Educational and Psychological Measurement

http://epm.sagepub.com/

Suppression Situations in Multiple Linear Regression

Gwowen Shieh

Educational and Psychological Measurement 2006 66: 435
DOI: 10.1177/0013164405278584

Educational and Psychological Measurement
Volume 66 Number 3
June 2006 435-447

10.1177/0013164405278584
http://epm.sagepub.com
hosted at
http://online.sagepub.com

Suppression Situations in Multiple Linear Regression

Gwowen Shieh
National Chiao Tung University

This article proposes alternative expressions for the two most prevailing definitions of suppression without resorting to the standardized regression modeling. The formulation provides a simple basis for the examination of their relationship. For the two-predictor regression, the author demonstrates that the previous results in the literature are incomplete and oversimplified. The proposed approach also allows a natural extension for multiple regression with more than two predictor variables. It is shown that the conditions under which both types of suppression can occur are not fully congruent with the significance of the partial F test. This implies that all the standard variable selection techniques—backward elimination, forward selection, and stepwise regression procedures—can fail to detect suppression situations. This also explains the controversial findings in the redundancy or importance of correlated variables in applied settings. Furthermore, informative visual representations of various aspects of these phenomena are provided.

Keywords: *coefficient of multiple determination; extra sum of squares; partial* F *test; suppressor variable; variable selection*

Multiple regression analysis is one of the most widely used of all statistical methods. One of the purposes of multiple regression is to investigate the relative importance of a number of predictor variables for their relationship with a response variable. The predictor variables are typically correlated among themselves, and therefore, there is no simple answer concerning how to assess their individual contribution. Several measures have been proposed such as t values, standardized regression coefficients, increments in R^2, and correlation coefficients. Related comments and discussions can be found in Bring (1995, 1996) and their references. Furthermore, the dominance analysis proposed by Budescu (1993) for relative importance of two predictors is based on the comparison of their added values of R^2 in all possible subset models. Also, see Azen and Budescu (2003) for direct extension and further details.

Author's Note: This work was partially supported by the National Science Council under Grant NSC-92-2118-M-009-009. The author thanks the referees for helpful comments that improved the presentation of this article. Please address correspondence concerning this article to Gwowen Shieh, Department of Management Science, National Chiao Tung University, Hsinchu, Taiwan 30050, R.O.C.; e-mail: gwshieh@mail.nctu.edu.tw.

In this article, we focus on the concept of suppression that occurred when comparing the contribution of an individual predictor variable with and without the presence of other predictor variables. Since Horst (1941) first discussed that a predictor variable can be totally uncorrelated with the response variable and still improves prediction by virtue of being correlated with other predictors, much discussion has been made concerning the concept of suppression in behavioral sciences. For detailed reviews of different approaches to defining suppression, see Conger (1974), Velicer (1978), Tzelgov and Henik (1981, 1991), Holling (1983), and Smith, Ager, and Williams (1992).

In this study, we are especially concerned with the definitions proposed by Conger (1974) and Velicer (1978) because they have drawn the most attention in both behavioral and statistical research. Essentially, Conger's (1974) definition is based on the standardized regression coefficient and simple correlation, whereas Velicer's (1978) definition is referred to as the squared multiple and simple correlations or equivalently the increment in R^2. As Pedhazur (1997) pointed out, the definition and interpretation of suppression, however, remain controversial. This is partly because the definitions of Conger (1974) and Velicer (1978) are fundamentally different with respect to model formulation. Specifically, the comparisons between these two definitions are restricted to the special case of the standardized regression model with two predictors as shown in Tzelgov and Henik (1981). Their results suggest that the cases of Conger's (1974) suppression situations subsume those under Velicer's (1978) definition. However, it is not clear exactly how and when such phenomena can happen with respect to the interrelation of the response and two predictor variables. Although the aforementioned articles intended to address different aspects of the two definitions, it seems that the arguments are not settled. The major problem is the lack of schematic approach to the examination and comparison of the two definitions. The needed approach should be general enough to lay the same basis for the comparability of the two definitions and at the same time should be precise enough to provide both concrete demonstration and visual representation for their similarities and differences.

