INSEAD

Faculty & Research

Stochastic Moderated Regression: An Efficient Methodology for Estimating Parameters in Moderated Regression

by

Hubert Gatignon
and
Joachim Vosgerau

2006/17/MKT
(Revised version of 2005/30/MKT)

Working Paper Series

Stochastic Moderated Regression:

An Efficient Methodology for Estimating Parameters in Moderated Regression

Hubert Gatignon

and

Joachim Vosgerau*

February 10, 2006

* Hubert Gatignon is the Claude Janssen Chaired Professor of Business Administration and Professor of Marketing at INSEAD, Boulevard de Constance, 77305 Fontainebleau Cedex, France. Tel: +33 (0)1 60 71 26 28, Fax: +33 (0)1 60 74 55, E-mail: Hubert.Gatignon@insead.edu.
Joachim Vosgerau is Assistant Professor of Marketing at the Tepper School of Business, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15201, USA. Tel: +1-412-268-9170, Fax: +1-412-268-7357, E-mail: vosgerau@andrew.cmu.edu

Stochastic Moderated Regression:

An Efficient Methodology for Estimating Parameters in Moderated Regression

Abstract:

In moderated regressions, the effect of a focal variable x_1 depends on the level of a moderator variable x_2. Moderation is estimated by introducing the product term of the two variables (x_1x_2) as an independent variable in the regression equation. Such moderator regressions often suffer from multicollinearity due to the usually high correlation between the product term and its components. We propose to recognize explicitly the stochastic nature of moderating effects to derive more efficient estimates of all the effects in the Stochastic Moderated Regression model (SMR). Using Monte-Carlo simulations, we assess the ability to extract better inference about these effects under different conditions of stochasticity at different levels of the moderating effect. In addition, because of the inability to remove the collinearity inherent in the model specification itself (having a product term and its components in the same model), we evaluate the impact of introducing (or removing) terms in the model specification on the significance of these effects.

Stochastic Moderated Regression:
An Efficient Methodology for Estimating Parameters in Moderated Regression

Many theories in the social sciences, and in marketing in particular, require the use of contingencies, such that one variable influences (moderates) the manner in which another variable exerts its impact on a criterion (dependent) variable of interest. For example, Bowman and Narayandas (2004) find that B2B customers' satisfaction with a vendor leads to stronger loyalty the smaller the customer is. To test such hypothesized moderating effects, the product term of the focal variable and the moderator variable is introduced in the regression equation (Saunders 1956; Sharma, Durand and Gur-Arie 1981).

OLS estimation in moderated regression treats the effects of the variables as if they were completely deterministic. Because of the collinearity inherent to the moderated regression specification, it is critical to estimate its parameters with the greatest level of statistical efficiency possible. [1] We propose to recognize the stochasticity of the moderating process: the Stochastic Moderated Regression model (SMR).[2] The additional stochastic information in the moderating equation in SMR can be used for more efficient estimation with Generalized Least Squares. In this paper we assess the extent to which the SMR model is able to extract better inference about all the effects specified in the model under different conditions of stochasticity and at different levels of the moderating effect.

---

[1] Mean-centering the variables has often been advocated as a means to reduce multicollinearity (Aiken and West 1991; Cohen and Cohen 1983; Jaccard, Turrisi and Wan 1990; Jaccard, Wan and Turrisi 1990; Smith and Sasaki 1979). Indeed, in extremely severe multicollinearity conditions, mean-centering can have an effect on the matrix inversion approximation algorithm that statistical software packages use (this did not occur in any of our simulations covering a wide range of correlations). However, the econometrics literature has documented that mean-centering does not affect multicollinearity at all (Belsley 1984, Judge et al. 1985). Even though the correlations between the mean-centered variables and their product term are much smaller than the correlations between the original raw variables and their product term (Cronbach 1987), the statistical inference is not affected by this linear transformation. The determinant of the covariance matrix of the independent variables (X'X) is the same after mean-centering, as are the residuals sum of squares, R^2s, F-values, the product term coefficient and its standard deviation (Kromrey and Foster-Johnson 1998; Dunlap and Kemery 1987). The coefficients of the component terms differ, though, as they are estimated at the mean and not at zero of the other component (Irwin and McClelland 2001). However, they are equivalent to the coefficients estimated on raw data, and their means can be easily transformed into those of the raw-data coefficients with a simple linear transformation, as well as the corresponding variances and covariances (Cohen 1978).

[2] We thank an anonymous reviewer for suggesting the terminology.

While improving the efficiency of the parameter estimates, SMR, like Ordinary Moderated Regression (OMR), is sensitive to model specification due to the collinearity inherent to the model specification: the product term x_1x_2 is inevitably correlated with the component variables x_1 and x_2, even when the two component variables x_1 and x_2 are uncorrelated.[3] Therefore, we also examine the impact of model specification on the ability to detect the various effects included in the model. For example, estimating a model with a moderating effect on data where no moderating effect is present drastically reduces the ability to detect significant effects (for all coefficients). We conclude that automatically estimating a fully-specified model is ill-advised, and admonish to carefully consider which effects should be specified in moderated regressions. [4]

Summarizing, we introduce the Stochastic Moderated Regression model. We assess the extent to which recognizing the stochastic nature of the moderating effect can improve the efficiency of the estimation in moderated regression models. Furthermore, because of the dramatic effects of model-inherent multicollinearity, we analyze the impact of model specification in moderated regressions.

A STOCHASTIC MODERATED REGRESSION MODEL

The structural relationship between the dependent variable y and the independent variables x_1 and x_2 is typically expressed as:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + u \quad (1)$$

A key feature of a moderating effect that is ignored in this expression is the stochastic nature of the moderating effect. This can be seen by expressing the moderator regression in

[3] We also consider the correlation between the component variables, which corresponds to the ill-conditioning of the data in contrast to the collinearity that is inherent to the model specification. In fact, we show that this data-driven collinearity has little impact on the statistical inference that can be made using such models, relative to the effect of model specification inherent collinearity.

[4] When moderating effects are estimated involving an interval-scaled focal variable, a fully specified model must be estimated because the product term implies that the effect of the moderator variable is only defined up to a constant (Irwin and McClelland 2001).

two equations, a ***response equation*** and a ***moderating equation***. The ***response equation*** for a single observation *i* expresses the effect of x_1 on y:

$$y_i = \beta_0 + \beta_{1i} x_{1i} + u_i \quad (2)$$

The only distinction of this Equation from a standard regression model is the subscript *i* associated with coefficient β_{1i}. This is because in a moderator regression or Moderated Multiple Regression (MMR), the effect of a focal variable x_1 is not constant but varies according to the level of a moderator variable x_2. We now formulate a ***moderating equation*** expressing the moderating role of x_2 on the effect of x_1 on y:[5]

$$\beta_{1i} = \alpha_0 + \alpha_1 x_{2i} + \varepsilon_i \quad (3)$$

Substituting the expression in Equation (3) into Equation (2) leads to:

$$y_i = \beta_0 + (\alpha_0 + \alpha_1 x_{2i} + \varepsilon_i) x_{1i} + u_i \quad (4)$$

By developing, the equation becomes

$$y_i = \beta_0 + \alpha_0 x_{1i} + \alpha_1 x_{1i} x_{2i} + u_i + \varepsilon_i x_{1i} \quad (5)$$

If there were no random term in Equation (3), Equation (5) would reduce to Equation (1), except for the term specifying the constant effect of x_2 on y, which is usually added to Equation (2). We will discuss this issue in a later section.

The coefficient α_1 in Equation (3) is directly interpretable in terms of a moderating effect of x_2, as it indicates by how much the effect of x_1 (i.e., β_1) changes when x_2 changes by one unit. Because of the varying nature of the effect of x_1 on y, we use the term "***constant effect***" for α_0, as it represents the ***constant portion of the effect*** of that variable.[6]

---

[5] Although we introduce a single moderating equation affecting a coefficient of the response equation, the approach can easily incorporate multiple moderating factors on multiple coefficients, as well as non-linear moderating effects (e.g., Gatignon 1993, 1984).

