Detecting Interaction Effects in Moderated Multiple Regression With Continuous Variables Power and Sample Size Considerations

Gwowen Shieh

Organizational Research Methods 2009 12: 510 originally published online 29 July 2008
DOI: 10.1177/1094428108320370

Organizational Research Methods
Volume 12 Number 3
July 2009 510-528

10.1177/1094428108320370
http://orm.sagepub.com
hosted at
http://online.sagepub.com

Detecting Interaction Effects in Moderated Multiple Regression With Continuous Variables

Power and Sample Size Considerations

Gwowen Shieh
National Chiao Tung University, Hsinchu, Taiwan

In view of the long-recognized difficulties in detecting interactions among continuous variables in moderated multiple regression analysis, this article aims to address the problem by providing feasible solutions to power calculation and sample size determination for significance test of moderating effects. The proposed approach incorporates the essential factors of strength of moderator effect, magnitude of error variation, and distributional property of predictor and moderator variables into a unified framework. Accordingly, careful consideration across different plausible and practical configurations of the prescribed factors is an important aspect of power and sample size computations in planning moderated multiple regression research. The performance of the suggested procedure and an alternative simplified method is illustrated with detailed numerical studies. The simulation results demonstrate that an acceptable degree of accuracy can be obtained using the recommended method in assessing moderated relationships.

Keywords: *interaction; moderator variable; moderating effect; power; sample size*

Despite the prevalent recognition and application of moderated multiple regression in management, psychology, education, and related disciplines, the efforts devoted to the detection of moderating effects are often futile because of insufficient statistical power. Consequently, this situation impedes theory development involving hypothesized moderating effects. To address this problem, various studies identified the conditions and factors pertaining to the power issues in moderated multiple regression, see Aguinis (1995, 2004) and the references therein. Specifically, Aguinis and Stone-Romero (1997), and Stone-Romero, Alliger, and Aguinis (1994) provided thorough treatments on the methodological artifacts and statistical implications associated with the effects of dichotomous moderators. Moreover, the corresponding empirical results are employed to develop approximate procedures and computer algorithms for estimating power using values of manipulated factors in Aguinis and Pierce (1998) and Aguinis, Pierce, and Stone-Romero (1994), respectively. Numerous investigations have been extended to the circumstance of categorical moderator variables with heterogeneity of error variance; see Aguinis, Beaty, Boik, and Pierce (2005) for a comprehensive and excellent review. However, unlike other

Author's Note: The author would like to thank the editor, Herman Aguinis, and the two anonymous reviewers for their valuable comments on earlier drafts of the article. This research was partially supported by National Science Council.

empirical examinations that depend exclusively on simulation results, a theory-based approximation to the power of detecting effects of categorical moderator variable was presented by Aguinis, Boik, and Pierce (2001). The essential features of their method are that the categorical moderator variable may take on more than two levels, and that the criterion variable and continuous predictor could be measured with error and/or truncation, and subgroup variance heterogeneity.

It should be noted that the notorious difficulty of detecting moderator effects is particularly true in observational studies with continuous variables. Arguably, the theoretical progress and computational considerations associated with the power analysis for the interaction effects of continuous predictor and moderator variables have received little attention in the context of moderated multiple regression. In addition, Baron and Kenny (1986) emphasized the importance of choosing a proper analytic procedure for testing moderation and considered four distinct cases for the moderator variable and predictor variable combination: categorical and categorical, categorical and continuous, continuous and categorical, and continuous and continuous. The method of Aguinis et al. (2001) for categorical moderator is not appropriate for calculating power in assessing moderator effect for the last situation that both moderator and predictor variables are continuous. Naturally, the special consideration of continuous moderator and predictor variables incurs the important notion of two different regression formulations. First, the conventional linear regression models fall under the fixed or conditional modeling framework in which the predictor configurations of the regression models are preset by the researcher. Therefore, the corresponding results would be specific to the particular values of the predictor variables and in general, are not applicable in other design settings. These should be the cases more likely to occur in experimental studies where the factors are under the control of the investigators. On the other hand, it is quite common in field studies that not only the values of response variables for each subject are just available after the observations are made, but the levels of predictor variables are also outcomes of the research, especially when the predictors are of continuous measures. Under such circumstances, it is more suitable to employ the random regression or unconditional setup. This subtle consideration is closely related to the implication of stochastic regressors commonly discussed in econometric texts such as Greene (2008) and Murray (2006). In practice, the tests of hypotheses and estimates of parameters are the same under both models. However, the distinction between the two modeling approaches becomes important when power and sample size calculations are to be made. See Cramer and Appelbaum (1978) and Sampson (1974) for further details about the intrinsic appropriateness and theoretical properties of fixed and random models. Similar emphasis and related implication can be found in Dunlap, Xin, and Myers (2004), Gatsonis and Sampson (1989), Mendoza and Stafford (2001), and Shieh (2006, 2007). Particularly, I rely primarily on the general result of Shieh (2007) for the distinct advantage of applying directly to the problems that arise in the detection of moderating effects.

In this article, I focus on the simple interaction models with criterion variable Y, continuous predictor variable X, continuous moderator variable Z, their cross-product term XZ, and error term ε in the formulation of $Y = \beta_I + X\beta_X + Z\beta_Z + XZ\beta_{XZ} + \varepsilon$ that have been extensively discussed in many moderated multiple regression applications. The moderator Z is essentially the second predictor variable hypothesized to moderate the $X - Y$

relationship. See Aiken and West (1991), Cohen, Cohen, West, and Aiken (2003) and Jaccard and Turrisi (2003) for general and illuminating expositions. To take account of the embedded randomness and variability of the predictor and moderator, the appropriate strategy is to consider the random regression setting. It has been demonstrated in McClelland and Judd (1993) that the joint distribution of the predictor and moderator is one of the deterministic factors of detecting moderating effects. Specifically, McClelland and Judd noted that the statistical power for the test of interactions is partly attributed to the extra variance of the product XZ after controlling for the main effects of X and Z or the unique variation in XZ that is not shared with either X or Z which, in turn, is determined entirely by the joint distribution of X and Z. Moreover, even though a useful yet complex expression was derived by McClelland and Judd for the extra variance of the product after controlling for the main effects of X and Z in terms of various variances and covariances of the predictor and moderator variables, no specific operational guideline was provided for power and sample size calculations to take account of the distinguishing underlying variability of the predictor and moderator variables. It will be shown later that failing to account for the variability of the predictor and moderator may distort power analysis and lead to a poor choice of sample size.

Regarding the distributional assumptions of the associated predictor and moderator variables, it is common to assume that the two continuous predictor and moderator variables have a joint bivariate normal distribution in illustrative and theoretical treatments of moderated multiple regression. For example, see McClelland and Judd (1993) and O'Connor (2006). However, it should be obvious that the product of two normally distributed variables does not have a normal distribution. Therefore, the established results for power analysis of fixed characteristics and multinormal settings do not apply in this application. Moreover, there are also many situations where the predictor and moderator variables are continuous, but the assumption of normality is completely unrealistic. Consequently, a general approach to encompassing both normal and nonnormal distributions of the predictor and moderator variables is essential to moderated multiple regression for performing power and sample size calculations in practical applications.