It should be noted that Velicer's (1978) definition of suppression in terms of the increment in R^2 possesses several important practical advantages relative to that of Conger (1974). First, R^2 and its partitions represent the proportions of variation accounted for by the predictor variables. Second, it solves the problem of incomparability that exists in the definition of Conger (1974). Third, the definition is always reciprocal so that the occurrence of suppression does not depend on the order of the predictors. Last, such formulation can be extended immediately from the special case of two predictors to the general p predictor case ($p > 2$). Interestingly, the type of suppression studied in statistical literature is in agreement with the definition of Velicer (1978). The geometric description, numerical example, and algebraic argument for the two-predictor regression have been given in Schey (1993); Neter, Kutner, Nachtsheim, and Wasserman (1996); and Sharpe and Roberts (1997), respectively. However, there is no extension beyond the two-predictor case. In relation to the notion of increase in R^2, the partial F test is the standard procedure for selecting important predictors. It should be extremely informative to clarify the relationship of both definitions of suppression with the partial F test in a general framework of multiple regression.

This article aims to provide alternative expressions of Conger's (1974) and Velicer's (1978) definitions of suppression that not only take into account the problem of comparability but also accommodate the extension to general multiple regression and permit thorough investigation of their relationship algebraically and graphically. In the next section, we provide the revised constructions of the two criteria of suppression and present the important details of the conditions under which different types of suppression can occur. In the third section, the concept of suppression is contrasted with the detection of important predictors in terms of the partial F test in variable selection. Finally, the fourth section contains some final remarks.

Two Definitions of Suppression

Consider the general linear regression model with response variable Y and p (≥ 2) predictor variables $X_1, \ldots, X_p$:

$$Y_i = \beta_0 + \sum_{j=1}^{p} X_{ij}\beta_j + \varepsilon_i, \qquad i = 1, \ldots, n, \tag{1}$$

where Y_i is the value of the response variable; $\beta_0, \beta_1, \ldots, \beta_p$ are parameters; $X_{i1}, \ldots, X_{ip}$ are the known constants of predictors $X_1, \ldots, X_p$; and ε_i are *iid* $N(0, \sigma^2)$ random variables. We are interested in the occurrence of suppression in the general linear regression model (Equation 1). First, consider the definition of suppression defined by Conger (1974) as follows. A suppression situation exists whenever

$$\hat{\beta}_j^{*2} > r_{Yj}^2, \tag{2}$$

for some j, $j = 1, \ldots, p$, where $\hat{\beta}_j^*$ is the least squares estimator of the standardized regression coefficient (beta weight, beta coefficient) and r is the coefficient of correlation between Y and X_j.

Next, Velicer (1978) defined a suppression situation in terms of the squared multiple and simple correlations:

$$R^2 > R_{Yh}^2 + r_{Yj}^2, \tag{3}$$

for some j, $j = 1, \ldots, p$, where R^2 is the squared multiple correlation coefficient of Y with ($X_1, \ldots, X_p$), and R_{Yh}^2 is the squared multiple correlation coefficient of Y with ($X_1, \ldots, X_{j-1}, X_{j+1}, \ldots, X_p$); that is, X_j is omitted from ($X_1, \ldots, X_p$).

At first sight, the two definitions given in Equations 2 and 3 may be fundamentally different. However, they are intertwined and are closely related. For the purpose of demonstrating their similarities and differences, we propose to consider two alternative formulations of suppression for providing important connection between Conger's (1974) and Velicer's (1978) definitions.

Definition 1. For regression model (1), a C-suppression situation exists if

$$\hat{\beta}_j^2 > \tilde{\beta}_j^2, \tag{4}$$

for some j, $j = 1, \ldots, p$, where $\hat{\beta}_j$ is the usual least squares estimator of the regression coefficient β_j in Equation 1 and $\tilde{\beta}_j$ is the least squares estimator of the slope coefficient for the simple regression model of response Y and predictor X_j. It is important to note that throughout this article, we assume regression model (1) is applicable.

It is well known that the least squares estimators of the standardized regression coefficients associated with model (1) can be written as $\hat{\beta}_j^* = \hat{\beta}_j(s_j / s_Y)$, with s_Y and s_j being the respective square root of s_Y^2 and s_j^2, where

$$s_Y^2 = \sum_{i=1}^{n}(Y_i - \overline{Y})^2 / (n-1), \qquad s_j^2 = \sum_{i=1}^{n}(X_{ij} - \overline{X}_j)^2 / (n-1),$$