[6] We prefer the term "constant effect" to "simple effect" because the latter is often interpreted in the literature as a "main" effect; this interpretation is inappropriate since it only reflects the estimate of the effect of the focal variable at one particular level of the moderator variable (zero when raw data are analyzed or mean value of the moderator if the data have been mean-centered). It is the constant parameter in the moderating equation (3). It is also the constant effect of the focal variable in the absence of a moderating effect (Jaccard, Turrisi and Wan 1990).

The conceptual role of the random term in the moderating equation

How useful is the random element in Equation (3)? It fundamentally expresses the stochastic nature of the moderating effect. This is appropriate when the theoretical basis for the moderating effect of x_2 is not exclusive of other possible variables, including randomness. The disturbance term ε_i recognizes that this equation is not fully deterministic. Conventional models without an error term in the moderating equation express the moderating effect as strictly deterministic, thereby implying that the moderating variable is the only factor that influences the moderation. However, such a restrictive model specification would be at odds with most theories, as they typically do not argue that a particular variable is the only moderating factor. In fact, such a deterministic model specification would depart from classical statistical inference theory, where the existence of random departures from a deterministic model is required as a source of unexplained variance.

Recognizing the stochasticity of the moderating effect brings two advantages. First, it provides additional information that can be used to provide more efficient estimates which can be especially useful when multicollinearity is present. Second, as pointed out by Gatignon and Hanssens (1987), the stochastic element provides information that may help discriminate between the moderating role of x_2 on x_1 and that of x_1 on x_2. Consider the error term in Equation (5) where x_2 moderates the impact of x_1 on y:

$$w_i = u_i + \varepsilon_i x_{1i} \tag{6}$$

If instead of x_2 moderating x_1, x_1 moderates the effect of x_2, the error term would become: [7]

$$w_i = u_i + \varepsilon_i x_{2i} \tag{7}$$

[7] The argument developed by Gatignon and Hanssens (1987) is even more powerful in the context of a non-linear response function associated with a linear functional form for the varying parameter function.

This difference in the error term structure provides information for assessing the causal direction of moderation, i.e., differentiating whether the moderating effect is due to x_1 moderating the effect of x_2 or to x_2 moderating the effect of x_1.

Estimation of model with a random term in the moderating equation

Stochastic moderated regression is a specific form of varying parameter models which have been introduced in the econometrics literature with the random coefficient model of Hildreth and Houck (1968). Judge et al. (1985) show that varying parameter models belong to the class of heteroscedastic error models, where the variance of y_i is a linear function of a set of variables.[8]

The error term in Equation (5) now shows heteroscedasticity with $V[u_i + x_{1i}\varepsilon_i] = \sigma_u^2 + x_{1i}^2\sigma_\varepsilon^2$. This implies that an estimated generalized least squares estimator (EGLS) will be asymptotically more efficient than OLS. The OLS estimator, which ignores the stochastic element (as in Ordinary Moderated Regression), remains unbiased but is not BLUE any longer (Judge et al. 1985).

While it does not reduce multicollinearity itself, the use of the stochastic information in the moderating equation reduces its effect to a certain extent. Additional information is the general solution to multicollinearity (Leamer 1978). The EGLS estimation takes into account the full information about the specification of a moderating effect, i.e., not only the deterministic relationship that explains the varying nature of a moderating effect but also its stochastic component as depicted in the model specified by Equations (2) and (3).

The improved efficiency is due to the information contained in the residuals and the particular form of heteroscedasticity implied by the moderating equation. However, in the

[8] Varying parameter models (random coefficient models are a special case) are not new to the Marketing literature. However, most applications are either time-varying parameters (e.g., Wildt and Winer 1983) or parameters that vary across sections within times-series of cross sections (e.g., Gatignon 1984, Gatignon and Hanssens 1987).

case of no moderating effect, the variance of the error term should be independent of the focal variable.

In summary, the random element of a moderating function is essential to fully describe the moderating process being studied and can be used to reduce the negative consequences of multicollinearity. In order to demonstrate the extent of the benefits of recognizing the stochastic nature of moderator effects, it is necessary to compare the SMR estimation with OLS estimation relative to the true data generating function. To that effect, we describe a Monte Carlo simulation in the next section. This simulation will also serve to assess the effects of model specification (due to model-inherent multicollinearity especially compared to data-driven collinearity).

THE MONTE CARLO SIMULATION DESIGN

In order to assess the extent of the benefits of SMR, we consider different levels of the strength of the moderating effect. We also manipulate orthogonally the noise in the regression corresponding to the error term in Equation (2), and the noise in the moderating equation corresponding to the error term in Equation (3) in order to evaluate the size of the error variance relative to the variance of the variables (fixed to one). Furthermore, we vary the sample size to evaluate the impact of the other factors relative to the value of additional data. And in order to study the impact of model specification and, in particular, the role of multicollinearity, data was generated with and without moderating effects. By estimating a model with a moderating term on these data sets, we can compare the likelihood of finding a significant moderating effect when in fact there is none (i.e., the moderating effect is set to 0) with the likelihood of finding a significant moderating effect when there is one (i.e., the moderating effect is set to values greater than 0). This analysis is performed for different levels of moderating effects. Finally, we vary the correlation between the two independent variables x_1 and x_2 to assess the role of data-driven multicollinearity (McClelland and Judd 1993)[9].

For each generated data set, an Ordinary Moderated Regression model (Equation (1)) was estimated with Ordinary Least Squares, and a Stochastic Moderated Regression model using Equations (2) and (3) was estimated with Estimated Generalized Least Squares.

Data was generated with SAS 9.1 for Windows. In order to simulate data sets with two independent variables and two error components simulating noise in the regression (u_i in Equation (2)) and noise in the moderating term (ε_i in Equation (3)), four random variables were generated with the *CALL VNORMAL* module (the seed was generated separately for each data set generation using the *RANNOR* function). The first two random variables x_1 and x_2 were generated with a mean of 10 and standard deviation of 1.[10] The latter two random variables representing noise in the regression were generated with a mean of 0 and standard deviation 1. The correlation of the two error-variables with all other variables was set to 0. The correlation between the two independent variables x_1 and x_2 was manipulated at five levels: 0, 0.2, 0.4, 0.6, and 0.8. Data was generated such that the resulting sample correlation between the independent variables x_1 and x_2 would remain positive (data generation was repeated in the infrequent cases of negative correlation until the resulting sample correlation was non-negative; this was necessary only when the target correlation between x_1 and x_2 was set to 0); it ranges from 0 to 0.92. The resulting correlations between the focal variable x_1 and the product term x_1x_2 range from 0.53 to 0. 98.

In order to investigate the impact of the moderating effect relative to the constant effect, we manipulate the size of the moderating effect while keeping the size of the constant

[9] This correlation impacts the collinearity in the data but is conceptually distinct from the model-inherent correlation between the product term and its components.

[10] We also ran simulations with larger values of the coefficients of the components and a wider range of values of the size of the moderating effect relative to the components' effects (larger as well as smaller); the general results reported hereafter were indeed unaffected. The specific impact on the conclusions of having means near zero is discussed in a latter section.

effect fixed. Specifically, the size of the impact of x_1 and x_2 without moderation ($\alpha_1 = 0$ in Equation (3)) was fixed at 1 (α_0 and β_2), and the intercept was fixed at 4 (β_0 in Equation (2)). The size of the moderating effect (α_1 in Equation (3)) was manipulated at six levels: 0, 0.2, 0.4, 0.6, 0.8, and 1.0. The first level (0) implies that there is no moderating effect in the generated data. For levels greater than 0, implying a moderating effect, noise within the moderating equation was manipulated at six levels by multiplying the random error variable (ε_i in Equation (3)) with the following constants: 0, 0.4, 0.8, 1.2, 1.6, 2.0. This provided moderating equation variances of 0, 0.16, 0.64, 1.44, 2.56, and 4.0, which correspond to explained variances of the moderating effect ranging from 0.99% to 86.2%.