Shieh (2007) has recently considered a unified approach to the determinations of power and sample size for random regression models with arbitrary distributional formulations of the stochastic explanatory variables. Note that the prescribed moderated multiple regression with continuous predictor and moderator variables can be viewed as a special case of the random regression models. Hence, the unified and pedagogical presentation of Shieh furnishes the basis for detailed examination and theoretical justification of the inferential procedures for moderated multiple regression analysis. In a continual effort to support the analytical development and improve the essence of research findings in moderated multiple regression, the general result of Shieh is specialized to provide explicit and useful computational formulas for power calculation and sample size determination here. Essentially, the impact of the methodological artifacts on statistical power for detecting moderating effects can be assessed directly without resorting to empirically derived formulas or results.

The rest of the article is organized as follows. In the next section, the fundamental theory and analytical results of the significance test of interactions and moderating effects in the context of moderated multiple regression with fixed and random configurations of predictor and moderator variables are described. Moreover, the emphasis is placed on the fundamental discrepancy between two approximate procedures in terms of power function and sample size determination. Then, numerical illustrations are presented to exemplify the critical differences between the two methods and the implementation of the accompanying SAS and R programs for power calculation and sample size determination. To further demonstrate the advantage of the proposed approach over the alternative simplified method, simulation studies are also conducted. Finally, some concluding remarks are given.

Moderated Multiple Regression

Consider the simple interaction model or moderated multiple regression model within the fixed modeling framework,

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

where Y_i is the value of the response variable Y, X_i and Z_i are the known constants of the predictor X and moderator Z, ε_i are *iid* $N(0, \sigma^2)$ random errors for $i = 1, \ldots, N$; and β_I, β_X, β_Z, and β_{XZ} are unknown parameters. For the purpose of detecting the moderator effect, the problem is naturally concerned with the least squares estimator $\hat{\beta}_{XZ}$ of β_{XZ}, and the distributional property of the corresponding test statistic t_{XZ} for the hypothesis H_0: $\beta_{XZ} = 0$ versus H_1: $\beta_{XZ} \neq 0$. The analytical formulation of t_{XZ} is described in Appendix A. If the null hypothesis H_0: $\beta_{XZ} = 0$ is true, the statistic t_{XZ} is distributed as $t(N-4)$, a central t distribution with $N-4$ degrees of freedom, and H_0 is rejected at the significance level α if $|t_{XZ}| > t_{N-4,\alpha/2}$, where $t_{N-4,\alpha/2}$ is the upper $100(\alpha/2)$th percentile of the t distribution $t(N-4)$. Moreover, the corresponding power function is,

$$P\{|t_{XZ}| > t_{N-4,\alpha/2}\} = P\{|t(N-4, \Lambda)| > t_{N-4,\alpha/2}\}, \tag{2}$$

where $t(N-4, \Lambda)$ is the noncentral t distribution with $N-4$ degrees of freedom and noncentrality parameter Λ given in Equation A2. The observed proportional reduction in error (PRE; see McClelland & Judd, 1993) or squared partial correlation of the moderator effect is a function of the statistic t_{XZ} as follows:

$$\text{PRE} = \frac{t_{XZ}^2}{1 + t_{XZ}^2}.$$

Here I restrict attention to the specific circumstance that both the predictor X and moderator Z are continuous variables. Owing to the nature of continuous measurements encountered in field research, the explanatory variables typically cannot be controlled and are only available after observation. Hence, to extend the concept and applicability to moderated multiple regression, the continuous predictor and moderator variables $\{(X_i, Z_i), i = 1, \ldots, N\}$ in Equation 1 are assumed to have a joint probability function $g(X_i, Z_i)$ with finite moments. It is assumed that the form of $g(X_i, Z_i)$ does not depend on any of the unknown parameters (β_I, β_X, β_Z, β_{XZ}) and σ^2. From a practical standpoint of providing generally useful and versatile solution without specifically confining to any particular joint probability function $g(X_i, Z_i)$, it is prudent to consider the large sample viewpoint described in Appendix B. The formulation is because of Shieh (2007). However, the simplified approximate distribution of t_{XZ} as a normal mixture of noncentral t distributions in Theorem 1 of Appendix B provides a theoretically more transparent representation than that of Shieh where the formulation is expressed in terms of a normal mixture of noncentral F distributions within the broader general linear hypothesis framework.

With the important justification of asymptotic properties given in Appendix B, the proposed approximate power function in the context of random regression is,

$$P\{|t_{XZ}| > t_{N-4,\alpha/2}\} = E_{W*}[P\{|t(N-4, \Lambda^*)| > t_{N-4,\alpha/2}\}], \tag{3}$$

where $t(N-4, \Lambda^*)$ is the noncentral t distribution with $N-4$ degrees of freedom and noncentrality parameter $\Lambda^* = \beta_{XZ}\{(N-1)W*/\sigma^2\}^{1/2}$ as defined in Equation B4, the expectation $E_{W^*}[\cdot]$ is taken with respect to the approximate normal distribution of W^* presented in Equation B2 with $W^* = 1/\{(N-1)M\}$, M is the (3, 3) element of $\mathbf{A}^{-1}$, $\mathbf{A} = \sum_{i=1}^{N}(\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})^{\mathrm{T}}$, $\bar{\mathrm{X}} = \sum_{i=1}^{N}\mathbf{X}_i/N$, and $\mathbf{X}_i = (X_i, Z_i, X_iZ_i)^{\mathrm{T}}$ is the 3×1 row vector for values of predictor X_i, moderator Z_i and their cross-product X_iZ_i for $i = 1, \ldots, N$. It should be noted that the numerical computation of approximate power in Equation 3 requires the evaluation of noncentral t cumulative distribution function and the one-dimensional integration with respect to a normal distribution. The suggested formulation of normal mixture of t distributions in Equation 3 is referred to as the NT approximation or approach for ease of exposition. This procedure is not as simple as using a z or t table, but it is not unreasonable in light of modern computing capabilities.

Consider the alternative simplified approximation to the distribution of t_{XZ} presented in Theorem 2 of Appendix B; one shall be content with the approximate power function,

$$P\{|t_{XZ}| > t_{N-4,\alpha/2}\} \doteq P\{|t(N-4, \lambda^*)| > t_{N-4,\alpha/2}\}, \tag{4}$$

where λ^* is a constant defined in Equation B6. For ease of reference this simplified t approximation is described as the ST approximation or method. In this case, the computation of approximate power of Equation 4 involves only the noncentral t cumulative distribution function as in the case of Equation 2. It can be justified that the two approximate distributions given in Theorems 1 and 2 of Appendix B are asymptotically equivalent as N goes to infinity. However, the finite-sample properties of the NT and ST approaches can be substantially different and the respective power functions defined in Equations 3 and 4 may yield markedly different results for small samples.

