This Manuscript is under Consideration for Publication in:
Methodology: European Journal of Research Methods for the Behavioral and Social Sciences

New Insights Into Two-predictor Suppression Effects by Relationship Simulation and Examining Three-dimensional Scatterplots

Morteza Nazifi

ORCiD: 0000-0002-0155-6743

University of Bojnord, Bojnord, North Khorasan Province, Iran

Postal Code: 9453155111

Corresponding Author e-mail: nazifi@ub.ac.ir

Hamid Fadishei

ORCiD: 0000-0002-6207-369X

University of Bojnord, Bojnord, North Khorasan Province, Iran

Postal Code: 9453155111

e-mail: fadishei@ub.ac.ir

This Manuscript is under Consideration for Publication in:
Methodology: European Journal of Research Methods for the Behavioral and Social Sciences

Author's Footnote

Morteza Nazifi is Ph.D. in Psychology, University of Bojnord, Bojnord, North Khorasan, Iran (e-mail: nazifi@ub.ac.ir); and Hamid Fadishei is Ph.D. in Computer Engineering, University of Bojnord, Bojnord, North Khorasan, Iran (e-mail: fadishei@ub.ac.ir). There is no funding agency and there are no relevant financial or non-financial competing interests to report. The authors thank the referees for helpful comments.

This Manuscript is under Consideration for Publication in:
Methodology: European Journal of Research Methods for the Behavioral and Social Sciences

Abstract

Two-predictor suppression situations continue to produce uninterpretable conditions in linear regression. This study introduces a software called Supsim that facilitates the study of suppression situations by randomly generating normally distributed vectors x_1, x_2, and y in a way that regressing y on both x_1 and x_2 creates numerous random two-predictor models (RTM's) with or without a suppressor variable effect. This study also introduces another tool called Supcalc that receives arbitrary r_{y1}, r_{y2} and n as inputs and calculates R^2, $\hat{\beta}_1$, $\hat{\beta}_2$ and their standard errors, semipartial and squared semipartial correlations, multicollinearity limits and ratios, and novel indexes of statistical control for two-predictor regression models. The Supsim provides users with 3D scatterplots of numerous RTM's with or without suppression. Comparing 3D scatterplots with regression surfaces of different suppression and non-suppression situations, this study provides important new insights into two-predictor suppression effects that help resolve the conceptual complexities of these situations. An important focus is to compare 3D scatterplots of some special enhancement situations, called Hamilton's extreme example, with those of redundancy situations. Such a comparison suggests that the basic mathematical concepts of two-predictor suppression situations need to be reconsidered in terms of statistical control function.

Keywords: Multicollinearity, Suppressor Variable, Hamilton's Extreme Example, Statistical Control Function, Relationship Simulation

This Manuscript is under Consideration for Publication in:
Methodology: European Journal of Research Methods for the Behavioral and Social Sciences

1. INTRODUCTION

1.1 The Challenge of Interpreting Two-predictor Suppression Situations

Two-predictor suppression effects remain among complex and confusing situations in linear regression research. It seems the confusion, controversy and dismay that surround the definition, explanation and interpretation of suppression situations (e.g., McFatter 1979; Holling 1983; Ludlow and Klein 2014) has been caused by impossibilities or contradictory results. When the inclusion of a second predictor say x_2 which is relatively highly correlated with x_1, in the regression equation leads to some kind of two-predictor suppression situation, possible contradictory results include: calculating a negative part of variance explained when partitioning R^2 (Cohen et al. 2003), finding opposite signs between the second predictor's zero-order correlation with y and its regression coefficient in the equation, observing situations in which one of the two predictors or both of them get a large regression coefficient in the equation despite showing “no or low” zero-order correlation with y, and finally finding situations in which $R^2 > r_{y1}^2 + r_{y2}^2$ (Hamilton 1987). Under the condition of $R^2 > r_{y1}^2 + r_{y2}^2$, Hamilton (1987) describes an even more challenging two-predictor suppression situation in which r_{y1} and r_{y2} are both close to 0 but R^2 and $|r_{12}|$ are both near 1. Given that research on these challenging two-predictor suppression effects requires access to some computerized simulation algorithm that can generate all these different situations, the authors develop and introduce a computerized algorithm called Supsim, a specialized software made available both as an open-source, command-line Python package (visit https://pypi.org/project/supsim/ for installing the package) and as a web-based JavaScript software (find screenshots from the user-interface of the web-based Supsim in panel B of Figure 1; also visit https://supsim.netlify.app/supsim to start working with Supsim). This algorithm enables researchers to easily generate numerous series of random data vectors x_1, x_2, and y so that one can create numerous situations with or without suppression by regressing *y* on both x_1 and x_2. The web-based Supsim also allows investigators to produce 3D scatterplots of these simulated random two-predictor models (RTM's). Before proceeding, a comprehensive definition of two-predictor suppression effects is needed to be used as a frame of reference.

1.2 Attempts to Provide a Comprehensive Definition of Various Suppression Situations

Friedman and Wall's interesting work (Friedman and Wall 2005) on two-predictor suppression effects is a promising approach that incorporates different definitions of suppression situations that has been presented so far. Holding arbitrary selected r_{y1} and r_{y2} constant and letting r_{12} vary over its possible limit (see inequality (1) below), Friedman and Wall (2005) show that four different regions with or without suppression can occur in a graph in which each region corresponds to some suppression or non-suppression situations defined previously by other leading researchers in this field. According to Friedman and Wall (2005), under conditions where r_{y1} and r_{y2} *both have positive signs,* and $r_{y1} > r_{y2}$, as it is common in the linear regression research, the regions on the graph, from left to right, include:

First, the "Region I: enhancement" (Friedman and Wall 2005) which is present under the following conditions:

- All r_{12}'s < 0
- $|\hat{\beta}_1| > |r_{y1}|$
- $R^2 > r_{y1}^2 + r_{y2}^2$
- And the signs of $\hat{\beta}_1$ and $\hat{\beta}_2$ are always similar to the signs of r_{y1} and r_{y2}, respectively.

Friedman and Wall's definition of enhancement in region I is equivalent to the definitions offered by Horst (1941) and Lynn (2003) for "classical suppression" (only when $r_{y2} = 0$), Conger (1974) for "reciprocal suppression", Cohen and Cohen (1975) and Lynn (2003) for "cooperative suppression", Currie and Korabinski (1984) , and Friedman and Wall (2005) for "enhancement", Shieh (2001) for "enhancement-synergism", Sharpe and Roberts (1997), and Velicer (1978) for "suppression", and finally Hamilton (1987) for "synergism";

Second, the "Region II: redundancy" (Friedman and Wall 2005) as an ordinary situation without suppression in which the following conditions hold:

- $0 \leq r_{12} \leq \frac{r_{y2}}{r_{y1}}$ (the ratio $\frac{r_{y2}}{r_{y1}}$ also has been called γ in the previous literature)
- $|\hat{\beta}_1| \leq |r_{y1}|$
- $R^2 \leq r_{y1}^2 + r_{y2}^2$

The redundancy region on Friedman and Wall's graph corresponds to Cohen and Cohen's (1975), Currie and Korabinski's (1984), Friedman and Wall's (2005), and finally Velicer's (1978) definitions of an ordinary situation without suppression;

Third, the "Region III: suppression" (Friedman and Wall 2005) in which the following conditions hold:

- $\gamma < r_{12} \leq \frac{2\,(r_{y1} \times r_{y2})}{r_{y1}^2 + r_{y2}^2}$ (the ratio $\frac{2\,(r_{y1} \times r_{y2})}{r_{y1}^2 + r_{y2}^2} = \frac{2\gamma}{1+\gamma^2}$)
- $|\hat{\beta}_1| > |r_{y1}|$
- $R^2 \leq r_{y1}^2 + r_{y2}^2$
- And in which r_{y2} and $\hat{\beta}_2$ are always of opposite signs.

The definition of the region III suppression on Friedman and Wall's graph is consistent with the definitions suggested by Darlington (1968) for "negative suppression", Cohen and Cohen (1975) and Currie and Korabinski (1984) for "net suppression", and finally Conger (1974), and Friedman and Wall (2005) for "suppression";

And finally, the "Region IV: enhancement" (Friedman and Wall 2005) in which the following conditions hold:

- all $r_{12}'s > \frac{2\gamma}{1+\gamma^2}$
- $|\hat{\beta}_1| > |r_{y1}|$
- $R^2 > r_{y1}^2 + r_{y2}^2$
- And in which r_{y2} and $\hat{\beta}_2$ are always of opposite signs.

Friedman and Wall's "region IV enhancement" is equivalent to the definitions suggested by Horst (1941) and Lynn (2003) for "classical suppression" (i.e., when r_{y2} = 0), Currie and Korabinski (1984) , and Friedman and Wall (2005) for "enhancement", Shieh (2001) for "enhancement-synergism", Darlington (1968) for "negative suppression", Cohen and Cohen (1975) for "net suppression", Conger (1974), Lynn (2003), Sharpe and Roberts (1997) and Velicer (1978) for "suppression", and finally Hamilton (1987) for "synergy".

It should be noted that in Friedman and Wall's graphs, when r_{y1} and r_{y2} *are of opposite signs*, the order of the regions described above becomes reverse. When reverse graph holds, from left to right, there are region IV (enhancement), region III (suppression), region II (redundancy), and region I (enhancement). When the latter is the case, region I covers any positive values of r_{12} (all r_{12}'s > 0), and regions II, III, and IV all are shifted to the negative side of the r_{12} axis. In addition, when r_{y2} = 0, a situation called "classical suppression", Friedman and Wall's graph has only two regions including, from left to right, region I enhancement, and region IV enhancement (Visit Friedman and Wall's online application https://steamtraen.shinyapps.io/suppressiongraphics/ to create the graphs).

Friedman and Wall (2005) believe that in order to get an accurate picture of two-predictor suppression effects each fixed pairs of r_{y1} and r_{y2} should be considered separately allowing r_{12} vary over its possible limit. They state that it is not the r_{12} per se but the combination of the three correlations that affect the sign change in $\hat{\beta}_2$. The possibility limit of r_{12}, when r_{y1} and r_{y2} are given, is defined by the following inequality (e.g., Neill 1973; Sharpe and Roberts 1997):

$$r_{y1} \times r_{y2} - \sqrt{(1 - r_{y1}^2)\,(1 - r_{y2}^2)} \le r_{12} \le r_{y1} \times r_{y2} + \sqrt{(1 - r_{y1}^2)\,(1 - r_{y2}^2)} \quad (1)$$

The limits imposed by the fact that the correlation matrix which r_{y1}, r_{y2}, and r_{12} come from must be nonnegative, definite (Neill 1973; Sharpe and Roberts 1997; Friedman and Wall 2005). The limits defined by inequality (1) imply that the possible interval of r_{12} can become very wide when both $|r_{y1}|$ and $|r_{y2}|$are close to 0 and it can also become very narrow when both $|r_{y1}|$ and $|r_{y2}|$ are near 1. Another important insight here is that the possible interval of r_{12} produced by non-negative definiteness restriction, also limits the value of R^2 between 0 and 1 while without limiting r_{12}, the R^2 values turn out to be far greater than 1 when large $|r_{12}|$ values beyond the allowed boundaries are used (see tables 1 through 3). Concentrating on the possibility interval of r_{12} is extremely important in understanding two-predictor suppression effects, because formulas of both R^2 and $\hat{\beta}_2$ (and $\hat{\beta}_1$ as well) are sensitive to the values of r_{12} as it is evident from formula (2) (Cohen et al. 2003) and formula (3) (Hamilton 1987; Cohen et al. 2003) below:

$$\hat{\beta}_2 = \frac{r_{y2} - r_{y1} r_{12}}{1 - r_{12}^2} \quad (2)$$

$$R^2 = \frac{r_{y1}^2 + r_{y2}^2 - 2 r_{y1} r_{y2} r_{12}}{1 - r_{12}^2} \quad (3)$$

As mentioned above, in Friedman and Wall's (2005) approach a fixed pair of r_{y1} and r_{y2} is selected arbitrarily to see what happens to the regression coefficients and R^2 values when r_{12}

vary over its possible limit. This approach beside its strengths has an important limitation because in this way one is completely unaware of the raw data and the 3D scatterplots related to each particular regression model. Hamilton (1987) does explain a method for generating artificial data vectors x_1, x_2, and y that are used in building regression models in which $R^2 > r_{y1}^2 + r_{y2}^2$, but he uses the data vectors x_1, x_2, and y only in drawing two-dimensional scatterplots and fails to explore 3D scatterplots of the resulting two-predictor models. The authors believe comparing 3D scatterplots of two-predictor regression models with or without suppression bear important new insights into the effects of multicollinearity on the results of linear regression models. In addition, little attention has been paid to the mechanisms of statistical control in redundancy situations compared to suppression situations in the previous research.