and, $\overline{Y}$ and $\overline{X}_j$ are the respective sample means of the Y and the X_j observations. Moreover, the coefficient of correlation between Y and X_j can be expressed as $r_{Yj} = \tilde{\beta}_j(s_j/s_Y)$ with respect to the simple linear regression for Y with X_j. Hence, the condition (2) of suppression proposed in Conger (1974) is equivalent to $\hat{\beta}_j^2 > \tilde{\beta}_j^2$, which is exactly the condition of C-suppression defined in Equation 4. Note that our formulation in Definition 1 is not limited to the case of standardized regression, and consequently Definition 1 subsumes Conger's definition as a special case. Essentially, the proposed definition of C-suppression in Equation 4 solves the comparability problem for the differences in the range of $\hat{\beta}_j^*$ and r_{Yj} raised by Velicer (1978) that $|r_{Yj}|$ is bounded by unity and $|\hat{\beta}_j^*|$ is not. Furthermore, it provides a useful connection with Velicer's definition of suppression shown next.

Along the same line of comparability issue, a little reflection should make one wary of the comparison of $\hat{\beta}_j$ and $\tilde{\beta}_j$ because of the differences in the associated variances. To permit comparison of the estimated regression coefficients $\hat{\beta}_j$ and $\tilde{\beta}_j$ in the same units, the adjustment with respect to the estimated variance is employed in the following formulation.

Definition 2. For regression model (1), a V-suppression situation exists if

$$t_j^2 > \tilde{t}_j^2, \tag{5}$$

for some j, $j = 1, \ldots, p$, where

$$t_j = \frac{\hat{\beta}_j}{\left\{\hat{V}(\hat{\beta}_j)\right\}^{1/2}} \text{ and } \tilde{t}_j = \frac{\tilde{\beta}_j}{\left\{\hat{V}(\tilde{\beta}_j)\right\}^{1/2}},$$

and $\hat{V}(\hat{\beta}_j)$ and $\hat{V}(\tilde{\beta}_j)$ are the estimated variance of $\hat{\beta}_j$ and $\tilde{\beta}_j$, respectively.

Under the model assumption (1), it can be shown that

$$\hat{V}(\hat{\beta}_j) = \frac{\hat{\sigma}^2}{(n-1)s_j^2(1-R_{jh}^2)} \text{ and } \hat{V}(\tilde{\beta}_j) = \frac{\hat{\sigma}^2}{(n-1)s_j^2}, \tag{6}$$

where $\hat{\sigma}^2 = SSE/(n-p-1)$ is the estimator of σ^2, SSE is the usual error sums of squares, and R_{jh}^2 is the squared multiple correlation coefficient of X_j with $(X_1, \ldots, X_{j-1}, X_{j+1}, \ldots, X_p)$. Note that both t_j and $\tilde{t}_j$ have a noncentral t distribution with $n - p - 1$ degrees of freedom for $\beta_j \neq 0$, $j = 1, \ldots, p$. Also, the estimators $\hat{\beta}_j$ and $\tilde{\beta}_j$ can be expressed as follows:

$$\hat{\beta}_j = \frac{r_{Y(j.h)}}{(1-R_{jh}^2)^{1/2}} \times \left(\frac{s_Y}{s_j}\right) \text{and } \tilde{\beta}_j = r_{Yj}\left(\frac{s_Y}{s_j}\right), \tag{7}$$

where $r_{Y(j,h)}$ is the semipartial correlation coefficient of Y with X_j and with X_j adjusted for $(X_1, \ldots, X_{j-1}, X_{j+1}, \ldots, X_p)$. Using the results in Equations 6 and 7, the condition (5) of V-suppression could be formulated as

$$r_{Y(j.h)}^2 > r_{Yj}^2, \tag{8}$$

and equivalently, $R^2 - R_{Yh}^2 > r_{Yj}^2$ for $r_{Y(j.h)}^2 = R^2 - R_{Yh}^2$. For more detailed discussions of partial and semipartial correlations, see Pedhazur (1997, chap. 7). Consequently, we can see that Definition 2 of V-suppression defined in Equation 5 is the same as Velicer's (1978) definition of suppression given in Equation 3. However, we believe that the revised expressions for C-suppression in Equation 4 and V-suppression in Equation 5 are more appealing than other approaches of conceiving the conceptual relationship of Conger's (1974) and Velicer's (1978) definitions of suppression. Furthermore, the mathematical relationship between Conger's (1974) and Velicer's (1978) definitions of suppression is facilitated by considering the alternative form of Equation 4 for C-suppression. Equation 7 enables us to rewrite Equation 4 as follows:

$$\frac{r_{Y(j.h)}^2}{1-R_{jh}^2} > r_{Yj}^2. \tag{9}$$

Because $r_{Y(j.h)}^2 > r_{Yj}^2$, implies $r_{Y(j.h)}^2(1-R_{jh}^2) > r_{Yj}^2$ for $R_{jh}^2 \in [0,1)$, it follows from Equations 8 and 9 that the occurrences of C-suppression subsume those of V-suppression as special cases. Comparing the expressions in Equations 9 and 8 for C- and V-suppression, respectively, we see that Equation 9 includes the extra term $(1-R_{jh}^2)$ and incurs the problem of incomparability in Conger's (1974) definition of suppression that was criticized by Velicer (1978).

