Fall 2020

A Monte Carlo Analysis of Ordinary Least Squares Versus Equal Weights

James Brewer Ayres
Western Kentucky University, James.Ayres335@topper.wku.edu

A MONTE CARLO ANALYSIS OF ORDINARY LEAST SQUARES VERSUS EQUAL WEIGHTS

A Thesis
Presented to
The Faculty of the Department of Psychological Sciences
Western Kentucky University
Bowling Green, Kentucky

In Partial Fulfillment
Of the Requirements for the Degree
Master of Science

By
J. Brewer Ayres

December 2020

A MONTE CARLO ANALYSIS OF ORDINARY LEAST SQUARES VERSUS EQUAL WEIGHTS

Date Recommended November 6, 2020

Reagan Brown Digitally signed by Reagan Brown
Date: 2020.11.16 17:39:51 -06'00'

Dr. Reagan Brown, Director of Thesis

Elizabeth Shoenfelt Digitally signed by Elizabeth Shoenfelt
Date: 2020.11.16 17:48:29 -06'00'

Dr. Elizabeth L. Shoenfelt

Katrina A. Burch Digitally signed by Katrina A. Burch
Date: 2020.11.17 09:13:28 -06'00'

Dr. Katrina A. Burch

Associate Provost for Research and Graduate Education

I dedicate this thesis to my parents, who have supported me no matter how difficult my challenges were.

I also want to thank my cohort—Ben, Brandi, Faith, Eli, and San—who became family during my short stay at WKU. I wish you all nothing but success in your future careers!

ACKNOWLEDGEMENTS

I would like to extend a special thanks to my Thesis Director, Dr. Reagan Brown, for all of his support and direction throughout my thesis writing process. His guidance made it possible for me to condense the capstone for a two-year program into a short year and a half. I would not have successfully defended without him!

TABLE OF CONTENTS

LIST OF TABLES

A MONTE CARLO ANALYSIS OF ORDINARY LEAST SQUARES VERSUS EQUAL WEIGHTS

J. Brewer Ayres December 2020 27 Pages

Directed by: Dr. Reagan Brown, Dr. Elizabeth Shoenfelt, and Dr. Katrina Burch

Department of Psychological Sciences Western Kentucky University

Equal weights are an alternative weighting procedure to the optimal weights offered by ordinary least squares regression analysis. Also called units weights, equal weights are formed by standardizing scores on the predictor variables and averaging these standardized scores to create a composite score. Research is limited regarding the conditions under which equal weights result in cross-validated R^2 values that meet or exceed optimal weights. In this study, I explored the effect of various predictor-criterion correlations, predictor intercorrelations, and sample sizes to determine the relative performance of equal and optimal weighting schemes upon cross-validation. Results indicated that optimally weighted predictors explained more criterion variance upon cross-validation as the variability in predictor-criterion correlations increased. Similarly, it appears that as predictor intercorrelations and sample size increase, optimally weighted predictors cross-validate to explain more criterion variance than equally weighted predictors. Implications and directions for future research are discussed.

Introduction

Regression analysis is a vital tool for research in the applied practice of industrial/organizational (I/O) psychology. As the cornerstone of predictive statistics, regression equations are equally applicable in studies ranging from predicting family wellbeing in hospitals (e.g., McAndrew et al., 2019) to the validation of statistical procedures and equations (e.g., Raju et al., 1999). Unfortunately, however, unavoidable sample differences cause a decrease in explained variance when a regression equation, developed with one sample, is applied to future samples (Pedhazur, 1982). To mitigate this reduction in explanatory power, researchers and practitioners should use the predictor weighting scheme that provides the best results in future groups of participants, not the original sample. This recommendation is especially important given that researchers and practitioners make their conclusions and recommendations for the benefit of *future* research and organizational processes. My study will highlight the importance of considering predictor-criterion correlations, predictor intercorrelations, and sample sizes to determine what type of regression analysis will produce the best result in subsequent samples.

Literature Review

Ordinary least squares (optimal weights) and equal weights are two foundational regression weighting techniques in I/O psychology. However, there is a dearth of research to indicate which technique will perform best in samples upon cross-validation. Advancements in technology make optimal weight calculations easy, but the field should not choose a statistical procedure simply because it is easier to conduct. Rather, researchers and practitioners should always implement the procedure that best fits their purposes. Failure to do so could (among other outcomes) produce invalid applications of the results of a study (e.g., selecting unqualified job applicants).