2. MATERIALS AND METHODS

2.1 Relationship Simulation or RTM Generation Algorithm

The idea of RTM generation algorithm is to facilitate the study of two-predictor suppression effects by generating numerous random functions (i.e., $y_o = f(x_1, x_2)$) and inserting errors into the outputs of those functions and then fitting an OLS regression surface to the resulting noisy data (y) to produce and classify numerous random, two-predictor models that some are affected by a suppressor variable while some are not. The proposed algorithm is illustrated by panel A of Figure (1). This iterative process starts by choosing two random vectors x_1 and x_2 so that the x_1 and x_2 show a specific amount of correlation with each other (r_{12}). Next, a random function is generated to produce y_o as a function of x_1 and x_2 and then a normally distributed noise e is added to y_o vector in order to generate a noisy data vector y (i.e., $y = y_o + e$). The distribution of the noise $e = N(\mu_e, \sigma_e)$ is controlled by the user before running the algorithm through selecting an A coefficient where $\mu_e = A\mu_{y_o}$ and $\sigma_e = A\sigma_{y_o}$.

User-provided constraints in this process, which all are inserted before running the algorithm, are imposed upon r_{y1}, r_{y2}, r_{12}, and R^2 enhancement. Otherwise all the required constraints are met, the current RTM shall be discarded and the current iteration shall be started again. When designing algorithm of Supsim, the authors noticed that one of the challenging constraints to meet was the specified amount of correlation between x_1 and x_2. The authors observed that satisfying this constraint requires an exhaustive search over a very large space of all possible RTMs which is not feasible in a reasonable time. In order to overcome this limitation and speed up the simulation, a random number generation method, suggested by Whuber (2017), is used that can generate numbers that are random, and at the same time, show a specific amount of correlation with another variable (Whuber 2017; for more details also visit the Supsim project website at https://supsim.netlify.app/). The next section explains the method for randomly generating correlated, normally distributed vectors x_1 and x_2.

2.2 The Method for Randomly Generating Correlated, Normally Distributed Vectors x_1 and x_2

According to Whuber's method (2017), the algorithm shown in panel A of Figure 1 first randomly chooses a normal vector x_1 and then chooses another normal, random vector *a* with the same length, mean, and standard deviation as x_1. Next it applies a transformation to *a* to calculate b in a way that the correlation between b and x_1 is set to the desired amount (r). Such a transformation is described in Equation (4) where *d* is the vector of residuals resulted from regressing a on x_1, σ_d represents the standard deviation of d, and σ_{x_1} represents the standard deviation of x_1. It should be noted that such a transformation changes the initial distribution properties in b vector. Therefore, in order to return b to a mean and a standard deviation similar to x_1, $x_2 = mb + n$ is used as the final random, correlated, normal vector, where $m = \sigma_{x_1} / \sigma_b$ and $n = \mu_{x_1} - m.\mu_b$.

$$b = r.\sigma_d.x_1 + d.\sigma_{x_1}.\sqrt{1-r^2} \tag{4}$$

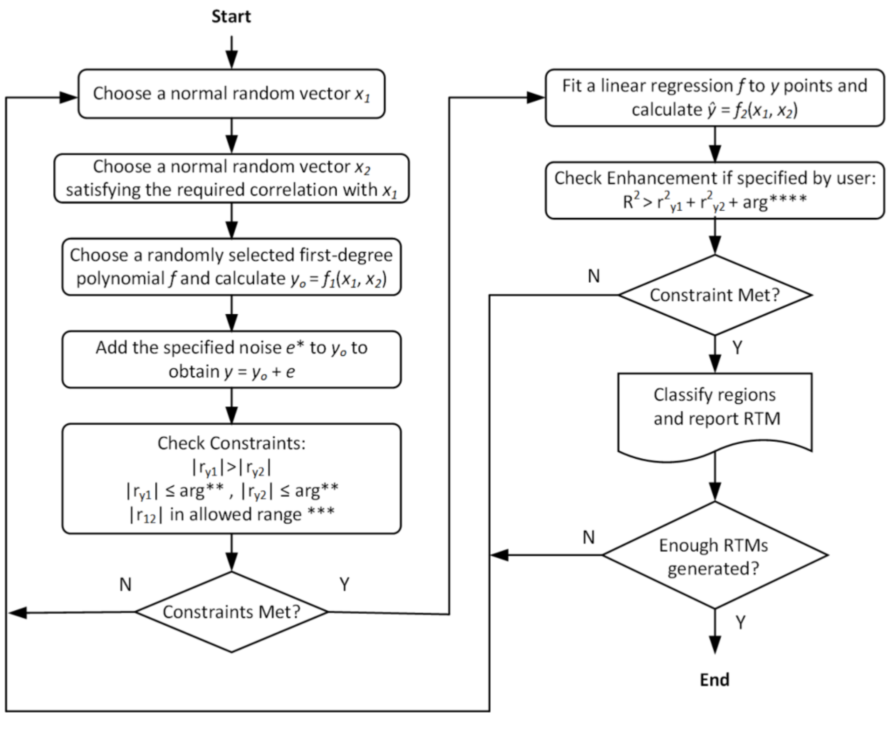
3. RESULTS

3.1 The Distribution of Large-Scale Samples of RTM's among Regions of Friedman and Wall's Graph

Random production of RTM's enables them to freely scatter among the regions of Friedman and Wall's graph so that it is possible to determine the probability of RTM's falling within each region (see Figure 2, panels A through C). With noise magnitude = 0.05, the distribution of a large-scale sample of RTM's (*N* = 10,000) is shown on Figure 2. The results show that the regular graph is filled by 49.55% of RTM's (see Figure 2, panel A), and another 49.7% of RTM's are scattered among four regions of the reverse graph (see Figure 2, panel B), and only 0.38% of RTM's are fallen within the two regions of the classical suppression graph (see Figure 2, panel C). Figure 2 also shows that the redundancy regions on both regular and reverse graphs are most likely to be filled by RTM's so that the total probability of "redundancy regions" is equal to 50.62%. The total probability of "region I enhancement" on all the three graphs is equal to 24.36%, and the total probability of "region IV enhancement" on all the three graphs is equal to 9.8%. And finally, the total probability of the "region III suppression" on both regular and reverse graphs is equal to 14.89%. An important point here is that for all random RTM's shown in Figure 2, when the noise magnitudes are close to 0, fit levels are near 1, as it is evident from R^2 values on all the three graphs (see Figure 2, panels A through C).

To further investigate the effects of noise on the way in which RTM's scatter among four regions of Friedman and Wall's graph, six more random large-scale samples of RTM's (all *N*'s = 10,000) were taken with different noise magnitudes including 0.2, 0.7, 1.3, 2.0, 2.7, and

3.0 to estimate the probabilities of each of the four regions being filled by RTM's as the noise magnitude is gradually grows to 3.0. Results showed that despite the increased noise magnitudes the above mentioned probabilities almost remain constant. However, as it is expected the fit levels decreased with increased noise magnitudes.



A: The Iterative Process of the RTM Generation Algorithm

B: Screenshots from the User-interface of the Web-based JavaScript Version of Supsim

Notes for Panel A:

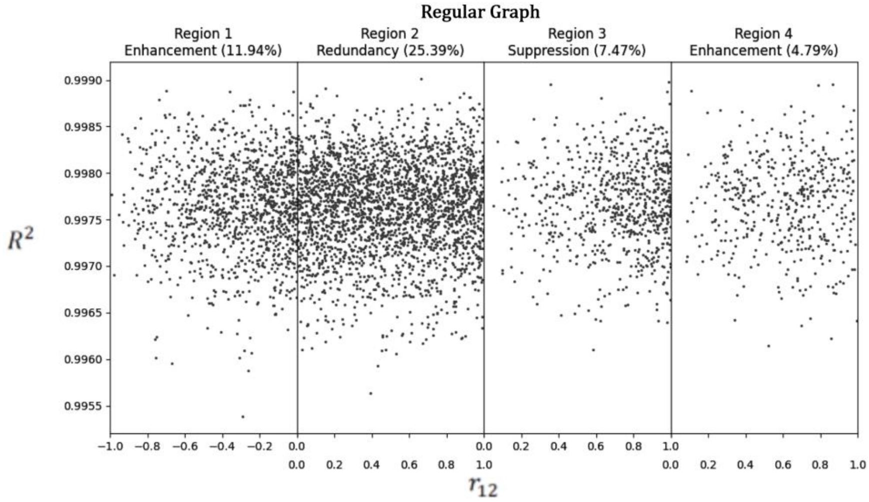
*: "e" is a distribution of errors of the same length as Yo (or original Y), while mean and standard deviation of "e" is determined arbitrarily by the user as a proportion of mean and standard deviation of Yo. "e" enables users to control the fit levels of the RTM's.

**: arguments (or arg's) are arbitrarily selected by the users to limit the magnitude of r_{y1} and r_{y2}. By using arg's, users control the amount of r_{y1} and r_{y2}.

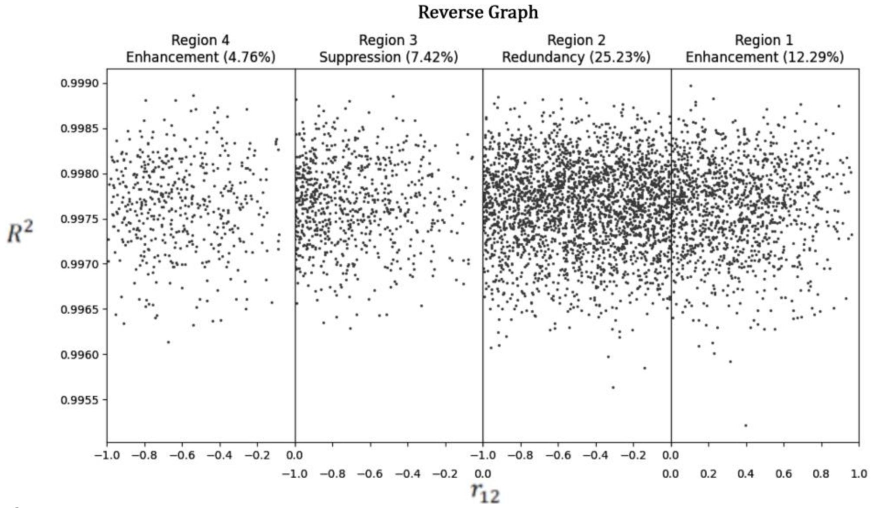
***: There are two kinds of "allowed range" for r_{12} in Supsim: first, the default allowed range is defined by "$r_{y1} \times r_{y2} - \sqrt{(1 - r^2_{y1})(1 - r^2_{y2})} \leq r_{12} \leq r_{y1} \times r_{y2} + \sqrt{(1 - r^2_{y1})(1 - r^2_{y2})}$"; Second, users are allowed to further limit the magnitude of r_{12} by selecting an arbitrary range between 0 and 1.

****: arg's about the amount of R^2 enhancement enable users to arbitrarily control the levels of R^2 enhancement by selecting a proportion between 0 and 1 (see Users Guide for Supsim for more details).

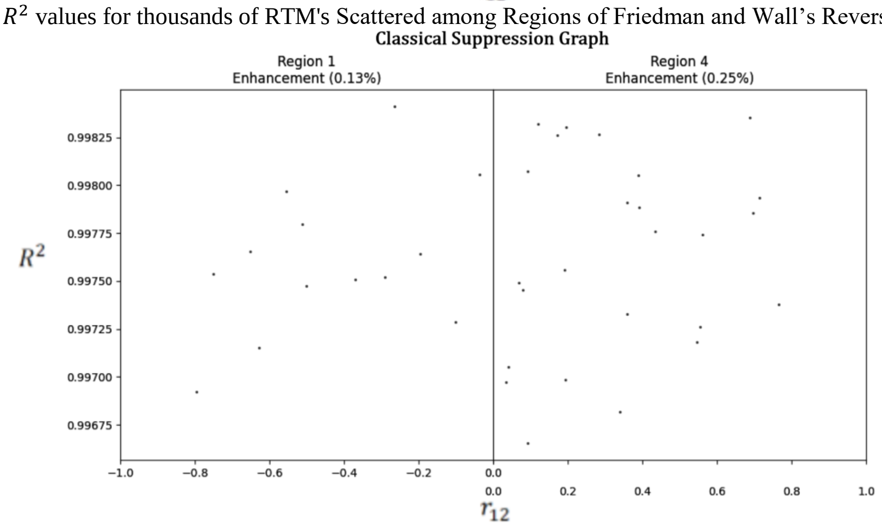
Figure 1: Flowchart of the RTM Generation Algorithm and Screenshots from the JavaScript version of Supsim



A: The R^2 values for thousands of RTM's Scattered among Regions of Friedman and Wall’s Regular Graph



B: The R^2 values for thousands of RTM's Scattered among Regions of Friedman and Wall’s Reverse Graph



C: The R^2 values for RTM's Scattered among Regions of Friedman and Wall’s Classical Suppression Graph

Figure 2: Distribution of a Large-Scale Sample of RTM’s (*N* = 10,000) among the Regions of Friedman and Wall’s Graph

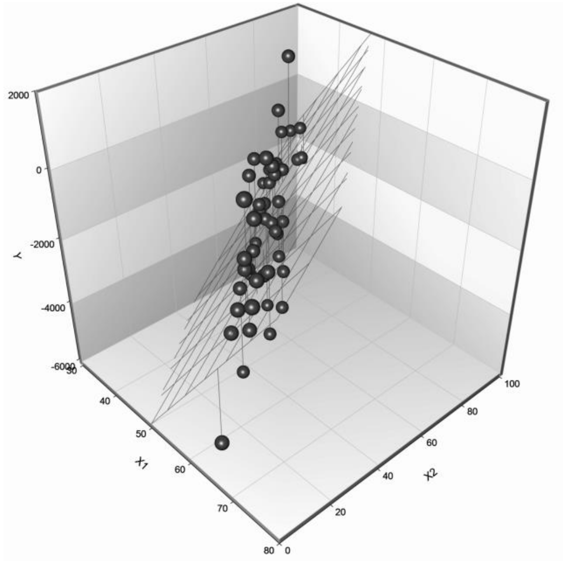
3.2 The Results of Case Studies on Unique RTM's

The "Supsim" allows users to set constraints on the magnitudes of r_{y1}, r_{y2}, r_{12}, noise, and the levels of R^2 enhancement to facilitate the production of single cases of RTM's with unique characteristics that are useful for specific purposes like case studies on unique RTM's (see Users Guide for Supsim for more details). The next section is devoted to case studies on unique RTM's created by fixed pairs of r_{y1} and r_{y2}. The authors primarily focus on the most challenging situation defined by Hamilton (1987) in which r_{y1} and r_{y2} are both close to 0 but R^2 and $|r_{12}|$ are both near 1 and then extend the discussion to other suppression situations.

