Are We Drawing the Correct Conclusions? Regression Analysis in the Nonprofit Literature

Robert Shearer & Truman Clark

Pepperdine University

ABSTRACT

Linear models are the most commonly used analytical tools in the nonprofit literature. Academics and practitioners utilize these models to test different hypotheses in support of their research efforts, seeking to find significant results that substantiate their theories. And yet the authors of this article have discovered a surprisingly large number of insignificant results in articles from established nonprofit journals. Insignificant hypotheses and Type II errors surely account for a number of these results, but the authors believe the majority of these results are due to a different cause, one that is detectable and preventable: multicollinearity.

RÉSUMÉ

Dans les articles sur les organismes sans but lucratif, les modèles linéaires sont les outils analytiques les plus communément utilisés. En effet, académiques et praticiens utilisent tous les deux ces modèles pour évaluer diverses hypothèses relatives à leurs recherches, espérant trouver des résultats significatifs pouvant confirmer leurs théories. Pourtant, les auteurs de cet article ont découvert un nombre surprenant de résultats non significatifs dans des articles de revues établies sur les organismes sans but lucratif. Des hypothèses non significatives et des erreurs du type II expliquent sûrement certains de ces résultats, mais les auteurs croient que la majorité des résultats ont une cause différente qui est détectable et évitable : la multicolinéarité.
.

KEYWORDS / MOTS CLÉS : Regression analysis; Multicollinearity; Statistical significance / Analyse de régression; Multicolinéarité; Signification statistique

ANSERJ

To be notified about new ANSERJ articles, **subscribe here.** / Afin d'être avisé des nouveaux articles dans ANSERJ, **s'inscrire ici .**

22

INTRODUCTION

Research into the nonprofit sector seeks to explain and provide insight into the complex dynamics faced by nonprofit organizations. Given the wide range of research topics in the field, it is not surprising that researchers and academics utilize a varied collection of analytical methods to aid them. The literature contains a wide range of both qualitative and quantitative methods, but of the papers that include quantitative analysis, one method is dominant: the regression model. Unfortunately, it appears that regression is widely misapplied in the literature, leading to insignificant results at an alarmingly high rate. This article reviews the use of regression in the nonprofit literature and explains why many of these insignificant results are likely due to multicollinearity (highly correlated predictor variables), an identifiable and correctable problem. We also review multicollinearity diagnostics and solutions in an attempt to show how these problems can be avoided in the future.

CURRENT STATE OF REGRESSION IN THE LITERATURE

A survey of the articles published during 2013 in two established nonprofit journals, *Nonprofit Management & Leadership* (NML) and *Nonprofit Voluntary Sector Quarterly* (NVSQ), finds that 29/79 (or 37%) of the articles included regression analysis. The purpose of regression in the nonprofit literature is almost exclusively to conduct multiple, simultaneous hypothesis tests. This use of regression, sometimes referred to as the full model approach, is common throughout the social sciences. Researchers using this approach propose a new theory with supporting hypotheses, collect data, fit a regression model testing the hypotheses, and then draw statistical inference from the regression's *t*-tests. Significant tests provide statistical support for the researcher's theory.

An example of this full model approach can be found in the work by Taedong Lee, Erica Johnson, and Aseem Prakash (2012), in which the authors studied the relationship between media independence and trust in non-governmental organizations (NGOs) among post-communist countries. They modelled public trust in NGOs as a function of thirteen predictor (or independent) variables: media independence, democracy, gross national income (GNI) per capita, gross domestic product (GDP) growth rate, ethnic freedom, aid, foreign direct investment (FDI) inflows, trade, the internet, European Union (EU) membership, Commonwealth of Independent States (CIS) membership, neighbourhood effect, and international non-governmental organization (INGO) presence. They obtained their data from the United States Agency for International Development (USAID). The response (dependent) variable, public trust in NGOs, is an ordinal measure with a range from 1 (low trust) to 7 (high trust).