Noise in the regression was manipulated at 8 levels by multiplying the random error variable (u_i in Equation (2)) with the following constants: 0.1, 0.7, 1.3, 1.9, 2.5, 3.1, 3.7, 4.3. This provided error variances of 0.01, 0.49, 1.69, 3.61, 6.25, 9.61, 13.69, and 18.49, corresponding to explained variances in the base case of no moderating effect from 9.76% to 99.5%. Data sets were generated with four differing sample sizes: 50, 175, 300, and 425. At each level of the manipulated factors, 10 data sets were generated. Thus, the total number of experimental conditions is 5 x 6 x 6 x 8 x 4 = 5760, and the total number of generated data sets is 5760 x 10 = 57600.

THE BENEFITS OF STOCHASTIC MODERATED REGRESSION

For each dataset generated and discussed above, two models with a moderating term were estimated: A model corresponding to Equation (1) was estimated with OLS, and a Stochastic Moderated Regression model corresponding to Equations (2) and (3) was estimated with GLS with the Hildreth and Houck (1968) method. Our criterion for measuring efficiency is based on the Generalized Mean Squared Error defined as $\mathrm{MSE}(\hat{\beta}) = (\hat{\beta} - \beta)(\hat{\beta} - \beta)' + \mathrm{V}[\hat{\beta}]$ (Judge et al. 1985). Even though the OLS parameter estimates are unbiased, they are unreliable due to multicollinearity. The first component of MSE measure reflects how close the estimated parameters are to their true value. The second component represents the variance of the parameter estimates. We compare the gains in efficiency by taking the difference in the trace of the Generalized Mean Squared Errors obtained from OLS and those obtained from GLS estimations (Table 1). This difference provides an overall measure of the improvements across all the model parameters.

Figure 1a plots the gains in efficiency obtained from SMR-GLS estimation relative to OLS as a function of the size of the noise in the moderating equation. As predicted, SMR is more efficient than OLS, and efficiency gains are greatest when there is a lot of noise in the moderating equation. The nature of the heteroscedasticity specified by Equation (5), $\mathrm{V}[u_i + x_{1i}\varepsilon_i] = \sigma_u^2 + x_{1i}^2\sigma_\varepsilon^2$, explains why this is the case. Efficiency gains depend on the ability to estimate the components of the variance, i.e., σ^2_u and σ^2_ε. This ability is a function of the relative size of the three components, σ^2_u, σ^2_ε, and the variance exhibited by x_{1i}. If the first component σ^2_u is large compared to the second part of the sum, there is little gain in efficiency, since the variance is basically homoscedastic. Also, if x_{1i} exhibits only a small variance relative to the noise of the response equation or of the moderating equation, the auxiliary regression corresponding to the estimation of the variance components (which characterize the heteroscedasticity pattern) cannot provide sufficient information. And as the variance of x_{1i} was fixed to one in our Monte Carlo simulation, we only need to consider how efficiency varies as a function of the noise in the moderating equation (see Figure 1a). The results show that, as expected, the gain in efficiency increases with the noise in the moderating equation.[11] Gains in efficiency also increase as the sample size decreases (Figure 1b). The smaller differences for large sample sizes follow from the increased

[11] A more complex model involving a larger number of moderating variables would result in an error variance which would be a linear function of the squares of these other parameters; the complexity of such a structure could potentially contribute to the benefits of SMR depending on the relative variances exhibited by these moderating variables.

statistical power obtained from large samples. So, SMR appears particularly useful in the typical research context where sample sizes are small and the moderating effect is noisy.

Considering separately the gains in efficiency of the intercept and the gains of the other parameters, the pattern remains the same as in Figures 1a and 1b; however, the magnitude is much larger for the intercept term (Table 1).[12] This difference is mostly due to the larger values corresponding to the intercept. However, in terms of percentages, improvements are comparable across coefficients. In 55% of the experimental conditions where a stochastic moderating effect is specified, SMR provides more efficient estimates for each parameter than OLS, and the average efficiency is improved by 9.7%, 7.6%, 8.7%, and 6.7%, for the intercept, the constant effects x_1 and x_2, and the moderating effect, respectively. In individual cases improvements can be quite dramatic: improvements of at least 10% to over 100% are obtained for the intercept estimate in 69% of the cases, for the estimate of the constant effect x_1 in 60% of the cases, for the estimate of the constant effect x_2 in 65% of the cases, and for the estimate of the moderating effect in 57% of the cases. Another substantial number of improvements occur in the range from 5% to 10% so that the percentage of cases where improvements exceed 5% are 87%, 79%, 83% and 76% respectively for the intercept, the constant effects x_1 and x_2, and the moderating effect. These improvements in efficiency are rather stable over the range of the size of the moderating effect. However, the improvement for the estimate of the constant effect of the moderating variable (x_2) is slightly greater than for the constant effect of the focal variable (x_1). This asymmetry results from the fact that the variance of the error term is a function of the square of the focal variable and not a function of the moderating variable. This allows discriminating between the moderating role and the constant effect of the moderating variable. But this discrimination is possible only to the extent that the focal and the moderating variables are not highly correlated.

[12] The individual diagonal element of the MSE matrix is then used to measure the efficiency gain for each corresponding parameter.

Indeed, as can be seen in Figure 1c, the difference in efficiency improvement between the effect of x_1 and the effect of x_2 disappears when these two variables are highly correlated (when the correlation approaches 0.8).

In summary, our Monte-Carlo study shows that SMR provides estimates that are generally more reliable (closer to their true value) and have a smaller standard deviation.

MODEL SPECIFICATION AND MODEL-INHERENT MULTICOLLINEARITY

As mentioned earlier, model-inherent collinearity occurs because the product term x_1x_2 is correlated with the component variables x_1 and x_2. As a consequence, collinearity can arise even when the two component variables x_1 and x_2 themselves are uncorrelated. This is exemplified in our Monte Carlo simulation. Even in the cases where the correlation between x_1 and x_2 was set to 0, the observed correlation between x_1 and the product term x_1x_2 still ranged from 0.53 to 0.89. This model-inherent collinearity bears the same negative consequences as data-driven multicollinearity (as indicated by the instability of the estimated coefficients reported in the prior section) but is conceptually different from the latter. It is therefore, critical to examine the impact of the model specification in this context.

In the following, we look at our Monte Carlo simulation cases where the moderating effect was set to 0 (i.e., there is no moderating effect in the generated data). We estimate a model with a moderating effect on this data and examine the likelihood of erroneously finding a significant moderating effect. We then examine how likely the model with the irrelevant moderating effect is to rightly pick up the constant effects in the data. We compare this likelihood with the likelihood of a model that does not contain the irrelevant moderating effect. Finally, we investigate the consequences of including other irrelevant components in the estimated model.

What are the chances of inferring the existence of moderating effects when none are present?

We first look at the simulation cases where there is only a constant effect for x_1 and x_2 but no moderating effect of x_2 on x_1 in the generated data (i.e., moderating effect size is set to 0). Notice that without a moderating effect in the generated data, SMR-GLS cannot detect heteroscedasticity (as it is caused by the moderating equation). Consequently, the results of OLS and SMR estimations are similar and we will report only OLS estimators.

Given that there is no moderating effect in the data, the proportion of significant moderating parameters by chance alone would be 5% at the .05 confidence level (Lehmann 2001). However, Ganzach (1997, p. 236) claims that "the coefficient of the product term in the regression may be significant even when there is no true interaction. The reason for this is that when the correlation between X and Z increases [here x_1 and x_2], so does the correlation between XZ and X." Ganzach's claim is in contrast to McClelland and Judd's (1993, p. 378) proposition that "given that a change in origin can always be found to ensure a zero covariance between the product and its components and given that such a change of origin does not alter the moderator statistical test, the covariance, if any between the components and their product, is in principle irrelevant for detecting moderator effects." Figure 2 shows the impact of multicollinearity as the correlation between x_1 and x_2 increases. Each bar represents the percentage of the estimated parameters that is significant (.05 confidence) for a given level of correlation between the focal and the moderator variables. Concentrating on the left side of the graph (A), the first group of bars from the left represents the intercept β_0 in Equation (1) at each of the five levels of correlation (increasing from left to right starting at 0 until 0.8). The second group corresponds to β_1, i.e., the coefficient of x_1. The third group corresponds to β_2, the coefficient of x_2, and finally the fourth group corresponds to the moderating effect β_3 (i.e., the object in question). The proportion of significant coefficients using a two-tailed test is 5.1%, and it is independent of the level of correlation between x_1 and x_2. This percentage is not statistically different from what would be expected by chance (5%).[13] Consequently, the model is consistent with the data generating function and the multicollinearity in the data does not affect the likelihood of finding a significant moderating effect when none is in the data. These results are the same when the data generating function does not include a constant effect for x_2, regardless of whether this constant effect is estimated or not (the two other groups of bars (B) and (C) respectively in Figure 2).