To understand the features of both types of suppression defined above, we begin by focusing on the case of two-predictor regression and then extend the discussion to the general multiple regression situations.

Two-Predictor Regression

For $p = 2$, model (1) reduces to

$$Y_i = \beta_0 + X_{i1}\beta_1 + X_{i2}\beta_2 + \varepsilon_i, \qquad i = 1, \ldots, n.$$

In this case, it follows from Equation 4 that a C-suppression situation exists if

$$\frac{(r_{Y2} - r_{12}r_{Y1})^2}{(1 - r_{12}^2)^2} > r_{Y2}^2, \tag{10}$$

or

$$\frac{(r_{Y1} - r_{12}r_{Y2})^2}{(1 - r_{12}^2)^2} > r_{Y1}^2, \tag{11}$$

where r_{12} is the coefficient of correlation between X_1 and X_2. To lay the basis for developing a simplified view and providing a concise visualization of the suppression situations, we define

$$\gamma = r_{Y2}/r_{Y1}.$$

Because the designation of X_1 and X_2 is arbitrary, as long as only one of r_{Y1} and r_{Y2} is zero, γ can be set as zero. The case in which both r_{Y1} and r_{Y2} are zero will be excluded because all the least squares estimators $\hat{\beta}_1, \hat{\beta}_2, \tilde{\beta}_1$ and $\tilde{\beta}_2$ are obviously zero without practical meaning. We are especially concerned with the cases in which X_1 alone, X_2 alone, or both are suppressors. By definition, predictors X_1 and X_2 are suppressors with respect to C-suppression if conditions in Equations 10 and 11 hold, respectively. In terms of the definition of γ, it can be shown that both predictors X_1 and X_2 are suppressors simultaneously if $(1 - r_{12}/\gamma)^2 > (1 - r_{12}^2)^2$ and $(1 - r_{12}\gamma)^2 > (1 - r_{12}^2)^2$. Alternatively, there is mutual or reciprocal C-suppression if

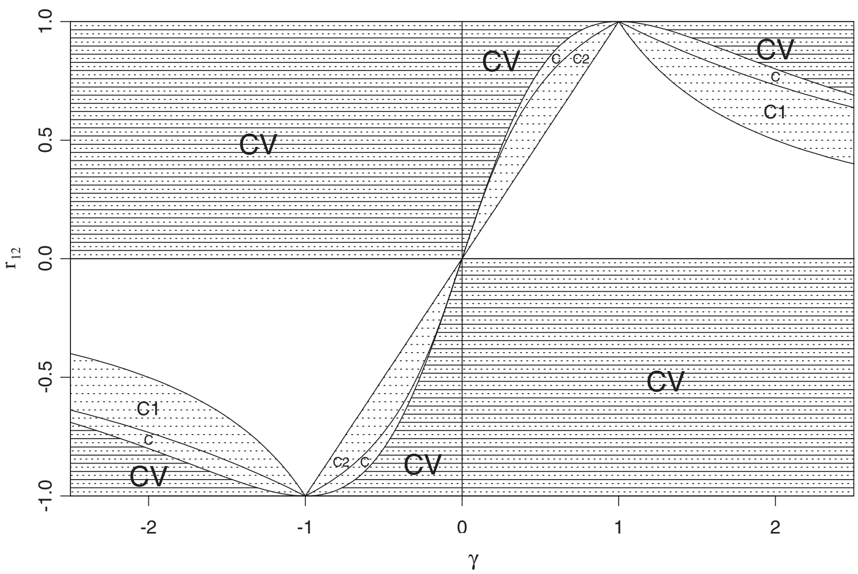
$$\gamma < r_{12}/(2 - r_{12}^2) \text{ or } \gamma > (2 - r_{12}^2)/r_{12} \text{ for } 0 < r_{12} < 1;$$

$$\gamma < (2 - r_{12}^2)/r_{12} \text{ or } \gamma > r_{12}/(2 - r_{12}^2) \text{ for } -1 < r_{12} < 0.$$