Regression Analyses

Regression models are a powerful means of forecasting outcomes prior to selecting a course of action. This statistical procedure has an extensive history; sources such as Dawes and Corrigan (1974) recounted Benjamin Franklin's use of regression by weighing pros and cons of various ideas, and then using the sum of these considerations to make the best choice. In Franklin's case, such regression analyses are considered normative, meaning they inform the best decision in a given situation. Dawes and Corrigan (1974) also highlighted that regression may be used as a descriptive tool, which allows researchers to represent an individual's behavior or standing on a construct (e.g., degree of emotional exhaustion; Bekker et al., 2005).

Regardless of the application, regression analysis functions the same way on a basic level. Whether using regression analysis in a normative or descriptive application, one needs a meaningful composite of the variables that affect an outcome. This composite, called predicted Y (i.e., Y'), is calculated with the following equation.

$$Y' = bX + a$$

Y' represents the criterion variable, which is the result of the predictor, X, the correlation coefficient, b, and the equation constant, a (Pedhazur, 1982). Researchers can then correlate Y' with actual scores on Y ($r_{YY'}$) to assess the relationship between predicted and actual scores (Pedhazur, 1982).

Regression equations may be expanded to account for as many variables as a researcher or practitioner desires. These larger regression equations can increase the ability to predict or describe complex behaviors, such as job performance (Guion, 1998). A multivariate regression composite is calculated with the following equation.

$$Y' = b_1X_1 + b_2X_2 \ldots + b_kX_k + a$$

Many of the variables in this equation are the same, but this model provides a composite score based on different partial correlation coefficients (b_k) for each predictor (X_k; Pedhazur, 1982). Furthermore, correlating Y' with Y results in a multivariate correlation, R, which transforms into R^2 when squared (Guion, 1998). These results are theoretically similar to the bivariate regression; the new notation simply denotes a multivariate analysis. In addition to providing the same benefits as a bivariate regression analysis, multivariate regression allows researchers to use a variety of weighting techniques to achieve the best prediction for their samples. Optimal weights and equal weights are two weighting techniques (Dawes & Corrigan, 1974), but the explanatory power of each technique can change drastically upon cross-validation.

Cross-Validation and Shrinkage

Guion (1998) highlighted the necessity of cross-validation in multiple regression analyses. The need for cross-validation is predicated on the fact that one sample of data may elicit a large, significant R^2, whereas the same prediction equation applied to data from a new sample results in a lower, possibly insignificant result. Sampling error is the cause of this reduction in predictive accuracy. Sampling error results in regression weights that are specific to the sample which they are derived from, but do not generalize to other samples from that same population.

Sampling error arises because aside from limited situations, researchers do not measure an entire population. Therefore, the distributional characteristics of a sample will deviate from the distributional characteristics of its parent population and subsequent samples. Wainer (1976) cited outliers as an example of sample-specific data, which could present issues to a researcher upon cross-validation.

To determine how well the results from a regression analysis generalize to different samples, the regression equation must be applied to a new, independent sample (Guion, 1998). The Society for Industrial and Organizational Psychology's (SIOP) *Principles for the Validation and Use of Personnel Selection Procedures* states that "Testing professionals [use] as unbiased an estimate as possible of the operational validity of the predictor in the population in which it is used" (American Psychological Association; APA, 2018, p. 14). This cross-validation process proceeds as follows. Scores from the new sample are inserted into the prediction equation that was derived from the original regression analysis, which results in a predicted criterion score that is the composite of the predictor scores. These composite scores are then correlated with the actual criterion scores in that sample. The resultant correlation (once squared) is the cross-validated R^2. The uncontrollable differences that are due to sampling error will result in a reduced (i.e., shrunken) R^2 (Guion, 1998).

