval estimation of interaction effects will generally be inadequate and misleading if they are based solely on the anticipated characteristics of the regressor variables. Thus, the simplified procedure based on a fixed modeling setup should not be applied indiscriminately. With increased computing power and the general availability of statistical software, computational simplicity is no longer an adequate criterion. An appropriate approach should involve all of the critical factors in sample size determination. Therefore, a more prudent strategy is to account for the stochastic behavior in regressor variables. This article gives explicit formulas for calculating the necessary sample size with respect to the considerations of expected confidence interval half-width and tolerance probability of interval half-width within a designated value. The proposed approaches have clear advantages in the flexibility of the joint distribution of predictor and moderator variables and the unification of a normal approximation for ease of computation. More important, the performance of the suggested methods appears to be remarkably good for the range of model specifications considered in the present article. The proposed methodology not only facilitates the advocated practice of interval procedures, but also further reinforces the potential usefulness of MMR analysis.

Author Note The author thanks the editor, Gregory Francis, and the two anonymous reviewers for their valuable comments on earlier drafts of the article.

Appendix

The properties of the confidence interval half-width It follows from the standard assumption in Eq. 1 for the simple interaction model that the half-width of the $100(1 - \alpha)\%$ confidence interval is $H = t_{N-4, \sigma/2} \left\{ \hat{\sigma}^2 / SSE_{XZ} \right\}^{1/2}$. Hence, the statistical property of H is determined by the joint distribution of $\hat{\sigma}^2$ and SSE_{XZ} . Since the distribution of $\hat{\sigma}^2 \sim \{ \sigma^2 / (N - 4) \} \chi^2(N - 4)$ does not depend on the predictor and moderator variables, $\hat{\sigma}^2$ and SSE_{XZ} are independent. The remaining issue in describing the feature of H is to attain the distribution of SSE_{XZ} . Note that the distribution of SSE_{XZ} is somewhat more complex, and an explicit expression generally does not exist. However, much of the complexity is considerably simplified if we consider the asymptotic phenomenon.

For ease of discussion, the moments of the explanatory vectors $\mathbf{X}_i = (X_i, Z_i, X_i Z_i)^T$ are defined as

$$\boldsymbol{\mu} = E[\mathbf{X}_i], \Sigma = E\left[(\mathbf{X}_i - \boldsymbol{\mu})(\mathbf{X}_i - \boldsymbol{\mu})^T \right]$$

and

$$\boldsymbol{\psi} = E\left[(\mathbf{X}_i - \boldsymbol{\mu})(\mathbf{X}_i - \boldsymbol{\mu})^T \otimes (\mathbf{X}_i - \boldsymbol{\mu})(\mathbf{X}_i - \boldsymbol{\mu})^T \right],$$

where the expectations are taken with respect to the joint probability density function $g(X_i, Z_i)$ of (X_i, Z_i) , and \otimes represents the Kronecker product. Analogous to the practical standpoint of Shieh (2009) for providing a generally useful and versatile solution without being specifically confined to any particular joint probability function $g(X_i, Z_i)$, we consider the large-sample distribution of $W^* = SSE_{XZ} / (N - 1)$,

$$W^* \sim N(\boldsymbol{\mu}_{W^*}, \boldsymbol{\sigma}_{W^*}^2), \tag{11}$$

where $\boldsymbol{\mu}_{W^*} = 1 / (\mathbf{c}^T \Sigma^{-1} \mathbf{c})$, $\boldsymbol{\sigma}_{W^*}^2 = v_{W^*}^2 / (N - 1)$, $v_{W^*}^2 = \boldsymbol{\mu}_{W^*}^4 \{ (\mathbf{c}^T \Sigma^{-1} \otimes \mathbf{c}^T \Sigma^{-1}) \boldsymbol{\psi} (\Sigma^{-1} \mathbf{c} \otimes \Sigma^{-1} \mathbf{c}) - \boldsymbol{\mu}_{W^*}^2 \}$, $\mathbf{c} =$

$(0, 0, 1)^T$ is a 3×1 row vector, and Σ and Ψ are defined above. It is noted in Shieh (2009) that the mean value μ_{W^*} is equivalent to the extra variance of the product XZ after controlling for X and Z . Therefore, it is more informative to express it as $\mu_{W^*} = (1 - \rho_{XZ}^2)\sigma_{XZ}^2$ here. As $SSE_{XZ} = (N - 1)W^*$ is the residual sum of squares for the regression of product term XZ on the X and Z variables, the values of both SSE_{XZ} and W^* are presumably nonnegative. It appears that the probability of negative W^* , $P(W^* < 0)$, is often small enough so that the large sample normal approximation of W^* is nearly adequate for practical purposes. Thus, the evaluations with respect to $g(X_i, Z_i)$ are transformed to and approximated by the corresponding assessments with respect to W^* .