Predictor X_1 is the only suppressor if $(1 - r_{12}/\gamma)^2 > (1 - r_{12}^2)^2$ and $(1 - r_{12}\gamma)^2 < (1 - r_{12}^2)^2$. Therefore, it is straightforward to show that the ranges for γ are

$$1/r_{12} < \gamma < (2 - r_{12}^2)/r_{12} \text{ for } 0 < r_{12} < 1;$$

Figure 1
The Regions of Different Types of Suppression



$$(2 - r_{12}^2)/r_{12} < \gamma < 1/r_{12} \text{ for } -1 < r_{12} < 0.$$

On the other hand, predictor X_2 is the only suppressor if $(1 - r_{12}/\gamma)^2 < (1 - r_{12}^2)^2$ and $(1 - r_{12}\gamma)^2 > (1 - r_{12}^2)^2$. Correspondingly, these two conditions yield the following ranges for γ:

$$r_{12}/(2 - r_{12}^2) < \gamma < r_{12} \text{ for } 0 < r_{12} < 1;$$

$$r_{12} < \gamma < r_{12}/(2 - r_{12}^2) \text{ for } -1 < r_{12} < 0.$$

Figure 1 presents the occurrence of C-suppression for combinations of r_{12} and γ. The dotted areas stand for the occurrence regions of C-suppression. The areas marked with "C" represent the occurrence of mutual C-suppression. Those areas marked with "C1" or "C2" represent the occurrences of single C-suppression with X_1 or X_2 as the only suppressor, respectively. We believe that Figure 1 can communicate the results of C-suppression more effectively than the respective Figure 1 in Conger (1974) or Tzelgov and Henik (1991), in which the identification of suppression was not directly related to correlations or was indirectly presented with selected values of γ.

For the occurrence of the V-suppression situation, it can be shown that the condition of Equation 5 reduces to

$$\frac{(r_{Y2}-r_{12}r_{Y1}^2)^2}{1-r_{12}^2} > r_{Y2}^2 \text{ or } \frac{(r_{Y1}-r_{12}r_{Y2})^2}{1-r_{12}^2} > r_{Y1}^2. \quad (12)$$

Alternatively, the condition of Equation 12 of V-suppression in terms of r_{12} and γ is

$$\gamma < (1-\sqrt{1-r_{12}^2})/r_{12} \text{ or } \gamma > (1+\sqrt{1-r_{12}^2})/r_{12} \quad \text{for } 0 < r_{12} < 1;$$

$$\gamma < (1+\sqrt{1-r_{12}^2})/r_{12} \text{ or } \gamma > (1-\sqrt{1-r_{12}^2})/r_{12} \quad \text{for } -1 < r_{12} < 0.$$

With two predictors (p = 2), it is easy to see that Equation 3 corresponds to the inequality between the coefficient of determination and the sum of two squared simple correlation coefficients: $R^2 > r_{Y1}^2 + r_{Y2}^2$ or the inequality between the extra sum of squares and the sum of squares for simple regression. Related comments and discussions can be found in Currie and Korabinski (1984), Hamilton (1987), Bertrand and Holder (1988), Schey (1993), Sharpe and Roberts (1997), and Shieh (2001). However, these articles do not cover the relationship between different definitions of suppression.

Whereas C-suppression may be single or mutual, it is important to note that V-suppression is always mutual (see Velicer, 1978). As a visual supplement, Figure 1 also presents the occurrence of V-suppression by areas covered in horizontal lines and marked with a "V." It can be seen from the plot that V-suppression is a subset of C-suppression, as pointed out in Tzelgov and Henik (1981). Nevertheless, it can be more precise that V-suppression is encompassed by mutual C-suppression as a special case, although their differences are marginal. This reveals that the figure in Tzelgov and Henik (1981) is oversimplified and questionable. Furthermore, it should be clear that our Figure 1 conceives the occurrences of different suppressions more effectively than Figure 2 of Tzelgov and Henik (1991).