To ensure that regression analyses do not suffer from a significant degree of shrinkage, researchers and practitioners have a few options. One method of reducing the degree of regression overfitting is by maximizing the ratio of study participants to predictor variables (Pedhazur, 1982). A simple way of operationalizing this statistical effect is by increasing the sample size. Sampling error inversely relates to sample size, as larger samples more accurately represent their population (Trochim & Donnelly, 2008). Aside from this ratio, predictors are more likely to work well if they are supported by a sound hypothesis (e.g., using a theoretical model to suggest certain predictors will perform well; Guion, 1998). These principles should work with both optimally and equally weighted regression, but it is important to understand the mathematic foundation of each option to fully understand how they might affect a particular application.

Ordinary Least Squares

Optimally weighted regression equations maximize the variance explained for a particular dataset. In other words, the regression coefficients are chosen to achieve the most accurate prediction *for that sample* (Pedhazur, 1982). This regression technique achieves such accuracy by assigning stronger weights to predictors that have stronger relationships with the criterion (Guion, 1998). These weights, or partial regression coefficients, are calculated with the following equation.

$$b_k = \frac{r_{YX_k} - r_{YX_{k+1}} \cdot r_{X_k X_{k+1}}}{1 - r^2_{X_k X_{k+1}}} \cdot \frac{S_Y}{S_{X_k}}$$

The above equation computes an unstandardized partial regression coefficient; standardized partial regression coefficients are obtained by deleting the standard deviations. An inspection of the equation reveals the following. First, a stronger correlation between the criterion and predictor (r_{YX_k}) will result in a stronger regression coefficient. Second, the correlations of other predictors with the criterion (r_{YX_k}) and among themselves ($r_{X_k X_{k+1}}$) will decrease the predictive power of the resulting partial regression coefficient. Moreover, unstandardized regression coefficients can be dramatically affected by their standard deviations (S_Y and S_{X_k}). The partial regression coefficient will increase as the standard deviation of the criterion increases, and the opposite is true as the standard deviation of the predictor increases.

As with partial regression coefficients, the total variance explained in the criterion variable is a function of the individual predictor-criterion correlations and the predictor intercorrelations. Pedhazur (1982) stated that uncorrelated predictors explain criterion variance equal to the sum of the explanatory power for each predictor (i.e., $R^2 = r_{X1}^2 + r_{X2}^2 + \cdots r_{Xk}^2$). For instance, if X_1 and X_2 are perfectly uncorrelated, and the explanatory power of these predictors are .25 and .30, respectively, then the total variance explained in the regression equation would be .55. However, explanatory power with intercorrelating predictors is not this simple. Intercorrelating predictors provide superfluous information by providing similar information on the criterion (Pedhazur, 1982). As further evidence to this point, it becomes impossible to use regression analyses with extreme predictor intercorrelations (e.g., $r_{X_k X_{k+1}} = 1.00$).

Equal Weights

Relative to optimal weights, equal weights are simpler to calculate and implement in a study. Raju et al. (1997) discussed two procedures for calculating equal weights. The first involves dividing each predictor observation by its standard deviation and then averaging all of the predictor quotients to form a composite. The second procedure, which will be used in my study, involves standardizing each of the predictors (i.e., *z* scores) and then calculating the mean of these standardized scores to form a composite (Guion, 1998; Raju et al., 1997). In either procedure, the composite is correlated with the criterion to determine R and R^2 (Guion, 1998; Raju et al., 1997).

Ordinary Least Squares versus Equal Weights

Assuming linear relations, equal weights cannot outperform the predictive power of optimal weights in the original sample. However, it is possible for equal weights to have a greater cross-validated R^2 than optimal weights. The better performance of equal weights relative to optimal weights occurs when an optimally weighted regression equation capitalizes upon chance distributional characteristics (Wainer, 1976). Critically, other samples may not reflect these distributional characteristics, resulting in greater shrinkage. Equal weights are not as strongly affected by sample specific characteristics (Cattin, 1980). Dawes and Corrigan (1974) documented that unit weights may be preferred when working with a changing population (e.g., changing employee pools), which is a highly salient issue in organizational activities such as personnel selection. Wainer (1976) has gone so far as to recommend using equally weighted predictors in all situations.

In a Monte Carlo study of various regression and cross-validation procedures, Raju et al. (1999) observed greater cross-validated R^2 values for equal weights across all sample sizes. Because Raju et al. (1999) investigated only one population dataset, other factors remain to be investigated.