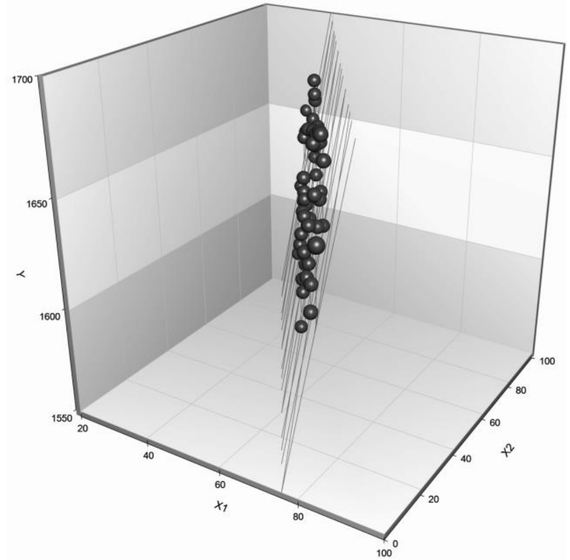
A Comparison among 3D Scatterplots of Unique RTM's drawn from Redundancy versus Enhancement/Suppression Regions

Sampling from the enhancement regions (i.e., regions I and regions IV on Friedman and Wall's graphs) with predetermined constraints on "R^2 enhancement" and "absolute values of r_{y1} and r_{y2}" was carried out by using Supsim to produce three unique RTM's with different proportions of R^2 enhancement (see Figure 3, panels A, C, and E) as well as two unique RTM's with matched absolute values of r_{y1} and r_{y2} but different proportions of R^2 enhancement (see Figure 4, panels A and C). Also sampling from redundancy region and region II suppression was performed by using Supsim to generate RTM's from redundancy and suppression regions (see Figure 3, panels B, D, and F, and Figure 4, panels B and D for more details). It should be noted that R^2 values in Figure 3 were matched between panels A and B, C and D, as well as E and F. The R^2 values also were matched between panels A and B, C and D in Figure 4. For RTM's drawn from enhancement regions, the specific constraints were "$|r_{y1}| > |r_{y2}|$", "$|r_{y1}| \leq 0.08$", "$|r_{y2}| \leq 0.08$", "$R^2 > (r_{y1}^2 + r_{y2}^2) + 0.05$". Different noise magnitudes were used in producing these RTM's (see Figure 3 and Figure 4 for more details). Then the 3D scatterplots with regression surfaces were drawn for each of the unique RTM’s in Figure 3 and Figure 4.

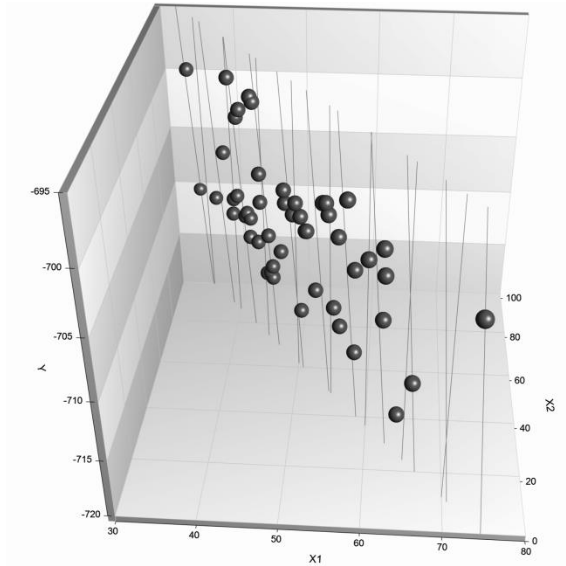
Sample Enhancement Situations
(R^2 values are matched between: A and B, C and D, E and F)



A: Region I enhancement (Enhancement = 0.11)
$\boldsymbol{R^2}$**= 0.119** r_{y1}= 0.08 r_{y2}= 0.008 r_{12}= -0.965
β_1= 1.322 β_2= 1.284; noise magnitude = 2.00

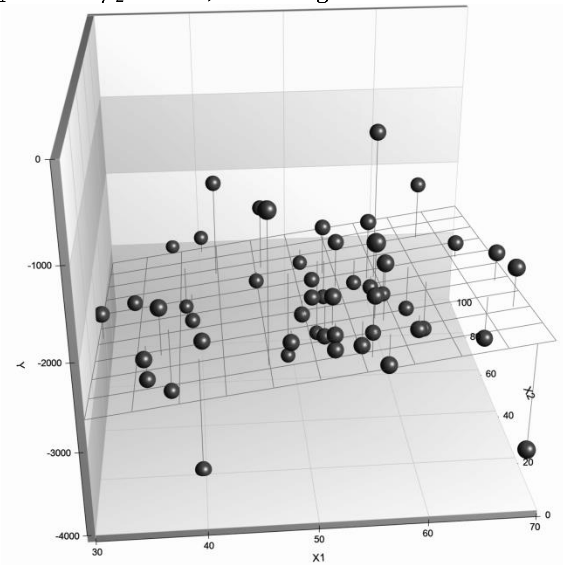


C: Region I enhancement (Enhancement =0.483)
$\boldsymbol{R^2}$**= 0.492** r_{y1}= 0.07 r_{y2}= 0.065 r_{12}= -0.981 β_1= 3.635 β_2= 3.632; noise magnitude = 1.00

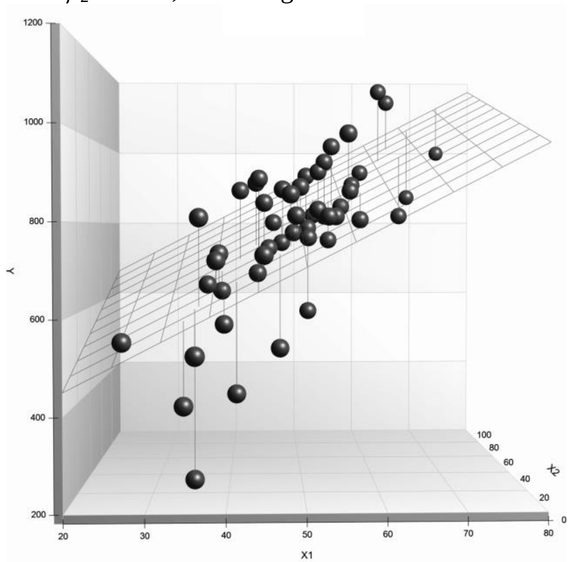


E: Classical Suppression (Enhancement = 0.995)
$\boldsymbol{R^2}$**= 0.999** r_{y1}= -0.056 r_{y2}= -0.00036 r_{12}= -0.996
β_1= -17.674 β_2= -17.647; noise magnitude = 0.04

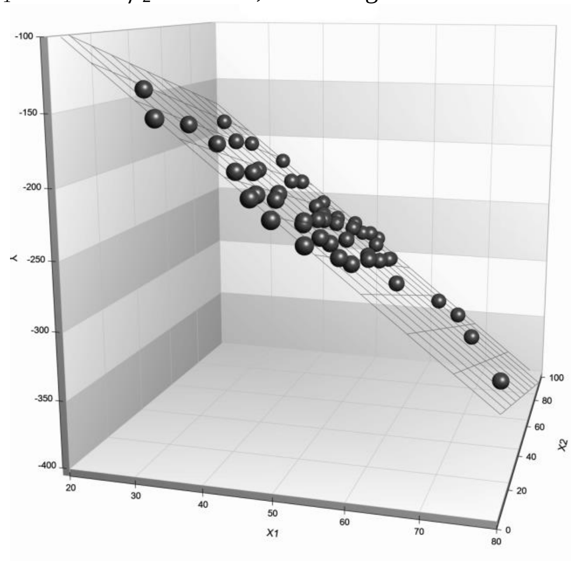
Sample Redundancy Situations
(R^2values are matched between: A and B, C and D, E and F)



B: Redundancy (RTM without Suppression)
$\boldsymbol{R^2}$**= 0.115** r_{y1}= 0.27 r_{y2}= -0.21 r_{12}= -0.212
β_1= 0.227 β_2= -0.209; noise magnitude = 2.00



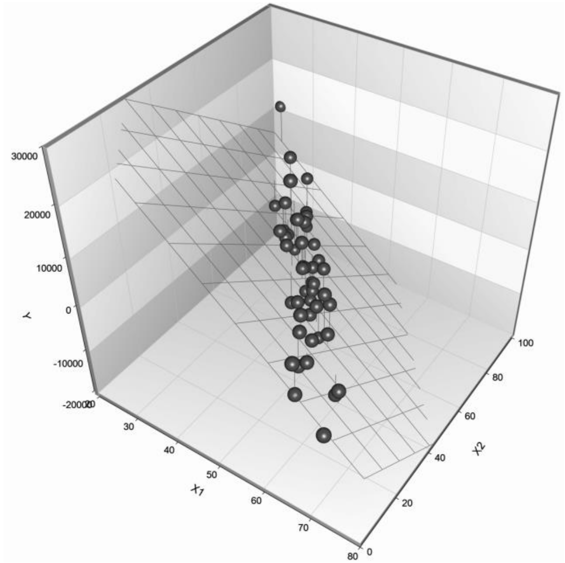
D: Redundancy (RTM without Suppression)
$\boldsymbol{R^2}$**= 0.49** r_{y1}= 0.688 r_{y2}= 0.657 r_{12}= 0.86 β_1= 0.47
β_2= 0.253; noise magnitude = 1.00



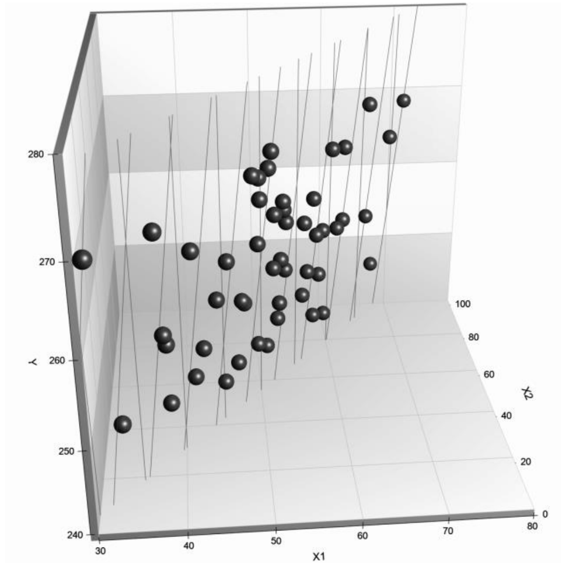
F: Redundancy (RTM without Suppression)
$\boldsymbol{R^2}$**= 0.998** r_{y1}= -0.856 r_{y2}= -0.548 r_{12}= 0.056
β_1= -0.837 β_2= -0.501; noise magnitude = 0.04

Figure 3: Matched Scatterplots from Enhancement Regions Compared to Redundancy Regions (Matched for $\boldsymbol{R^2}$)

RTMs from Enhancement Regions
(R^2 values are matched between: A and B, C and D)
(Correlations are Matched between A and C)

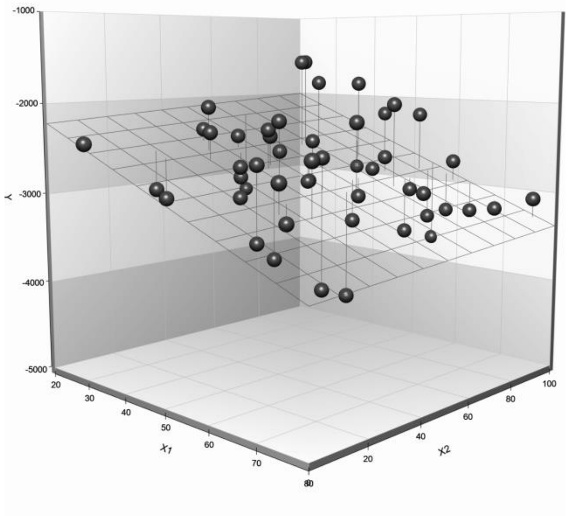


A: Region I enhancement (Enhancement = 0.121)
$\boldsymbol{R^2}$**=0.128** r_{y1}**= -0.07** r_{y2}**= -0.03** r_{12}= -0.956 β_1=-1.215 β_2= -1.194; noise magnitude= 2