According to the aforementioned theoretical results, the expected half-width can be simplified as

$$E[H] \doteq t_{N-4, \alpha/2} (N-1)^{-1/2} E[\hat{\sigma}] E[W^{*-1/2}] \quad (12)$$

It follows from the fact that $\hat{\sigma}^2$ has a multiple of a chi-square distribution that

$$E[\hat{\sigma}] = \sigma \{2/(N-4)\}^{1/2} \Gamma\{(N-4)/2\}.$$

Since there is no analytic expression for $E[W^{*-1/2}]$, the actual quantity needs to be evaluated with numerical integration with respect to the normal distribution in Eq. 11. In addition, the probability $P\{H < \omega\}$ for $\omega > 0$ can be rewritten as

$$P\{H < \omega\} \doteq P\{K < \omega^* W^*\},$$

Where $K = (N-4)\hat{\sigma}^2/\sigma^2 \sim \chi^2(N-4)$ and $\omega^* = \{(N-1)(N-4)\omega^2\}/t_{N-4, \alpha/2}^2 \sigma^2$. To permit computational simplifications, the approximation is alternatively expressed as

$$P\{H < \omega\} \doteq E[\Phi(\omega^* W^*)], \quad (13)$$

where $\Phi(c) = P\{K < c\}$, $c > 0$, is the cumulative density function of $\chi^2(N-4)$. Since all related functions of normal and chi-square distributions are readily embedded in contemporary statistical packages, the expressions for approximate values of $E[H]$ and $P\{H < \omega\}$ given in Eqs. 12 and 13, respectively, can be readily implemented.

References

- Aguinis, H. (1995). Statistical power problems with moderated multiple regression in management research. *Journal of Management*, 21, 1141–1158.
- Aguinis, H., Beaty, J. C., Boik, R. J., & Pierce, C. A. (2005). Effect size and power in assessing moderating effects of categorical variables using multiple regression: A 30-year review. *The Journal of Applied Psychology*, 90, 94–107.
- Aguinis, H., & Stone-Romero, E. F. (1997). Methodological artifacts in moderated multiple regression and their effects on statistical power. *The Journal of Applied Psychology*, 82, 192–206.
- American Psychological Association. (2001). *Publication manual of the American Psychological Association* (5th ed.). Washington, DC: Author.
- Cramer, E. M., & Appelbaum, M. I. (1978). The validity of polynomial regression in the random regression model. *Review of Educational Research*, 48, 511–515.
- Hahn, G. J., & Meeker, W. Q. (1991). *Statistical intervals: A guide for practitioners*. New York: Wiley.
- Kelley, K. (2008). Sample size planning for the squared multiple correlation coefficient: Accuracy in parameter estimation via narrow confidence intervals. *Multivariate Behavioral Research*, 43, 524–555.
- Kelley, K., & Maxwell, S. E. (2008). Sample size planning with applications to multiple regression: Power and accuracy for omnibus and targeted effects. In P. Alasuuta, L. Bickman, & J. Brammen (Eds.), *The Sage handbook of social research methods* (pp. 166–192). Newbury Park, CA: Sage.
- Krishnamoorthy, K., & Xia, Y. (2008). Sample size calculation for estimating or testing a nonzero squared multiple correlation coefficient. *Multivariate Behavioral Research*, 43, 382–410.
- Kupper, L. L., & Hafner, K. B. (1989). How appropriate are popular sample size formulas? *American Statistician*, 43, 101–105.
- McClelland, G. H., & Judd, C. M. (1993). Statistical difficulties of detecting interactions and moderator effects. *Psychological Bulletin*, 114, 376–390.
- O'Connor, B. P. (2006). Programs for problems created by continuous variable distributions in moderated multiple regression. *Organizational Research Methods*, 9, 554–567.
- Rencher, A. C. (2000). *Linear models in statistics*. New York: Wiley.
- Sampson, A. R. (1974). A tale of two regressions. *Journal of the American Statistical Association*, 69, 682–689.
- SAS Institute (2008). *SAS/IML User's Guide, Version 9.2*. Cary, NC: SAS Institute Inc.
- Shieh, G. (2009). Detecting interaction effects in moderated multiple regression with continuous variables: Power and sample size considerations. *Organizational Research Methods*, 12, 510–528.
- Shieh, G. (2010). Sample size determination for confidence intervals of interaction effects in moderated multiple regression with continuous predictor and moderator variables. *Behavior Research Methods*, 42, 824–835.
- Smithson, M. (2003). *Confidence intervals*. Thousand Oaks, CA: Sage.
- Steiger, J. H., & Fouladi, R. T. (1997). Noncentrality interval estimation and the evaluation of statistical models. In L. L. Harlow, S. A. Mulaik, & J. H. Steiger (Eds.), *What if there were no significance tests?* (pp. 222–257). Mahwah, NJ: Erlbaum.
- Stone-Romero, E. F., Alliger, G. M., & Aguinis, H. (1994). Type II error problems in the use of moderated multiple regression for the detection of moderating effects of dichotomous variables. *Journal of Management*, 20, 167–178.
- Wilkinson, L., & the American Psychological Association Task Force on Statistical Inference. (1999). *Statistical methods in psychology journals: Guidelines and explanations*. *American Psychologist*, 54, 594–604.