

The Dominance Analysis Approach for Comparing Predictors in Multiple Regression

Razia Azen

University of Wisconsin—Milwaukee

David V. Budescu

University of Illinois at Urbana—Champaign

A general method is presented for comparing the relative importance of predictors in multiple regression. Dominance analysis (D. V. Budescu, 1993), a procedure that is based on an examination of the R^2 values for all possible subset models, is refined and extended by introducing several quantitative measures of dominance that differ in the strictness of the dominance definition. These are shown to be intuitive, meaningful, and informative measures that can address a variety of research questions pertaining to predictor importance. The bootstrap is used to assess the stability of dominance results across repeated sampling, and it is shown that these methods provide the researcher with more insights into the pattern of importance in a set of predictors than were previously available.

The purpose of multiple regression (MR) is to predict or explain criterion (response) values from several well-selected predictors. Social scientists often use MR models to characterize the nature and degree of relationship between one response variable (Y) and multiple predictors (X_1, X_2, \dots, X_q). In this process, it makes sense to distinguish between two separate stages: model selection and predictor comparison (see Azen, Budescu, & Reiser, 2001; Budescu, 1993).

In the model selection stage one identifies the “correct” model, or a subset of p predictors from the full set of q predictors that provides the most adequate description of the response variable, Y . Pedhazur (1997) distinguishes between two general approaches to MR, referred to as *explanation* and *prediction*, that influence model selection. A researcher who uses the explanation approach is guided by prior theory or substantive research to identify the “true” predictors and determine the choice of model. In this case the re-

searcher’s goal is to confirm a specific theory. In contrast, a researcher who uses the prediction approach is simply attempting to find the most predictive model and is less concerned with whether the specific variables included in such a model are the true predictors. This can be achieved through one of numerous statistical model selection procedures (see, e.g., Cook & Weisberg, 1999; Hocking, 1976; Judge, Griffiths, Hill, Lutkepohl, & Lee, 1985; Miller, 1990; Neter, Kutner, Nachtsheim, & Wasserman, 1996; Thall, Russell, & Simon, 1997). In some cases researchers may use a combination of these two research goals by identifying one or several predictors based on theoretical considerations and additional predictors based on exploratory or predictive considerations. Regardless of the approach, once a model is selected it is treated as the correct model with all the implications of this assertion (e.g., that the sample estimates of the population parameters are unbiased).

The predictor comparison stage, often used for interpretation of the selected model, is the focus of this article. In this stage, researchers often wish to assess “predictor importance” by comparing the contributions made by the individual predictors in the selected model to prediction of the response. Although one may argue that if the chosen model is correct, then each and every predictor in it is equally important (for a discussion of this, see Azen et al., 2001), researchers nonetheless often wish to establish the relative importance of the predictors included in this model. The quest to determine the predictors’ relative importance is very natural in many contexts and with many theo-

Razia Azen, Department of Educational Psychology, University of Wisconsin—Milwaukee; David V. Budescu, Department of Psychology, University of Illinois at Urbana—Champaign.

Software for the computations presented in this article is available, in the form of an SAS macro, from Razia Azen and is available in the online version of this article in the PsycARTICLES database.

Correspondence concerning this article should be addressed to Razia Azen, Department of Educational Psychology, University of Wisconsin—Milwaukee, P.O. Box 413, Milwaukee, Wisconsin 53201. E-mail: azen@uwm.edu

ries. Naturally, the comparison method should fit the research questions as well as the purpose of the analysis, and this is discussed in detail later in this article (in the section *The Need for an Intuitive Measure of Importance*).

This article addresses the following general question: Given any pair of predictors from the chosen set, which would be more useful in predicting the criterion? For example, it makes sense to inquire first what variables predict a child's IQ and then determine their relative importance. One may start by considering a long list of correlated predictors reflecting such characteristics as the family's intellectual and educational background (e.g., the father's IQ, the father's education, the mother's IQ, the mother's education), the child's home (the parents' income, the number of older and younger siblings), and the child's community (socioeconomic status of the neighborhood, quality of the school attended, average class size, etc.). Once a parsimonious set of predictors is identified (according to the researcher's choice of a model selection process), the researcher may wish to rank the predictors according to their contribution in predicting the child's IQ. For example, if parents' income, quality of school, and mother's education are selected, and, of these, if parents' income is consistently a better predictor than each of the others both when considered by itself or with any subset of the others, the researcher can be quite confident that parents' income is the most dominant predictor of a child's IQ. Another example may involve social psychology researchers who are interested in predictors of satisfaction with life (e.g., Suh, Diener, Oishi, & Triandis, 1998). Variables thought to be predictive (based on theory or empirical findings) may include satisfaction with one's health, financial situation, family, housing, and self, and researchers want to determine which of these predictors is most important in predicting satisfaction with life. If, for example, satisfaction with self affects overall life satisfaction more strongly than any other predictor by itself as well as within the context of any subset of the remaining predictors, the answer is clear. If one then needed to choose to address one specific satisfaction domain for improving general life satisfaction, the choice would be clear. Taking this example further, if the researchers believe that on theoretical grounds satisfaction with self and family must be included as predictors, they may still want to ask: How do the remaining predictors compare in the context of these two given predictors? Is satisfaction with financial situation, housing, or health the most

dominant (useful) predictor of satisfaction with life after controlling for satisfaction with self and family? Of course, one may already have a specific theory regarding which predictor is more important, in which case the analysis can be used to confirm the theory's predictions rather than explore possible answers.

In this article we focus on the difficult task of measuring and interpreting predictor importance in the context of a selected set of correlated predictors. When correlated predictors are compared to each other, all sensible measures of relative importance will be affected by the other predictors (controlled for) in the model and, indirectly, by the predictors excluded from the model. There is currently no universally accepted definition of predictor importance, and the proper interpretation of the most commonly used measures (i.e., those provided automatically by popular statistical software programs) is often difficult, awkward, or subject to misinterpretation and misuse. In this article we propose an intuitive definition of predictor importance that leads to measures of importance that researchers can readily use and properly interpret in a variety of research contexts.

Importance Measures in MR

We begin with a partial review of some commonly used and recently advocated measures of predictor importance. Consider the standardized MR model in the population with p ($p \geq 1$) predictors and one criterion (i.e., Y, X_1, \dots, X_p are standardized to zero mean and unit variance so an intercept is not needed):

$$Y_j = \beta_1 X_{1j} + \beta_2 X_{2j} + \dots + \beta_p X_{pj} + e_j = \sum_{i=1}^p \beta_i X_{ij} + e_j \quad (1)$$

where β_i is the population standardized regression coefficient associated with the i th predictor; $e_j \sim N(0, \sigma^2)$; $i = 1, \dots, p$; $j = 1, \dots, n$. Suppose that we are operating on the population and that the model is correctly specified. Define the model's predicted values as

$$\hat{Y}_j = \sum_{i=1}^p \beta_i X_{ij},$$

and denote the squared correlation between the observed (Y) and predicted (\hat{Y}) values as $\rho_{Y\hat{Y}}^2$ (R^2 in the sample), a measure often used to quantify model fit and interpreted as the proportion of the variance in Y that can be reproduced, or accounted for, by the p predictors. Given a specific set of p correlated predic-

tors, how does one measure their individual relative contribution to prediction? Below we review a variety of importance measures. A summary of the geometrical interpretation of some importance measures is given in Bring (1996).

Importance Measures Based on Regression Coefficients

β_i . The regression coefficient β_i , associated with X_i in the MR Equation 1, represents the expected change in Y in standard deviation (SD) units associated with a change of 1 SD unit in X_i provided that all other predictors are held constant. In the special case of uncorrelated predictors, the i th (standardized) coefficient is equivalent to the correlation between the i th predictor and the criterion and provides an adequate measure of relative importance. Although a predictor whose coefficient is relatively large will presumably have a relatively larger effect on the prediction of the criterion, in the presence of correlated predictors this effect can only be interpreted when the effects of all other predictors are kept constant. Therefore, if the predictors are correlated, it may not be meaningful to think of the change in predictor X_i while all other predictors remain constant because a change in one predictor will most likely result in a change in all predictors correlated with it (e.g., Mosteller & Tukey, 1977).

Partially standardized β_i . Bring (1994) showed that removing the variable with the smallest standardized coefficient does not necessarily result in the smallest reduction in $\rho_{Y\hat{Y}}$. He suggests that the regular approach to standardization is inadequate because to standardize β_i , which measures the change in Y due to a change in X_i holding all other predictors constant, one uses the standard deviation s_i from the full sample and thus assumes that s_i is uniform at all levels of the other predictors. For example, using height and weight as predictors, the interpretation of the coefficient associated with height is valid only at a fixed level of weight, yet it is standardized using the standard deviation of heights from the full sample, which includes all weight levels. Therefore, Bring suggests that standardization be performed by using a partial standard deviation that is conditional on the "sub-samples" formed by holding variables constant. Bring further shows that when the regression coefficients are standardized by the partial standard deviations, the ratio between the t scores obtained for variable i and variable j is equivalent to the ratio between the standardized coefficients for variables i and j . The com-

parison of t values is useful because it is equivalent to the reduction in $\rho_{Y\hat{Y}}^2$ obtained by eliminating each of the variables in the ratio from the regression model.

δ^2 . Green, Carroll, and DeSarbo (1978) suggested another measure based on regression coefficients that uses Gibson's (1962) approach. Their importance measure consists of obtaining a set of orthonormal variables that best approximates the original set of predictors. The obtained (orthonormal) variables are standardized variables that are uncorrelated with each other. The squared regression coefficients from the regression of Y on the orthonormal variables are then computed. The proposed importance measure, δ^2 , decomposes each of these squared regression coefficients into original variable contributions while taking into account the correlations between each original variable and its corresponding orthonormal variable. The δ^2 values for all of the predictors sum to $\rho_{Y\hat{Y}}^2$. Johnson (2000) recently proposed a slight variation on this approach that produces more meaningful and interpretable results by correcting the way Green et al.'s measure is computed.

Importance Measures Based on Correlations

ρ_{YX_i} or $\rho_{YX_i}^2$. Perhaps the most obvious measure of importance is the simple product-moment correlation, ρ_{YX_i} , or the squared product-moment correlation, $\rho_{YX_i}^2$, between the criterion and each of the predictors. This value is equivalent to the simple standardized regression coefficient or the squared standardized regression coefficient, respectively, in the regression of Y on X_i alone. When all of the predictors are uncorrelated, the sum of the squared correlations is equal to the variation accounted for by all predictors; that is, $\sum_{i=1}^p \rho_{YX_i}^2 = \rho_{Y\hat{Y}}^2$, and each squared correlation constitutes a straightforward measure of the corresponding predictor's contribution to the total variance accounted for by the model. However, in the social sciences, predictor variables are rarely, if ever, uncorrelated, and the importance of the predictors might change dramatically once the intercorrelations between the predictors are taken into account. For example, consider a situation in which Y is to be modeled by three predictor variables, X_1 , X_2 , and X_3 . Suppose that X_1 and X_3 are each very highly correlated with Y , while X_2 is only moderately correlated with Y . Furthermore, suppose X_1 and X_3 are highly correlated with each other but are not highly correlated with X_2 . In this case, once X_1 is included in the regression model, X_3 will not make any additional contribution to the prediction of Y . Therefore, once X_1

is in the model, X_3 will be “less important” than X_2 , even though the correlation between Y and X_3 is higher than the correlation between Y and X_2 . Similarly, once X_3 is in the model, X_1 will appear to be less important than X_2 . To further complicate matters, it has been shown by several researchers (e.g., Shieh, 2001) that in a regression with two correlated predictors $\rho_{Y\hat{Y}}^2$ may exceed the sum of the squared correlations $\rho_{YX_1}^2 + \rho_{YX_2}^2$. Therefore, measures that take the predictors’ intercorrelations into account have been suggested to address importance in the case of correlated predictors.

Squared partial correlation. Kruskal (1987) suggested that sometimes the p predictors of a regression model may have some “natural” order by which they are to be entered into the model. If subscripts are assigned to the predictors according to this natural order, then X_1 enters first, followed by X_2 , and so on, up to X_p . In this case, X_1 accounts for a fraction of the variance in Y , which is measured by the squared correlation between X_1 and Y , $\rho_{YX_1}^2$. Of the remaining variance in Y (the variance not accounted for by X_1), a fraction $\rho_{YX_2 \cdot X_1}^2$ (the squared partial correlation between Y and X_2 after controlling for X_1) is accounted for by X_2 , and so on, until X_p accounts for a fraction $\rho_{YX_p \cdot X_1 X_2 \dots X_{(p-1)}}^2$ of whatever variance is unaccounted for by the other $p - 1$ predictors. The partial squared correlation coefficients thus represent measures of importance of the corresponding predictors relative to the specific ordering X_1, \dots, X_p . Kruskal further suggests that when a natural ordering is not available, this procedure can be repeated for all $p!$ orderings of the predictors and the importance measures taken as an average across the $p!$ measures obtained for each predictor. Lindeman, Merenda, and Gold (1980) advocated a similar procedure based on the squared semi-partial (part) correlations.

Information from squared partial correlation. Theil (1987) and Theil and Chung (1988) suggested taking Kruskal’s measure a step further. Because the unaccounted for variance, $1 - \rho_{Y\hat{Y}}^2$, is the product $(1 - \rho_{YX_1}^2)(1 - \rho_{YX_2 \cdot X_1}^2) \dots (1 - \rho_{YX_p \cdot X_1 X_2 \dots X_{(p-1)}}^2)$, taking the logarithm of each term turns this product into an additive decomposition; that is, the terms sum to $1 - \rho_{Y\hat{Y}}^2$. Furthermore, letting $I(x) = -\log_2(1 - x)$, $I(\rho_{Y\hat{Y}}^2) = I(\rho_{YX_1}^2) + I(\rho_{YX_2 \cdot X_1}^2) + \dots + I(\rho_{YX_p \cdot X_1 X_2 \dots X_{(p-1)}}^2)$, and $I(\rho_{Y\hat{Y}}^2)$ can be thought of as the total information on Y provided by the p predictors; that is, a larger value of $\rho_{Y\hat{Y}}^2$ results in a larger value of the function $I(\rho_{Y\hat{Y}}^2) = -\log_2(1 - \rho_{Y\hat{Y}}^2)$ and thus indicates that more “information” on Y is provided by the predictors. This allows

us to measure each predictor’s contribution to the total information, and, as Kruskal suggested, the average over all $p!$ orderings of the additive terms can be used when a natural ordering of the predictors cannot be determined.

Commonality analysis. Another variation on measuring importance through partial correlations is referred to as *commonality analysis* (Mayeske et al., 1969; Mood, 1969, 1971; Newton & Spurrell, 1967a, 1967b), a procedure by which the variance accounted for in the criterion is partitioned into two parts. The first is a unique part, attributable to the predictors individually, which is essentially the partial contribution of each predictor to the squared multiple correlation with the criterion. The second is a common part, attributable to a combination of the predictors, which is the contribution to the multiple correlation with the criterion that all of the predictors in the combination share. Presumably, variables with small uniquenesses and large commonalities can be eliminated with the smallest reduction in overall variance accounted for (Hedges & Olkin, 1981; Pedhazur, 1975, 1997). Some of the problems with this method are that higher order commonalities (in the case of more than two predictors) and negative commonality elements (e.g., in the presence of suppressor variables) are difficult to interpret (Pedhazur, 1997).