So, there is no evidence that the product term picks up part of the variance that would be explained by the constant effects of the focal or the moderator variable. Thus, researchers can be reassured that they are unlikely to report moderating effects when there are none.

Does the specification of a product term affect the estimates of the effects?

This question can be addressed by comparing the parameter estimates from a fully specified model including an irrelevant moderating effect and a constant-only-effect model that does not specify an irrelevant moderating effect. Again, as we are looking at cases where the data generating function does not include a moderating effect, we only report the results of the OLS estimates. The summary statistics of these parameter estimates are provided in Figure 2 (fully specified model A) and Figure 3 (constant-only-effects model). In Figure 3, each group of bars from left to right corresponds to the parameters β_0, β_1 and β_2 (the bars within each group correspond to each level of the correlation between x_1 and x_2). Figure 3 shows that most of the coefficients are significant (more than 90% of them are significant for both β_1 and β_2 in all but the highest correlation level of 0.8, in which case more than 80% are still significant). Comparing these significance levels with those of the fully specified model

[13] The critical value at $\alpha = .05$ is 5.44%.

(group A in Figure 2), it becomes apparent that the significance level is reduced dramatically from 90% to 30%. This reduction in significance levels of the constant effects is simply caused by including an irrelevant product term in the model (group A).

This demonstrates the tremendous amount of collinearity inherent in a model specification with a product term. Even though it does not affect the likelihood of detecting moderating effects beyond chance levels, specifying a product term when it is not needed substantially decreases the likelihood of detecting significant constant effects.

Consequently, we recommend performing the estimation of a model with potential moderating effects in two steps. The first step is to estimate the full model. Then, if the moderating effect is insignificant, a second step consists of re-estimating a model without the product term to estimate the constant effects of the two variables. These two steps correspond to the hierarchical testing procedure. However, if the coefficient of the product term is significant, coefficients of a model without a product term should not be interpreted, as they are biased due to model misspecification.

What is the impact of specifying the moderator variable as a constant effect in addition to the moderating effect?

The impact of model-inherent multicollinearity on the ability to find significant effects can also be analyzed in another way. As demonstrated, estimating a moderating effect when there is none in the data drastically diminishes the ability to find the constant effects. This is due to the introduction of a product term that is structurally correlated with its components. The same reasoning should hold for including irrelevant constant effects, too. Estimating a fully specified model (i.e., constant effect for x_1 and x_2 and a moderating effect of x_2 on x_1) on data that contains only a constant effect for x_1 but no constant effect for x_2 should lead to model-inherent collinearity. That is, specifying such an irrelevant constant effect for x_2 should also dramatically reduce the ability to find a significant constant effect for x_1.

In Figure 2, the two groups of bars (B) and (C) on the right side display significance levels for estimated regressions on data generated only with a constant effect for x_1. The middle group of bars (B) displays significance levels of an estimated regression without a constant effect of x_2. The rightmost group of bars (C) displays significance levels of an estimated regression with a constant effect of x_2.[14] If x_2 is not specified as having a constant effect, the constant effect of x_1 (and the intercept as well) is highly likely to be significant (although somewhat less so as the correlation between x_1.and x_2 increases, indicative of some increased collinearity of the data). However, when x_2 is specified as having a constant effect as well as a moderating effect, the percentage level of significant constant effects of x_1 is again considerably reduced from the 60%-90% range (depending on the correlation, as shown in Figure 2, group B) to the 30% level (group C in Figure 2).

So far, we have considered data that did not contain moderating effects. We now examine the same question when moderating effects *do* exist in the data. As GLS provides more efficient estimates than OLS when moderating effects are present (as demonstrated in section 3), we now report the results in terms of SMR-GLS estimates.

Figure 4 graphs the percentage of significant moderating effects and those of the other parameters. The graph shows that the percentage of significant moderating parameters β_3 increases, as one would expect, with the size of the true moderating effect (which varies from 0 to 1). However, the likelihood of finding a significant constant effect is much lower than the likelihood of finding a significant moderating effect (even though the true constant effect is greater than or equal to the moderating coefficient). In fact, these likelihoods are

[14] We do not consider the case where x_2 is in the data generating function but is not in the estimated model, as this is the typical problem of misspecification bias that has been analyzed in the econometrics literature.

extremely low, around the 10% level for any level of moderating effect, which is not far from the 5% that would be obtained just by chance.

Regarding the impact of including irrelevant constant effects, the analysis confirms our conclusion above. When a moderating effect exists without a constant effect for x_2, the moderating effect is not difficult to detect. The percentages are close to 100%, as can be seen from the middle group of bars (B) in Figure 4. However, model-inherent multicollinearity problems arise when the estimated equation contains a constant effect for x_2 (right-side group (C) of graphs in Figure 4). Similarly, the constant effect of x_1 is also better detected with 20% level in group (B) if the irrelevant x_2 variable is omitted vs. 10% level in group (C). So, the inclusion of irrelevant effects has the same consequences with or without moderating effects present in the data. Including an irrelevant constant effect for x_2 dramatically reduces the ability to find a significant effect for x_1 and the moderating effect. And, there is again no evidence that the constant effect of x_2 picks up any product term effect, regardless of the size of the moderating effect.

These results (with and without moderating effects) demonstrate the drastic impact of specifying irrelevant factors on the ability to find significant relevant effects. As the unnecessary inclusion of a moderating effect dramatically reduces the ability to find significant constant effects, so the unnecessary inclusion of a constant effect for the moderator variable reduces the ability to find a significant constant effect of the focal variable.

Because the unnecessary inclusion of a constant effect of the moderator variable yields these drastic consequences, it is unwise to automatically specify a constant effect of the presumed moderator variable in addition to the moderating effect being tested. Such automatic full-specification has been recommended based on interpretation of interaction term arguments (Cohen 1978; Irwin and McClelland 2001) and on the necessary condition of nested models required for hierarchical testing procedures (McClelland and Judd 1993). However, Equations (2) and (3) are clear indications that the coefficients are all interpretable in the absence of a constant effect of the moderator variable, and the residuals sum of squares of nested models (with and without a product term) can be compared without any problem (as well as other fit statistics based on the residuals sum of squares). So, it is not so much in terms of interpretability that one finds justification for automatic full-model specification but more appropriately in the misspecification-bias argument. If there is a constant effect for x_2 and no such variable is specified in the estimated model, estimated parameters will be biased because the missing variable will be correlated with the product term of the moderating effect, even in the case where x_1 and x_2 are uncorrelated.

So, it is not wise to automatically specify constant effects in addition to moderating effects if theory only justifies a moderating effect. Because of the impact an irrelevant factor has on the estimated relevant parameters, the absence of significance of a constant effect of x_2 cannot be used as a pre-test to decide whether or not to include x_2 (as shown above, any lack of significance could be due to multicollinearity). Consequently, reliance on theory (measurement and substantive) is essential to drive the decision to include a variable as a constant effect or not. Measurement theory applies through the scale of the focal variable (Irwin and McClelland 2001). If the focal variable is interval scaled instead of ratio scaled, it is only defined up to a constant which, multiplied by the moderator variable in the product term, generates a term including the moderator variable alone. Thus, the moderator variable must then be included as a control variable, although its interpretation is subject to caution as it can possibly include a scaling component.

While the directions of the effects obtained with our simulated data can be generalized, it should be pointed out that the extent to which the incorporation of irrelevant factors affects the estimated coefficients depends on the mean of the variables in the sample.

