MULTIPLE LINEAR REGRESSION VIEWPOINTS
VOLUME 18, NUMBER 1, FALL 1991

The Use of Regression Diagnostics to Improve Model Fit: A Case of Role Strain and Job Stress

Susan Tull Beyerlein and Michael Martin Beyerlein
University of North Texas

Abstract

This paper illustrates the importance of using regression diagnostics to improve model fit when using standard multiple regression statistical packages such as SASPC. This study examined the relationship between employee perceptions of their work environments and perceived job stress. The analysis was theory driven rather than exploratory in nature, and was performed using SASPC multiple regression procedures. Variables were coded to reduce possible collinearity. Various regression diagnostics were examined to detect the presence of outliers, influential observations, residual correlation, and collinearity (e.g., VIFs, DFFITS, the C_p criterion, HAT (leverage) values, and the Durbin-Watson test). These values, coupled with the various regression procedures yielded a final, best nine-variable model of R^2 = .48, significantly larger than the initial value of R^2 = .27. Future research in this area could be strengthened through 1) an examination of the path analytic and LISREL models in the literature that attempt to model indirect effects, 2) possible incorporation of select, higher-order terms from these studies, and 3) utilization of the regression diagnostic procedures outlined in this paper.

Introduction

Role conflict and role ambiguity are two stressors that have been linked to various health and physical outcomes. Role conflict involves conflicting task assignments initiated by superiors of equal rank and authority. Role ambiguity concerns the lack of clarity regarding job assignments, work objectives, and others' expectations. Kahn and others (1964) found that men who experience role conflict and role ambiguity on the job exhibit more tension and less job satisfaction than men whose roles are congruent or unambiguous. Research shows that role conflict correlates with a number of other outcomes including poor performance (Liddel & Slocum, 1976), poor peer relationships (French & Caplan, 1972), and turnover (Brief & Aldag, 1976; Hamner & Tosi, 1974). Role ambiguity has been linked to ineffective coping, as well as turnover.

Underutilization and job future ambiguity are two additional job stressors that have been shown to impact perceived job stress (Caplan, Cobb, French, Harrison, & Pinneau, 1980). Underutilization of abilities involves the lack of opportunity on the job to use skills and knowledge acquired in school or from previous experience and training. Job future ambiguity concerns levels of certainty regarding future career plans, opportunities for promotion, future value of current job skills, and future job responsibilities. These four variables, that is, role conflict, role ambiguity, underutilization, and job future ambiguity, plus years on the job and gender were chosen from a larger set of variables because of strong theoretical connections to stress, and after correlation analysis suggested they were the best set for predicting perceived job stress.

Method

The present study involved a survey of staff members at a large, southwestern university. Respondents were white collar workers in various clerical, secretarial and administrative positions. A total of 660, 14 page surveys were sent through the campus mail system, and 134 were returned for a response rate of 20.3 percent. Twenty-three cases were omitted because of missing data. The initial predictor variables used in the study were as follows: gender (D1), years on job (X1), role conflict (X2), role ambiguity (X3), underutilization (X4), and job future ambiguity (X5). The criterion variable was perceived job stress (Y). Gender (D1) was represented by dummy coding (i.e., 0 males, 1 females). Selected interaction terms were then created based on developed theory in the literature, that is, years on job times underutilization (X6), role conflict times role ambiguity (X7), and years on job times job future ambiguity (X8). Due to the fact that stress has often shown nonlinear relationships to other variables, several squared, higher order terms were included in the analysis, that is, role conflict (X2X2), role ambiguity (X3X3), underutilization (X4X4), and job future ambiguity (X5X5). Finally, all the predictor variables with the exception of gender (D1), were coded in order to reduce the likelihood of rounding errors in regression coefficients leading to collinearity (Mendenhall & Sincich, 1989, p. 343). Thus, to denote coded variables, "U" replaces "X" for all variables except D1 and the criterion variable Y.

Results

The analysis was performed using SASPC and involved a number of procedures. First, the initial set of predictors was included in the general regression procedure, PROC REG (i.e., D1, X1-X5). This analysis yielded an R^2 = .27. Next, this procedure was repeated with these variables and the additional interaction and higher order terms (i.e., D1, X1-X8, X2X2, X3X3, X4X4, X5X5). This yielded an R^2 = .34. The correlation procedure, PROC CORR, was also run at this point in order to obtain means and standard deviations for the predictor variables.