Furthermore, it is insightful to consider that the asymptotic mean μ_{W*} of W^* can be viewed as the extra variance of the product XZ after controlling for X and Z, denoted by $V[XZ|(X, Z)]$, as presented in Equation B3 of Appendix B. Moreover, it was demonstrated in the simulation studies of McClelland and Judd (1993) that the magnitude of $V[XZ|(X, Z)]$ or μ_{W*} plays an important role for the differences in statistical power performance between experiments and field studies. However, the impact of the joint distribution of X

and Z is more transparent and assessable with the aforementioned approximate power functions given in Equations 3 and 4. Accordingly, the effect size of the simplified t approximation in Equation 4 can be approximately quantified as,

$$f_{XZ} = \frac{\lambda^*}{N^{1/2}} \doteq \beta_{XZ}\left\{\frac{\mu_{W^*}}{\sigma^2}\right\}^{1/2}. \tag{5}$$

Owing to the proposed two-stage distribution approximation to the t statistic in Theorem 1 of Appendix B, the corresponding effect size is intrinsically determined by the normal approximation for W^* given in Equation B2. However, the actual effect size of the test statistic t_{XZ} is a function of the underlying distribution of Λ or W within the context of random regression.

For the purpose of sample size determination, the approximate power functions defined in Equations 3 and 4 can be employed to calculate the sample size needed to test hypothesis H$_0$: $\beta_{XZ} = 0$ versus H$_1$: $\beta_{XZ} \neq 0$ to attain the specified power for the chosen significance level α, parameter values (β_I, β_X, β_Z, β_{XZ}, σ^2), and moments of (X_i, Z_i) obtained with the joint probability distribution $g(X_i, Z_i)$. It can be readily seen that the prescribed power formulas Equations 2, 3, and 4 do not depend on the three coefficient parameters β_I, β_X, and β_Z, therefore they are entirely irrelevant to the power and sample size calculations for detecting moderating effect. Moreover, the mean values of the predictor, moderator and their product are not included in the power functions of Equations 3 and 4. Hence, the mean vector or first moments associated with the joint distribution of explanatory variables has no influence on the power or on the required sample size. An important aspect of computation is that a simple and standard iterative search is required to find the necessary sample size. These results will be applied later to implement varieties of power calculation and sample size determination for moderated multiple regression analysis. Finally, the corresponding hypothesis testing procedures and approximate power functions for the two one-sided tests of H$_0$: $\beta_{XZ} \leq 0$ versus H$_1$: $\beta_{XZ} > 0$, and H$_0$: $\beta_{XZ} \geq 0$ versus H$_1$: $\beta_{XZ} < 0$ and nonzero minimum effect can be readily established but the details are not given here.

Numerical Examples

For illustrative purposes, I present in this section the power and sample size calculations for detecting interaction effects in moderated multiple regression analysis based on a given set of pilot data. The following numerical assessment represents a typical research situation most frequently encountered in the planning stage of a study. The ultimate aim is to reveal the potential consequence of failing to account of the underlying stochastic property of the explanatory variables.

Suppose there are 40 pairs of observations for predictor variable X and moderator variable Z obtained from a pilot study. The values of (X, Z) listed in Table 1 represent random samples generated from a bivariate normal population with $\mu_X = \mu_Z = 0$, $\sigma_X^2 = \sigma_Z^2 = 1$ and correlation $\rho = 0.5$. According to the continuous characteristics of measurements X and Z, it is clear that the sample values in the subsequent study vary from one application to another. However, the observed configurations from the pilot study can be employed as an

Table 1
The Observed Values of Predictor Variable *X* and Moderator Variables *Z* of the Pilot Study

X	*Z*	*X*	*Z*	*X*	*Z*	*X*	*Z*	*X*	*Z*
0.11	−1.02	0.58	−0.46	0.27	0.51	0.64	0.35	0.76	0.48
0.98	0.26	−0.76	0.06	−0.18	0.15	0.78	0.70	0.18	0.47
−0.58	0.91	0.28	1.18	1.14	1.43	0.83	−0.86	−0.78	0.17
0.61	−0.17	0.08	0.74	−0.67	−1.70	1.52	0.32	0.18	0.85
0.04	2.06	1.08	−0.31	−0.15	−0.62	−0.50	0.79	−0.30	−0.02
0.60	0.56	−0.49	0.60	0.87	0.34	−0.29	−0.66	−1.04	1.30
0.14	−1.35	−1.12	−0.79	0.74	1.68	−0.69	−1.44	−0.80	−1.01
−3.21	−1.91	−0.42	−0.49	2.79	2.35	−0.47	−0.96	−0.77	−1.58

empirical approximation to the underlying joint distribution of *X* and *Z*. Moreover, it is shown next that the prescribed NT and ST approaches utilize the empirical features associated with the predictor and moderator variables in distinctive ways and, accordingly the two formulas lead to substantially different results in power and sample size calculations.

As a continued exposition of the numerical studies in McClelland and Judd (1993, p. 379), I follow their setup where the parameters of the moderated multiple regression are chosen as $\beta_I = 0$, $\beta_X = \beta_Z = \beta_{XZ} = 1$ and $\sigma^2 = 16$. It follows from the data in Table 1 that the approximate normal distribution of W^* has the mean $\hat{\mu}_{W^*} = 2.1030$ and variance $\hat{\sigma}^2_{W^*} = 54.5894$. Moreover, the approximate effect size defined in Equation 5 is $f_{XZ} = 0.3625$. As described in the preceding section, for planning future research based on current information, the sample sizes needed to obtain specified power at the significance level $\alpha = 0.05$ of the NT and ST methods are determined by the approximate power functions defined in Equations 3 and 4, respectively. The resulting sample sizes for NT method are 101 and 127 for power level 0.90 and 0.95, respectively. On the other hand, the corresponding sample sizes are 82 and 101 for the ST method. Obviously, the calculated sample sizes of the two procedures differ considerably for the given set of (*X*, *Z*) data. The discrepancies in terms of percentage of required size are $100\% \times (101 - 82)/101 = 18.81\%$ and $100\% \times (127 - 101)/127 = 20.47\%$ for the respective two power level 0.90 and 0.95. Furthermore, the sample size of 101 required for the ST method to attain power 0.95 yields instead the power level of 0.90 for the more sophisticated NT approach. Moreover, it can be shown that the presumably adequate sample size 82 reported by the ST procedure to reach 0.90 power value only gives the power level 0.84 with the NT approach. Thus, the differences in power performance are $0.90 - 0.84 = 0.06$ and $0.95 - 0.90 = 0.05$ for the two cases considered here. In short, the sample sizes calculated with the ST algorithm are comparatively smaller than those of the NT approach and the phenomenon shall continue to exist in other settings of random explanatory variables. The differences between the two approaches will be further examined and reinforced in the simulation study. The SAS/IML (SAS Institute, 2003) and R (R Development Core Team, 2006) programs used to perform these power and sample size calculations is provided in Appendices C and D, respectively. Users can easily identify the statements with this self-contained exposition and it only requires a slight modification of the program to accommodate their own specifications. Moreover, electronic copies of the programs are available on request and can also be downloaded from the website: www.ms.nctu.edu.tw/faculty/shieh.