Importance Measures Based on a Combination of the Regression Coefficients and the Correlations

$\beta_i \rho_{X_i Y}$. Pratt (1987) analyzed the well-known (e.g., Darlington, 1968) measure $\beta_i \rho_{X_i Y}$, which is the product of the standardized regression coefficient of the predictor and the predictor’s correlation with the criterion. The sum of this measure across all predictors is the multiple correlation coefficient; that is, $\sum_{i=1}^p \beta_i \rho_{X_i Y} = \rho_{Y\hat{Y}}^2$. Recently, this measure has been advocated by Thomas and Zumbo (1996) and Thomas, Hughes, and Zumbo (1998).

$\rho_{X_i \hat{Y}}$. Another measure that combines the correlation and regression coefficients is

$$\rho_{X_i \hat{Y}} = \frac{\rho_{X_i Y}}{\rho_{Y\hat{Y}}},$$

which is also known as the *structure coefficient* and recently has been advocated by Dunlap and Landis (1998) as well as by Courville and Thompson (2001).

The Need for an Intuitive Measure of Importance

The major weakness of most approaches reviewed is the lack of an intuitive and straightforward interpretation of the measures in a general context. The wide variety of measures may clearly result in different orderings of the predictors' importance, leading to confusion as to the meaning of importance. This often leads to misinterpretation, misuse, and abuse of these measures (Courville & Thompson, 2001, illustrate such problems in several published studies).

For example, students and researchers are often confused by results indicating that a regression coefficient is not statistically significant although the bivariate correlation of that predictor with the criterion is significant. This should not be surprising because the statistical test for the regression coefficient is performed in the presence of the effects of all other predictors, whereas the test for the bivariate correlation ignores all other predictors. Similarly, researchers are tempted to compare predictors on the basis of the significance or size of their corresponding regression coefficients (in the full model) and are confused when this comparison does not match the comparison made on the basis of the significance or size of the bivariate correlations between the predictors and the criterion (i.e., the significance of the simple regression coefficients), or of the significance and size of the coefficients in some subset model. The same can be said about the interpretation of partial or semipartial correlations. Although these measures are valuable, researchers are not necessarily well versed in their interpretation and thus have difficulties attempting to reconcile apparent contradictions among different measures and different subset models.

As with all statistical analyses, one's research question should dictate the most appropriate measure or procedure to use. Thus, if one's question concerns finding the single best predictor (i.e., the predictor that is most useful or predictive by itself), the simple correlation coefficient can be used as a measure of relative importance; if one's question concerns which predictor contributes most after accounting for the effects of all other predictors, the semipartial correlation or standardized regression coefficient (and associated *t* test) can be used; if one's question concerns which predictor contributes most to a given subset of predictors, then each predictor must be added individually to the base subset and its standardized regression coefficient (or associated *t* test) can be used; and so on. Questions of importance that encompass all

of these contexts, however, cannot be answered by any one of these measures. For example, suppose theory dictates that children's behavioral adjustment is predicted by their perceived social support. A researcher in this field may obtain measures on several sources of support, such as parental support, teacher support, peer support, general family support, and so on, with the goal of finding the most important source(s) of support in predicting adjustment. If the researcher finds that parental support contributes more highly to prediction than does peer support in the presence of all other support sources (using traditional measures), this does not indicate that parental support by itself is more predictive than peer support by itself or in the presence of any other subset of the remaining support sources. However, if parental support contributes more highly to prediction than does peer support, regardless of which set of the remaining predictors is present (controlled for), then parental support is shown to be an overall more dominant predictor than peer support. Thus, if researchers designing a treatment program to improve adjustment could focus on only one of these support sources, they would be more confident in focusing on parental support than on peer support. Similarly, a large university may wish to identify the predictors of academic success and then rank-order them according to their impact (see an example in the Discussion section). If it is found that one predictor (e.g., high school grade point average) is more important than another (e.g., Scholastic Aptitude Test score) singly as well as in the presence of any subset of the other predictors, the university can be quite confident in using that predictor (grade point average) over the other (Scholastic Aptitude Test score) to make admission decisions.

The questions of relative importance in the context of these examples are more general than those answered by traditional measures and, consequently, require an appropriately general measure of importance. Budescu (1993) argued that to translate the general concept of relative importance into a measurable and quantifiable statistical procedure there is a need to agree on the definition and meaning of *importance*.

Budescu (1993) developed a clear and intuitive definition that states that a predictor's importance reflects its contribution in the prediction of the criterion and that one predictor is "more important than another" if it contributes more to the prediction of the criterion than does its competitor at a given level of analysis. The level of analysis is the context of the subset model(s) that forms the basis for the compari-

son(s), and it is important because all measures of importance compare predictors within the context of a specific model (i.e., in the presence of a specific set of predictors). Thus, the level of analysis provides the context of comparison; for example, if the simple bivariate correlations of each predictor with the criterion are examined, the level of analysis ignores all other predictors (i.e., the context is the null model), but if the β coefficients (and possibly their associated t tests) are examined in the context of the full model (as is typical in statistical software output), the level of analysis includes all other predictors (i.e., the context is a subset model with $[p - 1]$ predictors). Further, if the comparison is carried out within a number of models (each containing some subset of the remaining predictors), the level of analysis includes all of these comparisons. Thus, the most specific level of analysis makes one comparison within the context of a specific model and the most general level of analysis makes all possible comparisons (one for each possible subset model).

In Budescu's (1993) approach, called *dominance analysis* (DA), the contribution to prediction was defined as the squared semipartial correlation, and the level of analysis was defined as all $2^{(p-2)}$ subset models for which the comparison of each pair of predictors is relevant. In other words, *one predictor is more important than another if it would be chosen over its competitor in all possible subset models where only one predictor of the pair is to be entered*. Thus, the quantification of importance depends on which set of predictors is involved in the analysis, a fact that is crucial for proper interpretation but often overlooked or misunderstood by researchers. The DA approach provides the most general context by taking into account all relevant subset models, where a relevant model is either any subset model that can be formed from the predictors or any subset model that is theoretically possible and of interest (see the *Constrained DA* section). One difficulty with DA is that dominance cannot always be established between every pair of predictors. In this article the original definition is refined and the levels of dominance extended so that dominance can be established in almost every case. An overview of DA including the definition of the levels of dominance follows.

Dominance Analysis

The standard measures reviewed earlier measure relative importance either for each variable alone

Table 1

Population Correlation Matrix for Hypothetical Example With Four Predictors

Variable	1	2	3	4	5
1. Y	—	0.6	0.3	0.4	0.5
2. X_1	0.6	—	0.8	0.1	0.3
3. X_2	0.3	0.8	—	0.1	0.1
4. X_3	0.4	0.1	0.1	—	0.5
5. X_4	0.5	0.3	0.1	0.5	—

(e.g., $\rho_{YX_i}^2$), in the presence of all $(p - 1)$ other predictors (e.g., β_i), or as an average across all $p!$ orderings of the predictors in the full model. DA is unique in that (a) it measures relative importance in a pairwise fashion, and (b) the two predictors are compared in the context of all $2^{(p-2)}$ models that contain some subset of the other predictors. As such, DA is more sensitive to the various importance patterns that can emerge in these cases and addresses a more general definition of importance.

For illustrative purposes, suppose that the correlation matrix shown in Table 1 was obtained for four predictors (X_1, X_2, X_3, X_4) used to predict a certain criterion. Several of the reviewed measures of importance are given for this example in Table 2. Note that the various measures result in a different ordering of relative importance¹ (i.e., β_i , semipartial ρ_i^2 and partial ρ_i^2 give the ordering X_1, X_2, X_3, X_4 ; ρ_{X_iY} , Kruskal's and Theil's average over all $p!$ orders and the structure coefficient ρ_{X_iY} give the ordering X_1, X_4, X_3, X_2 ; and $\beta_i \rho_{X_iY}$ implies X_1, X_3, X_4, X_2).

The additional contributions of each predictor to each subset model used in DA are shown in Table 3. We use the notation $\rho_{Y\mathbf{X}}^2$ to represent the proportion of variance in Y that is accounted for by the predictors in the model \mathbf{X} . The $2^4 = 16$ models and their corresponding $\rho_{Y\mathbf{X}}^2$ values are shown in the first two columns of Table 3; for example, $\rho_{YX_1X_3}^2 = 0.477$ represents the proportion of variance in Y that is accounted for by the model consisting of X_1 and X_3 . The *additional contribution* of a given predictor is measured by the increase in $\rho_{Y\hat{Y}}^2$ that results from adding that predictor to the regression model. Thus, the additional contributions of X_1 are computed as the increases in the proportion of variance accounted for when X_1 is added to each subset of the remaining

¹ Note that although all pairwise correlations are positive, β_2 is negative, making its interpretation very difficult.

Table 2
Measures of Importance in the Population for
Hypothetical Example With Four Predictors (Assuming
Standardized Variables)

Measure	Predictor (X_i)			
	X_1	X_2	X_3	X_4
β_i or B_i	0.905	-0.466	0.291	0.130
Semipartial ρ_i^2	0.246	0.071	0.061	0.010
Partial ρ_i^2	0.372	0.146	0.129	0.025
ρ_{X_iY}	0.6	0.3	0.4	0.5
Kruskal's (1987) average	0.362	0.115	0.132	0.144
Theil's (1987) average	0.651	0.177	0.206	0.233
$\rho_{X_i\hat{Y}}$	0.785	0.392	0.523	0.654
$\beta_i\rho_{X_iY}$	0.543	-0.140	0.117	0.065

predictors (i.e., the null subset $\{\cdot\}$, $\{X_2\}$, $\{X_3\}$, $\{X_4\}$, $\{X_2X_3\}$, $\{X_2X_4\}$, $\{X_3X_4\}$ and $\{X_2X_3X_4\}$). For example, the additional contribution of X_1 to the subset model $\{X_3\}$ is defined as $\rho_{Y(X_1X_3)}^2 - \rho_{YX_3}^2$. This is the difference between the proportion of variance in Y accounted for by both X_1 and X_3 and the proportion of variance in Y accounted for by X_3 alone. This value is also the squared semipartial correlation between X_1 and Y partialling out or controlling for the effect of X_3 on X_1 (i.e., $\rho_{Y(X_1X_3)}^2$; see Hays, 1994).

Similarly, the additional contributions of X_2 are the increases in the proportion of variance accounted for when X_2 is added to each subset of the remaining predictors (i.e., null subset $\{\cdot\}$, $\{X_1\}$, $\{X_3\}$, $\{X_4\}$, $\{X_1X_3\}$, $\{X_1X_4\}$, $\{X_3X_4\}$ and $\{X_1X_3X_4\}$). Note that in this example the only subset models to which *both* X_1 and X_2 make additional contributions are $\{\cdot\}$ (the null subset), $\{X_3\}$, $\{X_4\}$, and $\{X_3X_4\}$. Thus, the additional contribution of X_1 can be meaningfully compared with the additional contribution of X_2 only in the context of these common subset models. The additional contributions made by X_1 and X_2 to the null subset are simply $\rho_{YX_1}^2 = 0.360$ and $\rho_{YX_2}^2 = 0.090$, respectively. In the body of Table 3 these are the entries in the first row under the X_1 and X_2 columns, respectively. The additional contributions of X_1 and X_2 to the subset model $\{X_3\}$ are $\rho_{Y(X_1X_3)}^2 - \rho_{YX_3}^2 = 0.477 - 0.160 = 0.317$ and $\rho_{Y(X_2X_3)}^2 - \rho_{YX_3}^2 = 0.228 - 0.160 = 0.068$, respectively. In the body of Table 3 these are the entries in the row labeled X_3 under the X_1 and X_2 columns, respectively. For the model $\{X_4\}$, the additional contribution of X_1 is $\rho_{Y(X_1X_4)}^2 - \rho_{YX_4}^2 = 0.473 - 0.250 = 0.223$ and the additional contribution of X_2 is $\rho_{Y(X_2X_4)}^2 - \rho_{YX_4}^2 = 0.313 - 0.250 = 0.063$; the corresponding entries compared are in the row labeled X_4

under the X_1 and X_2 columns, respectively. For $\{X_3X_4\}$, the additional contribution of X_1 is $\rho_{Y(X_1X_3X_4)}^2 - \rho_{Y(X_3X_4)}^2 = 0.513 - 0.280 = 0.233$, the additional contribution of X_2 is $\rho_{Y(X_2X_3X_4)}^2 - \rho_{Y(X_3X_4)}^2 = 0.338 - 0.280 = 0.058$, and the entries compared are in the row labeled X_3X_4 under the X_1 and X_2 columns, respectively. Because the additional contribution of X_1 is larger than that of X_2 to each of these subset models, X_1 dominates X_2 (and has a higher relative importance). One predictor is said to *completely dominate* another if its additional contribution to each of the subset models that form the basis for comparison is greater than that of the other predictor.

This analysis can be repeated for each pair of predictors (the additional contributions of X_3 and X_4 are also given in Table 3, under the X_3 and X_4 columns, respectively). For example, to compare X_1 and X_3 , compare the columns of Table 3 labeled X_1 and X_3 : In the null model row, 0.360 (the additional contribution of X_1 to the null subset $\{\cdot\}$) is larger than 0.160 (the additional contribution of X_3 to the subset model $\{\cdot\}$); in the X_2 row, 0.360 (the additional contribution of X_1 to the subset model $\{X_2\}$) is larger than 0.138 (the additional contribution of X_3 to the subset model $\{X_2\}$); in the X_4 row, 0.223 (the additional contribution of X_1 to the subset model $\{X_4\}$) is larger than 0.030 (the additional contribution of X_3 to the subset model $\{X_4\}$); and in the X_2X_4 row, 0.210 (the additional contribution of X_1 to the subset model $\{X_2X_4\}$) is larger than 0.025 (the additional contribution of X_3 to the subset model $\{X_2X_4\}$). Therefore, X_1 completely dominates X_3 . Budescu (1993) shows that dominance is transitive, meaning that for every triplet of predictors, if X_i dominates X_j and X_j dominates X_k , then X_i must dominate X_k .

If one predictor's additional contribution is greater than the other's for some, but not all, of these subset models, complete dominance (and relative importance) cannot be established and is said to be undetermined. For example, the additional contribution of X_2 to the subset model $\{X_1\}$ (in the X_1 row and X_2 column of Table 3) is 0.090 and the additional contribution of 3 to the subset model $\{X_1\}$ (in the X_1 row and X_3 column of Table 3) is 0.117, suggesting that X_3 dominates X_2 . However, the additional contribution of X_2 to the subset model $\{X_4\}$ (in the X_4 row and X_2 column of Table 3) is 0.063, whereas the additional contribution of X_3 to the subset model $\{X_4\}$ (in the X_4 row and X_3 column of Table 3) is 0.030, suggesting that in this case X_2 dominates X_3 . Thus, complete dominance cannot be established between X_2 and

Table 3
Dominance Analysis in the Population for Hypothetical Example With Four Predictors

Subset model (\mathbf{X})	$\rho^2_{Y \cdot \mathbf{X}}$	Additional contribution of:			
		X_1	X_2	X_3	X_4
Null and $k = 0$ average	0	0.360	0.090	0.160	0.250
X_1	0.360		0.090	0.117	0.113
X_2	0.090	0.360		0.138	0.223
X_3	0.160	0.317	0.068		0.120
X_4	0.250	0.223	0.063	0.030	
$k = 1$ average		0.300	0.074	0.095	0.152
X_1X_2	0.450			0.124	0.073
X_1X_3	0.477		0.097		0.037
X_1X_4	0.473		0.051	0.041	
X_2X_3	0.228	0.346			0.110
X_2X_4	0.313	0.210		0.025	
X_3X_4	0.280	0.233	0.058		
$k = 2$ average		0.263	0.069	0.063	0.073
$X_1X_2X_3$	0.574				0.010
$X_1X_2X_4$	0.523			0.061	
$X_1X_3X_4$	0.513		0.071		
$X_2X_3X_4$	0.338	0.246			
$k = 3$ average		0.246	0.071	0.061	0.010
$X_1X_2X_3X_4$	0.584				
Overall average		0.292	0.076	0.095	0.121

Note. The column labeled $\rho^2_{Y \cdot \mathbf{X}}$ represents the variance in Y explained by the model appearing in the corresponding row. Columns labeled X_i contain the additional contributions to the explained variance gained by adding the column variable (X_i) to the row model. Blank cells indicate that data are not applicable.

