Behavior Research Methods
2010, 42 (1), 29-35
doi:10.3758/BRM.42.1.29

Two SPSS programs for interpreting multiple regression results

Urbano Lorenzo-Seva, Pere J. Ferrando, and Eliseo Chico
Universitat Rovira i Virgili, Tarragona, Spain

When multiple regression is used in explanation-oriented designs, it is very important to determine both the usefulness of the predictor variables and their relative importance. Standardized regression coefficients are routinely provided by commercial programs. However, they generally function rather poorly as indicators of relative importance, especially in the presence of substantially correlated predictors. We provide two user-friendly SPSS programs that implement currently recommended techniques and recent developments for assessing the relevance of the predictors. The programs also allow the user to take into account the effects of measurement error. The first program, MIMR-Corr.sps, uses a correlation matrix as input, whereas the second program, MIMR-Raw.sps, uses the raw data and computes bootstrap confidence intervals of different statistics. The SPSS syntax, a short manual, and data files related to this article are available as supplemental materials from http://brm.psychonomic-journals.org/content/supplemental.

Broadly speaking, in the behavioral sciences, multiple regression (MR) has two distinct applications (which are not mutually exclusive): prediction and explanation (Pedhazur, 1982). In a purely predictive study, interpretation of the MR results is relatively unimportant. However, in more theoretically oriented, explanation-based studies, it is very important to determine both the usefulness of the variables that were chosen as predictors and their relative importance. Standardized regression coefficients (also known as beta weights) are generally used to answer these questions, and they are routinely provided by the MR programs included in most statistical packages. In the context of studies related to prediction, relative weights and structure coefficients could be useful in identifying predictors that might be nearly equivalent, in which case other factors can be considered, such as the cost of obtaining each predictor.

It must be pointed out that Nunnally and Bernstein (1994) suggested that the definition of *importance* is not unique and that different situations might require different definitions of the word. One variable may explain the most variance, ignoring the other predictions (e.g., the largest squared validity); a second may contribute to the most unique variance (e.g., the largest beta weight), and a third may increment R^2 the most relative to a subset of predictors. As these three kinds of indexes answer different questions (and one is not better than the others), it is important for a statistical package to incorporate various measures of importance. In this way, the researcher can decide which measure should be inspected in each particular situation.

Standardized coefficients are context dependent (Courville & Thompson, 2001) and often do not work well for explanatory purposes, especially in the presence of substantially correlated predictors, in which case they can also become very unstable (Cooley & Lohnes, 1971; Johnson, 2000). To overcome these limitations, some authors suggest that additional indexes must be considered to augment interpretation (Cooley & Lohnes, 1971; Courville & Thompson, 2001), whereas others have developed new alternative measures for assessing the relative importance of the predictors (Budescu, 1993; Johnson, 2000, 2004). The programs described below incorporate both of these approaches and take into account the effects that the measurement error may have when one assesses the usefulness and relevance of the predictors.

SPSS Programs for Interpreting MR Results

Two user-friendly SPSS programs were created to implement the three approaches described above. Both programs run automatically from the SPSS (Norušis, 1991) syntax window, and the output can be configured. Specifically, these programs were developed on the basis of the MATRIX command language (see, e.g., Einspruch, 2003). It should be noted, however, that the user does not need to know how to program in this language in order to run our programs: It is only necessary to specify the values of some variables in order to adapt the syntax to the data at hand. In addition, our examples explain how these changes have to be made. Appendixes A and B show the extract of the code that can be modified by the user.

The first program, MIMR-Corr.sps, uses a correlation matrix as input: This matrix has to be defined explicitly by the user (as shown in Appendix A) in matrix `R`, where the first variable in the matrix is considered as the crite-

U. Lorenzo-Seva, urbano.lorenzo@urv.cat

rion variable. In addition, two pieces of information are needed. First, the name of the variable that corresponds to the criterion (i.e., the first variable in the matrix) has to be defined in the variable `vdname`, and the names of variables that correspond to the predictors (i.e., the rest of the variables in matrix `R`) have to be defined in the variable `pnames`. The second piece of information is the reliability of each predictor: It has to be defined in variable `rb_p`. If the reliabilities are not available, then a column of 1s (a 1 for each predictor) can be used instead.