Because of the prevalence of the fixed modeling setup and the close resemblance between the two power functions Equations 2 and 4, the researcher may unknowingly assume the ST procedure is the right tool for performing power assessment and planning necessary sample size within the context of random regression. It is well recognized that effect size has an important impact on statistical power performance and sample size determination. As shown above, the effect size f_{XZ} of ST method given in Equation 5 is a direct function of μ_{W*} or $V[XZ|(X, Z)]$, that is, the extra variance of the product XZ after controlling for X and Z. In the numerical demonstration just described for sample size planning in the detection of moderator effect, it is comprehensible that the more involved NT method not only utilizes μ_{W*} or $V[XZ|(X, Z)]$ in the formulation, but it also combines other distributional aspects into one unified framework. In viewing the indispensable role of the joint distribution of predictor and moderator along with the magnitudes of interaction coefficient parameter and model random error, it is advisable that researchers should have thorough understanding of the impact of each of these factors on statistical power and how they work as whole in the detection of moderating effect. The generality and accuracy of the proposed methodology for power and sample size calculations will be demonstrated in the subsequent section.

Simulation Studies

To evaluate the performance and reinforce the key concept of the proposed approach, further numerical investigations are performed for the detection of moderating effect in moderated multiple regression in this section. Because the considered approaches use large sample approximations, simulation studies are conducted to assess their adequacy for finite sample and robustness under various parameter specifications and power levels. For the sake of analytical tractability in derivation and primary focus in literature, the moderated multiple regression model with bivariate normal predictor and moderator variables is considered in this numerical examination. However, it is frequently the case that not only the distributions of the variables are not necessarily normal, often they are not even symmetrical, especially in small samples. Therefore, the simple interaction model with bivariate gamma predictor and moderator variables is examined as well.

Specifically, the following parameter values are fixed throughout the numerical investigation: $\beta_I = 0$, $\beta_X = \beta_Z = 1$ and $\sigma^2 = 16$. The magnitude of interaction coefficient parameter β_{XZ} is chosen as 0.50, 0.75, and 1 so as to represent reasonably the range of moderating effects that are possible in most studies. However, space limitations preclude reporting all details; only the outcomes associated with $\beta_{XZ} = 1$ are presented. For the joint bivariate distribution of predictor and moderator, the variables are standardized so that the means and variances of the predictor and moderator variables are $\mu_X = \mu_Z = 0$ and $\sigma_X^2 = \sigma_Z^2 = 1$. Moreover, the correlation parameter ρ between the predictor variable and moderator variable is set at the four levels of 0, 0.1, 0.5, and 0.9.

Table 2
Calculated Sample Sizes for the Detection of Moderating Effect With Bivariate Normal Predictor and Moderator Variables ($\beta_{XZ} = 1, \sigma^2 = 16, \sigma_X^2 = \sigma_Z^2 = 1, \alpha = .05$)

	The ST Method				The NT Method			
ρ	0.0	0.1	0.5	0.9	0.0	0.1	0.5	0.9
Power								
0.90	171	169	137	95	182	181	154	116
0.95	210	208	169	117	226	224	192	146

Table 3
Calculated Sample Sizes for the Detection of Moderating Effect With Bivariate Gamma Predictor and Moderator Variables ($\beta_{XZ} = 1, \sigma^2 = 16, \sigma_X^2 = \sigma_Z^2 = 1, \alpha = .05$)

	The ST Method				The NT Method			
ρ	0.0	0.1	0.5	0.9	0.0	0.1	0.5	0.9
Power								
0.90	174	150	119	93	203	194	165	120
0.95	214	185	147	115	255	246	211	151

With the specifications described above, the numerical study is conducted in two steps. First, under the selected values of coefficient parameters, error, distribution configurations of bivariate predictor and moderator distribution, the estimates of sample sizes required for testing the moderating effect of H_0: $\beta_{XZ} = 0$ versus H_1: $\beta_{XZ} \neq 0$ with significance level 0.05 and power = 0.90 and 0.95 are calculated. The computed sample sizes are presented in Tables 2 and 3 for the bivariate normal and gamma settings, respectively.

Preliminary inspection of the tables reveals the expected general relations: increase in sample sizes with increasing power level, and with decreasing effect size for both methods. As in the case of pilot study, the sample size associated with the ST procedure is less than that of the NT method for all cases in these tables.

In the second step, I continue the comparison by conducting simulation studies. The sample size N calculated by the NT approach is utilized as the benchmark to recalculate the approximate powers for both competing methods. Then, estimates of the true power associated with given sample size and parameter configuration are then computed through Monte Carlo simulation of 10,000 independent data sets. For each replicate, N sets of predictor and moderator values are generated from the designated bivariate normal or bivariate gamma distribution. The generations of bivariate normal random variables with $\mu_X = \mu_Z = 0$, $\sigma_X^2 = \sigma_Z^2 = 1$ and correlation $\rho = 0$, 0.1, 0.5 and 0.9 are straightforward. In contrast, the desirable bivariate gamma random variables are standardized version of (X, Z), where $(X, Z) = (X_1, Z_1)$, $(X_1 + G_1, Z_1 + G_1)$, $(X_1 + G_2, Z_1 + G_2)$ and $(X_1 + G_3,$

Table 4
Approximate Powers and Simulated Powers at Specified Sample Sizes for the Detection of Moderating Effect With Bivariate Normal Predictor and Moderator Variables ($\beta_{XZ} = 1$, $\sigma^2 = 16$, $\sigma_X^2 = \sigma_Z^2 = 1$, $\alpha = .05$)

		The ST Method		The NT Method	
N	Simulated Power	Approximate Power	Error	Approximate Power	Error
(i) $\rho = 0$ ($f_{XZ} = 0.2500$)					
182	0.8925	0.9184	−0.0259	0.9005	−0.0080
226	0.9469	0.9626	−0.0157	0.9506	−0.0037
(ii) $\rho = 0.1$ ($f_{XZ} = 0.2512$)					
181	0.8964	0.9195	−0.0231	0.9010	−0.0046
224	0.9476	0.9628	−0.0152	0.9503	−0.0027
(iii) $\rho = 0.5$ ($f_{XZ} = 0.2795$)					
154	0.8979	0.9314	−0.0335	0.9007	−0.0028
192	0.9470	0.9708	−0.0238	0.9505	−0.0035
(iv) $\rho = 0.9$ ($f_{XZ} = 0.3363$)					
116	0.9044	0.9486	−0.0442	0.9012	0.0032
146	0.9543	0.9811	−0.0268	0.9509	0.0034

$Z_1 + G_3$) for $\rho = 0$, 0.1, 0.5, and 0.9, respectively. The random variables X_1, Z_1, G_1, G_2, and G_3 have independent gamma distributions with $X_1 \sim$ gamma(3, 1), $Z_1 \sim$ gamma(5, 1), $G_1 \sim$ gamma(0.43, 1), $G_2 \sim$ gamma(3.94, 1), and $G_3 \sim$ gamma(35.89, 1). These values of predictor and moderator in turn determine the mean responses for generating N normal outcomes with the moderated multiple regression model. Next, the test statistic t_{XZ} is computed and the simulated power is the proportion of the 10,000 replicates whose absolute values $|t_{XZ}|$ exceed the critical value $t_{N-4, 0.025}$. The adequacy of the examined procedure for power and sample size calculation is determined by the error = simulated power− approximate power between the simulated power and approximate power computed earlier. The simulated power, approximate power, and error are summarized in Tables 4 and 5 for the bivariate normal and gamma settings, respectively.