X_3 —their relative importance changes depending on which subset model forms the basis of comparison. The results in Table 3 indicate that although X_1 completely dominates X_2 as well as X_3 and X_4 , complete dominance cannot be established between X_2 and X_3 , X_2 and X_4 , or X_3 and X_4 .

To reduce the incidence of undetermined dominance, we introduce two weaker levels of dominance. The first of these compares the additional contributions of each predictor to all subset models as before, but the measure ultimately used to compare the predictors is the average of the additional contributions to all subset models of a given model size. The model size is defined as the number of predictors included in the subset model and is denoted by k . If the average additional contribution within each model size is greater for one predictor than the other, then that predictor is said to *conditionally dominate* the other. These (four) comparisons are shown symbolically in Table 4 and numerically in the rows of Table 3 that

are labeled “ $k = (\text{number})$ average.” For example, the average contribution of X_1 to models of size 1 is computed as $[0.360 + 0.317 + 0.223]/3 = 0.300$, which is the entry in the “ $k = 1$ average” row under the X_1 column. Similarly, the average contribution of X_2 to the models of size 1 is $[0.090 + 0.068 + 0.063]/3 = 0.074$. Because the average contribution (within model size) of X_1 is greater than that of X_2 for each model size (i.e., $0.360 > 0.090$ for $k = 0$, $0.300 > 0.074$ for $k = 1$, $0.263 > 0.069$ for $k = 2$, and $0.246 > 0.071$ for $k = 3$), X_1 conditionally dominates X_2 . In addition, X_1 conditionally dominates X_3 and X_4 . However, as in the complete dominance comparisons, conditional dominance cannot be established between all other pairs of predictors.

The last level of dominance summarizes the additional contributions of each predictor to all subset models by averaging all the conditional values. In this example, this consists of averaging the four averaged entries in each column of Table 4. If this overall av-

Table 4

Formulas Used to Compute the Averaged Additional Contributions Within Model Size in the Population With Four Predictors

Subset model size (k)	Average additional contributions of:	
	X_1	X_2
0	$\rho_{Y \cdot X_1}^2$	$\rho_{Y \cdot X_2}^2$
1	$[(\rho_{Y \cdot X_1 X_2}^2 - \rho_{Y \cdot X_2}^2) + (\rho_{Y \cdot X_1 X_3}^2 - \rho_{Y \cdot X_3}^2) + (\rho_{Y \cdot X_1 X_4}^2 - \rho_{Y \cdot X_4}^2)]/3$	$[(\rho_{Y \cdot X_1 X_2}^2 - \rho_{Y \cdot X_1}^2) + (\rho_{Y \cdot X_2 X_3}^2 - \rho_{Y \cdot X_3}^2) + (\rho_{Y \cdot X_2 X_4}^2 - \rho_{Y \cdot X_4}^2)]/3$
2	$[(\rho_{Y \cdot X_1 X_2 X_3}^2 - \rho_{Y \cdot X_2 X_3}^2) + (\rho_{Y \cdot X_1 X_2 X_4}^2 - \rho_{Y \cdot X_2 X_4}^2) + (\rho_{Y \cdot X_1 X_3 X_4}^2 - \rho_{Y \cdot X_3 X_4}^2)]/3$	$[(\rho_{Y \cdot X_1 X_2 X_3}^2 - \rho_{Y \cdot X_1 X_3}^2) + (\rho_{Y \cdot X_1 X_2 X_4}^2 - \rho_{Y \cdot X_1 X_4}^2) + (\rho_{Y \cdot X_2 X_3 X_4}^2 - \rho_{Y \cdot X_3 X_4}^2)]/3$
3	$(\rho_{Y \cdot X_1 X_2 X_3 X_4}^2 - \rho_{Y \cdot X_2 X_3 X_4}^2)$	$(\rho_{Y \cdot X_1 X_2 X_3 X_4}^2 - \rho_{Y \cdot X_1 X_3 X_4}^2)$

eraged additional contribution is greater for one predictor than the other, that predictor is said to *generally dominate* the other. The corresponding values for this measure are shown in the last row of Table 3. For example, the general measure for X_1 is computed as $[0.360 + 0.300 + 0.263 + 0.246]/4 = 0.292$. In theory there is nothing preventing an unequal weighting of these averages (e.g., if each average is weighted by the number of models of a given size, the result would be the unweighted average of all additional contributions). However, using this equal weighting of the averages results in a measure that has the appealing property of summing (over all predictors) to $\rho_{Y \cdot Y}^2$ of the full model, or the proportion of variance in the criterion explained by all of the predictors. In the example, note that the sum is $0.292 + 0.076 + 0.095 + 0.121 = 0.584$, the value of $\rho_{Y \cdot Y}^2$ for the model containing all four predictors. This general dominance measure coincides with the average squared semipartial correlations across all $p!$ orders proposed in Lindeman et al. (1980) and has been advocated by Johnson (2000), who developed an efficient algorithm to compute what he calls a "relative weight" measure that approximates the general dominance measure very well. In terms of interpretation, the general measure represents the average difference in fit between all subset models (of equal size) that include X_i and those that do not include it. It is always possible to establish a general dominance ordering unless the general measure is identical for a pair of predictors. In this example, X_1 generally dominates all predictors, X_4 dominates X_3 and X_2 , and X_3 dominates X_2 (thus the ordering is X_1, X_4, X_3, X_2).

The three levels of dominance (complete, conditional, and general) are related to each other in a hierarchical fashion: Complete dominance implies conditional dominance, which, in turn, implies general dominance. However, for $p > 3$ the converse may not hold; that is, general dominance does not imply con-

ditional dominance and conditional dominance does not necessarily imply complete dominance.

Budescu (1993) provided several examples of DA with real and hypothetical data. For additional examples of applications see Behson (2002); Block (1995); Nickerson, Schwartz, Diener, and Kahneman (in press); Suh et al. (1998); and Weinberger (1995).

Suppressor Variable Case

To further illustrate the usefulness of DA we present the analysis in the presence of a suppressor variable and show the ability of DA to detect such a variable. A *suppressor* refers to a predictor whose inclusion in the regression model improves prediction through its correlation with other predictors rather than its direct correlation with the criterion. In other words, such variables "suppress" the variance in other predictors that is irrelevant to the criterion and thus enhance those predictors' predictive ability. The population correlation matrix in Table 5 includes three predictors and reflects a pattern in which X_2 acts as a "classic suppressor" according to Tzelgov and Henik's (1991) typology; that is, a predictor that is uncorrelated with the criterion but whose presence improves prediction because of its correlation with other predictors.

The contributions of a "regular" predictor averaged over all models of a given size (i.e., the conditional dominance measures) are expected to decrease monotonically as the model becomes more complex

Table 5
Example of a Population Correlation Matrix for Classical Suppression

Variable	1	2	3	4
1. Y	—	0.50	0.00	0.25
2. X_1	0.50	—	0.30	0.25
3. X_2	0.00	0.30	—	0.25
4. X_3	0.25	0.25	0.25	—

(larger) because of the correlations among the predictors. However, the contributions of a suppressor variable may deviate from this monotonic pattern—paradoxically, the predictor makes a more substantial contribution in more complex models. This result is shown in Table 6; while the average additional contributions (i.e., the conditional dominance measures) of X_1 and X_3 decrease with model size, k (from 0.250 to 0.239 to 0.234 in the case of X_1 , and from 0.062 to 0.042 to 0.026 in the case of X_3), these measures actually increase in the case of X_2 , the classic suppressor (from 0.000 to 0.014 to 0.034). Although not shown, similar patterns of results are observed in examples of negative and reciprocal suppression (as defined by Tzelgov & Henik, 1991).

Furthermore, as pointed out by an anonymous reviewer of this article, using any measure based on the correlation $\rho_{X,Y}$ (e.g., the product $\beta_i \rho_{X,Y}$ or the structure coefficient $\rho_{X_i,Y}$) as an index of predictor contribution would yield a value of 0 for the suppressor variable X_2 in the above example (because $\rho_{X_2,Y} = 0$). This would imply that X_2 contributes nothing overall, clearly a misleading interpretation. It thus appears that DA is more useful for detecting cases of suppression than other procedures.

DA With Other Measures of Fit

Some of the key features of DA are not restricted to the $\rho_{Y\hat{Y}}^2$ (or the sample R^2) metric. In fact, it is easy to

show that any measure of model fit that is a monotone (increasing or decreasing) function of the model's error sum of squares (SSE) would yield the same dominance pattern. In other words, if DA based on the squared multiple correlations shows that in a certain sample or population X_i dominates X_j (generally, conditionally, or completely), it must also be the case that DA of the same data set based on any other measure that is monotonically related to SSE would similarly show X_i to dominate X_j (generally, conditionally, or completely, respectively). A formal proof of this claim for three such measures (adjusted R^2 , Akaike's Information Criterion, and C_p) is given in Azen (2000).

Constrained DA: Examining Only Theoretically Meaningful Subsets

Constrained DA is a simple extension of DA for cases where a specific predictor (or set of predictors) is forced to be a part of the model. Such constraints may be imposed by theoretical or practical considerations that require the inclusion of some predictors. In this way, only theoretically meaningful subset models are used in the analysis and only relevant comparisons among the remaining predictors are conducted.

Using the illustrative example presented in Table 3, suppose that for theoretical reasons X_1 must be included in the regression model (i.e., subset models that do not include X_1 are of no practical or theoretical

Table 6
Dominance Analysis in the Population for the Classical Suppression Example

Subset model (\mathbf{X})	$\rho_{Y,\mathbf{X}}^2$	Additional contribution of:		
		X_1	X_2	X_3
Null and $k = 0$ average	0	0.250	0.000	0.062
X_1	0.250		0.025	0.017
X_2	0.000	0.275		0.067
X_3	0.062	0.204	0.004	
$k = 1$ average		0.239	0.014	0.042
X_1X_2	0.275			0.026
X_1X_3	0.267		0.034	
X_2X_3	0.067	0.234		
$k = 2$ average		0.234	0.034	0.026
$X_1X_2X_3$	0.300			
Overall average		0.241	0.016	0.043

Note. The column labeled $\rho_{Y,\mathbf{X}}^2$ represents the variance in Y explained by the model appearing in the corresponding row. Columns labeled X_i contain the additional contributions to the explained variance gained by adding the column variable (X_i) to the row model. The rows in bold show the average additional contributions that are conditional on model size. Blank cells indicate that data are not applicable.

interest). For example, Y may be an overall measure of satisfaction with life, X_1 may be satisfaction with self, and the other predictors may be satisfaction with housing, finances, and health. In this instance, it is possible that the researcher believes self-satisfaction must be in the model for theoretical reasons (e.g., self-satisfaction is theoretically necessary for satisfaction with other domains or self satisfaction is already known to be highly predictive of life satisfaction) and wishes to examine the dominance relationship among the remaining predictors. Those rows of Table 3 pertaining to subsets excluding X_1 are deleted, and the analysis proceeds as before to compare the predictors in this reduced (or constrained) table. To illustrate this analysis, Table 7 presents the DA results using only those subset models that include X_1 .

Unlike the full model analysis (Table 3), where only general dominance could be established among X_2 , X_3 , and X_4 , the results in this constrained case (Table 7) indicate that X_3 completely dominates X_4 . Interestingly, in the full analysis (Table 3) the general dominance ordering was X_1 , X_4 , X_3 , X_2 , but in this constrained case (Table 7) we find that the order among the three predictors of interest changes to X_3 , X_2 , X_4 . Also, recall that in the full analysis the general dominance measures sum to $\rho_{Y\hat{Y}}^2$ of the full model (0.584). In the constrained case where X_1 is forced to be part of the model, the sum of the general domi-

nance measures is $(0.078 + 0.087 + 0.059) = 0.224$ and the fit of the fixed predictor (X_1) is $\rho_{Y\cdot X_1}^2 = .360$, so the overall total $(0.224 + 0.360 = 0.584)$ is again equal to $\rho_{Y\hat{Y}}^2$ of the full model.

Inference

Given an MR model involving p predictors, one may fit all $2^p - 1$ subset regression models and determine, for each of the distinct $p(p - 1)/2$ pairs of predictors, whether one predictor dominates the other (for each definition of dominance: complete, conditional, and general). In this section the issue of generalizing dominance results beyond the particular observed sample is addressed; we describe how to estimate the standard error of dominance across repeated sampling and how to determine the confidence one can have in making inferences about the population from the observed sample.

If, ultimately, interest is simply in whether one predictor (X_i) dominates another (X_j), this can be summarized by a measure denoted D_{ij} , which can have one of three values: $D_{ij} = 1$ if X_i dominates X_j , $D_{ij} = 0$ if X_j dominates X_i , and $D_{ij} = 0.5$ if neither predictor dominates the other (dominance cannot be established between X_i and X_j). The bootstrap procedure is used to simulate the distribution of the D_{ij} values that would be obtained if the study were repeated many times,

Table 7
Constrained (X_1 Included in All Models) Dominance Analysis in the Population for the Hypothetical Example With Four Predictors

Subset model (\mathbf{X})	$\rho_{Y\cdot\mathbf{X}}^2$	Additional contribution of:		
		X_2	X_3	X_4
(X_1) and $k = 1$ average	0.360	0.090	0.117	0.113
$(X_1)X_2$	0.450		0.124	0.073
$(X_1)X_3$	0.477	0.097		0.037
$(X_1)X_4$	0.473	0.051	0.041	
$k = 2$ average		0.074	0.082	0.055
$(X_1)X_2X_3$	0.574			0.010
$(X_1)X_2X_4$	0.523		0.061	
$(X_1)X_3X_4$	0.513	0.071		
$k = 3$ average		0.071	0.061	0.010
$(X_1)X_2X_3X_4$	0.584			
Overall average		0.078	0.087	0.059

Note. The column labeled $\rho_{Y\cdot\mathbf{X}}^2$ represents the variance in Y explained by the model appearing in the corresponding row. Columns labeled X_i contain the additional contributions to the explained variance gained by adding the column variable (X_i) to the row model. Blank cells indicate that data are not applicable.

drawing a different random sample from the same population each time.

Bootstrap Procedure

The bootstrap is a procedure that was developed to empirically estimate the variability of a statistic whose theoretical distribution is unknown. The procedure “resamples” the data with replacement to generate an empirical estimate of the entire sampling distribution of the statistic of interest (Efron, 1979; Mooney & Duval, 1993). Starting with the n observations in the parent sample (obtained by the researcher), each resample or bootstrap sample is obtained by randomly drawing n observations with replacement from the parent sample. The resulting bootstrap sample thus consists of a set of n observations that are similar, but not (usually) identical, to those in the parent sample and can be treated as a random sample drawn from the same population that generated the parent sample. This process is repeated S times (where S is a large number) to generate a total of S bootstrap samples.