The authors, based upon previous literature and logical reasoning, expected the predictor variables to be significant, that is, each predictor alone had a relationship with public trust in NGOs. Several quotes from the article justify their expectations in Table 1.

Table 1: Justification of predictor variables

Predictor	Justification
Democracy	"NGOs need political space to function and democracies provide such political opportunities" (p. 21).
The internet	"… the internet is an especially important source of political, social, and economic information. The internet also allows NGOs to disseminate information on their activities … " (p. 21).
GNI per capita	"… wealth is critical for NGO sector development … " (p. 22).
Aid	"The NGO sector in post-communist Eurasian countries is highly dependent on foreign donors" (p. 22).

The authors were surprised to find the vast majority (11/13) of their predictors insignificant. Table 2 shows the output from their main regression model.
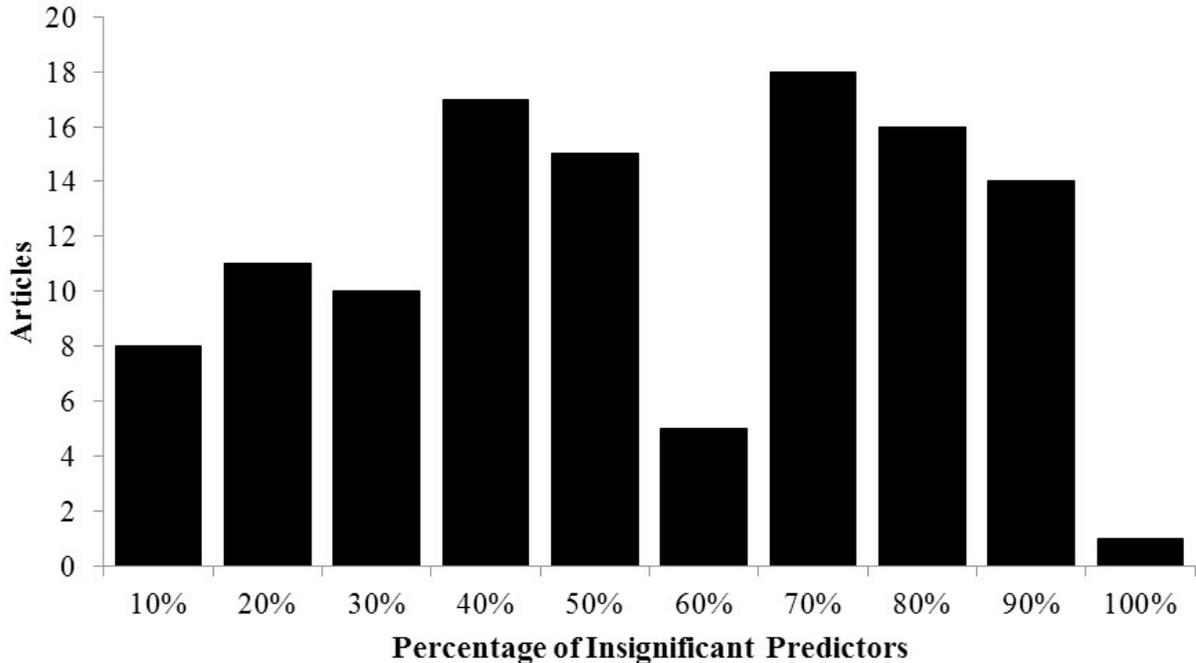
Table 2: Regression model output

Variable	Coefficient	Standard Error	*t*-Statistic
Media independence	0.337	0.076	4.32*
Democracy	0.013	0.018	0.72
GNI per capita	-0.181	0.132	-1.37
GDP growth rate	0.002	0.008	0.25
Ethnic fraction	0.473	0.416	1.14
Aid (% GNI)	-0.001	0.004	-0.25
FDI inflows	0.0002	0.009	0.02
Trade	0.002	0.002	1.00
Internet	-0.005	0.006	-0.83
EU membership	-0.078	0.200	-0.39
CIS membership	-0.123	0.181	-0.68
Neighborhood effect	0.092	0.100	0.92
INGO	0.432	0.130	3.32*

Note: Significant predictors ($\alpha = 0.01$) are identified with a *. Only two predictors (media independence and INGO) out of 13 were significant.