The second program, MIMR-Raw.sps, uses the raw data: The user has to have an active SPSS data file containing the criterion (as the first variable in the data file) followed by the predictors (i.e., the rest of the variables in the data file). As in the previous program, the reliability of each predictor can be specified in `rb_p`. Both programs provide essentially the same output. However, in addition to the outcome obtained using MIMR-Corr.sps, the second program computes the 95th percentile confidence intervals for all of the measures to be obtained using a bootstrap approach. The limits of the confidence interval are constructed by taking the upper and lower 2.5 percentiles of the empirical distribution that are based on 500 bootstrap trials. A variable (`k`) specifies the number of bootstrap samples to be used. Finally, it should be noted that, in both programs, a variable (`detailed`) specifies the amount of information to be presented in the output.

Additional measures. Courville and Thompson (2001) suggested that, for interpretative purposes, standardized coefficients should be complemented with structure coefficients. In MR, the structure coefficient for predictor X_j is the correlation between this predictor and the regression function ($\hat{Y}$; see Cooley & Lohnes, 1971). MIMR-Corr .sps provides the point-estimated structure coefficients obtained from the correlation matrix.

Apart from the additional measures, both programs provide the statistics that are usually reported in standard MR analysis. Our aim is to provide a self-contained program, so that the user can obtain all the necessary information for interpreting the results, with no need to use other programs. Thus, both programs provide the point-estimated beta weights, as well as the multiple R and R^2. In addition, MIMR-Raw.sps provides the unstandardized regression weights.

Measures of relative importance. At present, the techniques of choice for assessing the relative importance of correlated predictors are Budescu's dominance analysis (Budescu & Azen, 2004) and Johnson's relative weight procedure (Johnson, 2000, 2004). Both techniques are superior to the traditional measures in most respects and are almost always in perfect agreement (Budescu & Azen, 2004). However, relative weights are much easier to calculate and are the measures chosen in our programs. Essentially, the relative weight measures the proportionate contribution each predictor makes to R^2 (i.e., the relative importance) after correcting for the effects of the intercorrelations among predictors.

The computation of relative weights involves four steps: (1) transforming the original (correlated) predictors to obtain a set of orthogonal variables that have the highest correspondence to the original predictors; (2) regressing the criterion on the orthogonal variables; (3) relating the orthogonal variables back to the original predictors; and (4) combining the information to yield the set of corrected relative weights. For a full description of the technique, see Johnson (2000). The resulting weights range from 0 to 1, and their sum equals R^2. In our programs, relative weights are reported as percentages (i.e., they are divided by R^2 and multiplied by 100). Johnson (2004) explicitly suggested constructing confidence intervals using a bootstrap approach. He also found that, in most cases, the distribution of the relative weights was approximately normal (except when the weight was near zero) and suggested that confidence intervals can be safely estimated using the simpler normal approximation. In our program, however, we use the percentile approach described above. It requires more computing effort than the normal approach for obtaining accurate results, but is a valid approach regardless of what the weight is.

Measurement error. It is well known that the presence of measurement error attenuates correlations. If the predictors have different degrees of reliability, the correlations among them and with the criterion are differentially attenuated with respect to the correlations that would be obtained if the measures were totally free of error. Because measurement error distorts the correlations, and because these correlations are the starting point for most MR results, it follows that the beta weights, the structure coefficients, and the relative weights are all potentially affected by the presence of measurement error. The problem is quite important in behavioral sciences because, in this domain, most variables are measured with a substantial amount of error.

If the reliability estimates of the predictors' scores are available, the usual correction-for-attenuation formula (see, e.g., Nunnally, 1978) can be applied to correct the correlation matrix used as input. In our programs, all the correlations are corrected by taking into account only the unreliability of the predictors. MIMR-Corr.sps directly corrects the input correlation matrix, whereas MIMR-Raw.sps computes the correlation matrix from the raw data and then corrects it. By default, both programs use the corrected correlation matrix as the starting point if the vector of predictor reliabilities is provided as input. If no reliabilities are provided, both programs will use the uncorrected correlation matrix.

An Empirical Example

The use of the programs is illustrated with an empirical study carried out by our research group. The criterion was a satisfaction with life measure derived from the Satisfaction With Life Scale (Pavot & Diener, 1993), whereas the five chosen predictors were measures of dispositional optimism (OPTI) and pessimism (PESI) (Scheier, Carver, & Bridges, 1994) and three components of emotional intelligence (Salovey, Mayer, Goldman, Turvey, & Palfai, 1995): emotional attention (ATTEN), clarity (CLARI), and repair (REPAIR). The study was mainly explanative: Its main goal was to assess the relative importance of the measures chosen to predict satisfaction with life. The correlation matrix, the names of the variables, and the reliability of each predictor were specified as shown in Appendix A. Inspection of the correlation matrix suggests that all of the predictors are related to the criterion and are substantially correlated with each other.