In the case of MR, the purpose of resampling is to capture the random component of the regression model, and a distinction is thus made between regression models involving fixed predictors and those involving random predictors (Efron, 1979; Freedman, 1981; Mooney & Duval, 1993; Neter et al., 1996; Shao, 1996; Stein, 1996). When the levels of the predictors are random (rather than fixed), the bootstrapping process resamples the n cases (where each case consists of the criterion and its associated predictor values: $Y_j, X_{1j}, X_{2j}, \dots, X_{pj}; j = 1, 2, \dots, n$) and this was the procedure used in all the examples presented in this article.

Method

In each of the S bootstrap samples the dominance value (D_{ij}) obtained for a given pair of predictors (X_i and X_j) can have one of three values: 1 if X_i dominates X_j , 0 if X_j dominates X_i , or 0.5 if dominance cannot be established between X_i and X_j . For example, if the total number of bootstrap samples is 1,000 ($S = 1,000$), there will be 1,000 dominance values ($D_{ij}^1, D_{ij}^2, \dots, D_{ij}^{1000}$) for each pair of predictors (and each level of dominance). The average of these dominance values (over all S bootstrap samples) is written as

$$\bar{D}_{ij} = \frac{1}{S} \sum_{s=1}^S D_{ij}^s. \quad (2)$$

The average, \bar{D}_{ij} , represents the expected level of dominance of X_i over X_j in the population and is a

random variable that can take on any value between 0 and 1. The standard error (SE) of D_{ij} is obtained by computing the standard deviation of the dominance values over all S bootstrap samples, written as

$$SE(D_{ij}) = \sqrt{\frac{1}{S-1} \sum_{s=1}^S (D_{ij}^s - \bar{D}_{ij})^2}. \quad (3)$$

The standard error represents the expected variability of the dominance value over repeated sampling. The interpretation of \bar{D}_{ij} and the standard error is facilitated by three obvious reference points: \bar{D}_{ij} is 1 (and SE is 0) if, and only if, $D_{ij} = 1$ in all bootstrap samples. Conversely, \bar{D}_{ij} is 0 (and SE is 0) if, and only if, $D_{ij} = 0$ in all bootstrap samples. Finally, \bar{D}_{ij} is 0.5 if the distribution of dominance indices (D_{ij}^s) is symmetric in the sense that the number of cases in which X_i dominates X_j equals the number of cases in which X_j dominates X_i . In this case, the standard error depends on the number of indeterminate cases; it is maximized when there are no indeterminate cases (i.e., X_i dominates X_j in exactly one half of all bootstrap samples and X_j dominates X_i in the remaining half of the bootstrap samples), and it is minimized (i.e., SE is 0) if, and only if, $D_{ij} = 0.5$ in all bootstrap samples (i.e., all cases are indeterminate). Thus, the closer \bar{D}_{ij} is to one of the two end points (0 or 1), the stronger the case for a clear directional dominance, whereas the closer \bar{D}_{ij} is to 0.5, the stronger the case for indeterminacy.

Another way to evaluate the generality and robustness of the results (that is closer in spirit to the traditional hypothesis testing) is to measure the degree to which the dominance pattern obtained in the sample is reproduced in the S bootstrap samples (Thompson, 1994, refers to this as *internal replicability*). Suppose we found that, in the sample, X_3 completely dominates X_2 , and after observing the distribution of results in the S bootstrap samples, we determine that this result was replicated in 97% of the bootstrap samples. We conclude that the chance of reproducing the sample result is 0.97, and we can be 97% confident in concluding that X_3 dominates X_2 in the population. We refer to this value as the *reproducibility* of the result. Because interest is often in whether X_i dominates X_j (but not necessarily by how much), this approach allows the researcher to both state the probability that X_i will dominate X_j and determine a confidence level on that probability.

To illustrate the procedures, consider again the hypothetical example presented earlier with four predictors (X_1, X_2, X_3, X_4) and the (population) correlation

Table 8
Sample ($n = 100$) Correlation Matrix for Hypothetical Example With Four Predictors

Variable	1	2	3	4	5
1. Y	—	0.6087	0.2580	0.4735	0.4558
2. X_1	0.6087	—	0.7626	0.0886	0.2528
3. X_2	0.2580	0.7626	—	-0.0325	-0.0256
4. X_3	0.4735	0.0886	-0.0325	—	0.4878
5. X_4	0.4558	0.2528	-0.0256	0.4878	—

matrix shown in Table 1. As discussed earlier (Table 3), according to the complete dominance criterion it is expected that X_1 will dominate X_2 as well as X_3 and X_4 , but dominance would not be established between X_2 , X_3 , and X_4 . A random sample of $n = 100$ multivariate normal observations was generated from a population with this correlation matrix; the sample correlation matrix is presented in Table 8. This sample was then bootstrapped $S = 1,000$ times. The results are shown in Table 9: The first and second columns identify the two variables being compared; the third column is the value of D_{ij} in the sample; the fourth column is the average (\bar{D}_{ij}) value over the bootstrap samples; and the fifth column is the standard error (SE) of the D_{ij} values over the bootstrap samples. The next three columns describe the distri-

bution of D_{ij} over the S bootstrap samples; they show the proportion of bootstrap samples in which X_i dominated X_j or $D_{ij} = 1$ (P_{ij}), the proportion of samples in which X_j dominated X_i or $D_{ij} = 0$ (P_{ji}), and the proportion of samples in which dominance between X_i and X_j could not be established or $D_{ij} = 0.5$ (P_{noij}). The last column is the reproducibility of (or proportion of bootstrap samples that agree with) the sample results. Although each pair has two possible orders (e.g., for the pair X_1, X_2 , the two orders are $i = 1, j = 2$ and $i = 2, j = 1$), only one order is shown in the table because of the obvious redundancy ($D_{ij} + D_{ji} = 1$). To facilitate interpretation, the order of presentation (i then j) of each pair in the table was determined such that $P_{ij} > P_{ji}$ for complete dominance.² The complete dominance results show that, for this particular

Table 9
Hypothetical Example Sample Results: D_{ij} Values in the Sample ($n = 100$) and Their Means (\bar{D}_{ij}), Standard Errors, Probabilities, and Reproducibility Over $S = 1,000$ Bootstrap Samples

i	j	Sample D_{ij} ^a	\bar{D}_{ij}	$SE(D_{ij})$	P_{ij} ^b	P_{ji} ^c	P_{noij} ^d	Reproducibility
Complete dominance								
1	2	1.0	1.0000	0.000	1.000	0.000	0.000	1.000
1	3	1.0	0.9315	0.186	0.873	0.010	0.117	0.873
1	4	1.0	0.9680	0.122	0.936	0.000	0.064	0.936
2	4	0.5	0.5080	0.116	0.035	0.019	0.946	0.946
3	2	1.0	0.7475	0.260	0.505	0.010	0.485	0.505
3	4	1.0	0.7800	0.254	0.566	0.006	0.428	0.566
Conditional dominance								
1	2	1.0	1.0000	0.000	1.000	0.000	0.000	1.000
1	3	1.0	0.9325	0.190	0.879	0.014	0.107	0.879
1	4	1.0	0.9720	0.115	0.944	0.000	0.056	0.944
2	4	0.5	0.5100	0.145	0.052	0.032	0.916	0.916
3	2	1.0	0.8670	0.234	0.746	0.012	0.242	0.746
3	4	1.0	0.7990	0.251	0.604	0.006	0.390	0.604
General dominance								
1	2	1.0	1.0000	0.000	1.000	0.000	0.000	1.000
1	3	1.0	0.9630	0.189	0.963	0.037	0.000	0.963
1	4	1.0	1.0000	0.000	1.000	0.000	0.000	1.000
2	4	0.0	0.2640	0.441	0.264	0.736	0.000	0.736
3	2	1.0	0.9430	0.232	0.943	0.057	0.000	0.943
3	4	1.0	0.8480	0.359	0.848	0.152	0.000	0.848

^a $D_{ij} = 1 - D_{ji}$. ^b $P_{ij} = Pr(D_{ij} = 1)$. ^c $P_{ji} = Pr(D_{ij} = 0)$. ^d $P_{noij} = Pr(D_{ij} = 0.5)$.

sample, X_1 indeed dominates X_2 as well as X_3 and X_4 (i.e., $D_{12} = 1$, $D_{13} = 1$, $D_{14} = 1$), dominance could not be established between X_2 and X_4 (i.e., $D_{24} = 0.5$), but dominance was established for the remaining pairs, with X_3 dominating both X_2 and X_4 (i.e., $D_{32} = 1$ and $D_{34} = 1$). The \bar{D}_{ij} values reflect these results and refine them by demonstrating the expected value of D_{ij} over repeated sampling. For example, although both X_1 and X_3 dominate X_4 in the sample, the dominance of X_3 over X_4 is expected to be established less often ($\bar{D}_{34} = 0.780$) than the dominance of X_1 over X_4 ($\bar{D}_{14} = 0.968$). The standard error estimates give the expected amount of variability of the D_{ij} values around their expected values. More details are given by the three probability entries, one of which also serves as the reproducibility measure. Note, for example, that the proportion of undetermined outcomes was much higher between X_3 and X_4 ($P_{no34} = 0.428$) than between X_1 and X_4 ($P_{no14} = 0.064$). The reproducibility values show that the sample result $D_{34} = 1$ was obtained in only 56.6% of all bootstrap samples, whereas the sample result $D_{14} = 1$ was obtained in 93.6% of all bootstrap samples.

The conditional and general dominance results show a similar pattern while reducing the number of indeterminacies, as shown by the \bar{D}_{ij} values becoming closer to either 0 or 1 as the level of analysis becomes "weaker." For example, in the complete dominance case, $\bar{D}_{34} = 0.780$; in the weaker conditional dominance case, $\bar{D}_{34} = 0.799$; and in the still weaker general dominance case, $\bar{D}_{34} = 0.848$. This result is similarly demonstrated by the fact that the sample reproducibility of $D_{34} = 1$ increases from 56.6% for complete dominance, to 60.4% for conditional dominance, and to 84.8% for general dominance. The conclusion that can be drawn from these results is that the level of confidence in the dominance of X_3 over X_4 is high only in the general dominance case that uses a weaker standard. Conversely, consider the reproducibility of the indeterminate dominance relationship between X_2 and X_4 ($D_{24} = 0.5$), which decreases from 94.6% for complete dominance to 91.6% for conditional dominance. Moreover, X_4 generally dominates X_2 ($D_{24} = 0$) and the reproducibility of this result is 73.6%, indicating that the level of confidence in the established dominance relationship between X_2 and X_4 (in the general dominance case) is low in comparison to the level of confidence in the indeterminate relationship between X_2 and X_4 (in the complete and conditional dominance cases).

The pattern of results indicates an increase in the

reproducibility of $D_{ij} = 1$ and a decrease in the reproducibility of $D_{ij} = 0.5$ as the dominance level is relaxed (complete, conditional, general). Intuitively, this pattern of results makes sense—in the case of $D_{ij} = 1$, as the level of dominance is relaxed, more cases that satisfy this dominance relationship will occur, leading to higher reproducibility. On the other hand, in the case of $D_{ij} = 0.5$, the undetermined dominance relationship is more likely to be resolved as the level of dominance is relaxed, resulting in lower reproducibility of the undetermined dominance result.³

The \bar{D}_{ij} values are directly related to the P_{ij} values; in fact, $\bar{D}_{ij} = P_{ij} + \frac{1}{2}P_{noij}$. Thus, the \bar{D}_{ij} values account for the proportion of undetermined cases (P_{noij}), whereas the P_{ij} values do not. In addition, \bar{D}_{ij} can be interpreted as the expected value of D_{ij} , and P_{ij} can be interpreted as the probability that X_i dominates X_j .

Example: Satisfaction With Life

Suh et al. (1998) collected responses from nearly 7,000 college students in 41 countries on numerous variables pertaining to satisfaction with life. For this example, a subset of variables was selected and only responses from the United States ($n = 428$) were used. The criterion, Y , was the score on a satisfaction with life scale consisting of five items. The five predictors were responses to single items (not part of the global scale) pertaining to domain-specific reported levels of satisfaction. The domains were health (X_1), financial situation (X_2), family (X_3), housing (X_4), and self (X_5). The correlation matrix is shown in Table 10, and the values of several of the reviewed measures of importance are given in Table 11. The various measures indicate self and family (in that order) to be the two top predictors and health to be the least important predictor. The ordering of housing and finance is less clear and varies across the measures of importance.

² For example, suppose that of 1,000 cases, $D_{12} = 1$ occurred 600 times, $D_{12} = 0$ occurred 300 times, and $D_{12} = 0.5$ occurred 100 times. Then $P_{12} = 600/1000 = 0.6$, $P_{21} = 300/1000 = 0.3$, and $P_{no12} = 100/1000 = 0.1$. It would be redundant to write the results for both the order $i = 1, j = 2$ ($P_{ij} = P_{12} = 0.6$, $P_{ji} = P_{21} = 0.3$, $P_{no12} = 0.1$) and the order $i = 2, j = 1$ ($P_{ij} = P_{21} = 0.3$, $P_{ji} = P_{12} = 0.6$, $P_{no12} = 0.1$). Thus, given that for $i = 1$ and $j = 2$ $P_{ij} = 0.6 > P_{ji} = 0.3$, only this order would be presented.

³ In fact, because the undetermined relationship will usually become determined in the general dominance case, it probably makes sense to compare the reproducibility of the undetermined dominance result between the complete and conditional dominance levels only.

Table 10
Correlation Matrix for Satisfaction With Life Data Set ($n = 428$)

Variable	1	2	3	4	5	6
1. Y	—	0.2346	0.3637	0.4875	0.3162	0.6208
2. Health	0.2346	—	0.1775	0.1125	0.1510	0.3425
3. Finance	0.3637	0.1775	—	0.2614	0.2384	0.2899
4. Family	0.4875	0.1125	0.2614	—	0.2490	0.3818
5. Housing	0.3162	0.1510	0.2384	0.2490	—	0.1918
6. Self	0.6208	0.3425	0.2899	0.3818	0.1918	—

The D_{ij} values in the sample are shown in Table 12 along with the average and standard error of the D_{ij} values and the reproducibility of the sample results over $S = 1,000$ bootstrap samples. Examining the sample results, it appears that according to the complete dominance measure, X_2 (finance) dominates X_1 (health); dominance cannot be established between X_2 (finance) and X_4 (housing); X_3 (family) dominates X_1 (health), X_2 (finance), and X_4 (housing); X_4 (housing) dominates X_1 (health); and X_5 (self) dominates all of the other predictors. Thus, the overall ordering is self, family, finance/housing, and health. The conditional dominance results are identical, and general dominance establishes that X_2 (finance) dominates X_4 (housing).

The results over 1,000 bootstrap samples demonstrate the same pattern of results as in the sample and add information on the expected values and stability of these measures. For example, although all predictors dominate X_1 (health), X_4 (housing) completely dominates X_1 (health) less often ($\bar{D}_{41} = 0.7725$) than does X_2 (finance; $\bar{D}_{21} = 0.916$), X_3 (family; $\bar{D}_{31} = 0.9995$), or X_5 (self; $\bar{D}_{51} = 1.000$). The standard errors, however, indicate that these differences are quite small.

Use of lower levels of dominance reduce the number of undetermined cases and lead to expected values that become increasingly closer to either 0 or 1. Similarly, as the level of analysis gets weaker, the sample reproducibility increases because of the elimination of undetermined cases; for example, the dominance of X_2 (finance) over X_1 (health) is reproduced in 83.2% of cases for complete dominance, 95.4% of cases for conditional dominance, and 98.1% of cases for general dominance. Therefore, it appears that this dominance relationship is stable and highly reproducible. On the other hand, the indeterminacy between X_2 (finance) and X_4 (housing) is reproduced in only 47.3% of cases for complete dominance and 35.8% of cases for conditional dominance, whereas the general dominance of X_2 (finance) over X_4 (housing) was reproduced in 70.3% of cases. Therefore, one may be less confident in the indeterminacy results for complete and conditional dominance and relatively more confident in the established dominance of X_2 (finance) over X_4 (housing) using general dominance.