Several different explanations could account for the insignificant results. First, the prior literature was incorrect and a majority of the predictors have no relationship with public trust in NGOs. Second, random variation produced non-typical results. Third, and most likely, multicollinearity inflated the standard errors of many of the estimated coefficients in the model, resulting in insignificant p-values. Although we do not possess the data, it would not be surprising to find that democracy and internet availability are highly, positively correlated and that GDP growth rate and aid received are highly, negatively correlated.

Lee, Johnson, and Prakash (2012) are not alone in finding such unexpected results. The 37,2013 articles in NML and NVSQ that utilized regression contained 115 regression models. Only four of the 115, or less than 4 percent, contained no insignificant predictors. Several of the authors of the other 2013 articles expressed surprise at their results. Not only were relationships found to be insignificant when the literature would suggest otherwise, but often the sign of the coefficients were the opposite of what was expected. Now while insignificant results are to be expected, if only due to chance variation, the proportion of insignificant predictors across these articles was very surprising. The average article found nearly one half of its predictors insignificant. Figure 1 shows the distribution of insignificant predictors in the 2013 NVSQ and NML articles.

ANSERJ

To be notified about new ANSERJ articles, **subscribe here.** / Afin d'être avisé des nouveaux articles dans ANSERJ, **s'inscrire ici .**

24

## Figure 1: Distribution of Insignificant Predictors

Note: These two journals published 79 articles in 2013, 29 of which included linear models. The average article had 48 percent insignificant independent variables.

Multicollinearity

How then to explain the large percentage of insignificant results in such a large number of theoretically grounded analyses? We suggest that multicollinearity is the underlying cause. Multicollinearity arises when subsets of the independent variables are collinear, i.e., they are linear functions of one another. This creates serious problems in a regression model, as multicollinearity inflates the coefficient standard errors in the fitted regression model. Since the test statistic for significance in a regression model for predictor *i* is obtained from $t = \frac{\hat{\beta}_i}{S.E._{\hat{\beta}_i}}$, inflated standard errors deflate the *t*-statistic. Deflated *t*-statistics produce increased *p*-values, yielding insignificant predictors, even when the predictor alone is significant.

Consider the two linear regression models in Figure 2. Model A fits the response variable y as a function of x_1 and x_2. The correlation between x_1 and x_2 is close to zero, as both small and large values of x_1 occur with both small and large values of x_2. Model B fits the response variable y as a function of x_3 and x_4. The correlation between x_3 and x_4 is close to one, so that small values of x_3 only occur with small values of x_4 and large values of x_3 only occur with large values of x_4. The problems with multicollinearity emerge as one considers how the fitted models would change with different samples. Each new sample, with fixed (x_i,x_j), yields new values of y, resulting in a change in the orientation or tilt of the fitted model. Model A would change little, as the wide dispersion of (x_1,x_2) produces a flat, stable surface (similar to a table). This model will have small coefficient standard errors, which estimate the variability in the slopes of the model. Model B would change a lot, as the narrow dispersion of (x_3,x_4) produces a narrow, unstable surface (similar to a ridgeline). This model will have large coefficient standard errors.

Figure 2: Fitted regression models A and B
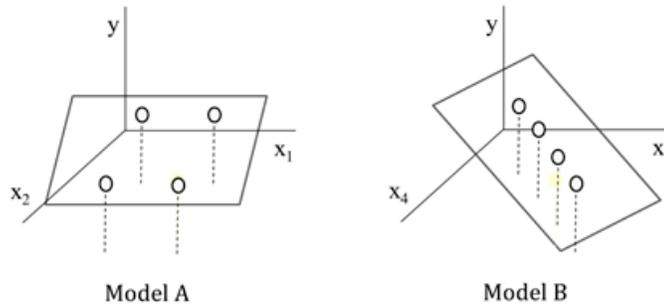
Model A   Model B

Note: Model A describes the response variable y as a linear function of uncorrelated explanatory variables x_1 and x_2. Model B describes the response variable y as a linear function of positively correlated explanatory variables x_3 and x_4. Model A produces a fitted model with small coefficient standard errors, while model B produced a fitted model with large coefficient standard errors.