Sample Size

As mentioned, sample size affects the amount of error in a study, and high degrees of error relate to instability in R^2. Consequently, researchers and practitioners may find it beneficial to consider how sample size affects the utility of their analyses. In a study of regression efficiencies, Schmidt (1971) found that optimal weights were not superior to equal weights upon cross-validation until samples met or exceeded 200 observations, and Dorans and Drasgow (1978) found that larger sample sizes (i.e., 120 observations) were required before optimal weights began to cross-validate as well as equal weights. Similarly, Claudy (1972) found that in small samples (i.e., 20 observations), equally weighted predictors produced the highest cross-validated population validity in 16 of his 18 generated populations. Furthermore, the population validities produced with the optimally weighted regression procedure had considerably more variance when there were fewer observations.

Predictor-Criterion Relationships

Another factor that researchers or practitioners should consider is the strength of the relationship between a predictor and its criterion. Claudy (1972) highlighted the value of predictor-criterion relationships after classifying his pre-generated populations by their characteristics. The equal weighting technique produced the highest population validities, regardless of sample size, in populations with low variability in predictor-criterion correlations and low to moderate (i.e., .00 to .40 predictor intercorrelations.). Smaller sample sizes (i.e., fewer than 50 observations) continued to perform better with equal weights in populations that retained low variability in predictor-criterion correlations but had predictor intercorrelations between -.20 and .00 or .40 and above (Claudy, 1972). However, larger sample sizes performed better with optimally weighted regression equations. Finally, Claudy (1972) reported that optimal weights performed best in populations with high variability in predictor-criterion correlations and predictor intercorrelations between -.30 and .40. Claudy (1972) closed with a discussion on the boundary condition that existed in conditions of 200 observations, wherein optimal weights based on smaller samples were overly complicated and less fruitful than simpler methods (e.g., equal weights).

Number of Predictors

When considering the number of predictors in a regression analysis, it is important to remember that parsimony is key; more is not necessarily better. A major advantage of multiple regression analyses, relative to bivariate regression, is the ability to include more predictors for increased explanation of criterion variance. However, at least with respect to optimally weighted regression weights, a major disadvantage with using a large number of predictors is that partial regression coefficients become less stable (Herzberg, 1967). Browne (2000) supported this point, finding that increasing the number of predictors benefitted the regression model to an extent, but additional predictors actually reduced the predictive power of the regression equation upon cross-validation. Regression analyses are prone to capitalizing upon chance distribution characteristics when there are many parameters (i.e., predictors) and the initial sample size is small (Browne, 2000). Therefore, researchers should maximize the ratio of study participants to predictor variables.

Although varying the number of predictors would be a valuable avenue of study, I should note that I will not assess the effects of this variable due to the multiplicative effect that it would have on my analyses. Furthermore, I implement various predictor-criterion correlations within each condition of this study, which presents methodological and explanatory issues for the retention of variables in smaller regression models.

Predictor Intercorrelation

The final factor that researchers and practitioners should consider is the correlation among predictors. The equation for partial regression coefficients indicates that stronger intercorrelations will result in lowered coefficients. Raju et al. (1999)

speculated that the observed superiority of equal weights over optimal weights upon cross-validation was due to the low and moderate predictor-criterion relationships; thus, future research should investigate predictive accuracy with varied predictor intercorrelations.

The Present Study

To follow in the path of Raju et al. (1999), my study will use Monte Carlo analyses to investigate factors that lead optimal weights to outperform equal weights upon cross-validation. Monte Carlo techniques have the benefit of allowing for relationships to be tested under a variety of conditions. Furthermore, Monte Carlo analyses can run the analyses many times to reduce the likelihood that the results are the product of sampling error. I make the following hypotheses:

Hypothesis 1: Optimal weights will have greater cross-validated R^2 values than will equal weights when predictor intercorrelations are high.

Hypothesis 2: Optimal weights will have greater cross-validated R^2 values than will equal weights when there is greater variability in bivariate predictor-criterion correlations.

Hypothesis 3: Optimal weights will have greater cross-validated R^2 values than will equal weights when sample sizes are large.

Method

Sample

The statistics program SAS University Edition® (SAS, 2020) was used to generate and analyze the datasets for this study. Scores were generated on five variables: a single criterion variable and four predictor variables. All variables were standardized in the population dataset.