The probability of dominance results shown in Table 12 can be used to gain more insight into the behavior of the dominance measures over repeated sampling. For example, using complete dominance,

Table 11
Various Measures of Importance for the Satisfaction With Life Data Set ($n = 428$)

Measure	Predictor (X_i)				
	Health	Finance	Family	Housing	Self
\hat{B}_i (raw coefficient)	0.0245	0.5519	1.2037	0.5789	2.2505
$\hat{\beta}_i$ (standardized coefficient)	0.0050	0.1334	0.2420	0.1347	0.4622
t value	0.13	3.58	6.29	3.69	11.51
r_{X_iY}	0.2346	0.3637	0.4875	0.3162	0.6208
Semipartial r_i^2	0.0000	0.0153	0.0471	0.0162	0.1578
Partial r_i^2	0.0000	0.0295	0.0857	0.0313	0.2389
Kruskal's (1987) average	0.0213	0.0699	0.1527	0.0570	0.3036
Theil's (1987) average	0.0314	0.1058	0.2428	0.0852	0.5267
$\hat{\beta}_i r_{X_iY}$	0.0012	0.0485	0.1180	0.0426	0.2869
$r_{X_i\hat{Y}}$	0.3327	0.5158	0.6914	0.4484	0.8804

Table 12

Results for Satisfaction With Life Example: D_{ij} Values in the Sample ($n = 428$) and Their Means (\bar{D}_{ij}), Standard Errors, Probabilities, and Reproducibility Over $S = 1,000$ Bootstrap Samples

i	j	Sample D_{ij}	\bar{D}_{ij} ^a	$SE(D_{ij})$	P_{ij} ^b	P_{ji} ^c	P_{noij} ^d	Reproducibility
Complete dominance								
2	1	1.0	0.9160	0.187	0.832	0.000	0.168	0.832
2	4	0.5	0.6405	0.335	0.404	0.123	0.473	0.473
3	1	1.0	0.9995	0.016	0.999	0.000	0.001	0.999
3	2	1.0	0.9605	0.151	0.930	0.009	0.061	0.930
3	4	1.0	0.9540	0.151	0.912	0.004	0.084	0.912
4	1	1.0	0.7725	0.249	0.545	0.000	0.455	0.545
5	1	1.0	1.0000	0.000	1.000	0.000	0.000	1.000
5	2	1.0	1.0000	0.000	1.000	0.000	0.000	1.000
5	3	1.0	0.9920	0.063	0.984	0.000	0.016	0.984
5	4	1.0	1.0000	0.000	1.000	0.000	0.000	1.000
Conditional dominance								
2	1	1.0	0.9735	0.127	0.954	0.007	0.039	0.954
2	4	0.5	0.6490	0.372	0.470	0.172	0.358	0.358
3	1	1.0	0.9995	0.016	0.999	0.000	0.001	0.999
3	2	1.0	0.9675	0.144	0.946	0.011	0.043	0.946
3	4	1.0	0.9645	0.138	0.934	0.005	0.061	0.934
4	1	1.0	0.9220	0.192	0.852	0.008	0.140	0.852
5	1	1.0	1.0000	0.000	1.000	0.000	0.000	1.000
5	2	1.0	1.0000	0.000	1.000	0.000	0.000	1.000
5	3	1.0	0.9980	0.032	0.996	0.000	0.004	0.996
5	4	1.0	1.0000	0.000	1.000	0.000	0.000	1.000
General dominance								
2	1	1.0	0.9810	0.137	0.981	0.019	0.000	0.981
2	4	1.0	0.7030	0.457	0.703	0.297	0.000	0.703
3	1	1.0	1.0000	0.000	1.000	0.000	0.000	1.000
3	2	1.0	0.9800	0.140	0.980	0.020	0.000	0.980
3	4	1.0	0.9940	0.077	0.994	0.006	0.000	0.994
4	1	1.0	0.9120	0.283	0.912	0.088	0.000	0.912
5	1	1.0	1.0000	0.000	1.000	0.000	0.000	1.000
5	2	1.0	1.0000	0.000	1.000	0.000	0.000	1.000
5	3	1.0	0.9980	0.045	0.998	0.002	0.000	0.998
5	4	1.0	1.0000	0.000	1.000	0.000	0.000	1.000

Note. The predictors are health (X_1), finance (X_2), family (X_3), housing (X_4), and self (X_5).

^a $D_{ij} = 1 - D_{ji}$. ^b $P_{ij} = Pr(D_{ij} = 1)$. ^c $P_{ji} = Pr(D_{ij} = 0)$. ^d $P_{noij} = Pr(D_{ij} = 0.5)$.

note that although X_4 (housing) dominates X_1 (health) with $\bar{D}_{41} = 0.7725$, the proportion of samples that result in undetermined dominance between X_4 (housing) and X_1 (health) is quite high ($P_{no41} = 0.455$), indicating that in almost half of all samples dominance between these predictors will be undetermined and that the reproducibility of the dominance relationship between these predictors is only slightly over 50%. This provides much more information on this dominance relationship than that provided by the sample results alone. Software for the computations presented in this article is available, in the form of an SAS Macro, from Razia Azen and is available in the

online version of this article in the PsycARTICLES database.

Discussion

Social scientists are often interested in comparing the predictors in an MR analysis to answer a variety of research questions in many contexts. The need for such a process is quite evident as researchers continue to be interested in finding ways to compare predictors and, unfortunately, often use measures that are not well suited for this purpose or not well understood, leading to misuse (Courville & Thompson, 2001; Pedhazur, 1997). DA is proposed as a method that can

address a wide variety of questions pertaining to predictor comparisons, from the very specific to the most general, in a simple and intuitive manner.

We believe that DA is superior to most other approaches of determining relative importance because of its more natural and intuitive interpretation, which allows a general approach to predictor comparisons. Unlike most other approaches, which start with a particular statistic (e.g., correlation coefficient) and try to “extract” its interpretation, DA starts with a clear definition of importance and identifies the measures that address the key question: *Is variable X_i more or less (or equally) important than variable X_j in predicting Y in the context of the predictors included in the selected model?* The definition of importance is straightforward— X_i is considered more important than X_j if it contributes more to the prediction of the response than does X_j ; in other words, if one had to choose only one of the two predictors, X_i is judged more important if it would be chosen over X_j .

The original definition of dominance (Budescu, 1993) was strict in its requirement that, to establish the dominance of one predictor over another, the additional contribution condition needs to be satisfied in each possible subset model. One important contribution of this article is the relaxation of this stringent criterion and the definition of the weaker versions of dominance that allow some violations of complete dominance as long as X_i dominates X_j on the average. The average can be taken either across all models of a fixed size (conditional dominance) or across all models (general dominance). As an analogy, consider the question: Is City A warmer than city B? The question can be answered in the affirmative fashion at three distinct levels that tolerate various amounts of inconsistency: (a) A is warmer than B if its temperature is higher every day of the year (complete); (b) A is warmer than B if its average monthly temperature is higher every month of the year but not necessarily every day (conditional); and (c) A is warmer than B if its average annual temperature is higher (general). We recommend that dominance be reported at the strongest level (complete) when it can be established, because it implies that the weaker levels of dominance have also been achieved. If dominance cannot be established at the complete level, we suggest reporting this; the fact that dominance is undetermined is useful in its own right. Researchers then have the option of proceeding by examining weaker levels of dominance to resolve any undetermined cases; the weakest (general) dominance level will almost always establish

dominance between each pair of predictors. However, if dominance could not be established at any level other than the general level of measurement, this finding would be indicative of a weaker form of dominance that holds only by averaging over all subset models.

The DA procedure can easily be used to answer either very specific or the most general questions that require predictor comparisons. For example, to identify the predictor whose removal will result in the smallest drop in the proportion of variance (of Y) accounted for, examine the rows of the DA table corresponding to subset models of size $p - 1$; this is equivalent to examining the t tests of the regression coefficients or the squared partial or semipartial correlations. To identify the predictor that will give the best prediction if used alone, examine the first row of the DA table (corresponding to the null subset model); this is equivalent to examining the simple bivariate correlation with Y . To identify the predictor that adds most to prediction after controlling for any specific subset of the remaining variables, use constrained DA by examining only the rows of the DA table that are relevant. To identify the importance of the predictors in the most general sense, which does not indicate a specific subset model as the basis for comparison, examine all possible subset models and all possible comparisons by using the full DA table as well as the various levels of dominance (complete, conditional, general). To identify suppression cases, one may also use the DA table, as illustrated earlier.

The intuitive appeal of the DA approach is exemplified in a recent report (Geiser & Studley, 2002), which examined the relative importance and contribution of high school grade point average (HSGPA) and Scholastic Aptitude Test (SAT) I and SAT II scores in predicting success in college as part of an effort to develop optimal and fair admission procedures for the University of California (UC) system. The authors analyzed the GPA of all ($n = 77,893$) UC freshmen for the period 1996–1999. To this end, they fitted all 7 (or $2^3 - 1$) distinct submodels, compared their overall fit (in terms of R^2), and examined the additional contribution (again, in terms of R^2) of each of the predictors to the various subset models (pages 5 and 6). For all practical purposes they performed a DA, although they never used this term nor made any explicit reference to this approach. It takes the authors three full paragraphs to summarize the conclusions of their various analyses and comparisons when, in fact, the results can be summarized by sim-

ply stating that HSGPA dominates SAT II generally and HSGPA and SAT II dominate SAT I completely. Table 13 presents the DA of these data. This report has recently been cited as one of the main reasons for the change in the nature of the SAT (Barnes, 2002) and illustrates how DA can be used to affect and shape policymaking.

Each of the predictor comparisons in DA is based on the same subset model; that is, the comparison of X_i and X_j is always conducted within the context of a given subset of the remaining predictors and thus the order of entry of the other predictors into the model is irrelevant. If one is interested in the effect of each predictor as it is added to the previously entered predictors, it is possible to perform a sequence of constrained DAs, but DA is not designed to generate or address one specific hierarchical order. DA was also not designed to address questions regarding path-analytic models; for example, DA would not be of help in studying a specific causal model in which X_1 predicts X_2 and X_2 predicts Y , nor would DA be appropriate for other path-analytic models such as mediation or indirect effect models.

Any of the problems that are inherent to poorly measured variables and their effect on the semipartial correlations (for examples, see Pedhazur, 1997) will clearly also affect the results of DA. Other potential

measurement problems that could lead to erroneous conclusions when comparisons are made—for example, the construct represented by one variable is better measured than the construct represented by another, or the values of one variable are more representative of the range of its values in the population than are those of another variable—are discussed in Cooper and Richardson (1986). If the validity of conclusions drawn from the appropriate use of other measures of importance is questionable because of poor data, poor sampling, or other research methodology problems, the validity of conclusions drawn from DA will be similarly adversely affected. DA is an alternative to other measures of importance that is based on substantive and interpretational considerations, but it will not alleviate any issues that are due to problematic data.

The DA approach to comparing predictors is based on a predictor's added predictive ability in the presence of other predictors. This article takes the perspective of researchers who are interested in deciding which of two predictors is more important without examining the magnitude of the advantage. In other words, we focus on a qualitative determination of importance rather than a quantitative measure. It is conceptually simple to generalize this process to any measure of model fit that is based on the SSE and to

Table 13
Dominance Analysis of Data From "UC and the SAT"

Subset model (X)	$R^2_{Y \cdot X}$	Additional contribution of:		
		HSGPA	SAT I	SAT II
Null and $k = 0$ average	0	0.154	0.133	0.160
HSGPA	0.154		0.054	0.068
SAT-I	0.133	0.075		0.029
SAT-II	0.160	0.062	0.002	
$k = 1$ average		0.068	0.028	0.048
HSGPA, SAT I	0.208			0.015
HSGPA, SAT II	0.222		0.001	
SAT-I, SAT II	0.162	0.061		
$k = 2$ average		0.061	0.001	0.015
HSGPA, SAT I, SAT II	0.223			
Overall average		0.094	0.054	0.074

Note. The column labeled $R^2_{Y \cdot X}$ represents the variance in Y explained by the model appearing in the corresponding row; the other columns contain the additional contributions to the explained variance gained by adding the column variable to the row model. Blank cells indicate that data are not applicable. HSGPA = high school grade point average; SAT = Scholastic Aptitude Test. Data are from "UC and the SAT: Predictive Validity and Differential Impact of the SAT I and SAT II at the University of California," by S. Geiser and R. Studley, 2002, *Educational Assessment*, 8, Table 2, p. 6. Copyright 2002 by Lawrence Erlbaum Associates. Adapted with permission.

extend this measure to other general linear models. Similar approaches to comparing predictors have been proposed in the statistics literature by Whittaker (1984) in the context of model selection and by Chevan and Sutherland (1991) in the context of predictor importance. The DA approach presented here is a general approach to comparing predictors that includes several other comparison approaches (e.g., simple bivariate correlations, squared semipartial correlations, regression coefficients) as special cases. DA allows the researcher to compare the relative importance of regression predictors to answer a variety of research questions in a variety of contexts and at different levels of measurement (i.e., complete, conditional, and general). The use of the bootstrap to assess the stability of dominance results across repeated sampling provides the researcher with more insights into the pattern of importance in a set of predictors than were previously available.

References

- Azen, R. (2000). Inference for predictor comparisons: Dominance analysis and the distribution of R^2 differences. *Dissertation Abstracts International B*, 61/10, 5616.
- Azen, R., Budescu, D. V., & Reiser, B. (2001). Criticality of predictors in multiple regression. *British Journal of Mathematical and Statistical Psychology*, 54, 201–225.
- Barnes, J. E. (2002, November 11). The SAT revolution. *U.S. News & World Report*, 133, 51–60.
- Behson, S. J. (2002). Which dominates? The relative importance of work–family organizational support and general organizational context on employee outcomes. *Journal of Vocational Behavior*, 61, 53–72.
- Block, J. (1995). On the relation between IQ, impulsivity, and delinquency: Remarks on the Lynam, Moffitt, and Stouthamer-Loeber (1993) interpretation. *Journal of Abnormal Psychology*, 104, 395–398.
- Bring, J. (1994). How to standardize regression coefficients. *The American Statistician*, 48, 209–213.
- Bring, J. (1996). A geometric approach to compare variables in a regression model. *The American Statistician*, 50, 57–62.
- 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.
- Chevan, A., & Sutherland, M. (1991). Hierarchical partitioning. *The American Statistician*, 45, 90–96.
- Cook, R. D., & Weisberg, S. (1999). *Applied regression including computing and graphics*. New York: Wiley.
- Cooper, W. H., & Richardson, A. J. (1986). Unfair comparisons. *Journal of Applied Psychology*, 71, 179–184.
- Courville, T., & Thompson, B. (2001). Use of structure coefficients in published multiple regression articles: β is not enough. *Educational and Psychological Measurement*, 61, 229–248.
- Darlington, R. B. (1968). Multiple regression in psychological research and practice. *Psychological Bulletin*, 69, 161–182.
- Dunlap, W. P., & Landis, R. S. (1998). Interpretations of multiple regression borrowed from factor analysis and canonical correlation. *Journal of General Psychology*, 125, 397–407.
- Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7, 1–26.
- Freedman, D. A. (1981). Bootstrapping regression models. *Annals of Statistics*, 9, 1218–1228.
- Geiser, S., & Studley, R. (2002). UC and the SAT: Predictive validity and differential impact of the SAT I and SAT II at the University of California. *Educational Assessment*, 8, 1–26.
- Gibson, W. A. (1962). Orthogonal predictors: A possible resolution of the Hoffman-Ward controversy. *Psychological Reports*, 11, 32–34.
- Green, P. E., Carroll, D., & DeSarbo, W. (1978). A new measure of predictor variable importance in multiple regression. *Journal of Marketing Research*, 15, 356–360.
- Hays, W. L. (1994). *Statistics* (5th ed.). Forth Worth, TX: Harcourt Brace.
- Hedges, L. V., & Olkin, I. (1981). The asymptotic distribution of commonality components. *Psychometrika*, 46, 331–336.
- Hocking, R. R. (1976). The analysis and selection of variables in linear regression. *Biometrics*, 32, 1–49.
- Johnson, J. W. (2000). A heuristic method for estimating the relative weight of predictor variables in multiple regression. *Multivariate Behavioral Research*, 35, 1–19.
- Judge, G., Griffiths, W. E., Hill, R. C., Lutkepohl, H., & Lee, T.-C. (1985). *The theory and practice of econometrics*. New York: Wiley.
- Kruskal, W. (1987). Relative importance by averaging over orderings. *The American Statistician*, 41, 6–10.
- Lindeman, R. H., Merenda, P. F., & Gold, R. Z. (1980). *Introduction to bivariate and multivariate analysis*. Glenview, IL: Scott Foresman.
- Mayeske, G. W., Wisler, C. E., Beaton, A. E., Weinfeld, F. D., Cohen, W. M., Okada, T., et al. (1969). *A study of our nation's schools*. Washington, DC: U.S. Department of Health, Education, and Welfare, Office of Education.
- Miller, A. J. (1990). *Subset selection in regression*. London: Chapman & Hall.