The closer the mean is to zero, the smaller the correlations between the product term and its components. Consequently, the impact of including irrelevant "constant" or moderating effects is less problematic when variables have zero means. This is good news indeed when the ratio scale measures have small means. However, this could also be interpreted as a justification for mean-centering the data (which transforms variables to a mean of zero). Unfortunately, this is not the case. Consider a regression equation with two mean-centered variables x_1 and x_2:

$$y = \alpha_0 + \alpha_1(x_1 - \bar{x}_1) + \alpha_2(x_2 - \bar{x}_2) + \alpha_3(x_1 - \bar{x}_1)(x_2 - \bar{x}_2) + \varepsilon \qquad (8)$$

which reduces to:

$$y = (\alpha_0 - \alpha_1\bar{x}_1 - \alpha_2\bar{x}_2 + \alpha_3\bar{x}_1\bar{x}_2) + (\alpha_1 - \alpha_3\bar{x}_2)x_1 + (\alpha_2 - \alpha_3\bar{x}_1)x_2 + \alpha_3 x_1 x_2 + \varepsilon \qquad (9)$$

The coefficient of x_2 contains two components, $\hat{\alpha}_2$ and $-\hat{\alpha}_3\bar{x}_1$, the latter component includes the mean of x_1. So the coefficient for x_2 obtained from mean-centered data is composed of the constant effect when x_1 is zero and a moderation effect when x_1 is at its mean. Insignificance of $\hat{\alpha}_2$ would merely indicate that the effect of x_2 is not statistically different from $-\hat{\alpha}_3\bar{x}_1$. So, insignificance of $\hat{\alpha}_2$ cannot be used to make any conclusion about the effect of x_2 because of the confounding of the constant effect of x_2 and a moderating effect at the mean value of the focal variable x_1. Another way to put it is that even if $\hat{\alpha}_2$ is not significant, the constant effect of x_2 is already included in the product of the mean-centered terms via the mean of x_1.

Hence, the conclusion about the existence of a constant effect in addition to the moderating effect should not be solely based on the insignificance of the corresponding coefficient, whether variables are mean-centered or not. Even though the coefficients obtained from mean-centered variables are less sensitive to the inclusion of an irrelevant constant effect of the moderating variable, their interpretation is questionable because they confound constant and moderating effects.

What is the impact of model-inherent collinearity compared to data-driven collinearity when moderating effects exist?

Contrary to prior convictions that a high correlation between x_1 and x_2 leads to (data-driven) multicollinearity which “may have an adverse effect on the estimated coefficients” (Morris, Sherman, and Mansfield 1986, p. 283), both Figure 2 and Figure 3 show that the negative impact of increasing correlations between x_1 and x_2 on the significance levels of the constant effects is rather small. In contrast, we have seen that the model-inherent collinearity introduced by specifying an unneeded product term (Figure 2 group A vs. Figure 3, and Figure 4) has a huge impact on the ability to make any inference about the constant effects. These results apply with and without a moderating effect in the data.

Figure 5, like Figure 2, graphs the percentage of significant coefficients as a function of the correlation between x_1 and x_2 (data contain moderating effects, aggregated across all levels of moderating effect sizes). The bars that correspond to the moderating effects are around the 5% level in Figure 2 (as discussed earlier) while somewhere around the 30% line in Figure 5 (group (A) and (C)), except when x_2 is not involved (the middle group of bars (B)), where it reaches the 90% level. Even in the case of extremely high correlation between x_1 and x_2 (i.e., 0.8), the percentage of significant moderating effects is above 80%. So, again the impact of data-driven collinearity on the ability to find significant constant effects is negligible compared to the impact of model-inherent collinearity.

One aspect of the graphs in Figure 5 is noteworthy, though. As noted by McClelland and Judd (1993), “increasing correlation between X and Z [here x_1 and x_2], with all else equal, improves the chances of detecting moderator effects,” (p. 380). This is indeed reflected in the results graphed in Figure 5 for the model including a constant effect of x_2, whether the data contain such an effect (group of bars (A) on the left side) or not (group of bars (C) on the right side). This effect of enhanced likelihood to detect moderating effects, however, is not very large, as it requires high correlations for no more than an average 5% difference. Nevertheless, it does correspond to the statistical explanations provided by McClelland and Judd (1993). However, it is intriguing that this effect is reversed if x_2 is not specified as a constant effect to be estimated (middle group of bars (B) in Figure 5). Furthermore, this reversal applies to all the coefficients (consistent with what happens in Figure 2 where there is no moderating effect in the data). This pattern is typical for data-driven multicollinearity, where increasing correlations decrease the percentage of significant coefficients. It appears that data-driven multicollinearity interacts with the particular form of model-inherent multicollinearity. This does not affect the moderating effect (McClelland and Judd 1993) but does impact the constant effects.

In summary, it is the model-inherent correlation between the product term and its components that dramatically decreases the significance of constant effects, although it affects much less the significance of moderating effects. The data-driven correlation between the focal variable and the moderator variable has differential consequences depending on the particular model specification. Data-driven collinearity slightly increases the significance of the moderating effect when this variable is also specified as having a constant effect (McClelland and Judd 1993). But it decreases the likelihood of finding significant moderating effects when x_2 is only specified as a moderating factor.

CONCLUSION

Theories in the social sciences, including marketing, increasingly involve sophisticated explanatory mechanisms that generate contingent predictions. In many cases, data are available in the quantity and quality that is necessary to test such predictions. Yet, the sophistication of econometric methods commonly employed to treat these data has not always been well matched to the task. We have proposed to use Stochastic Moderated Regression to estimate moderating effects and we have shown the conditions when it performs best.

We show that multicollinearity problems in moderated regressions arise not so much due to ill-conditioning of the data (especially the lack of independence between the focal and the moderator variables), but due to multicollinearity inherent in a model specification with a product term. This model-inherent multicollinearity has very little effect on moderating estimates. However, it has very strong effects on the ability to detect constant effects. Therefore, no simple rule (i.e., always include constant *and* product terms for all variables) can be used for model specification. Exploratory searches for moderating effects can be dangerous, as they are likely to dampen the significance of *all* parameter estimates, especially when a variable is included as a constant effect *and* as a moderating variable.

Stochastic Moderated Regression (SMR), which recognizes the stochastic nature of the moderating effect, can contribute to reduce the consequences of model-inherent multicollinearity. In contrast to OLS, GLS estimation of SMR makes use of the information about the error term structure, as implied by a conceptually appealing theory of moderating variables. Incorporating this information leads to more efficient estimation, which is badly needed for handling structurally introduced multicollinearity. These models have the promise of better detecting genuine moderating and constant effects, especially under the typical research conditions of small sample size and noisy data.

Table 1: **Generalized Mean Squared Errors of OLS and GLS-SMR Coefficients as a Function of the Noise in the Moderating Equation**

Coefficients' Distance from true Parameter (median)

Noise in Moderating Equation		β_0	β_1	β_2	β_3
0.4	OLS	423.81	4.413	4.362	0.0437
	SMR	*405.98*	*4.285*	*4.153*	*0.0427*
0.8	OLS	1347.79	14.058	13.904	0.1388
	SMR	*1247.73*	*13.238*	*13.043*	*0.1308*
1.2	OLS	2904.03	31.268	29.830	0.3124
	SMR	*2704.78*	*28.710*	*27.331*	*0.2850*
1.6	OLS	5089.42	52.356	52.737	0.5327
	SMR	*4777.03*	*49.489*	*49.156*	*0.4973*
2	OLS	7989.31	83.938	82.159	0.8433
	SMR	*7333.18*	*78.473*	*74.628*	*0.7680*

Variance (median)