The twelve variables (excluding D1 and Y) were then coded and analyzed using the general regression procedure, PROC REG with the INFLUENCE option, (i.e., U1-U8, U2U2, U3U3, U4U4, U5U5). This analysis yielded an R^2 = .31. The subsequent inclusion of D1 (gender) raised the R^2 value to .34. The DFFITS values were then examined in order to identify possible influential observations. The SAS User's Guide: Statistics (1985) describes the DFFITS statistic as "a scaled measure of the change in the predicted value of the ith observation (which is) calculated by deleting the ith observation" (p. 677). The difference, $y_i - y_{(i)}$, has been divided by its standard error so that the differences can be more easily compared. The investigator is interested in values that are considerably larger relative to the other differences in predicted values. For most purposes, a value of 1.0 is considered to be sufficiently large to warrant attention. In the present study, influence diagnostics revealed five DFFITS values greater than 1.0. These were subsequently deleted from the analysis leaving a remaining sample of n=108.

The regression procedure, PROC STEPWISE, was then utilized, specifically, the FORWARD, BACKWARD, and MAXR options. The PROC STEPWISE procedure is a good choice when there are a number of independent variables to consider. The various options do not always isolate the model with the highest R^2 but rather seek the best one-variable model, two-variable model, and so forth (SAS User's Guide: Statistics, 1985). The FORWARD option requests the forward selection technique, BACKWARD requests the backward elimination technique, and MAXR requests the maximum R^2 improvement technique. MAXR looks at all possible regression equations, however, as with the other options it outputs only the best models, for example, the best ten-variable, nine-variable, eight-variable models, and so forth.

After examining the output from the PROC STEPWISE analyses it was decided that the following was the best model: R^2 = .38, D1, U1, U2, U3, U4, U8, U3U3, U4U4, C_p = 7.87, with all variables significant at the 0.10 level. The C_p criterion is gleaned from the FORWARD and BACKWARD procedures (rather than MAXR) and is used to select the best subset model with a small total mean square error (C_p), and a value of C_p near p + 1, which indicates that slight or no bias exists [$E(C_p) \approx p + 1$]. In this case, the C_p value was slightly less than the number of parameters in the model (i.e., eight).

This model was then analyzed using the general regression procedure, PROC REG, with the VIF, P, R, DW, and INFLUENCE options. VIF prints variance inflation factors with the parameter estimates; variance inflation is the reciprocal of tolerance; P calculates predicted values from the estimated model and input data; R analyzes the residual and includes the Cook's D statistic which is an overall measure of influence for each observation, the standard errors of the predicted and residual values, and the studentized residual; DW calculates the Durbin-Watson statistic; INFLUENCE prints the following diagnostics used in the present study for each observation: the residual, studentized residual, HAT or leverage value (h_i), and the DFFITS statistic.

Examination of the plot RESID*PRED (residuals times predicted scores) revealed a value greater than +2 standard deviations, that is, a possible outlier. This value was subsequently deleted leaving a sample of n=105. A rerun of the general regression procedure, PROC REG, using the above best model yielded an R^2 = .41, and a Durbin-Watson, D = 2.21, suggesting the residuals were slightly negatively correlated (Mendenhall & Sincich, 1989, p. 307). However, calculation of the average leverage value, $\bar{h} = (k + 1)/n = .17$, and examination of the HAT values revealed four values greater than twice the average value, suggesting that these values were influential observations and should be eliminated from the data set. They were subsequently deleted leaving a final sample of n=101.

Examination of PROC STEPWISE options, that is, FORWARD, BACKWARD, and MAXR revealed significant gains in R^2 values. At this point, a nine-variable model was chosen as the best model for several reasons: 1) the C_p value was only slightly less than the number of predictors (Younger, 1985), whereas it was significantly larger for other models with similar R^2 magnitude; 2) all variables were significant at the 0.10 level; 3) there was a significant drop in R^2 using the BACKWARD procedure as one dropped to the eight variable models; and 4) the Durbin-Watson statistic was close to a value of two, suggesting minor residual correlation. Therefore, the best model chosen was as follows: R^2 = .48, D1, U1, U2, U3, U4, U7, U8, U3U3, U4U4, C_p = 8.85, DW = 2.21. Thus, the final model included the following variables: gender (D1), years on job (U1), role conflict (U2), role ambiguity (U3), underutilization (U4), role conflict times role ambiguity (U7), years on job times job future ambiguity (U8), and the squared, higher-order terms utilizing role ambiguity (U3U3) and underutilization (U4U4). The only conflicting evidence was the value of the variance inflation factors (VIFs) for U3, U4, U3U3, and U4U4. These values were greater than 10, whereas the VIFs for all other variables in the model were approximately 10 or less. VIFs greater than 10 indicate the presence of collinearity where, (VIF) = $1/(1-R^2_i)$, i = 1, 2, ..., k (Mendenhall & Sincich, 1989, p. 237). Values greater than 10 occurred only in those variables used both singularly and squared in the higher-order terms, making them obvious candidates for collinearity. In addition, Mendenhall and Sincich (1989) discuss the need to code the dependent, as well as the independent variables, in order to properly calculate VIFs (p. 236). The criterion variable, perceived job stress (Y), was not coded in this study. Finally, the R^2 = .48 representing the best model did not appear to be sufficiently large to indicate the presence of collinearity. Consistent with this finding, the standard errors of the individual beta parameters were not inflated, and the t-tests on the individual beta parameters were significant suggesting lack of evidence for collinearity (Mendenhall & Sincich, 1989, p. 236).