Model A, without inflated standard errors, will correctly identify insignificant predictors. Model B, with inflated standard errors, will identify significant predictors as insignificant due only to multicollinearity. Consider a third linear regression model with three predictor variables: x_1, x_2, and x_3. A sample of size *n* = 5 yields the following observations in Table 3.

Table 3: Sample data

x_1	1	2	3	4	5
x_2	1	10	3	4	5
x_3	2	6	4	2	8

Note that x_1 and x_2 only differ in the second observation. If $x_{2,2} = 2$, then the two predictors would be linearly dependent with a sample correlation coefficient of one. As $x_{2,2}$ decreases from 10 to 2, multicollinearity becomes more of an issue. Figure 3 shows the effect of multicollinearity on the coefficient standard error of x_2.

Figure 3: Impact of multicollinearity on coefficient standard error
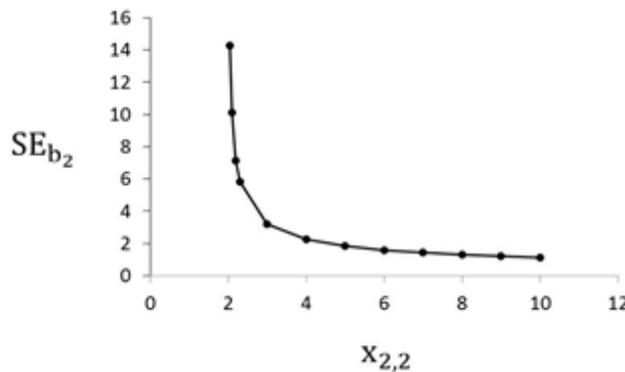
Note: The graph shows the increase in the coefficient standard error of x_2 as the value of $x_{2,2}$ decreases. As $x_{2,2}$ approaches two, the predictors x_1 and x_2 become more linearly dependent, introducing multicollinearity into the model.

As $x_{2,2}$ approaches 2, the standard error of its coefficient increases, increasing the *p*-value for x_2. At some point this inflated standard error will cause the *p*-value to exceed alpha (predetermined statistical significance), identifying x_2 as insignificant. Donald Farrar and Robert Glauber (1967) noted that, "the increase in sample standard errors for multicollinear regression coefficients virtually assures a tendency for relevant variables to be discarded incorrectly from regression equations" (p. 94). This is exactly what we fear is occurring in the nonprofit literature. Researchers fail to obtain significant findings not because their premises are wrong or the data was somehow flawed but rather because of multicollinearity.

SOLUTIONS

There are many steps that researchers can take to avoid the problems of multicollinearity in their regression models. First, limit the number of hypotheses tested. The danger of multicollinearity increases with the number of predictor variables. The understandable desire to test as many hypotheses as possible must be offset by the knowledge that the danger of insignificant results increases with each additional predictor. Second, include only predictor variables in the model that are uncorrelated. Several multicollinearity diagnostic tools are available to identify problems. The simplest approach to identify multicollinearity is through a correlation matrix. Most statisticians view correlations greater than 0.9 as problematic and would recommend removing one of the predictors from the model. Unfortunately, correlation matrices will not always identify multicollinearity, as it can also exist among sets of predictor variables. The most common approach to identify multicollinearity among subsets of predictors is through the use of variance inflation factors (VIF). These measure the increase in the variability of the regression coefficients due to the presence of multicollinearity in a data set. A VIF is obtained for each predictor variable by regressing that predictor against the other predictor variables. The R^2 from the resulting regression model then measures the proportion of the variability in the predictor that is explained by the other predictors, with R^2 values in excess of 0.80 identifying multicollinearity. Software is typically used to calculate the VIF, where $VIF = \frac{1}{1-R^2}$.