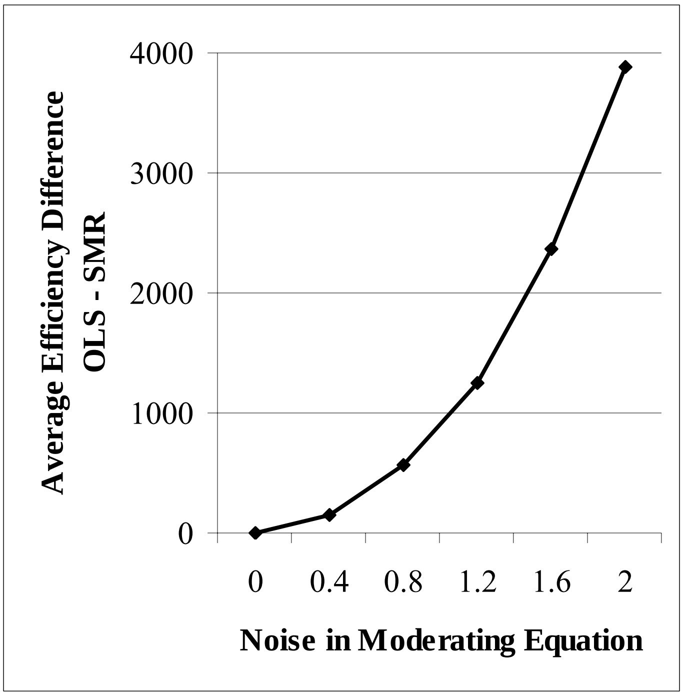
		β_0	β_1	β_2	β_3
0.4	OLS	848.20	8.571	8.551	0.0842
	SMR	*792.50*	*8.219*	*8.082*	*0.0817*
0.8	OLS	2627.22	26.390	26.464	0.2604
	SMR	*2435.47*	*25.197*	*24.805*	*0.2492*
1.2	OLS	5677.77	56.968	56.948	0.5617
	SMR	*5225.17*	*54.246*	*53.267*	*0.5398*
1.6	OLS	9867.41	99.257	99.316	0.9767
	SMR	*9066.89*	*94.245*	*92.476*	*0.9337*
2	OLS	15221.97	153.327	153.781	1.5051
	SMR	*13827.96*	*144.385*	*141.826*	*1.4344*

Generalized Mean Squared Error (median)

		β_0	β_1	β_2	β_3
0.4	OLS	1502.67	15.265	15.349	0.1518
	SMR	*1434.47*	*14.805*	*14.709*	*0.1477*
0.8	OLS	4739.46	48.312	48.590	0.4776
	SMR	*4393.04*	*45.784*	*45.092*	*0.4542*
1.2	OLS	10250.19	105.267	103.836	1.0362
	SMR	*9375.44*	*98.048*	*95.168*	*0.9705*
1.6	OLS	17454.35	179.321	176.909	1.7561
	SMR	*16187.78*	*169.078*	*165.470*	*1.6712*
2	OLS	27429.31	280.594	277.936	2.7651
	SMR	*25021.34*	*263.154*	*253.855*	*2.6036*

<u>Figure 1a:</u> Efficiency Difference between OLS and GLS-SMR Coefficients as a Function of Noise in Moderating Equation

(Positive values indicate the extent of increased efficiency of SMR vs. OLS)



<u>Figure 1b:</u> Efficiency Difference between OLS and GLS-SMR Coefficients as a Function of Sample Size

(Positive values indicate the extent of increased efficiency of SMR vs. OLS)

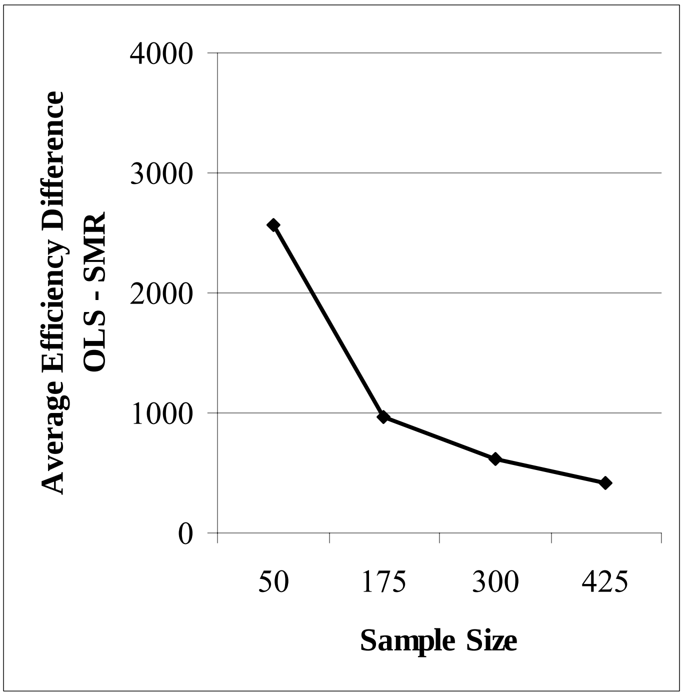


Figure 1c: **Average Percentage Improvement in Efficiency of SMR-GLS over OLS as a Function of Correlation between x_1 and x_2**

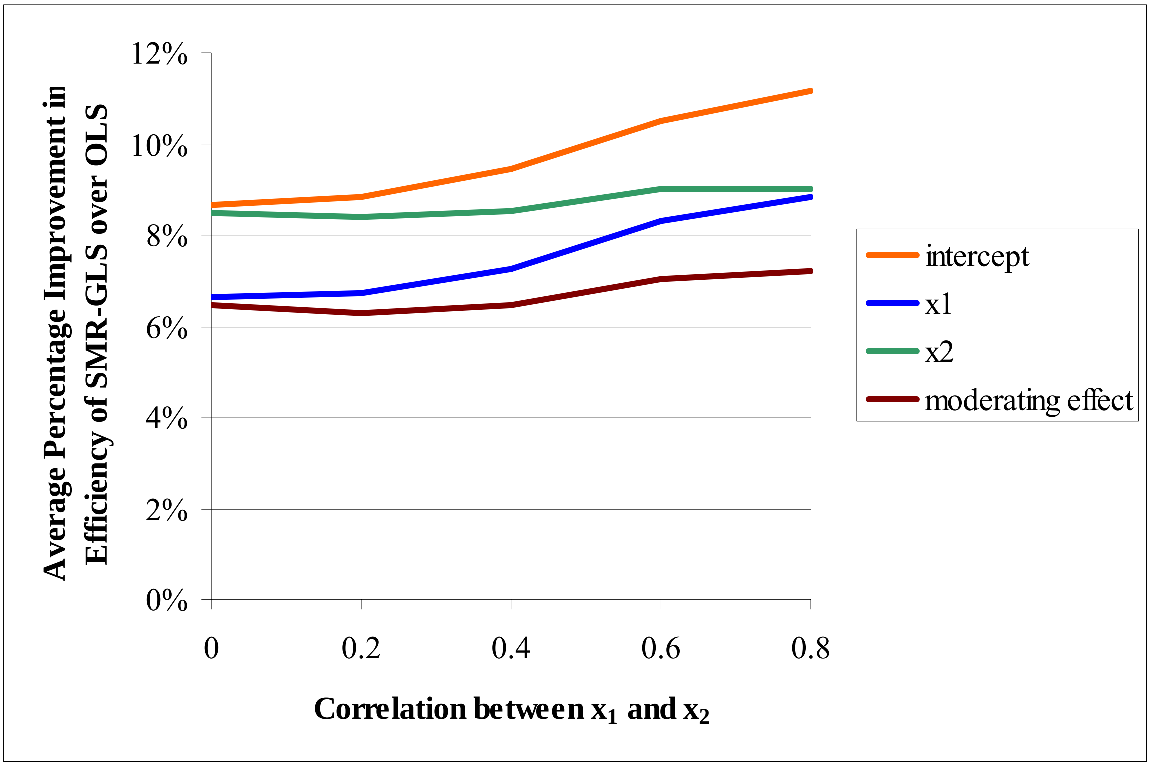


Figure 2: **Percentage of Significant Coefficients (OLS) as a Function of the Correlation between x_1 and x_2 when no Moderating Effect is Present**