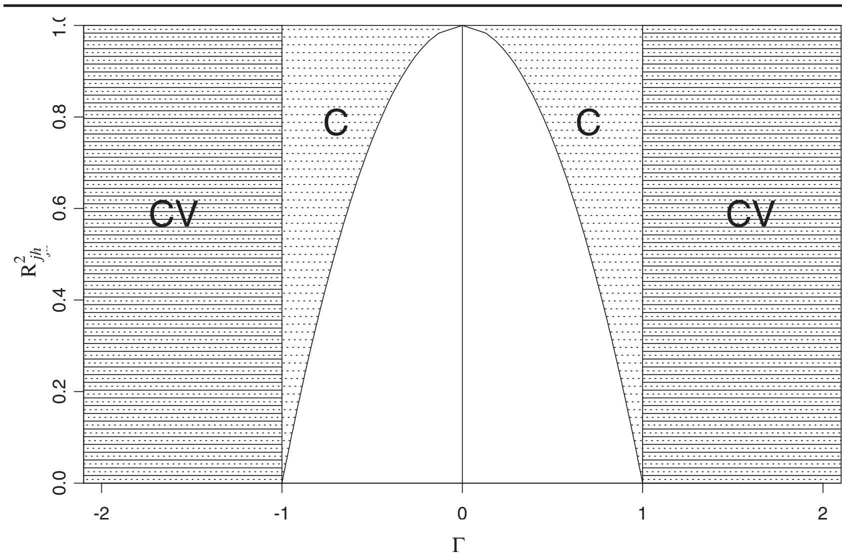
Multiple Regression

We consider the general setup of multiple regression with at least three predictors in the model. For the purpose of permitting clear and informative visual representation, we define

$$\Gamma = r_{Y(j.h)}/r_{Yj}.$$

It follows from Equation 9 that the condition of C-suppression is $\Gamma^2 > (1 - R_{jh}^2)$, whereas the condition of V-suppression in Equation 8 is $\Gamma^2 > 1$ with proper r_{Yj} and R_{jh}^2. The relation between both types of suppression is presented in Figure 2 for combinations of R_{jh}^2 and Γ, where "C" and "V" denote the C- and V-suppression situations, respectively. Similar results for Velicer's (1978) suppression situations have been shown in Smith et al. (1992), and they also extended the discussion to relations between two sets of predictors. However, the emphasis here is on the relation between different types of suppression. In addition, it should be noted that both definitions (4)

Figure 2
The Regions of C- and V-Suppression

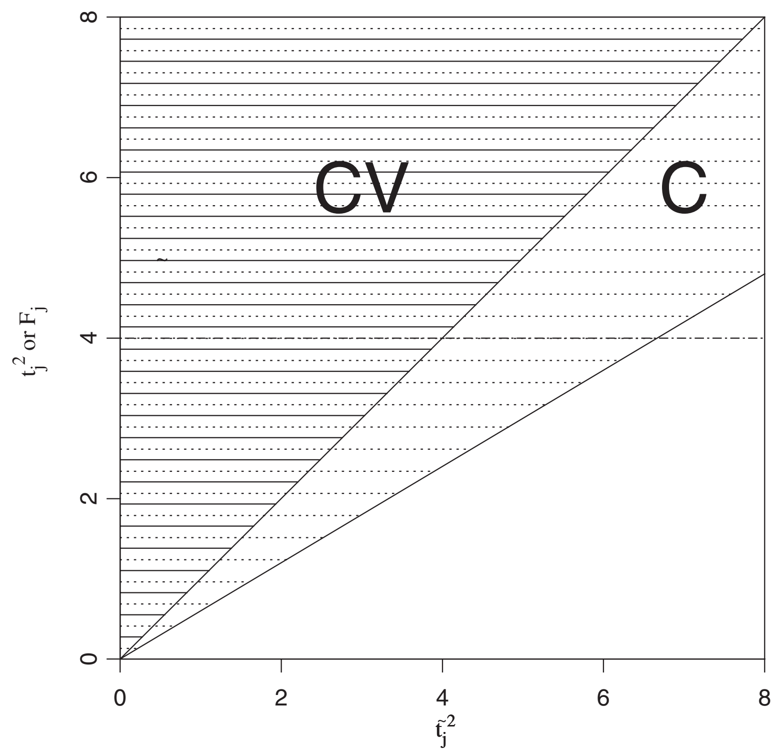


and (5) treat the $p - 1$ predictors ($X_1, \ldots, X_{j-1}, X_{j+1}, \ldots, X_p$) as a whole that leads to suppression.