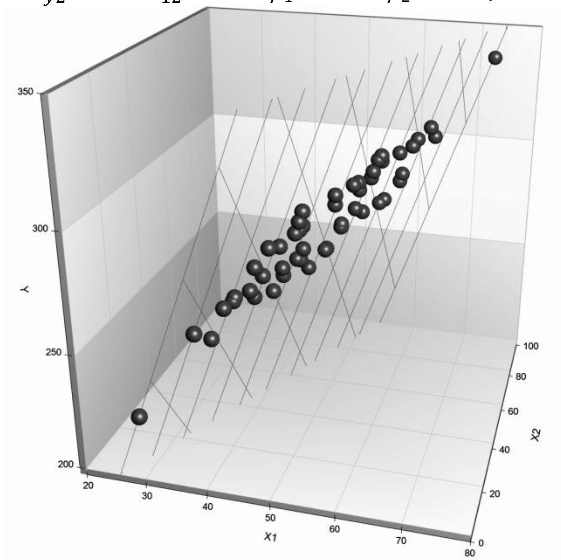


C: Region I Enhancement (Enhancement =0.99)
$\boldsymbol{R^2}$**= 0.997** r_{y1}**= 0.07** r_{y2}**= -0.03** r_{12}= 0.994 β_1= 9.48 β_2= -9.46; noise magnitude = 0.05

RTMs from Region III Suppression
(R^2 values are matched between: A and B, C and D)



B: Region III Suppression
$\boldsymbol{R^2}$**=0.128** r_{y1}= -0.349 r_{y2}= -0.116 r_{12}=0.523 β_1= -0.396 β_2=0.091; noise magnitude= 2



D: Region III Suppression
$\boldsymbol{R^2}$**=0.997** r_{y1}=0.901 r_{y2}=0.801 r_{12}=0.981 β_1= 3.07 β_2= -2.211; noise magnitude = 0.05

Figure 4: Matched Scatterplots of Enhancement Situations Compared to Region III Suppression (Matched for $\boldsymbol{R^2}$ or Zero-Order Correlations)

A comparison between 3D scatterplots of RTM's from enhancement regions (Figure 3, panels A, C, and E) and those of redundancy regions (Figure 3, panels B, D, and F) shows that contrary to RTM's from redundancy regions in which the patterns of scattered points are all consistent with their respective R^2 values, there are no consistency between the patterns of scattered points and the R^2 values for RTM's drawn from enhancement regions. For panels A, C, and E in Figure 3, the values of *y* are almost independent from the values of x_1 and x_2 which is evident from the scattered dots being almost orthogonal to the plane spanned by x_1 and x_2 in all the three scatterplots. Indeed, for panels A, C, and E while x_1 and x_2 are highly sensitive to each other's variability (i.e., all $|r_{12}|$'s ≥ 0.965) they are almost indifferent to the variability in *y*. Surprisingly, however, not only the three R^2 values are not near 0 but also they are considerably different from each other *as a function of different* $|r_{12}|$ *values* (the R^2 values are 0.119, 0.492, and 0.997 respectively for panels A, C, and E of Figure 3). Consider, for example, the scattered dots on Figure 3, panel E. The possibility interval of r_{12} here is -0.99841 to 0.99845, and the regression surface is almost parallel to the *y* axis and orthogonal to the plane spanned by x_1 and x_2. Indeed, both the regression surface and the scattered points on panel E of Figure 3 suggest that R^2 value *must be close to 0*, while the observed value of R^2 is 0.999 (i.e., near 1). Although, apparently the observed R^2 = 0.999 in panel E is calculated correctly, because the residuals here are near 0, and it is well known that R^2 has been defined as a function of residuals (Kvalseth 1985; Alexander, Tropsha, and Winkler 2015), but still something is wrong:

Panel E in Figure 3 is an extreme example of what first described by Hamilton (1987), a suppression situation with $R^2 > r_{y1}^2 + r_{y2}^2$ in which r_{y1} and r_{y2} are both close to 0 but R^2 and $|r_{12}|$ are both near 1. Hamilton (1987) shows that in this situation when R^2 = 1 and r_{y2} = 0 the following equality can be obtained from formula (3) above:

$$r_{12}^2 = 1 - r_{y1}^2 \quad (5)$$

Note that by moving "$- r_{y1}^2$" to the left side of equality (5) the following equality is obtained:

$$R^2 = r_{12}^2 + r_{y1}^2 = 1 \quad (6)$$

Readers see that under the condition of $R^2 > r_{y1}^2 + r_{y2}^2$ when R^2 = 1, $r_{y2} = 0$, and also r_{y1} is approximately close to 0 as it is the case in panel E of Figure 3, formula (3) tends to approximately substitute the value of r_{12}^2 for the value of R^2. Indeed, panel E shows a R^2 = 0.999 despite the fact that x_1 and x_2 only react strongly to each other's variability (r_{12}= -0.996), but they are almost indifferent to the variation in *y* as it is evident from $r_{y1}^2 + r_{y2}^2 = (-0.056)^2 + (-0.00036)^2 = 0.003$. In panel E of Figure 3 both the slope of the regression surface and the pattern of scattered points are almost parallel to the *y* axis and orthogonal to the plane spanned by both x_1 and x_2. Although regulating the slope in this way causes the residuals to be close to 0 and in turn the R^2 value to be close to 1, this cannot be correct because it simultaneously tends to approximately substitute the value of r_{12}^2 for the value of R^2 while the actual value of R^2 in this example *must be close to 0*. Therefore, the slope regulation appears to be extremely incorrect here which in turn dramatically affects the resulting values of the observed R^2, $\hat{\beta}_1$ and $\hat{\beta}_2$. The *slope regulation error* (SRE) here is evident from the inflation of the regression coefficients (IRC) which is very sever in this case compared to an equivalent model with the same values of $r_{\mathrm{y}1}$ and $r_{\mathrm{y}2}$ but r_{12}= 0. Given that in a model in which r_{12}= 0 then $r_{\mathrm{y}1} = \hat{\beta}_1$ and $r_{\mathrm{y}2} = \hat{\beta}_2$ while in cases where $r_{12} \neq 0$ both $|\hat{\beta}_1|$ and $|\hat{\beta}_2|$ deviate from the respective $|r_{\mathrm{y}1}|$ and $|r_{\mathrm{y}2}|$ values, the authors suggest quantifying the severity of IRC by a novel index that hereafter referred to as absolute beta-to-correlation ratio (or $|BC|$). The $|BC|$ is defined as follows:

$$|BC| = \left|\frac{\textit{the standardized regression coefficient}}{\textit{the respective zero-order correlation with "y"}}\right| \quad (7)$$

In Figure 3, panel E, the $|BC|$ for $\hat{\beta}_1$ equals 315.61 and it means that $|\hat{\beta}_1|$ is more than 315 times greater than $|\hat{\beta}_1|$ in an equivalent model with r_{12} = 0. And the $|BC|$ for $\hat{\beta}_2$ equals 49019.45 and it means that $|\hat{\beta}_2|$ is more than *49000 times greater than* $|\hat{\beta}_2|$ in an equivalent model with r_{12} = 0. In contrast, scatterplots from redundancy regions (see panels B, D, and F in Figure 3) show no sign of IRC. For example, in panel F of Figure 3, relatively large values of r_{y1} and r_{y2}, but not necessarily a large value of r_{12}, is needed to obtain a R^2 value as large as 0.998. in fact, the $|BC|$ ratios for those RTM's drawn from redundancy regions are always equal to or smaller than 1 indicating the absence of IRC as it is evident from panels B, D, and F in Figure 3.

Scatterplots on Figure 4 help further explain the issue of IRC in enhancement regions compared to region III suppression. Note that in Figure 3 panels A and B, as well as panels C and D have been matched for R^2 values. Panels A and C also have been matched for their zero-order correlations with *y*. The possible interval of r_{12} in both panels A and C of Figure 4 is between -0.995 and 0.9992. A comparison between the two enhancement situations in panels A and C reveals that to obtain a R^2 value of 0.128 a $|r_{12}| = 0.956$ is needed (see panel A of Figure 4). And then in panel C only a 0.038 increase in $|r_{12}|$ is needed to obtain a R^2 value of 0.997. Again, *y* is almost independent from both x_1 and x_2 in both panels A and C. But in panel A, the value of $|r_{12}| = 0.956$ is not strong enough to produce an orthogonal regression surface through generating large IRC's to obtain a R^2 value near 1. Indeed, panel A needs only a 0.038 increase in $|r_{12}|$ value to perform as strong as panel C of Figure 4 in enhancing the R^2 up to 0.997. The $|BC|$ ratios are 17.36 and 39.8 respectively for $\hat{\beta}_1$ and $\hat{\beta}_2$ in panel A of Figure 4 compared to 135.43 and 315.34 respectively for $\hat{\beta}_1$ and $\hat{\beta}_2$ in panel C of Figure 4. Similarly, IRC is always present in RTM's drawn from region III suppression (see panels B and D in Figure 4). For instance, the $|BC|$ ratios for panel B of Figure 4 are 1.135 and 0.784 respectively for $\hat{\beta}_1$ and $\hat{\beta}_2$, while they are more sever for panel D of Figure 4 as they are 3.41 and 2.76 respectively for $\hat{\beta}_1$ and $\hat{\beta}_2$.

So far the readers have seen that IRC may not occur in two-predictor models falling within redundancy regions while it is always present in models drawn from region III suppression, region I or region IV enhancement. These conclusions have already been verified by Friedman and Wall's (2005) definitions for each of the four regions on their graph.

As equality (6) and panel E in Figure 3 simultaneously show in Hamilton's (1987) extreme example an extreme SRE, which is evident from the sever standardized beta inflation (or IRC), occurs which is apparently related to the tendency of formula (3) to substitute the value of r_{12}^2 for the value of R^2 under the extreme condition defined by Hamilton (1987). By taking a closer look at underlying mathematical concepts of suppression situations and referring to the important issue of "statistical control" in two-predictor linear regression, the next section presents the results of further case studies on RTM's which warns researchers against the issue of IRC in suppression situations.

The Statistical Control Function: Quantifying the Statistical Control Part in RTM's Drawn from Redundancy Regions versus Those Drawn from Suppression Regions

The authors believe that comparing the mechanisms of statistical control between two-predictor regression models with or without a two-predictor suppression effect may provide important insights into *the effects of multicollinearity on the results of two-predictor linear regression models*. When a second predictor x_2 is entered into the regression equation, multicollinearity between x_1 and x_2 raises the issue of statistical control. To better understand the effects of multicollinearity the authors suggest equality (8) that can be derived from formula (3) by moving the terms "$1 - r_{12}^2$" from the denominator to the left side of the equation, multiplying them by R^2 and then moving the term "$-R^2 r_{12}^2$" to the right side:

$$R^2 = r_{y1}^2 + r_{y2}^2 - \left(2\, r_{y1} r_{y2} r_{12}\right) + R^2 r_{12}^2 \tag{8}$$

Equality (8) is important because it helps figure out the role of multicollinearity by partitioning R^2 into two parts: the first two-terms which are $r_{y1}^2 + r_{y2}^2$ is the collinearity-independent part (*CIP*), and the second two-terms which are $-(2\, r_{y1} r_{y2} r_{12}) + R^2 r_{12}^2$ is the collinearity-dependent part (*CDP*). It should be noted that the function of $-(2\, r_{y1} r_{y2} r_{12}) + R^2 r_{12}^2$ (or *CDP*), which is added to $r_{y1}^2 + r_{y2}^2$ (or *CIP*) when calculating R^2, is *to control for* the common variance explained jointly by x_1 and x_2 in cases of multicollinearity. An important insight here is that although when r_{12} = 0 the terms $-(2\, r_{y1} r_{y2} r_{12}) + R^2 r_{12}^2$ are equal to 0 but that condition is not always warrantied. Indeed, equality (8) shows that when redundancy is the case the R^2 formula tends to subtract a proportion of r_{12} from r_{y2}^2 to prevent the estimated value of R^2 from including the common variance explained jointly by x_1 and x_2. Therefore, the terms $-(2\, r_{y1} r_{y2} r_{12}) + R^2 r_{12}^2$ hereafter is called statistical control part (*SCP*). There is evidence that under the enhancement conditions, like those described by Hamilton (1987), the *SCP* can become positive. By obtaining equality (8) from formula (3), Hamilton (1987) argues that in cases where $R^2 > r_{y1}^2 + r_{y2}^2$, if r_{y2}= 0, and $R^2 = 1$then $r_{12}^2 = 1 - r_{y1}^2$ can be obtained from formula (3) (see equalities (5) and (6) above). In fact, by suggesting equality (5), Hamilton (1987) has been the first to unintentionally show that in extreme cases where $R^2 = 1$, r_{y2}= 0, and r_{y1} is also approximately near 0, formula (3) almost tends to substitute the value of r_{12}^2 for the value of R^2 and it explains why $|r_{12}|$ is also close to 1 under these conditions. Generally, when "enhancement" is the case the *SCP* is always positive adding a proportion of r_{12} to the value of r_{y2}^2 which in turn leads to the condition of

$R^2 > r_{y1}^2 + r_{y2}^2$. Such a function seems to the authors to be the opposite of the statistical control mechanism.