As an input to run the MIMR-Cor.sps program with this data set (see Appendix A), we needed the correlation matrix (where the first variable was the criterion and the other comprised the predictors), the labels for the criterion and the predictors, and the reliability of each predictor. As an input to run the MIMR-Raw.sps program (see Appendix B), we needed the raw data matrix (where the first variable was again the criterion and the other comprised the predictors) stored in an SPSS data file and the reliability of each predictor. Here we first discuss the results obtained without taking into account measurement error (i.e., the reliability of each predictor is set to 1: `compute rb_p = {1;1;1;1;1}`). For brevity, we only discuss the results obtained with the more complete MIMR-Raw.sps program.

The inspection of the outcome in Appendix C shows that the beta weights and the structure coefficients provide different information about the relevance of the predictors. The structure coefficients are substantially larger, and the rank order of magnitude (in absolute value) is different in both sets of measures. The beta weights are all relatively low. However, inspection of the signs of the upper and lower limits of the confidence intervals suggests that they are all statistically significant. If just the beta weights were to be interpreted, their values would suggest that all of the chosen predictors behave rather poorly. However, the multiple *R* estimate is reasonably good for a personality study, and, with the exception of ATTEN, the structure coefficients are rather high. Overall, ATTEN proves to be the weakest predictor. The remaining four measures must be considered as potentially useful predictors of satisfaction with life. However, the correlations among them show that their predictive power is shared among the predictors to a large degree (in a relatively arbitrary way) and that the corresponding standardized weights are low. This is clearly seen in the case of CLARI and REPAIR.

The standardized regression coefficients are context dependent and fail to take into account the direct effect of each predictor. On the other hand, the structure coefficients are reasonable estimates of relative importance but do not sufficiently reflect the effect of a particular predictor in the context of each of the other predictors (Johnson, 2000). The relative weights shown in Appendix C combine both effects: The estimates are the percentages of explained variance attributable to each of the five predictors. The rank order in this case is the same as that of the structure coefficients. However, the contributions are different. If both the direct effect and the context effect are taken into account, it seems clear that OPTI is the best predictor, followed at some distance by CLARI. In addition, PESI and REPAIR make similar contributions, and ATTEN is clearly the weakest predictor. Note also that the limits of the confidence intervals overlap as far as the contributions of the nonextreme predictors are concerned, which suggests that the results must be interpreted with some caution.

The reliability estimates of the scores in the five predictors (alpha coefficients) were available (see Appendix B). They were provided as input for MIMR-Raw.sps (i.e., the input correlation matrix was corrected for attenuation). The outcomes of this analysis are shown in Appendix D. The ranks of both the beta weights and the structure coefficients (in absolute values) are different from the ranks based on the uncorrected matrix, which indicates the differential effects of attenuation. As for the relative weights, correcting for unreliability also modifies the rank order of importance for the central predictors. Overall, correction seems to emphasize the differences between predictors (the best ones explain more and the worst ones explain less), which leads to a clearer interpretation.

Program Availability

Appendixes A and B show only a small portion of the SPSS code (namely, the code lines that can be modified by the user). The SPSS syntax, a short manual, and data files related to this article are available as supplemental materials from http://brm.psychonomic-journals.org/content/supplemental. Alternatively, these materials can be obtained free of charge by e-mail from urbano.lorenzo@urv.cat.

AUTHOR NOTE

The research reported in this article was partially supported by Grant 2009SGR1549 from the Catalan Ministry of Universities, Research and the Information Society, and by Grant PSI2008-00236/PSIC from the Spanish Ministry of Education and Science. Correspondence concerning this article should be sent to U. Lorenzo-Seva, Departament de Psicologia, Universitat Rovira i Virgili, Ctra. de Valls s/n 43007 Tarragona, Spain (e-mail: urbano.lorenzo@urv.cat).

REFERENCES

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.

Budescu, D. V., & Azen, R. (2004). Beyond global measures of relative importance: Some insights from dominance analysis. *Organizational Research Methods*, **7**, 341-350. doi:10.1177/1094428104267049

Cooley, W. W., & Lohnes, P. R. (1971). *Multivariate data analysis*. New York: Wiley.

Courville, T., & Thompson, B. (2001). Use of structure coefficients in published multiple regression articles: β is not enough. *Educational & Psychological Measurement*, **61**, 229-248.

Einspruch, E. L. (2003). *Next steps with SPSS*. London: Sage.

Johnson, J. W. (2000). A heuristic method for estimating the relative weight of predictor variables in multiple regression. *Multivariate Behavioral Research*, **35**, 1-19.