- Mood, A. M. (1969). Macro-analysis of the American educational system. *Operations Research*, 17, 770–784.
- Mood, A. M. (1971). Partitioning variance in multiple regression analyses as a tool for developing learning models. *American Educational Research Journal*, 8, 191–202.
- Mooney, C. Z., & Duval, R. D. (1993). *Bootstrapping: A nonparametric approach to statistical inference* (Sage University Paper Series on Quantitative Applications in the Social Sciences, Series No. 07-095). Newbury Park, CA: Sage.
- Mosteller, F., & Tukey, J. W. (1977). *Data analysis and regression: A second course in statistics*. Reading, MA: Addison-Wesley.
- Neter, J., Kutner, M. H., Nachtsheim, C. J., & Wasserman, W. (1996). *Applied linear statistical models* (3rd ed.). Homewood, IL: Irwin.
- Newton, R. G., & Spurrell, D. J. (1967a). A development of multiple regression for the analysis of routine data. *Applied Statistics*, 16, 51–64.
- Newton, R. G., & Spurrell, D. J. (1967b). Examples of the use of elements for clarifying regression analysis. *Applied Statistics*, 16, 165–172.
- Nickerson, C., Schwartz, N., Diener, E., & Kahneman, D. (in press). Zeroing on the dark side of the American Dream: A closer look at the negative consequences of the goal for financial success. *Psychological Science*.
- Pedhazur, E. J. (1975). Analytic methods in studies of educational effects. In F. N. Kerlinger (Ed.), *Review of research in education 3* (pp. 243–286). Itasca, IL: Peacock.
- Pedhazur, E. J. (1997). *Multiple regression in behavioral research* (3rd ed.). Forth Worth, TX: Harcourt.
- Pratt, J. W. (1987). Dividing the indivisible: Using simple symmetry to partition variance explained. In T. Pukkila & S. Puntanen (Eds.), *Proceedings of the Second International Tampere Conference in Statistics* (pp. 245–260). Tampere, Finland: University of Tampere.
- Shao, J. (1996). Bootstrap model selection. *Journal of the American Statistical Association*, 91, 655–665.
- Shieh, G. (2001). The inequality between the coefficient of determination and the sum of squared simple correlation coefficients. *The American Statistician*, 55, 121–124.
- Stein, R. (1996, July). *Re-sampling techniques: Jackknife and bootstrap*. Workshop conducted at the Inter-university Consortium for Political and Social Research, Ann Arbor, Michigan.
- Suh, E., Diener, E., Oishi, S., & Triandis, H. C. (1998). The shifting basis of life satisfaction judgments across cultures: Emotions versus norms. *Journal of Personality and Social Psychology*, 74, 482–493.
- Thall, P. F., Russell, K. E., & Simon, R. M. (1997). Variable selection in regression via repeated data splitting. *Journal of Computational and Graphical Statistics*, 6, 416–434.
- Theil, H. (1987). How many bits of information does an independent variable yield in a multiple regression? *Statistics & Probability Letters*, 6, 107–108.
- Theil, H., & Chung, C.-F. (1988). Information-theoretic measures of fit for univariate and multivariate linear regressions. *The American Statistician*, 42, 249–252.
- Thomas, D. R., Hughes, E., & Zumbo, B. D. (1998). On variable importance in linear regression. *Social Indicators Research: An International and Interdisciplinary Journal for Quality-of-Life Measurement*, 45, 253–275.
- Thomas, D. R., & Zumbo, B. D. (1996). Using a measure of variable importance to investigate the standardization of discriminant coefficients. *Journal of Educational and Behavioral Statistics*, 21, 110–130.
- Thompson, B. (1994). The pivotal role of replication in psychological research: Empirically evaluating the replicability of sample results. *Journal of Personality*, 62, 157–176.
- Tzelgov, J., & Henik, A. (1991). Suppression situations in psychological research: Definitions, implications, and applications. *Psychological Bulletin*, 109, 524–536.
- Weinberger, T. E. (1995). Determining the relative importance of compensable factors: The application of dominance analysis to job evaluation. *Compensations & Benefits Management*, 11, 17–23.
- Whittaker, J. (1984). Model interpretation from the additive elements of the likelihood function. *Applied Statistics*, 33, 52–64.

Received February 11, 2002

Revision received January 28, 2003

Accepted January 29, 2003 ■

E-Mail Notification of Your Latest Issue Online!

Would you like to know when the next issue of your favorite APA journal will be available online? This service is now available to you. Sign up at <http://watson.apa.org/notify/> and you will be notified by e-mail when issues of interest to you become available!

Dominance Probability Macro

This macro executes dominance analysis as described in “The Dominance Analysis Approach for Comparing Predictors in Multiple Regression,” by Razia Azen and David Budescu (*Psychological Methods*, 2003, Vol. 8, No. 2). Dominance analysis quantifies the importance of each predictor as its average increment to the model r-squared, across all possible submodel sizes. This macro was written by Razia Azen with valuable contributions from Robert Ceurvorst.

NOTE: This program is limited to at most 10 predictors!

----- INSTRUCTIONS: -----

1. In this macro, change the %macro dom line (which appears just below these instructions) according to the following directions:
%macro dom (data=_last_, dep=Y, indep='list of predictors', p=n_of_predictors,
 noprint=0, short=0, B=n_of_bootstraps, predtype='type_of_predictors',
 seed=random_number);

If defaults are used, then only p= OR indep= is necessary, e.g., "%dom(p=4)" will use the last data set created and operate on variables Y and X1-X4.

Either p= OR indep= is required -- not both. If both are specified, p is determined by counting the variables in the indep= list.

data= SAS dataset to be used. Default is last dataset created.

dep= Name of dependent variable. Default is Y.

indep= List of predictors in quotes.
OR

p= No. of predictors IF and ONLY IF they are named X1-Xp,
 in which case indep= is not required.

noprint=1 Suppresses printing of the input dataset.

short=1 Suppresses listing of each predictor's contributions to individual submodels.

B= The number of bootstrap runs. Default is 1000.

predtype= 'f' if the predictors are fixed;
 'r' if the predictors are random. Default is 'r'.

seed= Six digits. Default is 0 (uses clock value).

For example, with 4 predictors named X1, X2, X3, X4 and all default values:
(data=_last_, dep=Y, p=4, indep=, noprint=0, short=0, B=1000, predtype='r', seed=0);
Or, with 4 predictors named A, B, C, D and all default values:
(data=_last_, dep=Y, p=, indep='A B C D', noprint=0, short=0, B=1000, predtype='r', seed=0);

2. Save this macro, and add the following two lines (below) to the SAS program in which you've read the data set to be analyzed:

```
%include 'k:\sasmacro\uniDAmacro.sas'; *** CHANGE TO PATH WHERE MACRO IS SAVED ***;
%dom;
```

3. Run the program edited in step 2.

```
-- END OF INSTRUCTIONS --
*****/
option nosource;
/***** CHANGE ONLY THE LINE BELOW (See step 1 above)!!! *****/
%macro dom (data=_last_, dep=Y, p=, indep=' ', noprint=0, short=0, B=1000, predtype='r',
seed=0);
/*****/

%if &indep= and &p= %then %do;
    %put YOU MUST SPECIFY INDEP=list of predictors OR P=no. of predictors (IF THEY ARE NAMED
X1-Xp).;
    %goto done;
%end;

%else %if &indep ne %then %do;
    %if %index(&indep,%str('%'))>0 or %index(&indep,%str('%'))>0
        %then %let indep = %substr(&indep,2,%length(&indep)-2);
    %if %index(&indep,-) > 0 %then %expand;
    %let p=1;
    %do %while(%length(%scan(&indep,&p))>0);
        %let x&p = %scan(&indep,&p);
        %let p = %eval(&p+1);
    %end;
    %let p = %eval(&p-1);
%end;

%else %if &p > 0 %then %do;
    %let indep=;
    %do i=1 %to &p;
        %let indep=&indep x&i;
        %let x&i=x&i;
    %end;
%end;

%if &p>10 %then %do;
    %put THE MAXIMUM NUMBER OF PREDICTORS IS 10. YOU HAVE &P..;
    %goto done;
%end;

%LET DATANAME=&SYSLAST;
option nonotes;
DATA _NULL_; FILE PRINT LINESLEFT=WITH; CALL SYMPUT('WITH',WITH);
DATA _NULL_;
    FILE PRINT NOTITLES LINESLEFT=WITHOUT;
    CALL SYMPUT ('MTITLE',TRIM(LEFT(WITHOUT-&WITH+2)));
    CALL SYMPUT('NMODELS', 2**&p -1);
RUN;

data original; set &data;

%if &noprint=0 %then %do;
proc print;
title&mtitle 'Input data set (check to make sure it is correct)'; run;
%end;

data labels; set;
keep &indep;
length dvlabel $40;
call label(&dep,dvlabel);
```

```

call symput('dvl',trim(dvllabel));
if _n_=1 then stop;
run;

/*****
PART 1: regression/DA of the original data
*****/

title 'regression results for original data set';
option notes;
proc reg corr data=original outest=onereg (keep=_in_ _rsq_ &indep);
  model &dep=&indep / stb pcorr2 scorr2 ;
  model &dep=&indep / selection=adjrsq best=&nmodels %if &short=1 %then noprint;;
run; quit; option nonotes;

* order all subset models in lexicographical order;
data modelmat;
  set onereg;
  %if %substr(&sysver,1,1)=8 %then if _n_>1;;
  %do i=1 %to &p;
    &x&i=(&x&i>.);
  %end;
  if _IN_=. then delete;
  keep &indep _IN_ _RSQ_;
proc sort; by _IN_ %do i=1 %to &p; descending &x&i %end;; run;
proc print; title 'R-squared values for all subset models'; run;

/**** %if &short=0 %then %do;
  proc print; id _IN_ _RSQ_; format _RSQ_ 5.4; run;
%end; ****/

/*****
DOMINANCE ANALYSIS TABLE
*****/
option notes; title;
proc IML;
  reset noprint;
  start;
  * read the subset models matrix into DOM;
  use modelmat;
  read all into DOM;
  close modelmat;

* dom contains, for each subset model, 1/0 values for the predictors,
the number of predictors in (_IN_), and the model r2 (_RSQ_);

  p=ncol(dom)-2;
  ncol=ncol(dom);

* dom is rearranged to contain, for each subset model, the number of predictors
in and the r2 of the model, followed by 1/0 values for predictors in/out;
dom1=dom[,1:p]; dom2=dom[(p+1):ncol]; dom=dom2||dom1; free dom1 dom2;

* generate table of additional contributions;
null=J(1,ncol,0);
dom=null//dom;
nrow=nrow(dom);

```

```

* dom now contains a top row of zeros for the null model;
full=J(1,p,99); fullrsq=J(1,1,99);
reduced=J(1,p,99); redrsq=J(1,1,99);
contrib=J(nrow,p,0); * additional contributions matrix;

do i=1 to nrow;      * for each model;
  do j=1 to p;      * for each predictor;

    if dom(|i,j+2|)=0 then do; * if predictor is not in subset model;
      reduced=dom[i,3:ncol]; * the 1/0 row represents the reduced model;
      full=reduced; * the full model is same as reduced model;

      do k=1 to p;
        if k=j then full(|1,k|)=1; * add the jth predictor to the full model;
      end;

      do r=1 to nrow; * for each model;
        comp=dom[r,3:ncol]; * comp is the 1/0 row of dom;
        if comp=full then fullrsq=dom[r,2]; * r2 of row is fullrsq, or;
        if comp=reduced then redrsq=dom[r,2]; * r2 of row is redrsq;
      end;

      contrib(|i,j|)=fullrsq-redrsq; * contrib is r2 difference;
    end;

  else do;
    contrib(|i,j|)=.; * if predictor is in model, contrib is .;
  end;

end;

end;

contrib=dom||contrib;
contrib=contrib[1:nrow-1,];

cols = {IN RSQ %do i=1 %to &p; &&X&i %end; %do i=1 %to &p; CP&i %end;};

create rsqtable from contrib[colname=cols];
append from contrib;
close rsqtable;

finish;
run; quit;

option nonotes;

%if &short=0 %then %do;
  proc print data=rsqtable; id IN RSQ; format RSQ CP1-CP&p 5.4;
  title&mtitle 'Dominance Table: Additional Contributions of Predictors Across All Subset
Regression Models';
  %UNQUOTE(TITLE%EVAL(&MTITLE+1)) 'CPi indicates the additional contribution of predictor i
to the model r-square';
%end;

proc summary nway; var CP1--CP&p; class in;
  output out=avgcont (drop=_type_) mean=&indep;
run;
proc means noprint; var &indep;
  output out=meanc mean=;
run;