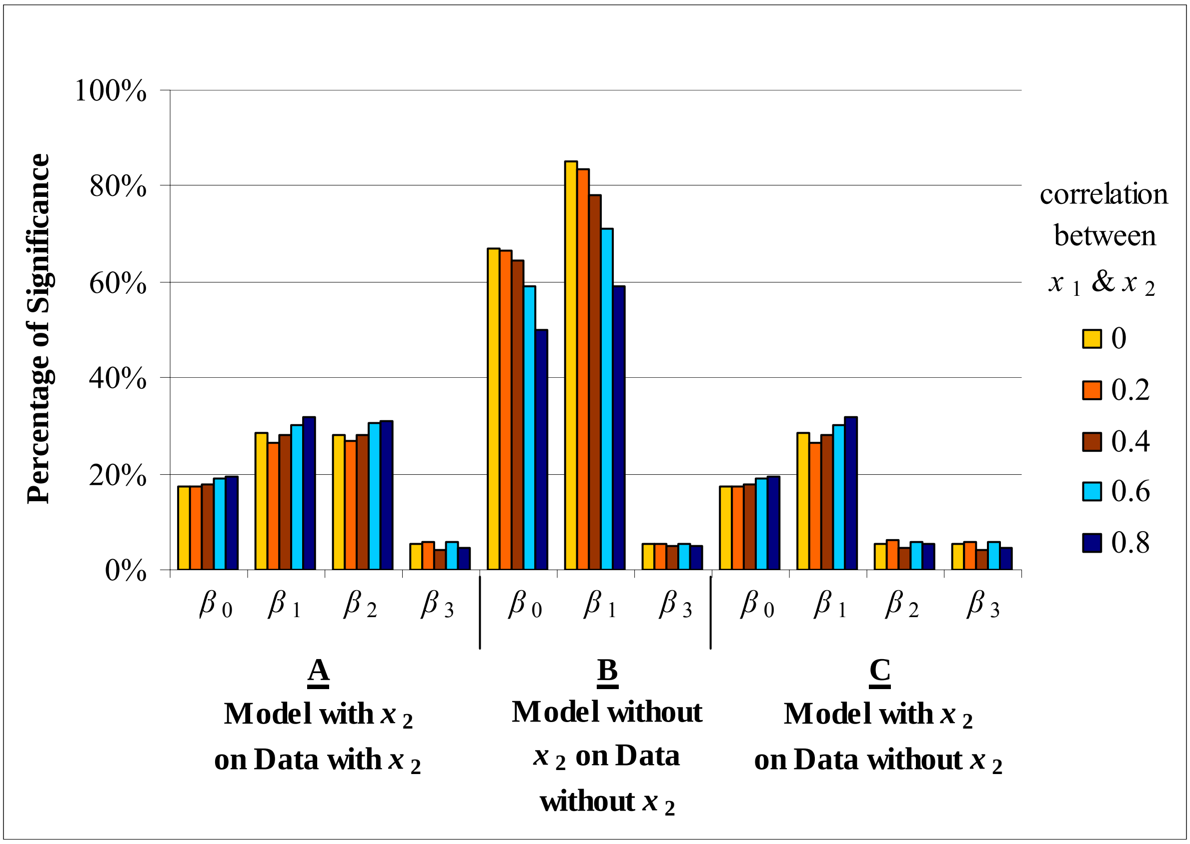


Figure 3: **Percentage of Significant Coefficients for Constant-Only-Effect Model as a Function of the Correlation between x_1 and x_2 when no Moderating Effect is Present**

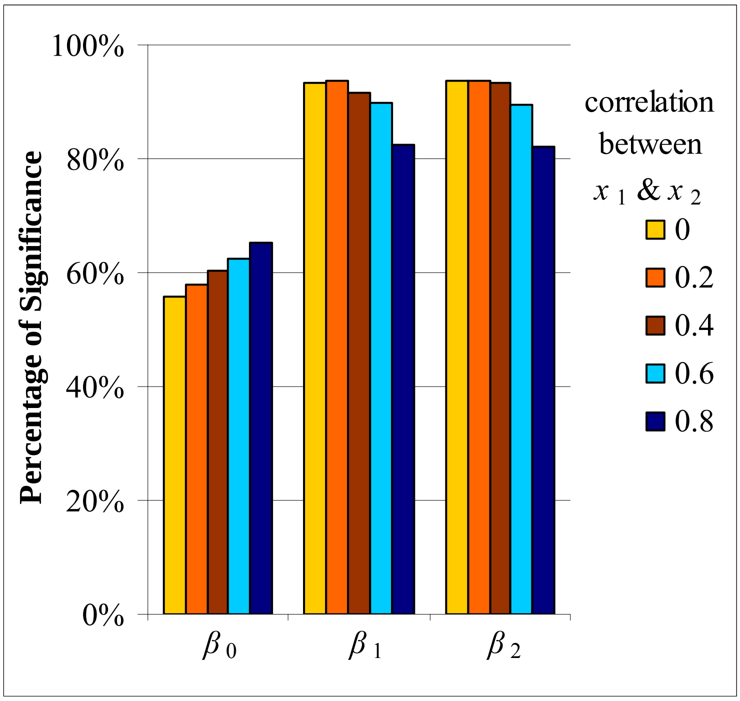


Figure 4: **Percentage of Significant Coefficients (SMR) as a Function of the True Moderating Effect**