Johnson, J. W. (2004). Factors affecting relative weights: The influence of sampling and measurement error. *Organizational Research Methods*, **7**, 283-299. doi:10.1177/1094428104266018

Norušis, N. J. (1991). *The SPSS guide to data analysis for SPSS/PC+*. Chicago: SPSS.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

Pavot, W., & Diener, E. (1993). Review of the Satisfaction With Life Scale. *Psychological Assessment*, **5**, 164-172.

Pedhazur, E. J. (1982). *Multiple regression in behavioral research: Explanation and prediction*. New York: Holt, Rinehart & Winston.

Salovey, P., Mayer, J. D., Goldman, S. L., Turvey, C., & Palfai, T. P. (1995). Emotional attention, clarity, and repair: Exploring emotional intelligence using the Trait Meta-Mood Scale. In J. W.

Pennebaker (Ed.), *Emotion, disclosure, and health* (pp. 125-154). Washington, DC: American Psychological Association.
SCHEIER, M. F., CARVER, C. S., & BRIDGES, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): A reevaluation of the Life Orientation Test. *Journal of Personality & Social Psychology*, **67**, 1063-1078.

SUPPLEMENTAL MATERIALS

The SPSS syntax, a short manual, and data files related to this article are available from http://brm.psychonomic-journals.org/content/supplemental. These materials can also be obtained free of charge by e-mail (urbano.lorenzo@urv.cat).

APPENDIX A
Extract of MIMR-Corr.sps That Shows the Code Lines Configurable by the User

```
**************** ENTER HERE YOUR CORRELATION MATRIX (Y X1 X2 ...Xm).
****************************    First variable:         criterion.
****************************    The rest of variables:  predictors.
compute R={
1.000,     .564,    -.423,    -.166,     .440,     .420;
 .564,    1.000,    -.356,    -.081,     .303,     .527;
-.423,    -.356,    1.000,     .198,    -.241,    -.244;
-.166,    -.081,     .198,    1.000,     .149,     .014;
 .440,     .303,    -.241,     .149,    1.000,     .323;
 .420,     .527,    -.244,     .014,     .323,    1.000
}.
************************* ENTER HERE THE NAME OF YOUR DEPENDENT VARIABLE.
compute vdname=
'SWLS'
}.
*********************** ENTER HERE THE NAMES OF YOUR PREDICTOR VARIABLES.
compute pnames={
'OPTI';
'PESI';
'ATTEN';
'CLARI';
'REPAIR'
}.
**************** ENTER HERE THE RELIABILITIES OF YOUR PREDICTOR VARIABLES.
compute rb_p={
       .63;
       .60;
       .89;
       .86;
       .83
}.
***************** ENTER HERE IF YOU WANT TO HAVE A DETAILED OUTPUT.
********************       detailed = 0    PRODUCES A SIMPLE OUTPUT.
********************       detailed = 1    PRODUCES A DETAILED OUTPUT.
compute detailed =0.
```

APPENDIX B
Extract of MIMR-Raw.sps That Shows the Code Lines Configurable by the User

```
************************** ENTER HERE THE RELIABILITIES OF YOUR PREDICTOR
VARIABLES.
compute rb_p={
       .63;
       .60;
       .89;
       .86;
       .83
}.
*********************** ENTER HERE IF YOU WANT TO HAVE A DETAILED OUTPUT.
****************************    detailed = 0    PRODUCES A SIMPLE OUTPUT.
****************************    detailed = 1    PRODUCES A DETAILED OUTPUT.
compute detailed = 0.
```

APPENDIX B (Continued)

```
************** ENTER HERE IF YOU WANT TO COMPUTE CONFIDENCE INTERVALS.
******************** k = 0   CONFIDENCE INTERVALS ARE NOT COMPUTED.
******************** k = 500 CONFIDENCE INTERVALS ARE COMPUTED, AND
                             500 BOOTSTRAP SAMPLES ARE USED
                            (OTHER VALUES CAN BE USED: FOR EXAMPLE,
                             k = 1000).
compute k = 500.
```

