Equations for Deriving Effect Sizes for Individual Predictors and Sets of Predictors under Specified Conditions in Multiple Regression Analysis: A Technical Report[1]

John L. Michela, Ph.D.[2]

Many published papers do not fully report effect sizes of predictor variables' influences on outcome variable scores. For example, many papers' multiple regression (MR) analyses report regression coefficients for individual predictors without providing each predictor's unique proportion of variance explained. This unique proportion may be obtained by squaring semi-partial correlation (*sr*) values that are available as an option in many statistical analysis programs.

This incompleteness of information occurs in many empirical reports concerning students' evaluations of teaching (SET). For example, in an MR analysis by Cannon and Cipriani (2022), some of the predictor variables were included so that they could serve as validating criteria for the SET rating response, and other variables were included to document distortion in SET responses due to halo bias. In an article in reply to Cannon and Cipriani, Michela (in press at the time of on-line posting of this report) argued that it is important to isolate the effect sizes of as many of the validating and non-validating predictors as possible, so that the extent of halo in SET response—the focus of the research by Cannon and Cipriani—could best be gauged.

Accordingly, this document provides equations for deriving effect sizes for individual predictors and sets of predictors under specified conditions in MR analyses. The effect size metric being obtained in each instance is either Pearson *r* or proportion of variance explained, expressed as R^2 or sr^2.

Some computations require availability of the overall *F*-ratio (i.e., for the complete MR equation). If only the overall R^2 of an MR analysis is provided, its corresponding *F*-ratio is obtained as:

$$F = ((N - k - 1) * R^2) / (k * (1 - R^2)). \qquad (1)$$

Here, N is the total number of survey respondents, and k is the total number of predictor variables in the overall MR equation.

[1] Deposited May 27, 2022, to the UWSpace document repository at the University of Waterloo.
This document also serves as Supplement 2 to Michela (in press).
[2] Department of Psychology, University of Waterloo, Waterloo, Ontario, N2L 3G1 Canada; jmichela@uwaterloo.ca

By algebraic transformation of equation (1), R^2 may be obtained from F as follows:

$$R^2 = (F * k) / ((N - 1) + (k * (F - 1))). \quad (2)$$

The adjusted R^2 provides a correction for the overfitting that occurs in sample data, thus providing a population estimate of R^2. It is calculated as:

$$R^2\_adj = 1 - (1 - R^2) * ((N - 1) / (N - k - 1)). \quad (3)$$

If R^2_adj is available but R^2 is not, the latter is recoverable from this transformation of (3):

$$R^2 = 1 - (1 - R^2\_adj) * ((N - k - 1) / (N - 1)). \quad (4)$$

When individual predictors' unstandardized regression weights (b_i) are given along with the overall R^2 of the MR equation, their effect sizes are obtainable as sr^2 values from:

$$sr^2_i = (F_i * (1 - R^2)) / (N - k - 1). \quad (5)$$

For use in equation (5), F_i may be estimated from the square of the predictor's t-ratio, because

$$t^2 = F. \quad (6)$$

In practice, the table of coefficients from an MR analysis is likely to show some combination of regression coefficients (bs), standard errors (SEs) or t-ratios that is sufficient for obtaining the necessary quantities to insert into equation (5).

A Pearson correlation, r, generally is recoverable from an MR analysis when there are two linear predictors (as in the case of Model 1 in Cannon and Cipriani's paper). Given an overall R^2, and having calculated sr^2_2 (that is, sr^2 for the second of two predictors),

$$r_1 = \text{SQRT}(R^2 - sr^2_2), \quad (7)$$

and vice versa for r_2 and sr^2_1.

The unique R^2 for a set of predictors (R^2_change) is recoverable when the F-ratio unique to that set (F_change) has been provided along with R^2 for the full equation (R^2_full):

$$R^2\_change = (F\_change * k\_change * (1 - R^2\_full)) / (N - k\_full - 1). \quad (8)$$

Equation (8) is an algebraic transformation of a conventional equation for the F-ratio (F_change) concerning addition or subtraction of a set of k predictors (k_change) in an MR model:

$$F = ((N - k\_full - 1) * R^2\_change) / (k\_change * (1 - R^2\_full)). \quad (9)$$

References and Sources

Cannon, E., & Cipriani, G. P. (2022). Quantifying halo effects in students' evaluation of teaching. *Assessment and Evaluation in Higher Education*, *47* (1), 1-14. DOI:10.1080/02602938.2021.1888868

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kerlinger, F. N. & Pedhazur, E. J., (1973). *Multiple regression in behavioral research*. New York, NY: Holt Rinehart & Winston.

Michela, J. L. (in press). Toward understanding and quantifying halo in students' evaluation of teaching. *Assessment and Evaluation in Higher Education*.

Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics* (3rd ed.). New York: HarperCollins.