As can be seen from the results, the approximate powers of the NT method are almost identical to 0.90 or 0.95, whereas the approximate power associated with the ST method are marginally greater than 0.90 or 0.95. Moreover, the performance of the large sample approximations improves with the power levels because large errors generally occur with smaller power level for both ST and NT methods. However, the situation for the ST method is much more pronounced than that of the NT approach. It appears that the ST method yielded acceptable results with absolute errors smaller than 0.02 for some cases in Tables 4 and 5, but absolute error are as large as 0.0442 and 0.0499 in Tables 4 and 5 for $\rho = 0.9$ and 0.5, respectively, when the simulated power nears 0.90. In contrast, the absolute errors of the NT approach are uniformly smaller than 0.02 for all the cases considered here. The largest error is 0.0179 which is associated with the case of $\rho = 0.5$ and simulated power nears 0.90 in Table 5. In addition, for $\rho = 0.1$, 0.5, and 0.9, the discrepancies of the

Table 5
Approximate Powers and Simulated Powers at Specified Sample Sizes for the Detection of Moderating Effect With Bivariate Gamma Predictor and Moderator Variables ($\beta_{XZ} = 1$, $\sigma^2 = 16$, $\sigma_X^2 = \sigma_Z^2 = 1$, $\alpha = .05$)

		The ST Method		The NT Method	
N	Simulated Power	Approximate Power	Error	Approximate Power	Error
(i) $\rho = 0$ ($f_{XZ} = 0.2500$)					
203	0.9033	0.9397	−0.0364	0.9004	0.0029
255	0.9586	0.9762	−0.0176	0.9506	0.0080
(ii) $\rho = 0.1$ ($f_{XZ} = 0.2665$)					
194	0.9160	0.9585	−0.0425	0.9009	0.0151
246	0.9609	0.9862	−0.0253	0.9502	0.0107
(iii) $\rho = 0.5$ ($f_{XZ} = 0.2997$)					
165	0.9191	0.9690	−0.0499	0.9012	0.0179
211	0.9677	0.9912	−0.0235	0.9500	0.0177
(iv) $\rho = 0.9$ ($f_{XZ} = 0.3402$)					
120	0.9133	0.9587	−0.0454	0.9021	0.0112
151	0.9591	0.9859	−0.0268	0.9502	0.0089

NT procedure in Table 5 for the predictor and moderator with a bivariate gamma distribution are generally larger than those in Table 4 with a bivariate normal distribution. According to these findings, the ST method is always outperformed by the NT approach. Although the accuracy of both approaches improve with smaller $\beta_{XZ} = 0.50$ and 0.75, the general pattern of results is similar to those reported herein for $\beta_{XZ} = 1$. Overall, the performance of the NT method appears to be reasonably good for the range of model specifications considered here. Therefore, the accurate power formula Equation 3 can be employed to calculate the achieved power level with available sample size so that the possible low statistical power problem for detecting moderating effect can be recognized in advance and naturally, alternative action or analysis may be considered. In planning future research, the procedure can be inverted to determine the minimum sample size required to attain adequate nominal power with specified model configurations and significance level. Thus, the difficulty in detecting moderating effect because of low statistical power can be alleviated with appropriate calculation of sample size.

Note that the effect sizes f_{XZ} are also presented in Tables 4 and 5. Although it is rather obvious that effect size is an increasing function of β_{XZ}, it appears that the effect size and $V[XZ|(X, Z)]$ increase with increasing $|\rho|$ as well. In other words, if X and Z are highly correlated, then the three explanatory variables X, Z, and XZ tend to have a strong linear dependence among them. It was previously stated that the multicollinearity has an adverse effect on the power of the test H_0: $\beta_{XZ} = 0$, see Morris, Sherman, and Mansfield (1986). However, it was empirically shown in Dunlap and Kemery (1987) that it is not the case. Nonetheless, the numerical results in Tables 4 and 5 reveal that if sample size and other factors remain constant, the power is an increasing function of $|\rho|$.

Conclusions

Despite the existing low statistical power problem, moderated multiple regression has remained as one of the important research methods that are applicable to a wide range of fields including education, management, psychology, and other social sciences. Notably, moderation analysis has been the main focus in a variety of management subdisciplines, including organizational behavior, human resources management, operation management, and strategic management. From the methodological viewpoint, the lack of a full range of accessible and accurate statistical methods is a severe dilemma and major setback to the advance of moderation research. In addition to the considerations of methodological artifacts and statistical implications that have been identified and discussed, this article purports to provide feasible formulas for power calculation and sample size determination in moderated multiple regression with continuous predictor and moderated variables. As reported in numerous methodological investigations, the distributional property of the predictor and moderator variables is a key factor affecting power of moderated multiple regression. The essential and distinct notions of fixed and random modeling formulations are greatly emphasized and their discrepancies in power analysis are closely evaluated. In particular, the suggested procedure entails the large sample theory and results in a versatile formulation that possesses the flexibility in the joint distribution of predictor and moderator variables. The primary reason of inducing such scheme and asymptotic approximation is because of the lack of a proper and exact procedure that accounts for the nature of all possible continuous joint distributions of predictor and moderator variables. The numerical assessments suggest that the formula and algorithm for power and sample size calculations are accurate enough for practical purposes. Specifically, the simplified ST method gives reliable results for sufficiently large sample sizes when the predictor and moderator variables have a joint bivariate normal distribution. Alternatively, the NT approach is not seriously affected by mild departures from the normality assumption for the predictor and moderator variables. However, when the predictor and moderator variables of interest are extremely long tailed or heavily skewed, they can be transformed to a more appropriate scale before applying the proposed procedure. Consequently, researchers are advised to comprehend fully the underlying features of predictor and moderator variables and synthesize the information into study designs.