The authors think that a statistical control function is inherent in formula (3) which if carefully quantified can help explain the causes of suppression situations. Readers know that when $r_{12} = 0$, $R^2 = r_{y1}^2 + r_{y2}^2$, while in cases where $r_{12} \neq 0$ then the value of R^2 deviates from the value of $r_{y1}^2 + r_{y2}^2$. This explains why many texts (e.g., Cohen et al. 2003; Darlington and Hayes 2017) suggest the following formulas:

$$R_{y.12}^2 = r_{y1}^2 + sr_2^2 \tag{9}$$

$$sr_2 = \frac{r_{y2} - r_{y1}r_{12}}{\sqrt{1 - r_{12}^2}} \tag{10}$$

Where sr_2 is the semipartial correlation of x_2 with y and sr_2^2 is its squared value representing a proportion of the total variance in y explained by x_2 over and above the variance explained by the previous predictor(s) in the model. In fact, when calculating R^2, sr_2^2 is used instead of r_{y2}^2 *to prevent* R^2 *from including* the common variance explained jointly by x_1 and x_2 in cases of multicollinearity (i.e., when $r_{12} \neq 0$). Here again if $r_{12} = 0$ then $sr_2^2 = r_{y2}^2$, while if $r_{12} \neq 0$ then sr_2^2 deviates from r_{y2}^2. Indeed, sr_2^2 in formula (9) can be divided into two parts:

$$sr_2^2 = r_{y2}^2 + SCP \tag{11}$$

And formula (9) can be rewritten as follows:

$$R_{y.12}^2 = r_{y1}^2 + r_{y2}^2 + SCP \tag{12}$$

Therefore equality (11) gives another simple method for quantifying *SCP*:

$$SCP = sr_2^2 - r_{y2}^2 \tag{13}$$

As a result when r_{y1}, r_{y2} and r_{12} are known the statistical control also can be defined as a function of the combination of three zero-order correlations:

$$SCP = f(r_{y1}, r_{y2}, r_{12}) = \left(\frac{r_{y2} - r_{y1}r_{12}}{\sqrt{1 - r_{12}^2}}\right)^2 - r_{y2}^2 \tag{14}$$

Readers know that the first term in function (14) is equal to sr_2^2, and therefore function (14) is identical to equality (13).

As the readers may guess, there is also a collinearity-dependent part (*CDP_B*) in both $\hat{\beta}_1$ and $\hat{\beta}_2$ formulas that help explain the reason why regression coefficients become inflated in suppression situations. The following equalities can be obtained from $\hat{\beta}_1$ and $\hat{\beta}_2$ formulas (see formula (2) above):

$$\hat{\beta}_1 = r_{y1} - r_{y2}r_{12} + \hat{\beta}_1 r_{12}^2 \tag{15}$$

$$\hat{\beta}_2 = r_{y2} - r_{y1}r_{12} + \hat{\beta}_2 r_{12}^2 \tag{16}$$

Similarly, equalities (15) and (16) each partition the respective standardized regression coefficient into two parts: the first term which is r_{y1} (or r_{y2}) is the collinearity-independent part (CIP_B) and the second two-terms which are "$-r_{y2}r_{12} + \hat{\beta}_1 r_{12}^2$" (or "$-r_{y1}r_{12} + \hat{\beta}_2 r_{12}^2$") is the collinearity-dependent part (CDP_B). The collinearity-dependent part in $\hat{\beta}_1$ hereafter is represented by CDP_{B1} and the collinearity-dependent part in $\hat{\beta}_2$ is represented by CDP_{B2}. Here again, when a redundancy situation holds the function of adding CDP_B values to the values of r_{y1} or r_{y2} is to penalize the regression coefficients for multicollinearity. However, the term "penalty" can be used strictly for CDP_{B1} and CDP_{B2} values as long as no kind of two-predictor suppression effect exists in the model, because over the redundancy regions, where the effects of suppressor variables are absent, the signs of CDP_{B1} and CDP_{B2} are constantly opposite to the signs of r_{y1} and r_{y2}, respectively, making them to constantly produce $\hat{\beta}_1$ and $\hat{\beta}_2$ values smaller than or equal to r_{y1} and r_{y2}. In contrast, in region III suppression and both in region I and region IV enhancement, the signs of *CDP*$_{B1}$ are always similar to the signs of r_{y1} adding progressively larger proportions of r_{12} to r_{y1} to produce inflated $\hat{\beta}_1$ values as $|r_{12}|$ increases to its maximum value. Interestingly, over both region III suppression and region IV enhancement always $|CDP_{B2}|$'s $> |r_{y2}|$ and the signs of CDP_{B2} values are always the opposite of r_{y2} making them to produce inflated $\hat{\beta}_2$ values of the opposite sign compared to r_{y2}. Therefore, over region III suppression and region IV enhancement, *CDP*$_{B2}$ subtracts progressively larger proportions of r_{12} from r_{y2} as $|r_{12}|$ increases to its maximum value. Finally, in region I enhancement the signs of *CDP*$_{B2}$ values are always similar to the sign of r_{y2} adding progressively larger proportions of r_{12} to r_{y2} to produce inflated $\hat{\beta}_2$ values as $|r_{12}|$ increases to its maximum value.

To verify these observations, the authors have developed another simple but important tool called the "suppression calculator" (or Supcalc) by using Microsoft Excel 2010 (visit https://supsim.netlify.app/ to download the Supcalc) that allows researchers to examine the effects of the levels of multicollinearity (or r_{12}) on R^2, $\hat{\beta}_1$, $\hat{\beta}_2$, sr_2, sr_2^2, *SCP*, *CDP*$_{B1}$, and *CDP*$_{B2}$ while holding an arbitrary r_{y1} and r_{y2} constant. The Supcalc also calculates the multicollinearity ratios including "γ" and "$\frac{2\gamma}{1+\gamma^2}$" as well as the lower and the upper limits of r_{12} for each fixed pair of r_{y1} and r_{y2}.

Consider, three different fixed pairs of r_{y1} and r_{y2} which are selected arbitrarily to represent medium, near 0, and large values of $r_{y1}^2 + r_{y2}^2$. The three pairs can be: (-0.6, -0.5), (0.0005, 0.0003), and (0.95, -0.9). Results of calculations using Supcalc are presented in tables 1 through 3. To further discuss the mechanisms of statistical control also for the pair (-0.6, -0.5) all the values of R^2, $\hat{\beta}_1$, and $\hat{\beta}_2$ against different r_{12} values are plotted in panels A through C of Figure 5.

For the pair (-0.6, -0.5) the possibility interval of r_{12} is -0.39282 ≤ r_{12} ≤ 0.9928203. Table 1 and panels A through C in Figure 5 show that when the minimum allowed value of r_{12} is used that is r_{12} = -0.39282 then the results of calculations done by Supcalc shows that R^2 = $r_{y1}^2 + sr_2^2$= $(-0.6)^2 + 0.64$ = 1, $\hat{\beta}_1 = -0.942$, $\hat{\beta}_2 = -0.87, sr_2$= -0.8, sr_2^2 = 0.64, *SCP* = $sr_2^2 - r_{y2}^2$= 0.64 – 0.25 = 0.39, CDP_{B1} = -0.342, CDP_{B2} = -0.37. Because the latter case is a region I enhancement situation, the sign of *SCP* is positive and both the signs of CDP_{B1} and CDP_{B2} are similar to the sign of r_{y1} and r_{y2} meaning that *SCP* plays a role opposite to statistical control and both CDP_{B1} and CDP_{B2} add some proportions of r_{12} to r_{y1} and r_{y2}, instead of penalizing them for multicollinearity making them to produce inflated $|\hat{\beta}_1|$ and $|\hat{\beta}_2|$ values which are respectively 1.57 and 1.74 times greater than $|\hat{\beta}_1|$ and $|\hat{\beta}_2|$ in an equivalent model with r_{12} = 0. Panel A in Figure 5, also shows that in this example where the minimum allowed r_{12} is applied $SCP = 1 - (r_{y1}^2 + r_{y2}^2) = 0.39$. In contrast, table 1 shows that if r_{12} = 0 then R^2 = $r_{y1}^2 + r_{y2}^2$= $(-0.6)^2 + (-0.5)^2$ = 0.61, $\hat{\beta}_1 = r_{y1} = -0.6$, $\hat{\beta}_2 = r_{y2} = -0.5, sr_2 = r_{y2}$ = -0.5, $sr_2^2 = r_{y2}^2$ = 0.25, *SCP* = $sr_2^2 - r_{y2}^2$= 0.25 – 0.25 = 0, CDP_{B1} = 0, CDP_{B2} = 0 (see also panels A through C in Figure 5). An interesting insight here is that where r_{12} = 0 and R^2 = $r_{y1}^2 + r_{y2}^2$, the redundancy region usually begins with r_{12} = 0, whereas in special cases where $\frac{r_{y2}}{r_{y1}}$ is close to 1 and $r_{y1}^2 + r_{y2}^2$ > 1 then the redundancy region begins with a non-zero value of r_{12} which is the subject of latter discussion. If $r_{12} = \gamma = \frac{-0.5}{-0.6}$ = 0.833333333 then R^2 = $r_{y1}^2 + sr_2^2$= $(-0.6)^2 + 0$ = 0.36, $\hat{\beta}_1 = r_{y1} = -0.6$, $\hat{\beta}_2 = 0, sr_2$= 0, sr_2^2 = 0, *SCP* = $sr_2^2 - r_{y2}^2$= 0 – 0.25 = – 0.25, CDP_{B1} = 0, CDP_{B2} = $-(r_{y2}) = 0.5$. Readers know that $r_{12} = \gamma$ is the end-point of the redundancy region in which the statistical control function removes the entire part of x_2 by estimating $\hat{\beta}_2 = 0$ and $SCP = -r_{y2}^2$, thus panel A in Figure 5 shows that $SCP = -0.25$. In fact, when $r_{12} = \gamma$ linear regression model assumes that any explained variance in y related to x_2 is in common with x_1 and therefore x_2 have no specific contribution to add to the explained variance in y. If $r_{12} = \frac{2\gamma}{1+\gamma^2}$ = 0.983606557 then $R^2 = r_{y1}^2 + r_{y2}^2 = (-0.6)^2 + (-0.5)^2 = 0.61$, $\hat{\beta}_1 = -3.327$, $\hat{\beta}_2 = 2.773$, $|sr_2| = |r_{y2}| = 0.5$, $sr_2^2 = r_{y2}^2 = 0.25$, $SCP = sr_2^2 - r_{y2}^2 = 0.25 – 0.25 = 0$, CDP_{B1} = -2.727, CDP_{B2} = 3.2726 (also see panels A through C in Figure 5). Although in the latter case SCP is 0 and again $R^2 = r_{y1}^2 + r_{y2}^2$, contrary to situations where r_{12} =0, CDP_{B1} and CDP_{B2} here are quite large creating inflated $\hat{\beta}_1$ and $\hat{\beta}_2$ with $|\hat{\beta}_1|$ being 5.545 times greater than $|\hat{\beta}_1|$ in an equivalent model with $r_{12} = 0$ and $|\hat{\beta}_2|$ being 5.546 times greater than $|\hat{\beta}_2|$ in an equivalent model with $r_{12} = 0$. Another important insight here is that as $|r_{12}|$ increases beyond the value of $|\gamma|$ the statistical control mechanism is weakened gradually so that by $|r_{12}| = \left|\frac{2\gamma}{1+\gamma^2}\right|$ the penalty level against multicollinearity reaches 0 (i.e., $SCP = 0$; see panels A through C in Figure 5). Finally, if the maximum allowed value of r_{12} is used that is r_{12} = 0.992820323 then $R^2 = r_{y1}^2 + sr_2^2 = (-0.6)^2 + 0.64 = 1$, $\hat{\beta}_1 = -7.24$, $\hat{\beta}_2 = 6.6881$, sr_2= 0.79999861, $sr_2^2 = 0.64$, $SCP = sr_2^2 - r_{y2}^2 = 0.64 – 0.25 = 0.39$, CDP_{B1} = -6.64, CDP_{B2} = 7.1881. Again here $SCP = 1 - (r_{y1}^2 + r_{y2}^2) = 0.39$ but both CDP_{B1} and CDP_{B2} show that the IRC error here is much more sever so that $|\hat{\beta}_1|$ and $|\hat{\beta}_2|$ are respectively 12.07 and 13.376 times greater than $|\hat{\beta}_1|$ and $|\hat{\beta}_2|$ in an equivalent model with $r_{12} = 0$.