APPENDIX C
Outcome When Measurement Error Is Not Corrected

```
Correlations Between Dependent Variable and Original Predictors
         SWLS    OPTI    PESI    ATTEN   CLARI  REPAIR
SWLS    1.000    .564   -.423   -.166    .440    .420
OPTI     .564   1.000   -.356   -.081    .303    .527
PESI    -.423   -.356   1.000    .198   -.241   -.244
ATTEN   -.166   -.081    .198   1.000    .149    .014
CLARI    .440    .303   -.241    .149   1.000    .323
REPAIR   .420    .527   -.244    .014    .323   1.000

**************************************************************************
Regression coefficients
**************************************************************************

Multiple R.
Point-Estimate and bootstrap 95% confidence interval (Lower and Upper)

 P. Estim   Lower   Upper
     .680    .629    .725

R square.
Point-Estimate and bootstrap 95% confidence interval (Lower and Upper)

 P. Estim   Lower   Upper
     .463    .396    .526

Regression coefficients (b).
Point-Estimate and bootstrap 95% confidence interval (Lower and Upper)

        SWLS   Lower   Upper
OPTI    .346    .271    .427
PESI   -.174   -.238   -.103
ATTEN  -.051   -.074   -.027
CLARI   .107    .077    .135
REPAIR  .037    .007    .064

Standardized regression coefficients (Beta).
Point-Estimate and bootstrap 95% confidence interval (Lower and Upper)

        SWLS   Lower   Upper
OPTI    .350    .273    .426
PESI   -.177   -.239   -.106
ATTEN  -.146   -.218   -.075
CLARI   .279    .200    .351
REPAIR  .104    .021    .183

Structure coefficients.
Point-Estimate and bootstrap 95% confidence interval (Lower and Upper)
```

APPENDIX C (Continued)

	SWLS	Lower	Upper
OPTI	.829	.760	.885
PESI	-.621	-.712	-.520
ATTEN	-.244	-.370	-.105
CLARI	.647	.543	.731
REPAIR	.617	.513	.704

****
Johnson's Relative Weights
****

Relative contribution to Multiple R (reported as percentages).
Point-Estimate and bootstrap 95% confidence interval

	SWLS	Lower	Upper
OPTI	36.9	28.3	45.5
PESI	18.0	11.6	25.1
ATTEN	5.3	1.7	10.5
CLARI	24.7	16.5	33.0
REPAIR	15.2	9.5	21.9

****

APPENDIX D
Outcome When Measurement Error Is Corrected

Correlations Between Dependent Variable and Original Predictors
corrected for attenuation

	SWLS	OPTI	PESI	ATTEN	CLARI	REPAIR
SWLS	1.000	.711	-.546	-.176	.474	.461
OPTI	.711	1.000	-.578	-.109	.412	.729
PESI	-.546	-.578	1.000	.271	-.335	-.346
ATTEN	-.176	-.109	.271	1.000	.171	.016
CLARI	.474	.412	-.335	.171	1.000	.382
REPAIR	.461	.729	-.346	.016	.382	1.000

****
Regression coefficients
****

Multiple R.
Point-Estimate and bootstrap 95% confidence interval (Lower and Upper)

P. Estim	Lower	Upper
.764	.707	.829

R square.
Point-Estimate and bootstrap 95% confidence interval (Lower and Upper)

P. Estim	Lower	Upper
.584	.499	.687

Regression coefficients (b).
Point-Estimate and bootstrap 95% confidence interval (Lower and Upper)

	SWLS	Lower	Upper
OPTI	.549	.429	.665
PESI	-.289	-.409	-.176
ATTEN	-.058	-.084	-.029
CLARI	.124	.092	.155
REPAIR	.044	.014	.081

APPENDIX D (Continued)

```
Standardized regression coefficients (Beta).
Point-Estimate and bootstrap 95% confidence interval (Lower and Upper)

        SWLS   Lower   Upper
OPTI     .619    .419    .966
PESI    -.115   -.255    .079
ATTEN   -.117   -.209   -.025
CLARI    .248    .133    .359
REPAIR  -.123   -.386    .010

Structure coefficients. Point-Estimate and bootstrap 95% confidence interval
(Lower and Upper)

        SWLS   Lower   Upper
OPTI     .930    .872    .961
PESI    -.714   -.819   -.568
ATTEN   -.230   -.346   -.098
CLARI    .620    .505    .708
REPAIR   .603    .502    .702

**************************************************************************
Johnson's Relative Weights
**************************************************************************

Relative contribution to Multiple R (reported as percentages). Point-Estimate
and bootstrap 95% confidence interval

        SWLS   Lower   Upper
OPTI    44.0    33.9    53.2
PESI    19.9    11.6    28.8
ATTEN    4.0     1.5     8.1
CLARI   19.1    12.0    27.2
REPAIR  12.9     9.4    18.1

**************************************************************************
```

(Manuscript received May 13, 2009;
revision accepted for publication July 21, 2009.)