In view of the usual approximate nature of advance research planning, it is a difficult task to assess the robustness of the proposed approach for selected configuration of the predictor and moderator variables, magnitude of moderating effect, and error variance. Hence, it is good practice to consider a range of design variations to provide guidance about the achieved power levels and required sample sizes for the study. Typical sources like previously published research, successful pilot study, and subject matter expertise can offer plausible and reasonable planning values for the vital model characteristics. Nonetheless, the suggested procedure will yield accurate power estimation and sample size calculation provided that all the required information is properly specified. Hopefully, the presented results can be utilized to improve the verification and identification of moderator effects so that researchers can further extend theoretical models to incorporate more complex and rich moderated relationships in future applications of moderation multiple regression.

Appendix A
The Distribution of t_{XZ}

It follows from the standard assumption in Equation 1 that the test statistic t_{XZ} can be expressed as,

$$t_{XZ} = \frac{\hat{\beta}_{XZ}}{\{\hat{\sigma}^2 M\}^{1/2}}, \tag{A1}$$

where $\hat{\sigma}^2$ is the usual unbiased estimator of σ^2, M is the (3, 3) element of A^{-1}, $\mathrm{A} = \sum_{i=1}^{N} (\mathrm{X}_i - \bar{\mathrm{X}})(\mathrm{X}_i - \bar{\mathrm{X}})^{\mathrm{T}}$, $\bar{\mathrm{X}} = \sum_{i=1}^{N} \mathrm{X}_i/N$, and $\mathrm{X}_i = (X_i, Z_i, X_iZ_i)^{\mathrm{T}}$ is the 3×1 row vector for values of predictor X_i, moderator Z_i, and their cross-product X_iZ_i for $i = 1, \ldots, N$. It is well known under the conditional setup that t_{XZ} has a noncentral t distribution $t(N-4, \Lambda)$ with $N-4$ degrees of freedom and noncentrality parameter Λ, where,

$$\Lambda = \frac{\beta_{XZ}}{\{\sigma^2 M\}^{1/2}}. \tag{A2}$$

Appendix B
The Approximate Distributions of t_{XZ} Under Random Regression Modeling

The moments of the explanatory vectors $\mathrm{X}_i = (X_i, Z_i, X_iZ_i)^{\mathrm{T}}$ are defined as $\mu = E[\mathrm{X}_i]$, $\Sigma = E[(\mathrm{X}_i - \mu)(\mathrm{X}_i - \mu)^{\mathrm{T}}]$, $\Psi = E[(\mathrm{X}_i - \mu)(\mathrm{X}_i - \mu)^{\mathrm{T}} \otimes (\mathrm{X}_i - \mu)(\mathrm{X}_i - \mu)^{\mathrm{T}}]$, $E[\cdot]$ denotes the expectation 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. For notational simplicity, μ and Σ are expressed as,

$$\mu = \begin{bmatrix} \mu_X \\ \mu_Z \\ \mu_{XZ} \end{bmatrix} \text{ and } \Sigma = \begin{bmatrix} \sigma_X^2 & C(X,Z) & C(X,XZ) \\ C(X,Z) & \sigma_Z^2 & C(Z,XZ) \\ C(X,XZ) & C(Z,XZ) & V(XZ) \end{bmatrix},$$

where σ_X^2, σ_Z^2, and $V(XZ)$ are the variances of X_i, Z_i, and cross-product X_iZ_i, respectively, and $C(X, Z)$, $C(X, XZ)$ and $C(Z, XZ)$ denote the covariances among the random variables X_i, Z_i, and X_iZ_i. It follows from the standard asymptotic result (Muirhead, 1982, Corollary 1.2.18) and algebraic operation that,

$$(N-1)^{1/2}[\mathrm{vec}(\mathbf{S}) - \mathrm{vec}(\Sigma)] \;\dot{\sim}\; N_{p^2}(0_{p^2}, \Psi - \mathrm{vec}(\Sigma) \cdot \mathrm{vec}(\Sigma)^{\mathrm{T}}), \tag{B1}$$

where $\mathrm{S} = \sum_{i=1}^{N} (\mathrm{X}_i - \bar{\mathrm{X}})(\mathrm{X}_i - \bar{\mathrm{X}})^{\mathrm{T}}/(N-1)$ and $\mathrm{vec}(\cdot)$ is a matrix operator that arranges the columns of a matrix into one long column. According to the formulations of A and M presented in Equation A1 for the test statistic t_{XZ}, both $\mathrm{A} = (N-1)\mathrm{S}$ and M are random variables within the random regression framework and therefore t_{XZ} has a noncentral t distribution with random noncentrality Λ. Nonetheless, it is worthwhile to note that the reciprocal of M, denoted by W $(= 1/M)$, can be viewed as the error sum of squares from regressing XZ on X and Z when both predictor X and moderator Z are continuous variables. Ultimately, the actual distributional property of t_{XZ} depends on the distribution of W. With the prescribed asymptotic result in Equation B1, it can be shown that,

$$W^* \dot{\sim} N(\mu_{W^*}, \sigma^2_{W^*}), \tag{B2}$$

where $W^* = W/(N-1)$, $\mu_{W^*} = 1/(c^T\Sigma^{-1}c)$, $\sigma^2_{W^*} = \mu^4_{W^*}\{(c^T\Sigma^{-1} \otimes c^T\Sigma^{-1})\Psi(\Sigma^{-1}c \otimes \Sigma^{-1}c) -\mu^{-2}_{W^*}\}/\,(N-1)$, and $c = (0, 0, 1)^T$ is a 3×1 row vector. Denote the extra variance of the product XZ after controlling for X and Z by $V[XZ|(X, Z)]$. Then, it can be shown by applying the algebraic manipulation that μ_{W^*} is equivalent to $V[XZ|(X, Z)]$ and can be written as,

$$\mu_{W^*} = V(XZ) - [\,C(X,XZ \quad C(Z,XZ)]\begin{bmatrix} \sigma^2_X & C(X,Z) \\ C(X,Z) & \sigma^2_z \end{bmatrix}^{-1}\begin{bmatrix} C(X,XZ) \\ C(Z,XZ) \end{bmatrix}. \tag{B3}$$

See Equation 2 of McClelland and Judd (1993) for an alternative expression.

Accordingly, the noncentrality parameter Λ in Equation A2 associated with the noncentral t distribution of t_{XZ} is expressed as Λ^* to emphasize the stochastic property of both predictor X and moderator Z:

$$\Lambda^* = \beta_{XZ}\left\{\frac{(N-1)W^*}{\sigma^2}\right\}^{1/2} \tag{B4}$$

And W^* approximately follows a normal distribution given in Equation B2. Therefore, the suggested approximate distribution of t_{XZ} under the random regression setting is completely specified in the following theorem.

Theorem 1. Consider the linear regression model in Equation 1, and (X_i, Z_i) are independent and identically distributed with finite moments, $i = 1, \ldots, N$. The t_{XZ} statistic defined in Equation A1 has the following approximate two-stage distribution,

$$t_{XZ}|W^* \sim t(N-4, \beta_{XZ}\{(N-1)W^*/\sigma^2\}^{1/2}) \text{ and } W^* \dot{\sim} N(\mu_{W^*}, \sigma^2_{W^*}), \tag{B5}$$

where μ_{W^*} and $\sigma^2_{W^*}$ are given in Equation B2.