Now consider the pair (0.0005, 0.0003) in which r_{y1} and r_{y2} are deliberately selected extremely close to 0 to represent the extreme example described by Hamilton (1987). The possible interval of r_{12} here is $-0.99999968 \leq r_{12} \leq 0.99999998$ and although both r_{y1} and r_{y2} are extremely close to zero of course this is not a classical suppression situation, because both

γ and $\frac{2\gamma}{1+\gamma^2}$ are non-zero (see Table 2). Table 2 show that if r_{12} = -0.99999968 then R^2 = $r_{y1}^2 + sr_2^2$= 0.00000025 + 0.99999975 = 1, $\hat{\beta}_1$ = 1250, $\hat{\beta}_2$ = 1249.9999, sr_2 = 0.999999875, sr_2^2 = 0.99999975, *SCP* = $sr_2^2 - r_{y2}^2$= 0.99999975 – 0.00000009 = 0.99999966, CDP_{B1} = 1249.9995, CDP_{B2} = 1249.9996. Here again $SCP = 1 - (r_{y1}^2 + r_{y2}^2) =$ 1- 0.00000034 = 0.99999966 and an extremely sever IRC error occurs in which $|\hat{\beta}_1|$ and $|\hat{\beta}_2|$ are respectively 2,500,000 and 4,100,000 times greater than $|\hat{\beta}_1|$ and $|\hat{\beta}_2|$ in an equivalent model with r_{12} = 0. As an instance of Hamilton's extreme example (1987), it is also obvious that R^2 here is approximately replaced with the value of r_{12}^2. Similarly, if r_{12} reaches its maximum value that is r_{12}= 0.99999998 then R^2 = $r_{y1}^2 + sr_2^2$= 0.00000025 + 0.99999975 = 1, $\hat{\beta}_1$ = 5000.00021, $\hat{\beta}_2$= -4999.99981, $sr_2 = -0.999999881$, sr_2^2 = 0.99999975, *SCP* = $sr_2^2 - r_{y2}^2$= 0.99999975 – 0.00000009 = 0.99999966, CDP_{B1} = 4999.99971, CDP_{B2} = - 5000.0001. Although again $SCP = 1 - (r_{y1}^2 + r_{y2}^2) =$ 1- 0.00000034 = 0.99999966 the CDP_B values indicate an even more extreme IRC error here so that $|\hat{\beta}_1|$ and $|\hat{\beta}_2|$ are respectively 10,000,000 and 16,666,666 times greater than $|\hat{\beta}_1|$ and $|\hat{\beta}_2|$ in an equivalent model with r_{12} = 0. Finally, predictably if r_{12} = 0 then R^2 = $r_{y1}^2 + r_{y2}^2$= 0.00000025 + 0.00000009 = 0.00000034, $\hat{\beta}_1 = r_{y1}$ = 0.0005, $\hat{\beta}_2 = r_{y2} = 0.0003$, $sr_2 = r_{y2}$ = 0.0003, $sr_2^2 = r_{y2}^2$ = 0.00000009, *SCP* = $sr_2^2 - r_{y2}^2$= 0.00000009 – 0.00000009 = 0, CDP_{B1} = 0, and CDP_{B2} = 0.

Table 1: Supcalc Calculations for $r_{y1} = -0.6, r_{y2} = -0.5, n = 25$

γ	= 0.833333333	**Lower limit of r_{12}**	= -0.39282
$2\gamma/1+\gamma^2$	= 0.983606557	**Upper limit of r_{12}**	= 0.9928203

Range* of r_{12}	R^2	$\hat{\beta}_1$	$\hat{\beta}_2$	sr_2^2	SCP	CDP_{B1}	CDP_{B2}	$SE\hat{\beta}$'s
Max=*0.992820323*	*1.000*	*-7.240*	*6.688*	*0.640*	*0.390*	*-6.640*	*7.188*	*0.000*
0.99	0.80402	-5.28	4.72	0.444	0.194	-4.68	5.22	0.669
ratio=*0.983606557*	*0.610*	*-3.327*	*2.773*	*0.250*	*0.000*	*-2.727*	*3.273*	*0.738*
0.90	0.36842	-0.79	0.21	0.008	-0.241	-.189	0.710	0.389
γ = 0.833333333	*0.360*	*-0.600*	*0.000*	*0.000*	*-0.250*	*0.000*	*0.500*	*0.309*
0.80	0.36111	-0.56	-0.06	0.001	-0.249	0.044	0.44	0.284
0.70	0.37255	-0.49	-0.16	0.013	-0.237	0.11	0.34	0.236
0.60	0.39063	-0.47	-0.22	0.031	-0.219	0.131	0.28	0.208
0.50	0.41333	-0.47	-0.27	0.053	-0.197	0.133	0.23	0.189
0.40	0.44048	-0.48	-0.31	0.080	-0.17	0.123	0.19	0.174
0.30	0.47253	-0.49	-0.35	0.113	-0.137	0.105	0.148	0.162
0.20	0.51042	-0.52	-0.40	0.150	-0.099	0.079	0.10	0.152
0.10	0.55556	-0.56	-0.44	0.196	-0.054	0.044	0.055	0.143
0.00	*0.61000*	*-0.60*	*-0.50*	*0.250*	*0.000*	*0.000*	*0.000*	*0.133*
-0.10	0.67677	-0.66	-0.57	0.317	0.067	-0.056	-0.065	0.122
-0.20	0.76042	-0.73	-0.65	0.400	0.15	-0.129	-0.146	0.106
-0.30	0.86813	-0.82	-0.75	0.508	0.258	-0.224	-0.247	0.081
Min=*-0.392820323*	*1.000*	*-0.942*	*-0.870*	*0.640*	*0.390*	*-0.342*	*-0.370*	*0.000*
-0.40	1.01190	-0.95	-0.88	0.652	0.4	-0.352	-0.381	-
-0.50	1.21333	-1.13	-1.07	0.853	0.6	-0.533	-0.567	-
-0.60	1.51563	-1.41	-1.34	1.156	0.9	-0.806	-0.844	-
-0.70	2.01961	-1.86	-1.80	1.660	1.4	-1.26	-1.304	-
-0.80	3.02778	-2.78	-2.72	2.668	2.41	-2.17	-2.22	-
-0.90	6.05263	-5.53	-5.47	5.693	5.44	-4.92	-4.97	-
-0.99	60.50251	-55.03	-54.97	60.143	59.9	-54.89	-54.48	-

Table 2: Supcalc Calculations for $r_{y1} = 0.0005, r_{y2} = 0.0003, n = 25$

γ	= 0.6	**Lower limit of r_{12}**	= -0.99999968
$2\gamma/1+\gamma^2$	= 0.882352941	**Upper limit of r_{12}**	= 0.99999998

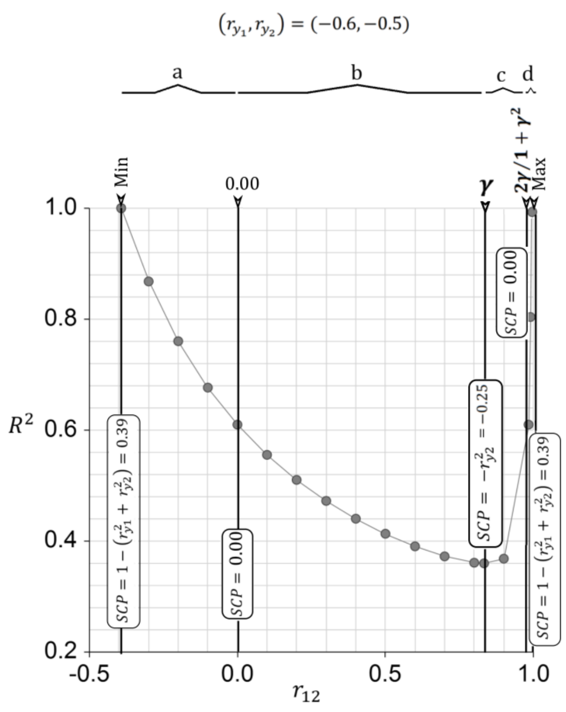
Range* of r_{12}	R^2	$\hat{\beta}_1$	$\hat{\beta}_2$	sr_2^2	SCP	CDP_{B1}	CDP_{B2}	$SE\hat{\beta}$'s
Max=0.99999998	1.000	5000.0	-4999.99	0.99999975	0.99999966	4999.99971	-5000.0	0.000
0.99	0.0000022	0.01020	-0.009799	0.0000019	0.0000018	0.0097010	-0.0100990	1.51133817
0.90	0.0000004	0.00121	-0.000789	0.0000001	0.0000000	0.0007105	-0.0010895	0.48911589
ratio=*0.882352941*	0.00000034	0.0010625	0.0006375	0.00000009	0.0000000	0.0005625	-0.0009375	0.45305145
0.80	0.0000003	0.00072	-0.000278	0.0000000	-0.000000	0.0002222	-0.0005778	0.35533447
0.70	0.0000003	0.00057	-0.000098	0.0000000	-0.000000	0.0000686	-0.0003980	0.29854067
γ =0.60	0.0000003	0.00050	0.000000	0.0000000	-0.000000	0.0000000	-0.0003000	0.26650086
0.50	0.0000003	0.00047	0.000067	0.0000000	-0.000000	-0.0000333	-0.0002333	0.24618295
0.40	0.0000003	0.00045	0.000119	0.0000000	-0.000000	-0.0000476	-0.0001810	0.23262102
0.30	0.0000003	0.00045	0.000165	0.0000000	-0.000000	-0.0000495	-0.0001352	0.22349505
0.20	0.0000003	0.00046	0.000208	0.0000000	0.0000000	-0.0000417	-0.0000917	0.21759704
0.10	0.0000003	0.00047	0.000253	0.0000001	0.0000000	-0.0000253	-0.0000475	0.21427475
0.00	*0.0000003*	*0.00050*	*0.000300*	*0.0000001*	*0.0000000*	*0.0000000*	*0.0000000*	*0.21320068*
-0.10	0.0000004	0.00054	0.000354	0.0000001	0.0000000	0.0000354	0.0000535	0.21427474
-0.20	0.0000004	0.00058	0.000417	0.0000002	0.0000001	0.0000833	0.0001167	0.21759702
-0.30	0.0000005	0.00065	0.000495	0.0000002	0.0000001	0.0001484	0.0001945	0.22349502
-0.40	0.0000005	0.00074	0.000595	0.0000003	0.0000002	0.0002381	0.0002952	0.23262099
-0.50	0.0000007	0.00087	0.000733	0.0000004	0.0000003	0.0003667	0.0004333	0.24618290
-0.60	0.0000008	0.00106	0.000938	0.0000006	0.0000005	0.0005625	0.0006375	0.26650079
-0.70	0.0000011	0.00139	0.001275	0.0000008	0.0000007	0.0008922	0.0009745	0.29854056
-0.80	0.0000016	0.00206	0.001944	0.0000014	0.0000013	0.0015556	0.0016444	0.35533424
-0.90	0.0000032	0.00405	0.003947	0.0000030	0.0000029	0.0035526	0.0036474	0.48911520
-0.99	0.0000320	0.04005	0.039950	0.0000318	0.0000317	0.0395503	0.0396497	1.51131562
Min=-0.99999968	1.000	1250.00	1249.9999	0.99999975	0.99999966	1249.99955	1249.99965	0.000

Note: *SCP* = statistical control part; CDP_{B1}= collinearity-dependent part of β_1; CDP_{B2} = collinearity-dependent part of β_2; $SE\hat{\beta}$'s = standard errors of $\hat{\beta}$'s; Min = minimum allowed value of r_{12}; Max = maximum allowed value of r_{12}; ratio = $2\gamma/1+\gamma^2$; *: The possibility interval of r_{12} is highlighted in gray in r_{12} columns;

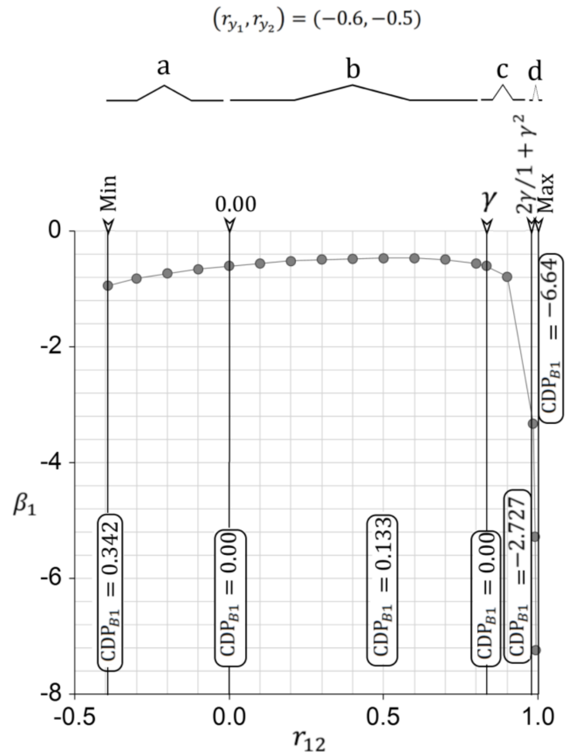
Table 3: Supcalc Calculations for $r_{y1} = 0.95, r_{y2} = -0.9, n = 25$