Suppression and Variable Selection

In multiple regression, variable selection procedures are commonly used to identify the legitimate variables and discard those that are not useful. All the backward elimination, forward selection, and stepwise regression procedures are the typical algorithms for selecting the best subset of predictor variables. These procedures determine whether a predictor should be added to or deleted from the candidate set of predictor variables according to the significance or nonsignificance of the partial *F* test at each step. As shown in the previous section, the contribution of a predictor can be enhanced in the presence of other predictors for the intercorrelation or multicollinearity among them. Hamilton (1987) pointed out the inability of the forward selection technique to detect important predictors and recommended backward elimination as a more robust alternative for such seemingly paradoxical situations. Unlike these automatic search procedures, the all-possible-subsets regression procedure leads to the identification of a limited number of promising models. Then the final model can be selected according to a specified criterion in conjunction with the number of parameters involved. The notion of such a model selection process is somewhat different from that of the stepwise procedures described next.

Figure 3
The Regions of Suppression and Partial *F* Test



At each stage of the variable selection procedures with predictors $(X_1, \ldots, X_p)$, the partial F test statistic can be written as $F_j = t_j^2$, where F_j follows the F distribution with $(1, n-p-1)$ degrees of freedom or F^1_{n-p-1}, and t_j is defined in Equation 5, which follows a t distribution with $n-p-1$ degrees of freedom under the hypothesis $\beta_j = 0$. The predictor X_j is retained with the subset of predictors $(X_1, \ldots, X_{j-1}, X_{j+1}, \ldots, X_p)$ if $F_j > F(1, n-p-1, \alpha)$, where $F(1, n-p-1, \alpha)$ is the $(1-\alpha)$ percentile of F^1_{n-p-1}. From the previous results, the respective condition of C- and V-suppression can be expressed as

$$t_j^2 / \tilde{t}_j^2 < 1 - R_{jh}^2 \quad \text{and} \quad t_j^2 / \tilde{t}_j^2 > 1.$$

Obviously, the detection of an important predictor with the partial F test in all variable selection procedures is not completely compatible to the occurrences of both definitions of suppression. These phenomena are presented in Figure 3 for $R_{jh}^2 = 0.4$ and $F(1, n-p-1, \alpha) = t_j^2 = 4$. As in Figure 2, "C" and "V" denote the C- and V-suppression situations, respectively. The area below the dashed horizontal line represents the nonsignificant cases of the partial F test. Therefore, any of the variable selection pro-

Table 1
Numerical Illustration

Predictor X_j	X_1	X_2	X_3
$\hat{\beta}_j$	0.000763	1.192383	4.719083
$\tilde{\beta}_j$	–0.000388	2.559390	7.079554
$\hat{\beta}_j^*$	0.131528	0.391241	0.576485
r_{Yj}	–0.0670	0.8398	0.8648
r_{Yj}^2	0.0045	0.7052	0.7480
R_{Yh}^2	0.8020	0.7671	0.7103
t_j	1.20	2.12	3.08
$\tilde{t}_j$	0.63	7.88	8.12
F_j (p value)	1.44 (0.2480)	4.51 (0.0497)	9.51 (0.0071)
C-suppression ($\hat{\beta}_j^2 > \tilde{\beta}_j^2$)	Yes	No	No
Conger's (1974) definition ($\hat{\beta}_j^{*2} > r_{Yj}^2$)	Yes	No	No
V-suppression ($t_j^2 > \tilde{t}_j^2$)	Yes	No	No
Velicer's (1978) definition ($R^2 > R_{Yh}^2 + r_{Yj}^2$)	Yes	No	No
The significance of partial F test at $\alpha = .05$	No	Yes	Yes

cedures can fail to uncover the C- or V-suppression when there is a significant partial F test. On the other hand, it is possible to have either types of suppression even if the partial F test is nonsignificant. This observation is generally true for all $R_{jh}^2 \in [0, 1)$ and $F(1, n - p - 1, \alpha) > 0$. Hence, the inclination in Hamilton (1987) that backward elimination is satisfactory for those not wanting to miss enhancement or synergism is open to question. Note that his illustrations of $R^2 > r_{Y1}^2 + r_{Y2}^2$ is equivalent to the notion of V-suppression for $p = 2$ as shown in the Two-Predictor Regression section. Furthermore, Velicer's (1978) claim that his definition of suppression is consistent with stepwise regression procedures is doubtful. To exemplify these findings, we consider the problem given in Kleinbaum, Kupper, Muller, and Nizam (1998, pp. 126-127) in which a sociologist used data from 20 cities to investigate the relationship between the homicide rate per 100,000 city population (Y) and the following three independent variables: the city's population size (X_1), the percentage of families with yearly income less than \$5,000 (X_2), and the rate of unemployment (X_3). The numerical results are summarized in Table 1 for each variable. Note that the overall squared multiple correlation coefficient of Y with (X_1, X_2, X_3) is $R^2 = .8183$. In conclusion, the partial F test of a single parameter with respect to the increase in R^2 is not a profound indicator of the suppression situations in multiple regression.