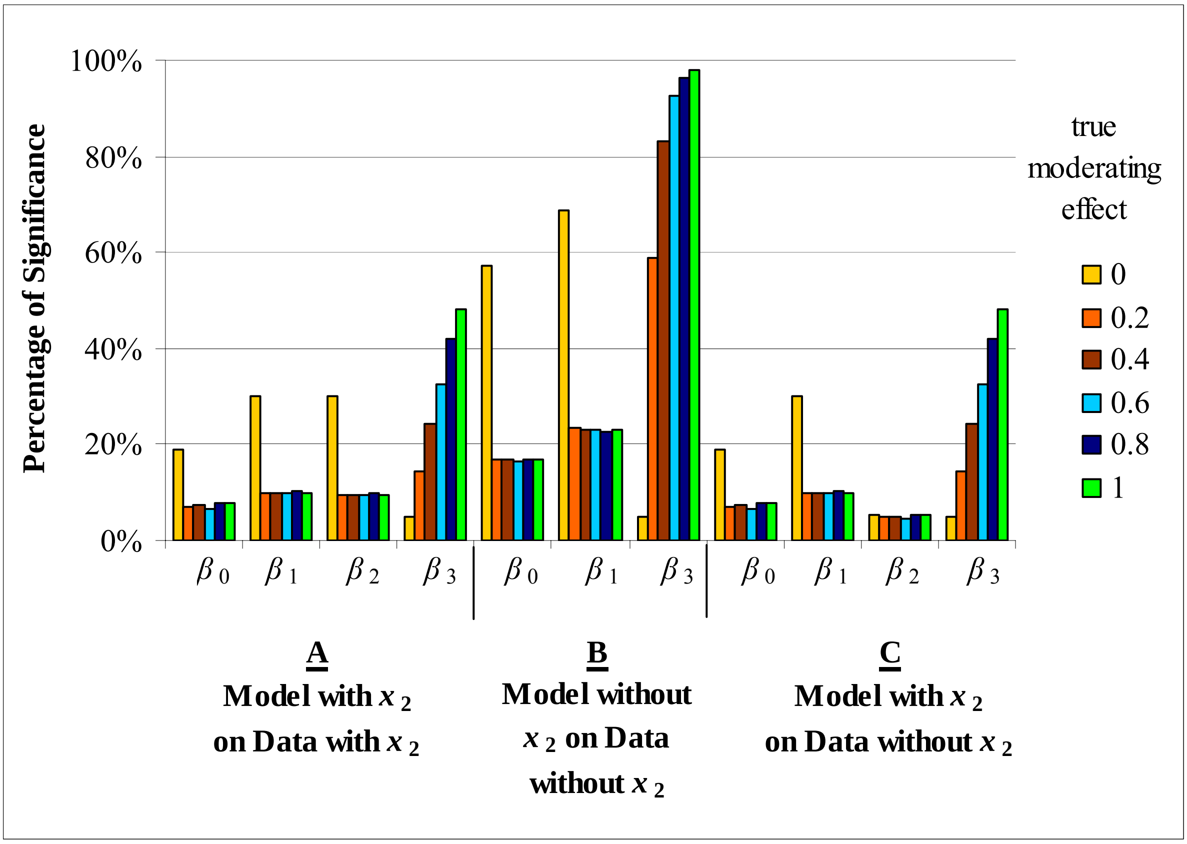
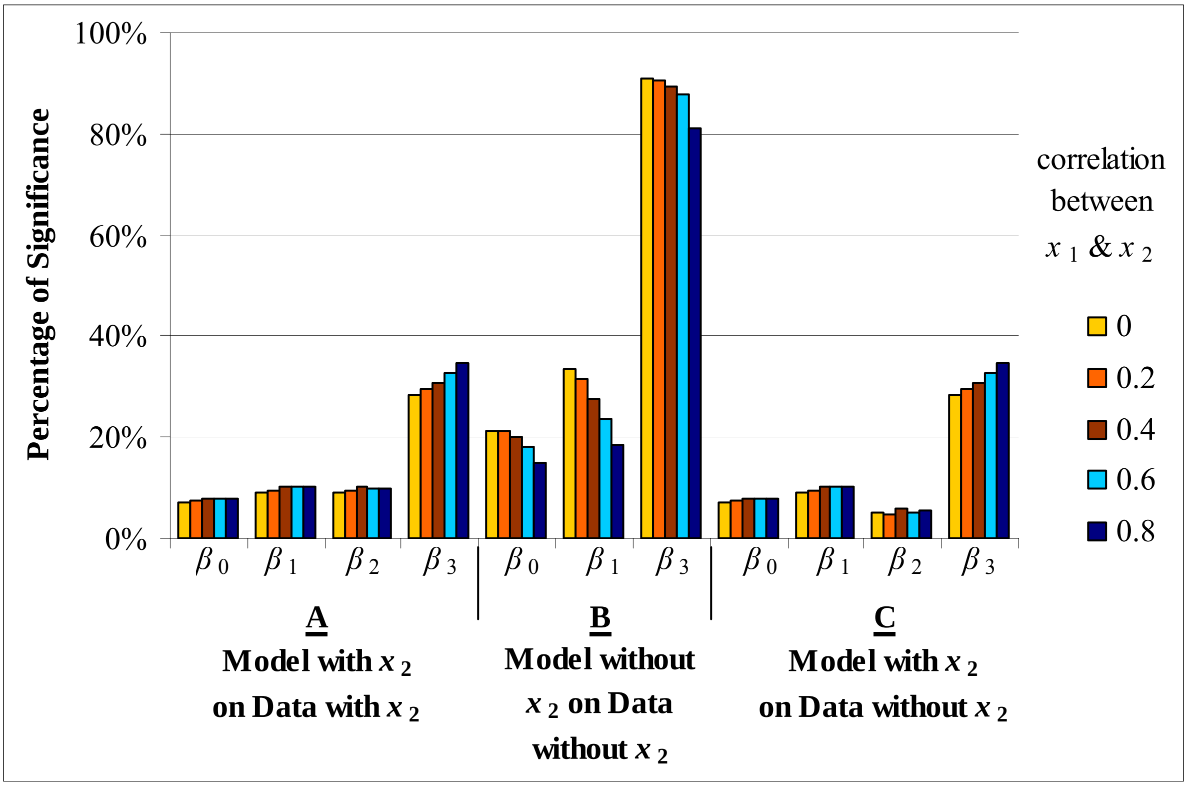


Figure 5: **Percentage of Significant Coefficients (SMR) as a Function of the Correlation between x_1 and x_2 in the Presence of Moderating Effects**



References

Aiken, Leona S., and Stephen G. West (1991), *Multiple Regression: Testing and Interpreting Interactions.* Thousand Oaks, CA, US: Sage Publications

Belsley, David A. (1984), "Demeaning Conditioning Diagnostics Through Centering," *The American Statistician*, 38, 2, 73-77.

Bowman, Douglas, and Das Narayandas (2004), "Linking Customer Management Effort to Customer Profitability in Business Markets," *Journal of Marketing Research*, 41 (4), 433-447.

Cohen, Jacob (1978), "Partialed Products are Interactions; Partialed Powers are Curve Components," *Psychological Bulletin*, 85, 4, 858-66

---- and Patricia Cohen (1983), *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences.* (2nd ed.), NJ, Hillsdale: Lawrence Earlbaum Associates

Cronbach, Lee J. (1987), "Statistical Tests for Moderator Variables: Flaws in Analyses Recently Proposed," *Psychological Bulletin*, 102 (3), 414-17.

Dunlap, William P. and Edward R. Kemery (1987), "Failure to Detect Moderating Effects: Is Multicollinearity the Problem?," *Psychological Bulletin*, 102 (3), 418-20.

Ganzach (1997), "Misleading Interaction and Curvilinear Terms." *Psychological Methods*, 2 (3), 235-47.

Gatignon, Hubert (1984), "Competition as a Moderator of the Effect of Advertising on Sales," *Journal of Marketing Research*, 21, 4 (November), 387-98.

---- (1993), "Marketing Mix Models," in Jehoshua Eliashberg and Gary L. Lilien, Eds., *Handbooks in Operations Research and Management Science*, Amsterdam, The Netherlands: Elsevier Science Publishers B.V.

---- and Dominique M. Hanssens (1987), "Modeling Marketing Interactions with Application to Salesforce Effectiveness," *Journal of Marketing Research*, 24 ,3 (August), 247-57.

Hildreth, Clifford and James P. Houck (1968), "Some Estimators for a Linear Model with Random Coefficients," *Journal of the American Statistical Association*, 63, 584-95.

Irwin, Julie R. (2001), "Methodological and Statistical Concerns of the Experimental Behavioral Researcher," *Journal of Consumer Psychology*, 10 (1&2), 97-98.

---- and Gary H. McClelland (2001), "Misleading Heuristics and Moderated Multiple Regression Models," *Journal of Marketing Research*, 38 (1), 100.

Jaccard, James, Robert Turrisi, and Choi K. Wan (1990), *Interaction Effects in Multiple Regression*. Newbury Park, CA: SAGE Publications, Inc.

----, Choi K. Wan, and Robert Turrisi (1990), "The Detection and Interpretation of Interaction Effects between Continuous Variables in Multiple Regression," *Multivariate Behavioral Research*, 25, 4, 467-78.

Judge, George G., W. E. Griffiths, R. Carter Hill, Helmut Lutkepohl, and Tsoung-Chao Lee (1985), *The Theory and Practice of Econometrics*. New York, NY: John Wiley and Sons.

Kromrey, Jeffrey D. and Lynn Foster-Johnson (1998), "Mean Centering in Moderated Multiple Regression: Much Ado About Nothing," *Educational and Psychological Measurement*, 58, 1, 42-67.

Leamer, Edward E. (1978), *Specification Searches*. New York: Wiley.

Lehmann, Donald R (2001), "Methodological and Statistical Concerns of the Experimental Behavioral Researcher," *Journal of Consumer Psychology*, 10 (1&2), 90-91.

McClelland, Gary H. and Charles M. Judd (1993), "Statistical Difficulties of Detecting Interactions and Moderator Effects," *Psychological Bulletin*, 114 (2), 376-90.

Morris, James H., J. Daniel Sherman, and Edward R. Mansfield (1986), "Failures to Detect Moderating Effects with Ordinary Least Squares-Moderated Multiple Regression: Some Reasons and Remedy," *Psychological Bulletin*, 99 (2), 282-88.

Saunders, David R. (1956), "Moderator Variables in Prediction," *Educational and Psychological Measurement*, 16, 209-22.

Sharma, Subhash, Richard M. Durand, and Oded Gur-Arie (1981), "Identification and Analysis of Moderator Variables," *Journal of Marketing Research*, 18 (August), 291-300.

Smith, Kent W. and M. S. Sasaki (1979), "Decreasing Multicollinearity: a Method for Models with Multiplicative Functions," *Sociological Methods and Research*, 8 (1), 35-56.

Wildt, Albert R. and Russ S. Winer (1983), "Modeling and Estimation in Changing Market Environments, *Journal of Business*, 56 (July), 365-388.