Note that under the null hypothesis H_0: $\beta_{XZ} = 0$, the null distribution of t_{XZ} remains as $t(N-4)$ under random settings. The values of W^* involved in the noncentrality parameter of t distribution for t_{XZ} in Equation B5 are presumably nonnegative. In the normal approximation of W^* with large samples, the probability of negative W^*, $P(W^* < 0)$, is often small enough so that the normal approximation is nearly adequate. It was found that there is essentially no practical difference in the adequacy for power and sample size calculations by replacing negative values of W^* with 0.

For comparative purpose, consider the even stronger asymptotic result for S that S converges in probability to Σ. As direct consequences of this property, W^* and Λ^* can be approximated by μ_{W^*} and λ^*, respectively, where,

$$\lambda^* = \beta_{XZ}\left\{\frac{(N-1)\mu_{W^*}}{\sigma^2}\right\}^{1/2} \tag{B6}$$

Then, it leads to the alternative and simplified approximation to the distribution of t_{XZ} summarized in the next theorem.

Theorem 2. Consider the linear regression model in Equation 1, and (X_i, Z_i) are independent and identically distributed with finite moments, $i = 1, \ldots, N$. The t_{XZ} statistic defined in Equation A1 has the following approximate distribution,

$$t_{XZ} \dot{\sim} t(N-4, \lambda^*),$$

where λ^* is given above in Equation B6.

Appendix C
The SAS IML program for Power and Sample Size Calculations

```
PROC IML;PRINT ‘MODERATED MULTIPLE REGRESSION’;
PRINT ‘H0: BETAXZ = 0 VS H1: BETAXZ <> 0—TWO-TAIL TEST’;

*REQUIRED USER SPECIFICATIONS PORTION;
ALPHA = 0.05;POWER = 0.90;BETAXZ = 1;SIGMASQ = 16;
XZ = {
0.11 −1.02, 0.58 −0.46, 0.27 0.51, 0.64 0.35, 0.76 0.48,
0.98 0.26, −0.76 0.06,−0.18 0.15, 0.78 0.70, 0.18 0.47,
−0.58 0.91, 0.28 1.18, 1.14 1.43, 0.83 −0.86, −0.78 0.17,
0.61 −0.17, 0.08 0.74, −0.67 −1.70, 1.52 0.32, 0.18 0.85,
0.04 2.06, 1.08 −0.31, −0.15 −0.62, −0.50 0.79, −0.30 −0.02,
0.60 0.56, −0.49 0.60, 0.87 0.34, −0.29 −0.66, −1.04 1.30,
0.14 −1.35, −1.12 −0.79, 0.74 1.68, −0.69 −1.44, −0.80 −1.01,
−3.21 −1.91, −0.42 −0.49, 2.79 2.35,−0.47 −0.96, −0.77 −1.58};
*END OF REQUIRED USER SPECIFICATIONS;
XE = XZ||(XZ[,1]#XZ[,2]);
N = NROW(XE);
XC = XE-J(N, 1,1)*XE[:,];
H = J(3,3,0);HH = H@H;
DO I = 1 TO N;
H = H + XC[I,]’*XC[I,];
HH = HH + (XC[I,]’*XC[I,])@(XC[I,]’*XC[I,]);
END;
SIGM = H/N;
PSI = HH/N;
ISIGM = INV(SIGM);
MUW = 1/ISIGM[3,3];
VW = ISIGM[3,];
VARW = (MUW##4)#((VW@VW)*PSI*(VW’@VW’)-MUW##(−2));
LFXZ = BETAXZ#SQRT(MUW/SIGMASQ);
PRINT ALPHA POWER N BETAXZ SIGMASQ;
PRINT MUW VARW LFXZ[FORMAT = 6.4];
*FOR NUMERICAL INTEGRATION;
NUMINT = 1000;
COEVEC = ({1}||REPEAT({4 2},1,NUMINT/2−1)||{4 1})’;
INT = PROBIT(0.999995);
INTERVAL = 2#INT/NUMINT;
ZVEC = ((INTERVAL#(0:NUMINT)) + (-INT))’;
WZPDF = (INTERVAL/3)#COEVEC#PDF(‘NORMAL’,ZVEC, 0,1);
```

```
*ST APPROACH;
STPOWER = 0;M = 5;
DO WHILE(STPOWER < POWER);
M = M + 1;
TCRIT = TINV(1-ALPHA/2,M-4);
STPOWER = 1-CDF('T',TCRIT, M-4,SQRT(M)#LFXZ);
END;
NST = M;
*NT APPROACH;
NTPOWER = 0;M = MAX(NST-10,5);
DO WHILE(NTPOWER < POWER);
M = M + 1;
TCRIT = TINV(1-ALPHA/2,M-4);
WVEC = SQRT(VARW/(M-1))#ZVEC + MUW;WVEC = WVEC#(WVEC > 0);
NTPOWER = 1-WZPDF'*CDF('T',TCRIT,M-4,BETAXZ#SQRT((M-1)#WVEC/
SIGMASQ));
END;
NNT = M;
*RECALCULATE NTPOWER FOR SAMPLE SIZE NST;
M = NST;
TCRIT = TINV(1-ALPHA/2,M-4);
WVEC = SQRT(VARW/(M-1))#ZVEC + MUW;WVEC = WVEC#(WVEC > 0);
NTPOWER_NST = 1-WZPDF'*CDF('T',TCRIT,M-4,BETAXZ#SQRT((M-1)#WVEC/
SIGMASQ));
DN = NNT-NST;
DNTPOWER = NTPOWER-NTPOWER_NST;
PRINT NST STPOWER[FORMAT = 6.4] NNT NTPOWER[FORMAT = 6.4];
PRINT DN NTPOWER_NST[FORMAT = 6.4] DNTPOWER[FORMAT = 6.4];
QUIT;
```