```



```

data avgcont; set labels meanc (in=y) avgcont;
  if y then in=999;
proc print double; id IN; var &indep;
  format &indep 5.4 IN %if &p<11 %then 1.0; %else 2.0;;
title&mttitle 'Dominance Analysis: Overall Average Contributions of Predictors (First
Row)';
%UNQUOTE(TITLE%EVAL(&MTITLE+1)) 'And Average Contributions to Models of Each Size
(Remaining Rows)'; run;

proc transpose prefix=size out=meanc (rename=(SIZE999=OVERALL _NAME_=VAR)); id IN; var
&indep; run;
proc sort; by descending overall;
title&mttitle 'Dominance Analysis: Average Predictor Contributions Overall and to Models of
Each Size';
proc print; id _character_; format _numeric_ 5.4; run;

TITLE&MTITLE;
%DONE: option notes _last_=&syslast;
run;

/*****
      DOMINANCE MATRICES (Sample)
*****/

Proc IML;
start;
use modelmat;
  read all into damat;
close modelmat;
Dcsample=J(&p,&p,0); Dasample=J(&p,&p,0); Dgsample=J(&p,&p,0);
RSQ=damat[,ncol(damat)];

/*****
      Complete dominance      *****/
/*****/
do i=1 to &p-1;
  do j=i+1 to &p;

    * define matrix of constants ("comp") to determine complete dominance between
    each pair of predictors;

    comp=J(2**(&p-2),2**&p-1,0);

    * Xh is any subset model that does not include i and j;
    * Xh contains the columns of damat that do not involve i or j;
    Xh=J(nrow(damat),&p-2,99);
    Xhcol=0;                * index column number in Xh;
    do h=1 to &p;
      if (i ^= h & j ^= h) then do;    * find non i,j columns;
        Xhcol=Xhcol+1;                * update column number in Xh;
        Xh[,Xhcol] = damat[,h];      * assign column to Xh;
      end;
    end;

    * contrast rows (subsets) representing XiXh and XjXh;
    comprow=1;
    do r=1 to 2**&p-2;                * for each pair of rows;
      do s=r+1 to 2**&p-1;

```

```

        if Xh[r,]=Xh[s,] then do;          * if the rows of Xh are the same;
                                           * and if i and j are contrasted;
        if (damat[r,i]=1 & damat[r,j]=0 & damat[s,i]=0 &
            damat[s,j]=1) then do;
            comp[comprow,r]=1;
            comp[comprow,s]=-1;
            comprow=comprow+1;
        end;
    end;

end;    * do r loop;
end;    * do s loop;

*** Determine the complete Dij value ***;

cdiffij=comp*RSQ;

zero=J(nrow(cdiffij), ncol(cdiffij),0);
ijdom=nrow(cdiffij);

* obtain complete dominance matrices;
* undetermined case (all differences are zero);
if cdiffij=zero then do;
    Dcsample[i,j]=0.5;
    Dcsample[j,i]=Dcsample[i,j];
end;
* else, check signs of difference elements;
if cdiffij ^= zero then do;
    nonneg=0; nonpos=0;
    do k=1 to nrow(cdiffij);
        if cdiffij[k,]>=0 then nonneg=nonneg+1;
        if cdiffij[k,]<=0 then nonpos=nonpos+1;
    end;
    * dominance case;
    if nonneg=ijdom then Dcsample[i,j]=1; else Dcsample[i,j]=0;
    if nonpos=ijdom then Dcsample[j,i]=1; else Dcsample[j,i]=0;
    * undetermined case (differences have different signs);
    if (Dcsample[i,j]=0 & Dcsample[j,i]=0) then do;
        Dcsample[i,j]=0.5;
        Dcsample[j,i]=Dcsample[i,j];
    end;
end;

/***** Conditional dominance *****/
/***** Conditional dominance *****/
* define matrix of constants ("avg") to determine average (within model
size) dominance between each pair of predictors;

subsets=damat;
* subsets is a matrix that consists of p+2 columns (and
the null model in the top row): 1/0 for x1 - xp, in, Rsq;

avg=J(&p,2**&p-1,0);
do c=1 to &p;          * for each model size;
    num=0;             * index number of subsets per size;
    do m=1 to 2**&p-1; * for each model, determine number of models
                        of size is c-1 to which the ith predictor makes

```

```

                                an additional contribution;
    if (subsets[m,&p+1]=c-1 & subsets[m,i]=0) then num=num+1;
end;
if num=0 then num=1;      * for the case of the null model (size=0);
do r=1 to 2**&p-1;      * determine subsets to contrast;
  * consider additional contributions for model size (c);
  if (subsets[r,&p+1]=c-1 | subsets[r,&p+1]=c) then do;
    * average contribution of models that include i but not j;
    if (subsets[r,i]=1 & subsets[r,j]=0) then avg[c,r]=1/num;
    * average contribution of models that include j but not i;
    if (subsets[r,i]=0 & subsets[r,j]=1) then avg[c,r]=-1/num;
  end;
end;
end;

*** Determine the average Dij value ***;

adiffij=avg*RSQ;

zero=J(nrow(adiffij), ncol(adiffij),0);
ijdom=nrow(adiffij);

* obtain average dominance matrices;
* undetermined case (all differences are zero);
if adiffij=zero then do;
  Dasample[i,j]=0.5;
  Dasample[j,i]=Dasample[i,j];
end;
* else, check signs of difference elements;
if adiffij ^= zero then do;
  nonneg=0; nonpos=0;
  do k=1 to nrow(adiffij);
    if adiffij[k,]>=0 then nonneg=nonneg+1;
    if adiffij[k,]<=0 then nonpos=nonpos+1;
  end;
  * dominance case;
  if nonneg=ijdom then Dasample[i,j]=1; else Dasample[i,j]=0;
  if nonpos=ijdom then Dasample[j,i]=1; else Dasample[j,i]=0;
  * undetermined case (differences have different signs);
  if (Dasample[i,j]=0 & Dasample[j,i]=0) then do;
    Dasample[i,j]=0.5;
    Dasample[j,i]=Dasample[i,j];
  end;
end;

/*****
/*****      General dominance      *****/
/*****
* define matrix of constants ("glob") to determine global (overall average)
dominance between each pair of predictors;

glob=J(1,2**&p-1,99);
do g=1 to 2**&p-1;
  glob[g]=avg[+,g]/&p;
end;

*** Determine the global Dij value ***;

gdiffij=glob*RSQ;

```

```

zero=J(nrow(gdiffij), ncol(gdiffij),0);
ijdom=nrow(gdiffij);

* obtain global dominance matrices;
* undetermined case (all differences are zero);
if gdiffij=zero then do;
  Dgsample[i,j]=0.5;
  Dgsample[j,i]=Dgsample[i,j];
end;
* else, check signs of difference elements;
if gdiffij ^= zero then do;
  nonneg=0; nonpos=0;
  do k=1 to nrow(gdiffij);
    if gdiffij[k,]>=0 then nonneg=nonneg+1;
    if gdiffij[k,]<=0 then nonpos=nonpos+1;
  end;
* dominance case;
if nonneg=ijdom then Dgsample[i,j]=1; else Dgsample[i,j]=0;
if nonpos=ijdom then Dgsample[j,i]=1; else Dgsample[j,i]=0;
* undetermined case (differences have different signs);
if (Dgsample[i,j]=0 & Dgsample[j,i]=0) then do;
  Dgsample[i,j]=0.5;
  Dgsample[j,i]=Dgsample[i,j];
end;
end;

end; *(i loop);
end; *(j loop);

create Dcsample from Dcsample;
  append from Dcsample;
close Dcsample;
create Dasample from Dasample;
  append from Dasample;
close Dasample;
create Dgsample from Dgsample;
  append from Dgsample;
close Dgsample;

* create a file with sample dominance values;
* variables: i, j, cell value;
Sample_dc = J(&p*(&p-1), 3, 0); Sample_da = J(&p*(&p-1), 3, 0);
Sample_dg = J(&p*(&p-1), 3, 0); row=0;
do i=1 to nrow(Dcsample);
  do j=1 to ncol(Dcsample);
    if i ^= j then do;
      row=row+1;
      Sample_dc[row,1]=i; Sample_dc[row,2]=j; Sample_dc[row,3]=Dcsample[i,j];
      Sample_da[row,1]=i; Sample_da[row,2]=j; Sample_da[row,3]=Dasample[i,j];
      Sample_dg[row,1]=i; Sample_dg[row,2]=j; Sample_dg[row,3]=Dgsample[i,j];
    end;
  end;
end;
create sample_dc from Sample_dc[colname={i j dij}];
  append from Sample_dc;
close sample_dc;
create sample_da from Sample_da[colname={i j dij}];

```

```

    append from Sample_da;
close sample_da;
create sample_dg from Sample_dg[colname={i j dij}];
    append from Sample_dg;
close sample_dg;

finish;
run;

/*****
*
*      PART 2: Bootstrapping Loop
*****/
* input:  data set containing Y, data set containing X;
data depvar;
    set original;
    keep &dep;
run;

data indepvar;
    set original;
    keep &indep;
run;

* steps:  1. resample, 2. compute r-squared vector, 3. repeat B times;
option ls=70 nonotes nosource nosource2;
proc IML;
reset noprint; * nolog;
start;

*A. Generate a vector of B seeds (seedvec);
/*****
/* The vector seedvec contains B random seeds from a uniform          */
/* distribution, generated using the initial value (seed) specified    */
/* above. The vector is then saved to an external file, seeds,        */
/* which is subsequently read into the SAS dataset (seeds).            */
*****/

seedvec=J(&B,1,0);
do i=1 to &B;
    seedvec(|i,|) = uniform(&seed)*100000;
end;

create seeds from seedvec;
    append from seedvec;
close seeds;

finish;
run; quit;

/*=====*/
/*      start 1 to B loop                                           */
/*=====*/

%do r = 1 %to &B;

proc IML;
reset noprint; * nolog;
start;

```



```

* obtain Aij and Dij for each ij pair;

* read in the original data set (into DATA);
use depvar;
  read all into YData;
close depvar;
use indepvar;
  read all into XData;
close indepvar;
DATA = YData || XData;
Y = YData;
X = XData;

* read in the seeds vector (into seedvec);
use seeds;
  read all into seedvec;
close seeds;

n=nrow(DATA);
r=ncol(DATA);

*** check input;
if &r=1 then do;
  print 'Check that the data were read properly!!!';
  first5Y = J(1,5,0); first5Y=Y[1:5,];
  first5X = J(&p,5,0); first5X=X[1:5,];
  print 'The number of observations used is:' n;
  print 'The total number of variables used is:' r;
  print 'The dependent values for the first five cases are:' first5Y;
  print 'The predictor values for the first five cases are:' first5X;
  print 'You requested' &B 'bootstrap samples';
  print 'The bootstrap procedure is now running... PLEASE WAIT...';
end;

/*****
/* For bootstrapping run i, the i-th entry of the vector seedvec          */
/* (resample) is used to generate a vector (integers) whose entries      */
/* consist of the integers 1 to n sampled randomly, with replacement,    */
/* from a uniform distribution.                                           */
*****/
integers=J(n,1,0);

* Specify seed for resampling (resample);
resample=J(1,1,0);
resample=seedvec(|&r,|);

* generate a vector (integers) containing the integers 1 to n sampled
randomly, with replacement, from a uniform distribution whose initial seed is
resample;
Do i=1 to n;
  integers(|i|)=int(uniform(resample)*(n-1)+1);
end;
*print &r integers;

%if &predtype='r' %then %do;
/*=====*/
/***                PAIR RESAMPLING BOOTSTRAP                ***/
/*=====*/
/*****

```

```

/* The i-th entry in integers is an integer (j) between 1 and n which */
/* indicates the entry (row) number of the data matrix to be taken */
/* as the i-th entry (row) in the matrix obsboot. This is repeated */
/* for i = 1 to n. */
/*****
* obsboot is a matrix of randomly sampled observations;
obsboot=J(n,r,0);
Do i=1 to n;
  j=integers(|i, |);
  obsboot(|i, |)=DATA(|j, |);
end;

* convert bootstrapped data to external file;
*****;
create bootdata from obsboot[colname={&dep &indep}];
append from obsboot;
close bootdata;
run;

%end;
/*===== END of Pair Resampling =====*/

%if &predtype='f' %then %do;
/*=====*/
/***          RESIDUAL RESAMPLING BOOTSTRAP          ***/
/*=====*/
/*****
/* From the matrix DATA, whose first column represents Y and the */
/* remaining columns represent X (as specified above), the vector of */
/* Least-squares predicted values, YHat, and the vector of the */
/* Least Squares residuals, Res, are computed. */
/*****

* fit a model with an intercept to obtain YHat and Res;
ONE=J(n,1,1);
X1=ONE|X;
YHat=X1*INV(X1`*X1)*X1`*Y;
Res=YHat-Y;

/*****
/* The i-th entry in integers is an integer between 1 and n (j) which */
/* indicates the entry (row) number of the residuals vector, Res, */
/* to be added to the i-th predicted value, Y-Hat. This is repeated */
/* for i = 1 to n. The vector Yboot consists of the initial predicted */
/* values (YHat) with the randomly selected residuals added. The */
/* matrix Fixdata consists of Yboot and the original X matrix. */
/*****

* Resboot is a vector of randomly sampled residuals;
Resboot=J(n,1,0);
Do i=1 to n;
  j=integers(|i, |);
  Resboot(|i, |)=Res(|j, |);
end;

* add random residuals to Y-hat;

```

```

Yboot=YHat+Resboot;

* an n by r matrix of bootstrapped Y and fixed Xs;
Fixdata=Yboot|X;

* convert bootstrapped data to external file;
*****;
create bootdata from obsboot[colname={&dep &indep}];
append from obsboot;
close bootdata;
run;

%end;
/*===== END of Residual Resampling =====*/

finish;
run; quit;

* produce bootstrap data sets;

/* %%%%%%%%%%%%%%%*/
/*          Computing Dij          */
/* %%%%%%%%%%%%%%%*/

* regression of bootstrapped data;
*****;

data resample;
set bootdata;
run;

proc reg corr data=resample noprint outest=regfile (keep=_in_ _rsq_ &indep);
  model &dep=&indep / selection=adjrsq;
run; quit; option nonotes;

* order all subset models in lexicographical order;
data bootmat;
set regfile;
%if %substr(&sysver,1,1)=8 %then if _n_>1;;
%do i=1 %to &p;
  &x&i=(&x&i>.);
%end;
if _IN_=. then delete;
keep &indep _IN_ _RSQ_;
run;
proc sort; by _IN_ %do i=1 %to &p; descending &x&i %end;; run;

* obtain Aij and Dij for each ij pair;
proc IML;
reset noprint; * nolog;
start;
/*****
*          COMPUTE DOMINANCE MATRICES (resample)
*****/
use bootmat;
read all into damat;
close bootmat;

```

```

* damat is a matrix that consists of p+2 columns: 1/0 for x1 - xp, in, Rsq;

use Dcsample;
  read all into sampleDc;
close Dcsample;
use Dasample;
  read all into sampleDa;
close Dasample;
use Dgsample;
  read all into sampleDg;
close Dgsample;

* at first run, initialize frequency matrices and sample pattern counter;
if &r =1 then do;
  freqDc1=J(&p, &p, 0); freqDc0=J(&p, &p, 0); freqDcn=J(&p, &p, 0);
  freqDal=J(&p, &p, 0); freqDa0=J(&p, &p, 0); freqDan=J(&p, &p, 0);
  freqDg1=J(&p, &p, 0); freqDg0=J(&p, &p, 0); freqDgn=J(&p, &p, 0);
  patcount=J(1,3,0); *columns correspond to dominance types;
end;

* at subsequent runs retrieve files as sum matrices;
if &r ^= 1 then do;
use sumDc1;
  read all into freqDc1;
close sumDc1;
use sumDc0;
  read all into freqDc0;
close sumDc0;
use sumDcn;
  read all into freqDcn;
close sumDcn;

use sumDal;
  read all into freqDal;
close sumDal;
use sumDa0;
  read all into freqDa0;
close sumDa0;
use sumDan;
  read all into freqDan;
close sumDan;

use sumDg1;
  read all into freqDg1;
close sumDg1;
use sumDg0;
  read all into freqDg0;
close sumDg0;
use sumDgn;
  read all into freqDgn;
close sumDgn;

use patfreq;
  read all into patcount;
close patfreq;

end;

* r-squared vector is the last column of damat (without null model);
RSQ=damat[,ncol(damat)];