γ	= -0.94736842			**Lower limit of r_{12}**	= -0.991106576			
$2\gamma/1+\gamma^2$	= -0.99854015			**Upper limit of r_{12}**	= -0.718893424			
Range* of r_{12}	R^2	$\hat{\beta}_1$	$\hat{\beta}_2$	sr_2^2	*SCP*	CDP_{B1}	CDP_{B2}	$SE\hat{\beta}$'s
0.99	171.126	92.513	-92.487	170.22	169.41	91.563	-91.587	-
0.90	17.113	9.263	-9.237	16.211	15.401	8.313	-8.337	-
0.80	8.557	4.639	-4.611	7.654	6.844	3.689	-3.711	-
0.70	5.705	3.098	-3.069	4.802	3.992	2.148	-2.169	-
0.60	4.279	2.328	-2.297	3.376	2.566	1.378	-1.397	-
0.50	3.423	1.867	-1.833	2.521	1.711	0.917	-0.933	-
0.40	2.853	1.560	-1.524	1.950	1.140	0.610	-0.624	-
0.30	2.446	1.341	-1.302	1.543	0.733	0.391	-0.402	-
0.20	2.140	1.177	-1.135	1.238	0.428	0.227	-0.235	-
0.10	1.903	1.051	-1.005	1.000	0.190	0.101	-0.105	-
0.00	1.713	0.950	-0.900	0.810	0.000	0.000	0.000	-
-0.10	1.557	0.869	-0.813	0.655	-0.155	-0.081	0.087	-
-0.20	1.428	0.802	-0.740	0.525	-0.285	-0.148	0.160	-
-0.30	1.318	0.747	-0.676	0.416	-0.394	-0.203	0.224	-
-0.40	1.224	0.702	-0.619	0.322	-0.488	-0.248	0.281	-
-0.50	1.143	0.667	-0.567	0.241	-0.569	-0.283	0.333	-
-0.60	1.073	0.641	-0.516	0.170	-0.640	-0.309	0.384	-
-0.70	1.011	0.627	-0.461	0.108	-0.702	-0.323	0.439	-
Max=-0.718893424	*1.000*	*0.627*	*-0.449*	*0.098*	*-0.712*	*-0.323*	*0.451*	*0.000*
-0.80	0.957	0.639	-0.389	0.054	-0.756	-0.311	0.511	0.074
-0.90	0.913	0.737	-0.237	0.011	-0.799	-0.213	0.663	0.144
γ = -0.94736842	*0.9025*	*0.950*	*0.000*	*0.000*	*-0.810*	*0.000*	*0.900*	*0.208*
-0.99	0.985	2.965	2.035	0.082	-0.728	2.015	2.935	0.186
Min=-0.991106576	*1.000*	*3.276*	*2.347*	*0.098*	*-0.712*	*2.326*	*3.247*	*0.000*

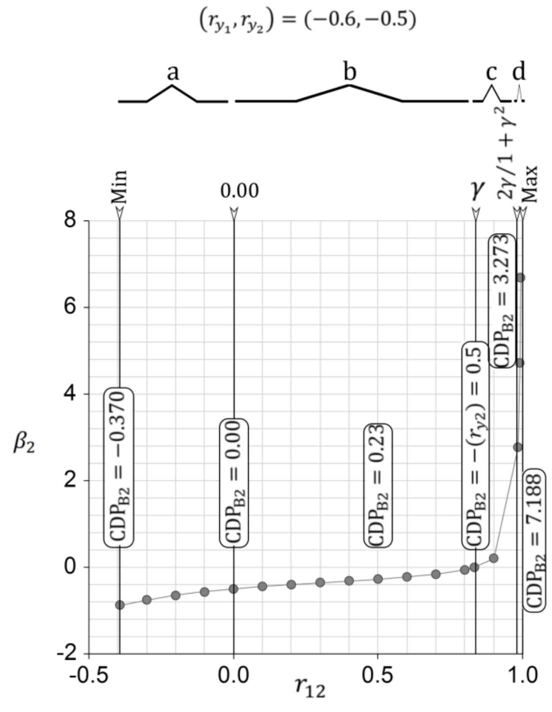
Note: *SCP* = statistical control part; CDP_{B1} = collinearity-dependent part of β_1; CDP_{B2} = collinearity-dependent part of β_2; $SE\hat{\beta}$'s = standard errors of $\hat{\beta}$*'s*; Min = minimum allowed value of r_{12}; Max = maximum allowed value of r_{12}; ratio = $2\gamma/1+\gamma^2$; *: The possibility interval of r_{12} is highlighted in gray in r_{12} columns;



A: Changes in R^2 due to changes in both r_{12} and *SCP*



B: Changes in $\hat{\beta}_1$ due to changes in both r_{12} and CDP_{B1}



C: Changes in $\hat{\beta}_2$ due to changes in both r_{12} and CDP_{B2}

a: Region I: Enhancement: When calculating the R^2 value, *SCP* adds progressively greater proportions of r_{12} to r_{y2}^2 as r_{12} approaches its minimum value.

b: Region II: Redundancy: *SCP* penalizes R^2 for multicollinearity by subtracting progressively greater proportions of r_{12} from r_{y2}^2 as r_{12} approaches γ.

c: Region III: Suppression: *SCP* subtracts progressively smaller proportions of r_{12} from r_{y2}^2 as r_{12} approaches until the penalty level against multicollinearity reaches 0 by $r_{12} = 2\gamma/1+\gamma^2$.

d: Region IV: Enhancement: When calculating the R^2 value, *SCP* adds progressively greater proportions of r_{12} to r_{y2}^2 as r_{12} approaches its maximum value.

a: Region I: Enhancement: When calculating $\hat{\beta}_1$, CDP_{B1} adds progressively greater proportions of r_{12} to r_{y1} to create inflated $\hat{\beta}_1$ values as r_{12} approaches its minimum value. The signs of CDP_{B1} and r_{y1} are always similar in this region.

b: Region II: Redundancy: CDP_{B1} penalizes $\hat{\beta}_1$ for multicollinearity by subtracting different proportions of r_{12} from r_{y1} when calculating $\hat{\beta}_1$. When $r_{12} = 0.00$ or $r_{12} = \gamma$ the penalty level against multicollinearity always is 0 and therefore $\hat{\beta}_1 = r_{y1}$. CDP_{B1} and r_{y1} are always of opposite signs in this region.

c: Region III: Suppression: CDP_{B1} adds progressively greater proportions of r_{12} to r_{y1} to create inflated $\hat{\beta}_1$ values as r_{12} approaches $2\gamma/1+\gamma^2$. The signs of CDP_{B1} and r_{y1} are always similar in this region.

d: Region IV: Enhancement: CDP_{B1} adds progressively greater proportions of r_{12} to r_{y1} to create inflated $\hat{\beta}_1$ values as r_{12} approaches its maximum value. The signs of CDP_{B1} and r_{y1} are always similar in this region.

a: Region I: Enhancement: When calculating $\hat{\beta}_2$, CDP_{B2} adds progressively greater proportions of r_{12} to r_{y2} to create inflated $\hat{\beta}_2$ values as r_{12} approaches its minimum value. The signs of CDP_{B2} and r_{y2} are always similar in this region.

b: Region II: Redundancy: CDP_{B2} penalizes $\hat{\beta}_2$ for multicollinearity by subtracting progressively greater proportions of r_{12} from r_{y2} as r_{12} approaches γ. CDP_{B2} and r_{y2} are always of opposite signs in this region.

c: Region III: Suppression: Always $|CDP_{B2}| > |r_{y2}|$, CDP_{B2} and r_{y2} are always of opposite signs in this region, and CDP_{B2} subtracts progressively greater proportions of r_{12} from r_{y2} as r_{12} approaches $2\gamma/1+\gamma^2$. Therefore, CDP_{B2} creates inflated $\hat{\beta}_2$ values of the opposite sign with respect to r_{y2}.

d: Region IV: Enhancement: Always $|CDP_{B2}| > |r_{y2}|$, CDP_{B2} and r_{y2} are always of opposite signs in this region, and CDP_{B2} subtracts progressively greater proportions of r_{12} from r_{y2} as r_{12} approaches its maximum value. CDP_{B2} creates inflated $\hat{\beta}_2$ values of the opposite sign.

Figure 5: Comparing the Statistical Control Mechanisms among Suppression and Non-Suppression Situations

Readers already know that in classical suppression conditions the redundancy region and the region III suppression may not occur. Friedman and Wall (2005) also describe another condition in which the opposite is true that is only the redundancy region and the region III suppression may occur while enhancement is impossible and this is what happens with the pair (0.95, -0.9). According to Friedman and Wall (2005) if γ is close to 1 and $r_{y1}^2 + r_{y2}^2 \geq 1$ the redundancy region holds over a large interval of r_{12}, the region III suppression occurs in a small range and no area remains for enhancement. For the pair (0.95, -0.9) the possible interval of r_{12} is -0.991106576 ≤ r_{12} ≤ -0.718893424. Here the redundancy region begins with r_{12} = -0.718893424 which is the greatest possible value of r_{12}. Therefore, If r_{12} = -0.718893424 then $R^2 = r_{y1}^2 + sr_2^2 = (0.95)^2 + 0.0975 = 1$, $\hat{\beta}_1 = 0.6271$, $\hat{\beta}_2 = -0.449203$, sr_2= -0.3122499, $sr_2^2 = 0.0975$, *SCP* = $sr_2^2 - r_{y2}^2$= 0.0975 – 0.81 = -0.7125, CDP_{B1} = -0.32293, CDP_{B2} = 0.4508. Because in this example $r_{y1}^2 + r_{y2}^2 = 1.7125$ is greater than 1, the *SCP* penalizes r_{y2}^2 by 0.7125 to prevent the resulting value of R^2 from including the shared variance explained jointly by x_1 and x_2. The signs of both CDP_{B1} and CDP_{B2} are also opposite to the signs of r_{y1} and r_{y2} to penalize $\hat{\beta}_1$ and $\hat{\beta}_2$ for multicollinearity. And If r_{12} = γ = -0.947368421 then $R^2 = r_{y1}^2 + sr_2^2 = (0.95)^2 + 0.000 = 0.9025$, $\hat{\beta}_1 = r_{y1} = 0.95$, $\hat{\beta}_2 = 0.000$, sr_2= 0.000, $sr_2^2 = 0.000$, *SCP* = $sr_2^2 - r_{y2}^2$= 0.000 – 0.81 = -0.81, CDP_{B1} = 0.000, CDP_{B2} = 0.9. The redundancy region ends at $r_{12} = \gamma$ and here the model removes the entire part of x_2 by estimating $\hat{\beta}_2 = 0.000$ and *SCP* = -0.81. Interestingly, while CDP_{B1} = 0.000 leads to $\hat{\beta}_1 = r_{y1}$, $|CDP_{B2}|$ is equal to $|r_{y2}|$ though it shows the opposite sign of r_{y2} and as a result $\hat{\beta}_2 = 0.000$. Finally, Table 3 shows that at the left most of r_{12} axis where r_{12} = -0.991106576, $R^2 = r_{y1}^2 + sr_2^2 = (0.95)^2 + 0.0975 = 1$, $\hat{\beta}_1 = 3.276$, $\hat{\beta}_2 = 2.3465$, sr_2= 0.3122499, $sr_2^2 = 0.0975$, *SCP* = $sr_2^2 - r_{y2}^2$= 0.0975 – 0.81 = -0.7125, CDP_{B1} = 2.3256, CDP_{B2} = 3.2465. Although the *SCP* value here is equal to the *SCP* value in the previous redundancy situation, the fact that the values of both $|CDP_{B1}|$ and $|CDP_{B2}|$ are highly greater than $|r_{y1}|$ and $|r_{y2}|$ indicate that a suppression effect is present that produces inflated $|\hat{\beta}_1|$ and $|\hat{\beta}_2|$ values being respectively 3.45 and 2.61 times greater than $|\hat{\beta}_1|$ and $|\hat{\beta}_2|$ in an equivalent model with r_{12} = 0.

4 DISCUSSION

The concept of two-predictor suppression effects has been the subject of debate over terminology (Friedman and Wall 2005), definition, and interpretation (Mendershausen 1939; Horst 1941; Meehl 1945; Conger and Jackson 1972; Conger 1974; Cohen and Cohen 1975; Velicer 1978; Tzelgov and Henik 1991; Sharpe and Roberts 1997; Shieh 2001; Lynn 2003). One point of agreement has been the approach chosen by researchers who agree that a suppressor variable showing “no or low” correlation with the criterion variable *y* but is correlated with another significant predictor x_1 can be included in the regression equation to increase the predictive validity of x_1 and for this reason they consider suppressor variables useful and even desirable for situations where the purpose of the study is prediction (Horst 1941; Conger and Jackson 1972; Tzelgov and Henik 1991; Pedhazur 1997; Cohen et al. 2003; Friedman and Wall 2005; Watson et al. 2013; Darlington and Hayes 2017). On the other hand, some texts have warned researchers against multicollinearity and suggest some “rules of thumb” to limit the magnitude of multicollinearity between predictor variables specially when the purpose of the study is "theoretical explanation" (e.g., Cohen et al. 2003), they argue that highly correlated predictor variables, when simultaneously included in the regression equation, cause “instabilities” in different meanings: first, increased standard errors as a function of high multicollinearity may cause “instability” in estimating the regression coefficients (Neter et al. 1996; Fox 1997; Cohen et al. 2003); second, computational inaccuracies are more likely to occur in calculating the inverses of matrices with highly correlated variables (Cohen and Cohen 1983); and third, high levels of r_{12} can lead to rapid increase in $\hat{\beta}_2$, a condition in which "the interpretation of regression coefficients may become problematic" (Cohen et al. 2003). Friedman and Wall (2005) argue against the latter texts by presenting evidence that show the standard errors (*SE's*) of regression coefficients do not increase steadily with increasing multicollinearity and there are cases in which low standard errors are coincident with high multicollinearity and that *SE's* of regression coefficients always become 0 when the multicollinearity for each given pair of r_{y1} and r_{y2} reaches its absolute maximum values (see table 1 through 3 above). They also argue that the issue of computational accuracy is no longer problematic for the latest generations of regression algorithms (Friedman and Wall 2005). And finally, Friedman and Wall (2005) conclude that when regressing *y* on two predictors there are no limits on multicollinearity except those warranty a nonnegative definite matrix. Although Friedman and Wall's observation concerning *SE's* of regression coefficients is quite correct, their final conclusion that there is no limits on multicollinearity except nonnegative, definiteness limitation is incorrect. Similarly, as Cohen et al. (2003) observed, it is true that there is a rapid increase in $\hat{\beta}_2$ at high levels of r_{12}, but their agreement to use the suppressor variables to increase R^2 in cases where the main purpose of the study is increasing the predictive validity is misleading. As noted earlier in the introduction section, two important aspects of two-predictor suppression effects have been overlooked that have led researchers to misleading conclusions: first, failure to examine 3D scatterplots of suppression and non-suppression situations; and second, insufficient attention to the important issue of statistical control mechanisms in non-suppression compared to suppression situations. Taking into consideration these two important aspects, this study achieved significant findings as follows:

First, a closer look at the terms in R^2, $\hat{\beta}_1$, and $\hat{\beta}_2$ formulas indicates that these formulas consist of two separate parts (see Equality 7 above): the collinearity-independent part (*CIP*) and the collinearity-dependent part (*CDP*). The *CDP* terms in R^2, $\hat{\beta}_1$, and $\hat{\beta}_2$ formulas are associated with statistical control mechanisms, and therefore should be quantified and examined separately.