Appendix D
The R Program for Power and Sample Size Calculations

```
function () {
#REQUIRED USER SPECIFICATIONS PORTION
alpha < −0.05
power < −0.90
betaxz < −1
sigmasq < −16
xzvec < −c( 0.11, −1.02, 0.58, −0.46, 0.27, 0.51, 0.64, 0.35, 0.76, 0.48, 0.98, 0.26, −0.76,
0.06, −0.18, 0.15, 0.78, 0.70, 0.18,
0.47, −0.58, 0.91, 0.28, 1.18, 1.14, 1.43, 0.83, −0.86, −0.78, 0.17, 0.61, −0.17, 0.08, 0.74,
−0.67, −1.70, 1.52, 0.32, 0.18,
0.85, 0.04, 2.06, 1.08, −0.31, −0.15, −0.62, −0.50, 0.79, −0.30, −0.02, 0.60, 0.56, −0.49,
0.60, 0.87, 0.34, −0.29, −0.66, −1.04, 1.30, 0.14, −1.35,
−1.12, −0.79, 0.74, 1.68, −0.69, −1.44, −0.80, −1.01, −3.21, −1.91, −0.42, −0.49, 2.79,
2.35, −0.47, −0.96, −0.77, −1.58)
#END OF REQUIRED USER SPECIFICATION
xz < -matrix(xzvec, length(xzvec)/2,2,byrow = TRUE)
xe < -cbind(xz, xz[,1]*xz[,2])
```

```
n < -nrow(xe)
cmean < -apply(xe, 2,mean)
xc < -xe-matrix(rep(cmean, 40),40,3,byrow = TRUE)
h < -matrix(rep(0,9),3,3)
hh < -h %x% h
for (i in 1:n)
{
xci < -xc[i,,drop = FALSE]
h < -h + (t(xci)%*%(xci))
hh < -hh + (t(xci)%*%(xci))%x%(t(xci)%*%(xci))
}
sigm < -h/n
psi < -hh/n
isigm < -solve(sigm)
muw < -1/isigm[3,3]
vw < -isigm[3,,drop = FALSE]
varw < -(muw\^4)*(((vw)%x%(vw))%*%psi%*%(t(vw)%x%t(vw))-muw\^(-2))
lfxz < -betaxz*sqrt(muw/sigmasq)
print(c(alpha, power, n,betaxz, sigmasq),digits = 4)
print(c(muw, varw, lfxz),digits = 4)
#for numerical integration
numint < -1000
coevec < -c(1,rep(c(4,2),numint/2–1),4,1)
int < -qnorm(0.999995,0,1)
interval < -2*int/numint
zvec < -interval*seq(0,numint) + (-int)
wzpdf < -(interval/3)*coevec*dnorm(zvec, 0,1)
#st approach
stpower < -0
m < -5
while (stpower < power){
m < -m + 1
tcrit < -qt(1-alpha/2,m-4,0)
stpower < -1-pt(tcrit, m-4,sqrt(m)*lfxz)
}
nst < -m
#nt approach
ntpower < -0
m < -max(nst-10,5)
while (ntpower < power){
m < -m + 1
tcrit < -qt(1-alpha/2,m-4,0)
wvec < -sqrt(varw/(m-1))*zvec + muw
wvec < -wvec*(wvec>0)
ntpower < -
1-sum(wzpdf*pt(tcrit, m-4,betaxz*sqrt((m-1)*wvec/sigmasq)))
}
nnt < -m
#recalculate ntpower for sample size nst
m < -nst
tcrit < -qt(1-alpha/2,m-4,0)
wvec < -sqrt(varw/(m-1))*zvec + muw
wvec < -wvec*(wvec>0)
```

```
ntpower_nst < - 1-sum(wzpdf*pt(tcrit, m-4,betaxz*sqrt((m-1)*wvec/sigmasq)))
dn < -nnt-nst
dntpower < -ntpower-ntpower_nst
print(c(nst, stpower, nnt, ntpower),digits = 4)
print(c(dn, ntpower_nst, dntpower),digits = 4)
}
```

References

Aguinis, H. (1995). Statistical power problems with moderated multiple regression in management research. *Journal of Management*, *21*, 1141-1158.

Aguinis, H. (2004). *Regression analysis for categorical moderators*. New York: Guilford.

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. *Journal of Applied Psychology*, *90*, 94-107.

Aguinis, H., Boik, R. J., & Pierce, C. A. (2001). A generalized solution for approximating the power to detect effects of categorical moderator variables using multiple regression. *Organizational Research Methods*, *4*, 291-323.

Aguinis, H., & Pierce, C. A. (1998). Statistical power calculations for detecting dichotomous moderator variables with moderated multiple regression. *Educational and Psychological Measurement*, *58*, 668-676.

Aguinis, H., Pierce, C. A., & Stone-Romero, E. F. (1994). Estimating the power to detect dichotomous moderators with moderated multiple regression. *Educational and Psychological Measurement*, *54*, 690-692.

Aguinis, H., & Stone-Romero, E. F. (1997). Methodological artifacts in moderated multiple regression and their effects on statistical power. *Journal of Applied Psychology*, *82*, 192-206.

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173-1182.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.

Cramer, E. M., & Appelbaum, M. I. (1978). The validity of polynomial regression in the random regression model. *Review of Educational Research*, *48*, 511-515.

Dunlap, W. P., & Kemery, E. R. (1987). Failure to detect moderating effects: Is multicollinearity the problem? *Psychological Bulletin*, *102*, 418-420.

Dunlap, W. P., Xin, X., & Myers, L. (2004). Computing aspects of power for multiple regression. *Behavior Research Methods, Instruments & Computers*, *36*, 695-701.

Gatsonis, C., & Sampson, A. R. (1989). Multiple correlation: Exact power and sample size calculations. *Psychological Bulletin*, *106*, 516-524.

Greene, W. H. (2008). *Econometric analysis* (6th ed.). Upper Saddle River, NJ: Pearson.

Jaccard, J., & Turrisi, R. (2003). *Interaction effects in multiple regression* (2nd ed.). Thousand Oaks, CA: Sage.

McClelland, G. H., & Judd, C. M. (1993). Statistical difficulties of detecting interactions and moderator effects. *Psychological Bulletin*, *114*, 376-390.

Mendoza, J. L., & Stafford, K. L. (2001). Confidence interval, power calculation, and sample size estimation for the squared multiple correlation coefficient under the fixed and random regression models: A computer program and useful standard tables. *Educational and Psychological Measurement*, *61*, 650-667.

Morris, J. H., Sherman, J. D., & Mansfield, E. R. (1986). Failures to detect moderating effects with ordinary least squares-moderated multiple regression: Some reasons and a remedy. *Psychological Bulletin*, *99*, 282-288.

Muirhead, R. J. (1982). *Aspects of multivariate statistical theory*. New York: John Wiley.

Murray, M. P. (2006). *Econometrics: A modern introduction*. Upper Saddle River, NJ: Pearson.

O'Connor, B. P. (2006). Programs for problems created by continuous variable distributions in moderated multiple regression. *Organizational Research Methods*, *9*, 554-567.

R Development Core Team. (2006). R: A language and environment for statistical computing [Computer software and manual]. Available from the R Project for Statistical Computing Web site, http://www.r-project.org

Sampson, A. R. (1974). A tale of two regressions. *Journal of the American Statistical Association*, *69*, 682-689.

SAS Institute (2003). *SAS/IML user's guide*, *Version 8*. Carey, NC: Author.

Shieh, G. (2006). Exact interval estimation, power calculation and sample size determination in normal correlation analysis. *Psychometrika*, *71*, 529-540.

Shieh, G. (2007). A unified approach to power calculation and sample size determination for random regression models. *Psychometrika*, *72*, 347-360.

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.

Gwowen Shieh is a professor of management science at National Chiao Tung University. His current research interests focus on sample size methodology and research methods.