```

```

*print &r RSQ;

Dc=J(&p,&p,0); Da=J(&p,&p,0); Dg=J(&p,&p,0);

/***** Complete dominance *****/
do i=1 to &p-1;
  do j=i+1 to &p;

    * define matrix of constants ("comp") to determine complete dominance between
    each pair of predictors;

    comp=J(2**(&p-2),2**&p-1,0);

    * Xh is any subset model that does not include i and j;
    * Xh contains the columns of damat that do not involve i or j;
    Xh=J(nrow(damat),&p-2,99);
    Xhcol=0; * index column number in Xh;
    do h=1 to &p;
      if (i ^= h & j ^= h) then do; * find non i,j columns;
        Xhcol=Xhcol+1; * update column number in Xh;
        Xh[,Xhcol] = damat[,h]; * assign column to Xh;
      end;
    end;

    * contrast rows (subsets) representing XiXh and XjXh;
    comprow=1;
    do r=1 to 2**&p-2; * for each pair of rows;
      do s=r+1 to 2**&p-1;
        if Xh[r,]=Xh[s,] then do; * if the rows of Xh are the same;
          * and if i and j are contrasted;
          if (damat[r,i]=1 & damat[r,j]=0 & damat[s,i]=0 &
              damat[s,j]=1) then do;
            comp[comprow,r]=1;
            comp[comprow,s]=-1;
            comprow=comprow+1;
          end;
        end;
      end;
    end; * do r loop;
  end; * do s loop;

*** Determine the complete Dij value ***;

cdiffij=comp*RSQ;

zero=J(nrow(cdiffij), ncol(cdiffij),0);
ijdom=nrow(cdiffij);

* obtain complete dominance matrices;
* undetermined case (all differences are zero);
if cdiffij=zero then do;
  Dc[i,j]=0.5;
  Dc[j,i]=Dc[i,j];
end;
* else, check signs of difference elements;

```



```

if cdiffij ^= zero then do;
  nonneg=0; nonpos=0;
  do k=1 to nrow(cdiffij);
    if cdiffij[k,]>=0 then nonneg=nonneg+1;
    if cdiffij[k,]<=0 then nonpos=nonpos+1;
  end;
  * dominance case;
  if nonneg=ijdom then Dc[i,j]=1; else Dc[i,j]=0;
  if nonpos=ijdom then Dc[j,i]=1; else Dc[j,i]=0;
  * undetermined case (differences have different signs);
  if (Dc[i,j]=0 & Dc[j,i]=0) then do;
    Dc[i,j]=0.5;
    Dc[j,i]=Dc[i,j];
  end;
end;

```

```

/*****
/*****      Conditional dominance      *****/
/*****
/*****
* define matrix of constants ("avg") to determine average (within model
size) dominance between each pair of predictors;

subsets=damat;
* subsets is a matrix that consists of p+2 columns (and
the null model in the top row): 1/0 for x1 - xp, in, Rsq;

avg=J(&p,2**&p-1,0);
do c=1 to &p;          * for each model size;
  num=0;              * index number of subsets per size;
  do m=1 to 2**&p-1;   * for each model, determine number of models
                        of size is c-1 to which the ith predictor makes
                        an additional contribution;
    if (subsets[m,&p+1]=c-1 & subsets[m,i]=0) then num=num+1;
  end;
  if num=0 then num=1; * for the case of the null model (size=0);
  do r=1 to 2**&p-1;   * determine subsets to contrast;
    * consider additional contributions for model size (c);
    if (subsets[r,&p+1]=c-1 | subsets[r,&p+1]=c) then do;
      * average contribution of models that include i but not j;
      if (subsets[r,i]=1 & subsets[r,j]=0) then avg[c,r]=1/num;
      * average contribution of models that include j but not i;
      if (subsets[r,i]=0 & subsets[r,j]=1) then avg[c,r]=-1/num;
    end;
  end;
end;

*** Determine the average Dij value ***;

adiffij=avg*RSQ;

zero=J(nrow(adiffij), ncol(adiffij),0);
ijdom=nrow(adiffij);

* obtain average dominance matrices;
* undetermined case (all differences are zero);
if adiffij=zero then do;
  Da[i,j]=0.5;
  Da[j,i]=Da[i,j];

```

```

end;
* else, check signs of difference elements;
if adiffij ^= zero then do;
  nonneg=0; nonpos=0;
  do k=1 to nrow(adiffij);
    if adiffij[k,]>=0 then nonneg=nonneg+1;
    if adiffij[k,]<=0 then nonpos=nonpos+1;
  end;
* dominance case;
if nonneg=ijdom then Da[i,j]=1; else Da[i,j]=0;
if nonpos=ijdom then Da[j,i]=1; else Da[j,i]=0;
* undetermined case (differences have different signs);
if (Da[i,j]=0 & Da[j,i]=0) then do;
  Da[i,j]=0.5;
  Da[j,i]=Da[i,j];
end;
end;

/***** General dominance *****/
/*****
* define matrix of constants ("glob") to determine global (overall average)
dominance between each pair of predictors;

glob=J(1,2**&p-1,99);
do g=1 to 2**&p-1;
  glob[,g]=avg[+,g]/&p;
end;

*** Determine the global Dij value ***;

gdifffij=glob*RSQ;
*print 'global dominance comparisons';
*print i j gdifffij;

zero=J(nrow(gdifffij), ncol(gdifffij),0);
ijdom=nrow(gdifffij);

* obtain global dominance matrices;
* undetermined case (all differences are zero);
if gdifffij=zero then do;
  Dg[i,j]=0.5;
  Dg[j,i]=Dg[i,j];
end;
* else, check signs of difference elements;
if gdifffij ^= zero then do;
  nonneg=0; nonpos=0;
  do k=1 to nrow(gdifffij);
    if gdifffij[k,]>=0 then nonneg=nonneg+1;
    if gdifffij[k,]<=0 then nonpos=nonpos+1;
  end;
* dominance case;
if nonneg=ijdom then Dg[i,j]=1; else Dg[i,j]=0;
if nonpos=ijdom then Dg[j,i]=1; else Dg[j,i]=0;
* undetermined case (differences have different signs);
if (Dg[i,j]=0 & Dg[j,i]=0) then do;
  Dg[i,j]=0.5;
  Dg[j,i]=Dg[i,j];
end;
end;

```

```

end;

end; *(i loop);
end; *(j loop);

**** compare current D matrices to sample D matrices ****;
if Dc=sampleDc then patcount[,1]=patcount[,1]+1;
if Da=sampleDa then patcount[,2]=patcount[,2]+1;
if Dg=sampleDg then patcount[,3]=patcount[,3]+1;
create patfreq from patcount;
  append from patcount;
close patfreq;

thisDc1=J(&p,&p,0); thisDc0=J(&p,&p,0); thisDcn=J(&p,&p,0);
thisDal=J(&p,&p,0); thisDa0=J(&p,&p,0); thisDan=J(&p,&p,0);
thisDg1=J(&p,&p,0); thisDg0=J(&p,&p,0); thisDgn=J(&p,&p,0);
do i=1 to &p;
  do j=1 to &p;
    if i ^=j then do;
      if Dc[i,j]=1 then thisDc1[i,j]=1;
      if Dc[i,j]=0 then thisDc0[i,j]=1;
      if Dc[i,j]=0.5 then thisDcn[i,j]=1;
      if Da[i,j]=1 then thisDal[i,j]=1;
      if Da[i,j]=0 then thisDa0[i,j]=1;
      if Da[i,j]=0.5 then thisDan[i,j]=1;
      if Dg[i,j]=1 then thisDg1[i,j]=1;
      if Dg[i,j]=0 then thisDg0[i,j]=1;
      if Dg[i,j]=0.5 then thisDgn[i,j]=1;
    end;
  end;
end;

**** add current D matrices to previous D matrices ****;
freqDc1 = freqDc1+thisDc1;
freqDal = freqDal+thisDal; freqDg1 = freqDg1+thisDg1;
freqDc0 = freqDc0+thisDc0;
freqDa0 = freqDa0+thisDa0; freqDg0 = freqDg0+thisDg0;
freqDcn = freqDcn+thisDcn;
freqDan = freqDan+thisDan; freqDgn = freqDgn+thisDgn;

create sumDc1 from freqDc1;
  append from freqDc1;
close sumDc1;
create sumDc0 from freqDc0;
  append from freqDc0;
close sumDc0;
create sumDcn from freqDcn;
  append from freqDcn;
close sumDcn;

create sumDal from freqDal;
  append from freqDal;
close sumDal;
create sumDa0 from freqDa0;
  append from freqDa0;
close sumDa0;

```

```

create sumDan from freqDan;
  append from freqDan;
close sumDan;

create sumDg1 from freqDg1;
  append from freqDg1;
close sumDg1;
create sumDg0 from freqDg0;
  append from freqDg0;
close sumDg0;
create sumDgn from freqDgn;
  append from freqDgn;
close sumDgn;

*at last run create summary tables containing the following columns:
i, j, f=1, f=0, f=0.5 pij pji;
if &r=&B then do;
  row=1;
  summary_c=J(&p*(&p-1), 8, -9);
  summary_a=J(&p*(&p-1), 8, -9);
  summary_g=J(&p*(&p-1), 8, -9);
  do i=1 to &p;
    do j=1 to &p;
      if i^=j then do;
        summary_c[row,1]=i; summary_c[row,2]=j;
        summary_c[row,3]=freqDc1[i,j];
        summary_c[row,4]=freqDc0[i,j]; summary_c[row,5]=freqDcn[i,j];
        summary_c[row,6]=freqDc1[i,j]/&B; summary_c[row,7]=freqDc0[i,j]/&B;
        summary_c[row,8]=freqDcn[i,j]/&B;
        summary_a[row,1]=i; summary_a[row,2]=j;
        summary_a[row,3]=freqDa1[i,j];
        summary_a[row,4]=freqDa0[i,j]; summary_a[row,5]=freqDan[i,j];
        summary_a[row,6]=freqDa1[i,j]/&B; summary_a[row,7]=freqDa0[i,j]/&B;
        summary_a[row,8]=freqDan[i,j]/&B;
        summary_g[row,1]=i; summary_g[row,2]=j;
        summary_g[row,3]=freqDg1[i,j];
        summary_g[row,4]=freqDg0[i,j]; summary_g[row,5]=freqDgn[i,j];
        summary_g[row,6]=freqDg1[i,j]/&B; summary_g[row,7]=freqDg0[i,j]/&B;
        summary_g[row,8]=freqDgn[i,j]/&B;
        row=row+1;
      end;
    end;
  end;
create csummary from summary_c[colname={i j f1 f0 fn Pij Pji pno}];
  append from summary_c;
close csummary;
create asummary from summary_a[colname={i j f1 f0 fn Pij Pji pno}];
  append from summary_a;
close asummary;
create gsummary from summary_g[colname={i j f1 f0 fn Pij Pji pno}];
  append from summary_g;
close gsummary;
create patsum from patcount[colname={complete conditional general}];
  append from patcount;
close patsum;
end;

finish;
run; quit;

```

```
%end; ***** end the 1 to B loop *****;
```

```

/*****
/*      PART 3: Probabilities and reproducibilities      */
*****/
* input:  nine p by p matrices of COUNTS;
* stpes:  convert to probabilities, compute test statistics / CIs, plots;
* output: tables;
data average;
  set asummary;
  dominance='conditional';
  pair=i*10+j;
run;
data complete;
  set csummary;
  dominance='complete';
  pair=i*10+j;
run;
data global;
  set gsummary;
  dominance='general';
  pair=i*10+j;
run;

data merged;
  set complete average global;
run;
/* proc print data=merged; title 'results'; run; */
data results; set merged;
  array f[3] f1 f0 fn;
  do domtype=1 to 3;
    domfreq = f[domtype];
    output;
  end;
run;
data results; set results;
  if domtype=1 then Dij=1;
  if domtype=2 then Dij=0;
  if domtype=3 then Dij=0.5;
run;
proc sort data=results; by dominance pair; run;
proc means data=results noprint;
  var Dij; freq domfreq;
  by dominance pair;
  output out=meansout mean=Dbar STD=SE;
run;

data samplea;
  set Sample_da; pair=i*10+j; dominance='conditional';
run;
data samplec;
  set Sample_dc; pair=i*10+j; dominance='complete';
run;
data sampleg;
  set Sample_dg; pair=i*10+j; dominance='general';
run;
data sample;

```



```

set samplec samplea sampleg;
run;

***** SUMMARY OUTPUT *****;
* variables: dominance-type, i, j, p(iDj), p(jDi), p(noD), Dbar, SE,
sample-value, pattern-prob;
*** merge sample, results and meansout data sets by dominance and pair;
proc sort data=sample; by dominance pair; run;
data results; set results; if domtype=1; run;
proc sort data=results; by dominance pair; run;
proc sort data=meansout; by dominance pair; run;
data printout;
  merge results sample meansout;
  by dominance pair;
run;

data table;
  set printout(drop=domtype domfreq _TYPE_ _FREQ_);
/* i_dom_j=pij; j_dom_i=pji; no_dom=pno; */
  Pijno=pno;
  Dij_mean = dbar;
  Dij_se = se;
  if dij=1 then reprod = pij;
  if dij=0 then reprod = pji;
  if dij=0.5 then reprod = pijno;
  format Dij_se F6.3;
run;
data table; set table(drop = f1 f0 fn pno dbar); run;
proc sort; by dominance pair; run;
proc print data=table;
  title 'summary of results: all pairs';
  title2 'Pij, Pji, Pijno are dominance probabilities and';
  title3 'reprod are reproducibility values';
  var dominance i j Dij Dij_mean dij_SE Pij Pji Pijno reprod;
run;
data plots; set table;
  if Pij ge Pji;
  format Dij_se F6.3;
run;
proc print data=plots;
  title 'summary of results: pair arranged by dominant predictor';
  title2 'Pij, Pji, Pijno are dominance probabilities and';
  title3 'reprod are reproducibility values';
  var dominance i j Dij Dij_mean dij_SE Pij Pji Pijno reprod;
run;

%mend dom;

/*****
  The following macros expand a variable list containing
  hyphens into a list specifying each individual variable.
*****/

%MACRO EXPAND;
  %LET LENGTH = %LENGTH(&INDEP);  %LET TEMP=;
  %DO _INDEX_ = 1 %TO &LENGTH;
    %LET ITEM = %SCAN(&INDEP,&_INDEX_,%QUOTE( ));

```

```

      %IF %LENGTH(&ITEM) EQ 0 %THEN %GOTO DONE;
      %IF %INDEX(&ITEM,-) > 0 %THEN %EXPANDED;
      %LET TEMP = &TEMP &ITEM;
      %END; %DONE: %LET INDEP = &TEMP;
%MEND EXPAND;

%MACRO EXPANDED;
  %LET DASH = %INDEX(&ITEM,-);
  %DO I = %EVAL(&DASH-1) %TO 1 %BY -1;
    %LET ALPHANUM = %SUBSTR(&ITEM,&I,1);
    %DO II = 0 %TO 9; %IF &ALPHANUM EQ &II %THEN %GOTO FOUND; %END;
    %GOTO DONE;
  %FOUND: %END;
  %DONE: %LET PREFX = %SUBSTR(&ITEM,1,&I);
  %LET LOWER = %SUBSTR(&ITEM,%EVAL(&I+1),%EVAL(&DASH-&I-1));
  %LET UPPER = %SUBSTR(&ITEM,%EVAL(&DASH+&I+1));
  %LET ITEM=; %DO II = &LOWER %TO &UPPER; %LET ITEM = &ITEM &PREFX.&II;
  %END;
%MEND EXPANDED;

```