Second, the *CDP* terms in R^2 formula act differently in redundancy and suppression regions in terms of statistical control. While the *SCP*, or the sum of the *CDP* terms in R^2 formula, is always negative in redundancy regions penalizing R^2 for multicollinearity, the penalty level of *SCP* decreases progressively in region III suppression, and as a result *SCP* subtracts progressively smaller proportions of r_{12} from r_{y2}^2 as r_{12} approaches $2\gamma/1+\gamma^2$. At $2\gamma/1+\gamma^2$ point the penalty level of *SCP* against multicollinearity reaches 0. Beyond the $2\gamma/1+\gamma^2$ ratio in region IV enhancement, *SCP* becomes positive and adds progressively greater proportions of r_{12} to r_{y2}^2 as r_{12} approaches its absolute maximum value (see Figure 5 panel A). When r_{y1} and r_{y2} have similar signs, all r_{12}'s < 0 create the region I enhancement (or reciprocal suppression) but when r_{y1} and r_{y2} are of opposite signs, all r_{12}'s > 0 produce the region I enhancement (or reciprocal suppression). The *SCP* again is positive in region I enhancement adding progressively greater proportions of r_{12} to r_{y2}^2 as r_{12} approaches its absolute maximum value. For example, panel A in Figure 5 shows that *SCP* is positive and equal to $1-(r_{y1}^2+\ r_{y2}^2)$ both at the upper and the lower limits of r_{12} for the pair (-0.6, -0.5) whereas if r_{12} = 0, *SCP* = 0; if r_{12} = γ, *SCP* = $-(r_{y2}^2)$; and if r_{12} = $2\gamma/1+\gamma^2$, again *SCP* = 0. According to these findings, the authors suggest renaming the regions suggested by Friedman and Wall (2005) in terms of their statistical control functioning. Therefore, the following labels are suggested: "region I: statistical anti-control", "region II: statistical control", "region III: statistical de-control", and "region IV: statistical anti-control" respectively for "region I: enhancement", "region II: redundancy", "region III: suppression", and "region IV: enhancement". In fact, the aim of these "relabeling" is to show that "correct statistical control" can only occur in "region II: redundancy" while in region III: suppression, *SCP* gradually removes the statistical control against multicollinearity until by the $2\gamma/1+\gamma^2$ point *SCP* = 0; in region I as well as region IV enhancement areas, *SCP* acts against the purpose of statistical control adding progressively greater proportions of r_{12} to r_{y2}^2 until by the maximum absolute values of r_{12} in both directions SCP becomes equal to $1-(r_{y1}^2+r_{y2}^2)$. Therefore, an important conclusion here is that all different two-predictor suppression effects are different kinds of "statistical control dysregulation". These findings emphasize that no proportions of r_{12} can replace the variance explained in *y*, and the results produced by two-predictor suppression effects are completely erroneous and misleading.

Third, the *CDP* terms in formulas of both $\hat{\beta}_1$ and $\hat{\beta}_2$ also function differently in redundancy and suppression regions. The signs of both CDP_{B1} and CDP_{B2} values in redundancy regions are always opposite to the signs of r_{y1} and r_{y2} and they always subtract different proportions of r_{12} from r_{y1} and r_{y2} to penalize $\hat{\beta}_1$ and $\hat{\beta}_2$ values for multicollinearity and to produce $\hat{\beta}_1$ and $\hat{\beta}_2$ values which are always smaller than or equal to r_{y1} and r_{y2}. In contrast, in region III suppression the signs of CDP_{B1} values are always similar to the sign of r_{y1} adding progressively greater proportions of r_{12} to r_{y1} to produce inflated $\hat{\beta}_1$ values as r_{12} approaches $2\gamma/1+\gamma^2$ whereas the signs of CDP_{B2} values are always opposite to the sign of r_{y2} in region III suppression but always $|CDP_{B2}| > |r_{y2}|$ which in turn produces inflated $\hat{\beta}_2$ values of the opposite sign compared to r_{y2}. Similarly, in region IV enhancement the signs of CDP_{B1} values are always similar to the sign of r_{y1}creating inflated $\hat{\beta}_1$ values as r_{12} approaches its absolute maximum value whereas the signs of CDP_{B2} values again are always opposite to the sign of r_{y2} but always $|CDP_{B2}| > |r_{y2}|$ producing inflated $\hat{\beta}_2$ values of the opposite sign with respect to r_{y2}. In region I enhancement, the signs of both CDP_{B1} and CDP_{B2} values are always similar to the signs of r_{y1} and r_{y2} adding gradually greater proportions of r_{12} to the zero-order correlations to create progressively more inflated $\hat{\beta}_1$ and $\hat{\beta}_2$ values as r_{12} approaches its absolute maximum value (see Figure 5 Panels B and C). These findings show that the statistical control mechanisms can correctly adjust the slope of the regression surface only and only in redundancy regions, while the slope of the regression surface unjustifiably increases in all the three different types of suppression situations so that the regression surface sharply cuts the plane spanned by both x_1 and x_2, a condition that can be called "slope regulation error" (SRE) (see Figure 3 and Figure 4). Again these findings emphasize that no proportions of r_{12} can replace the values of regression coefficients, and therefore the slope regulations affected by two-predictor suppression effects are completely erroneous and misleading.

This study presents a complete account of two-predictor suppression effects in terms of the important issue of statistical control that expands the previous knowledge and resolves the complexities. The authors have also developed some useful applications and tools that help simulate and examine all different kinds of two-predictor suppression effects and boost further research. These findings also provide important implications for the issue of "effect size" in linear regression and can change the educational contents and materials of the topic of two-predictor suppression effects in linear regression modeling.

This study also involves important limitations. First, the case studies and examples in this study do not include situations where there are more than two predictors in the model. Second, only continuous quantitative variables have been included, and further investigation on regression with categorical variables or with a combination of continuous and categorical variables remains to be carried out. The implications for the issue of "effect size" also need to be investigated in future. And finally an important question is that how can these findings and software be best incorporated into educational contents and materials.

References

Alexander, D. L. J., Tropsha, A., and Winkler, D. A. (2015), "Beware of R(2): Simple, Unambiguous Assessment of the Prediction Accuracy of QSAR and QSPR Models," *Journal of chemical information and modeling*, 55 (7), 1316-1322. DOI: 10.1021/acs.jcim.5b00206.

Cohen, J., and Cohen, P. (1975), *Applied multiple regression/correlation analysis for the behavioral sciences*, Oxford, England: Lawrence Erlbaum.

--- (1983), *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences* (2nd ed.), Hillsdale, N.J: Lawrence Erlbaum Associates.

Cohen, J., Cohen, P., West, S., and Aiken, L. (2003), *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences* (Third ed.), Mahwah,NJ: Lawrence Erlbaum Associates.

Conger, A. J. (1974), "A revised definition for suppressor variables: A guide to their identification and interpretation," *Educational and Psychological Measurement*, 34 (1), 35-46. DOI: 10.1177/001316447403400105.

Conger, A. J., and Jackson, D. N. (1972), "Suppressor Variables, Prediction, and the interpretation of Psychological Relationships," *Educational and Psychological Measurement*, 32 (3), 579-599. DOI: 10.1177/001316447203200303.

Currie, I., and Korabinski, A. (1984), "Some Comments on Bivariate Regression," *Journal of the Royal Statistical Society. Series D (The Statistician)*, 33 (3), 283-293. DOI: 10.2307/2988232.

Darlington, R. B. (1968), "Multiple regression in psychological research and practice," *Psychological Bulletin*, 69 (3), 161-182. DOI: 10.1037/h0025471.

Darlington, R. B., and Hayes, A. F. (2017), *Regression Analysis and Linear Models: Concepts, Applications, and Implementation*, New York: The Guilford Press.

Fox, J. (1997), *Applied regression analysis, linear models, and related methods*, Thousand Oaks, CA, US: Sage Publications, Inc.

Friedman, L., and Wall, M. (2005), "Graphical Views of Suppression and Multicollinearity in Multiple Linear Regression," *The American Statistician*, 59 (2), 127-136. DOI: 10.1198/000313005X41337.

Hamilton, D. (1987), "Sometimes R 2 > r 2 yx 1 + r 2 yx 2 : Correlated Variables are Not Always Redundant," *The American Statistician*, 41 (2), 129-132. DOI: 10.1080/00031305.1987.10475460.

Holling, H. (1983), "Suppressor Structures in the General Linear Model," *Educational and Psychological Measurement*, 43 (1), 1-9. DOI: 10.1177/001316448304300101.

Horst, P. (1941), *The prediction of personal adjustment: A survey of logical problems and research techniques, with illustrative application to problems of vocational selection, school success, marriage, and crime,* New York, NY, US: Social Science Research Council.

Kvalseth, T. O. (1985), "Cautionary Note about R²," *The American Statistician*, 39 (4), 279-285. DOI: 10.2307/2683704.

Ludlow, L., and Klein, K. (2014), "Suppressor Variables: The Difference Between 'is' Versus 'Acting As'," *Journal of Statistics Education*, 22 (2), null-null. DOI: 10.1080/10691898.2014.11889703.

Lynn, H. S. (2003), "Suppression and Confounding in Action," *The American Statistician*, 57 (1), 58-61. DOI: 10.1198/0003130031090.

McFatter, R. M. (1979), "The Use of Structural Equation Models in Interpreting Regression Equations Including Suppressor and Enhancer Variables," *Applied Psychological Measurement*, 3 (1), 123-135. DOI: 10.1177/014662167900300113.

Meehl, P. E. (1945), "A Simple Algebraic Development of Horst's Suppressor Variables," *The American Journal of Psychology*, 58 (4), 550-554. DOI: 10.2307/1417770.

Mendershausen, H. (1939), "Clearing Variates in Confluence Analysis," *Journal of the American Statistical Association*, 34 (205), 93-105. DOI: 10.2307/2279167.

Neill, J. J. (1973), "Tests of the Equality of Two Dependent Correlations," Doctoral Dissertation, University of California, Ann Arbor, Ann Arbor, MI: University Microfilms No.: 74-7671.

Neter, J., Kutner, M., Nuchtsheim, C., and Wasserman, W. (1996), *Applied Linear Statistical Models (4th ed)*, Chicago: Irwin.

Pedhazur, E. (1997), *Multiple Regression in Behavioral Research: Explanation and Prediction*, Wadsworth: Thomson Learning.

Sharpe, N. R., and Roberts, R. A. (1997), "The Relationship among Sums of Squares, Correlation Coefficients, and Suppression," *The American Statistician*, 51 (1), 46-48. DOI: 10.1080/00031305.1997.10473587.

Shieh, G. (2001), "The Inequality Between the Coefficient of Determination and the Sum of Squared Simple Correlation Coefficients," *The American Statistician*, 55 (2), 121-124. DOI: 10.1198/000313001750358437.

Tzelgov, J., and Henik, A. (1991), "Suppression Situations in Psychological Research: Definitions, Implications, and Applications," *Psychological Bulletin*, 109 (3), 524-536. DOI: doi: 10.1037/0033-2909.109.3.524.

Velicer, W. F. (1978), "Suppressor Variables and the Semipartial Correlation Coefficient," *Educational and Psychological Measurement*, 38 (4), 953-958. DOI: 10.1177/001316447803800415.

Watson, D., Clark, L. A., Chmielewski, M., and Kotov, R. (2013), "The value of suppressor effects in explicating the construct validity of symptom measures," *Psychological assessment*, 25 (3), 929-941. DOI: 10.1037/a0032781.

Whuber. 2017. Generate a random variable with a defined correlation to an existing variable(s). Stack Exchange Inc.