Predictor variables with VIF values in excess of five (R^2 values greater than 0.80) should not be included the final regression model. Third, include only control variables that are not correlated with the explanatory variables. Multicollinearity between control variables is not a problem, as hypotheses are not based on these variables. Fourth, utilize pre-fit variable reduction techniques in inferential models. Combine nonprofit theory, subject matter expertise, and multicollinearity diagnostics (correlation matrices and VIF calculations) to identify appropriate explanatory variables. Fifth, utilize stepwise techniques to remove multicollinearity in predictive models and cross validate the model to remove bias. Parsimonious models tend to yield the best predictions, leading to the use of model selection procedures (e.g., backward elimination, forward selection, or stepwise) to identify a "best" subset of predictors for inclusion in the model. These selection procedures identify the most significant explanatory variables and remove multicollinearity (both necessary for predictions), but also come with well-documented problems: biased parameter estimates (Bradbury, Freckleton, Stephens, & Whittingham, 2006), sample dependent model selection (Bradbury et al., 2006), inflated R^2 (Austin & Tu, 2004, Bradbury et al., 2006), and biased mean square error (Hurvich & Tsai, 1990). Two model validation procedures allow for the measure of the extent of these problems: 1) testing the consistency of the model by refitting the selected model form with the training data and 2) measuring the accuracy of the predictive model by using it to forecast the test data response values. As John Tukey (1980) noted, "Oftentimes, confirmation requires a new unexplored set of data" (p. 821).

CONCLUSION

The nonprofit academic community provides a valuable resource for the greater nonprofit community: research into the complex environment in which nonprofits operate. Unfortunately, much of the quantitative analysis conducted by using regression models yields insignificant results that provide our community little by way of insight. Random chance should lead to a proportion of insignificant predictors slightly above the significance level of the test, yet regression models in the literature find nearly half the predictors insignificant. Moreover, these insignificant results are not due to a poor theoretical understanding of the problems faced by the nonprofit community, nor by faulty data collection methods. Rather, the problem appears to be multicollinearity. Fortunately, methods exist to both identify and remove the sources of multicollinearity in our regression models. Application of these methods will allow us to obtain more significant results, and, as a consequence, allow us to better inform the nonprofit community.

REFERENCES

Austin, P., & Tu, J. (2004). Automated variable selection methods for logistic regression produced unstable models for predicting acute myocardial infarction mortality. *Journal of Clinical Epidemiology, 57*, 1138–1146.

Bradbury, S., Freckleton, R., Stephens, P., & Whittingham, M. (2006). Why do we still use stepwise modelling in ecology and behavior? *Journal of Animal Ecology, 75*(5), 1182–1189.

Farrar, D., & Glauber, R. (1967). Multicollinearity in regression analysis: The problem revisited. *The Review of Economics and Statistics, 49*(1), 92–107.

Hurvich, C., & Tsai, C. (1990). The impact of model selection on inference in linear regression. *The American Statistician, 44*(3), 214–217.

Lee, T., Johnson, E., & Prakash, A. (2012). Media independence and trust in NGOs: The case of postcommunist countries. *Nonprofit and Voluntary Sector Quarterly, 41*(8), 8–35.

Tukey, J. (1980). Methodological comments focused on opportunities. In Lyle V. Jones (Ed.), *The collected works of John W. Tukey* (819-867). Murray Hill, NJ: Bell Laboratories.

ABOUT THE AUTHORS / L'AUTEURS

Robert Shearer is Assistant Professor of Decision Science at Pepperdine University. Email: robert.shearer@pepperdine.edu

Truman Clark is a Business Administration Major at Pepperdine University. Email: truman.clark@pepperdine.edu