Conclusions

In social studies, it is often the case that many of the variables are highly correlated. According to previous results, we wish to stress to users of multiple linear regression that the contribution of a variable can be enhanced by the presence of other variables. In general, it is not recommended that one discard variables that are highly correlated with the variables to be retained in the best subset.

In view of the discrepancy between different suppression definitions in the behavioral literature, the proposed C- and V-suppression are more appealing for several reasons. They simplify, clarify, and expand the existing formulations. More important, it leads to results that are not well known. With the added information of distinguishing between mutual and single suppressions algebraically and graphically, we are able to discern the complexities of the definitions of Conger (1974) and Velicer (1978). This study permits new insights into their definitions of suppression as to how they occur and when they differ. Although our presentation is concerned exclusively with the suppression situations in multiple regression, it can be applied easily to other designs (ANOVA and ANCOVA) and multivariate models.

We also present new characteristics about the contribution of each variable in multiple regression. Recognition of the relationship between suppression situations and the partial F test helps clarify the issue of variable selection. This information should be useful in screening existing measurements and designing new ones. In fact, the occurrence of suppression and the significance of a partial F test can be employed simultaneously in the stepwise technique of variable selection. Further investigation and verification of this combined approach under a variety of different applications would be useful.

References

Azen, R., & Budescu, D. V. (2003). The dominance analysis approach for comparing predictors in multiple regression. *Psychological Methods*, *8*, 129-148.

Bertrand, P. V., & Holder, R. L. (1988). A quirk in multiple regression: The whole regression can be greater than the sum of its parts. *The Statistician*, *37*, 371-374.

Budescu, D. V. (1993). Dominance analysis: A new approach to the problem of relative importance of predictors in multiple regression. *Psychological Bulletin*, *114*, 542-551.

Bring, J. (1995). Variable importance by partitioning R^2. *Quality and Quantity*, *29*, 173-189.

Bring, J. (1996). A geometric approach to compare variables in a regression model. *The American Statistician*, *50*, 57-62.

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

Currie, I., & Korabinski, A. (1984). Some comments on bivariate regression. *The Statistician*, *33*, 283-293.

Hamilton, D. (1987). Sometimes $R^2 > r_{yx1}^2 + r_{yx2}^2$, correlated variables are not always redundant. *The American Statistician*, *41*, 129-132.

Holling, H. (1983). Suppressor structures in the general linear model. *Educational and Psychological Measurement*, *43*, 1-9.

Horst, P. (1941). The prediction of personal adjustment. In *Social Science Research Council bulletin* (Vol. 48). New York: Social Science Research Council.

Kleinbaum, D. G., Kupper, L. L., Muller, K. E., & Nizam, A. (1998). *Applied regression analysis and other multivariate methods* (3rd ed.). Pacific Grove, CA: Brooks/Cole.

Neter, J., Kutner, M. H., Nachtsheim, C. J., & Wasserman, W. (1996). *Applied linear statistical models* (4th ed.). Homewood, IL: Irwin.

Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction* (3rd ed.). New York: Holt, Rinehart & Winston.

Schey, H. M. (1993). The relationship between the magnitudes of SSR(x_2) and SSR($x_2|x_1$): A geometric description. *The American Statistician*, *47*, 26-30.

Sharpe, N. R., & Roberts, R. A. (1997). The relationship among sums of squares, correlation coefficients, and suppression. *The American Statistician*, *51*, 46-48.

Shieh, G. (2001). The inequality between the coefficient of determination and the sum of squared simple correlation coefficients. *The American Statistician*, *55*, 121-124.

Smith, R. L., Ager, J. W., & Williams, D. L. (1992). Suppressor variables in multiple regression/correlation. *Educational and Psychological Measurement*, *52*, 17-29.

Tzelgov, J., & Henik, A. (1981). On the differences between Conger's and Velicer's definitions of suppressor. *Educational and Psychological Measurement*, *41*, 1027-1031.

Tzelgov, J., & Henik, A. (1991). Suppression situations in psychological research: Definition, implications, and applications. *Psychological Bulletin*, *109*, 524-536.

Velicer, W. F. (1978). Suppressor variables and the semipartial correlation coefficient. *Educational and Psychological Measurement*, *38*, 953-958.