Discussion and Recommendations

Of several hundred studies of stress examined by the authors, it appears that none have used the regression diagnostics discussed in this paper, suggesting that the results of models presented in the literature may be weaker than necessary. Results of the present study illustrate that the use of the various regression diagnostics can improve best model fit considerably. In addition, it should be obvious that investigators cannot depend solely on regression selection options such as MAXR, FORWARD and BACKWARD when searching for the best subset regression model. Options such as MAXR will provide R^2 values for all generated models, however, the final decision as to which is the best model cannot be made without the C_p statistic and other values, for example, regression diagnostics such as HAT values, Cook's D, results of jackknifing procedures such as deleted residuals, DFFITS, DFBETAS, and the Durbin-Watson statistic which are available under the FORWARD and BACKWARD options of the PROC STEPWISE procedure.

The FORWARD and BACKWARD options offer different best models. That is, they each output best models based on the particular programmed criteria embedded in their respective routines, with the R^2 as the salient criterion. However, a strong R^2 value is not inequivocably the last word on model fit. For example, if two models with similar R^2 values are examined, it may be that the model with the slightly lower R^2 will better satisfy the other criteria discussed above and will thus be the better choice overall. Therefore, the investigator needs to utilize the power of these routines coupled with intelligent decision making regarding the various procedures. Coding variables reduces the likelihood of collinearity, and outputting regression diagnostics enables the investigator to experiment with dropping outliers and influential observations to see how their absence affects the variance accounted for by the overall model. In summary, there is nothing automatic about the process. SASPC and other packages will provide the mathematics, but it remains the responsibility of the investigator to examine the output carefully to arrive at truly the best model.

The analysis of the present study was theory-driven rather than exploratory in nature. In other words, because of the authors' preference for confirmatory modelling techniques, a limited number of interaction and higher-order terms were chosen based on the literature. However, the literature is replete with more complex models, that is, path analyses and LISREL models that attempt to model indirect effects. Thus, future analyses could be improved by studying the literature in more depth to arrive at other plausible variables and higher-order terms. Possible variables to be included in additional studies involve two general categories, that is, 1) job design facets such as autonomy, responsibility, feedback, task significance, task wholeness, leadership style; and 2) moderator variables consisting of personality characteristics and other demographics, such as Type-A, locus of control, and growth need strength. In addition, existing studies could be strengthened through replication and utilization of the regression diagnostics detailed in the present study.

References

Brief, A. P., & Aldag, R. J. (1976). Correlates of role indices. Journal of Applied Psychology, 61, 468-472.

Caplan, R. D., Cobb, S., French, J. R. P., Harrison, R. V., & Pinneau, S. R., Jr. (1980). Job demands and worker health. Ann Arbor, MI: Institute for Social Research.

French, J. R. P., & Caplan, R. D. (1972). Organizational stress and individual stress. In A. J. Marrow (Ed.), Failure of success. New York: AMACOM.

Hamner, W. C., & Tosi, H. L. (1974). Relationship of role conflict and role ambiguity to job involvement measures. Journal of Applied Psychology, 39, 497-499.

Kahn, R. L., Wolfe, D. M., Quinn, R. P., Snoek, J. D., & Rosenthal, R. A. (1964). Occupational stress: Studies in role conflict and ambiguity. New York: Wiley.

Liddell, W. W., & Slocum, J. W. (1976). The effects of individual-role compatibility upon group performance: An extension of Schutz's FIRO theory. Academy of Management Journal, 19, 413-426.

Matteson, M. T., & Ivancevich, J. M. (1987). Controlling work stress: Effective human resource and management strategies. San Francisco: Jossey-Bass.

Mendenhall, W., & Sincich, T. (1989). A second course in business statistics: Regression analysis (3rd ed.). San Francisco: Dellen Publishing Company.

SAS Institute, Inc. (1985). SAS User's Guide: Statistics (5th ed.). Cary, NC: Author.

Younger, M. S. (1985). A first course in linear regression (2nd ed.). Boston: Duxbury Press.