Design

I tested ten different populations with two different sample sizes. The populations were the result of five predictor-criterion conditions and two intercorrelation conditions. Each of these three variables (predictor-criterion correlation, predictor intercorrelation, and sample size) are explained below.

Predictor-criterion correlations

The predictor-criterion correlations were set as follows.

Condition 1: Four moderate. r_{xy} = .30 for all four predictors.

Condition 2: Half strong, half weak. r_{xy} = .40 for two predictors and r_{xy} = .20 for two predictors.

Condition 3: Half very strong, half very weak. r_{xy} = .50 for two predictors and r_{xy} = .10 for two predictors.

Condition 4: One strong, three weak. r_{xy} = .40 for one predictor and r_{xy} = .20 for three predictors.

Condition 5: Three strong, one weak. r_{xy} = .40 for three predictors and r_{xy} = .20 for one predictor.

Predictor Intercorrelation

To address the effect of intercorrelation among predictor variables, I tested two levels of correlation among the predictor variables, moderate (.30) and strong (.50).

Sample Size

Increases in sample size decrease the effect of sampling error, which subsequently improves predictor weights while decreasing the detrimental effects of regression overfitting. To address this effect, I implemented two sample sizes in each of the conditions. These sample size conditions included 150 and 200 observations. It is well-established (e.g., Claudy, 1972; Dorans & Drasgow, 1978; Schmidt, 1971) that equal weights are superior in smaller sample sizes. Therefore, I chose to implement larger samples to better understand how my study's factors affected regression analyses when optimal weights could be expected to start cross-validating as well as equal weights.

Each observation consisted of a criterion score as well as four predictor scores. Each population consisted of one million cases. The ten population correlation matrices are listed in Appendix A. Each population was sampled 1,000 times. Composite scores for the four predictors were computed two different ways in each condition, via optimal weights and equal weights.

Cross-Validation Analysis

Empirical cross-validation of the optimally weighted and equally weighted composites occurred in two steps. First, predictor scores from the population were applied to both prediction equations to generate predictor composite scores. Second, these composite scores were correlated with the actual scores on the criterion in the population to determine the cross-validated R (and R^2). Results were averaged across 1,000 replications.

Results

Tables 1 and 2 report the average predictive power of the initial (R^2_{OLS} and R^2_{EW}) and cross-validated ($R^2_{OLS,CV}$ and $R^2_{EW,CV}$) regression models across the ten population matrices. Although my hypotheses did not test the relative performance of optimal versus equal weights within the initial (i.e., derivation) sample, it is worth examining the predictive power of these two weighting schemes. Unsurprisingly, in the derivation sample, optimally weighted regression analyses outperformed the equally weighted alternative in every condition, regardless of sample size. In some conditions, the difference in predictive power between the analyses was trivial (e.g., .016 in Condition 1), but in other conditions, the difference was quite large (e.g., .238 in Condition 5).

Table 1

Average Predictive Power of Initial and Cross-Validated Regression Models with a Sample Size of 150

	R^2_{OLS}	$R^2_{OLS,CV}$	R^2_{EW}	$R^2_{EW,CV}$
Condition 1	0.210	0.174	0.194	0.189
Condition 2	0.166	0.128	0.148	0.144
Condition 3	0.263	0.232	0.193	0.189
Condition 4	0.241	0.209	0.147	0.144
Condition 5	0.430	0.406	0.192	0.189
Condition 6	0.473	0.453	0.145	0.144
Condition 7	0.195	0.158	0.135	0.131
Condition 8	0.181	0.143	0.106	0.100
Condition 9	0.318	0.287	0.261	0.258
Condition 10	0.274	0.241	0.200	0.196

Note. R^2_{OLS} = initial R^2 with optimally weighted predictors; $R^2_{OLS,CV}$ = cross-validated R^2 with optimally weighted predictors; R^2_{EW} = initial R^2 with equally weighted predictors; $R^2_{EW,CV}$ = cross-validated R^2 with equally weighted predictors.

Overall, with samples of 150 observations (Table 1), optimally weighted predictors (R^2_{OLS}) explained, on average, 10.3% more criterion variance than equally weighted predictors. The 200-observation sampling condition (Table 2) reflects a similar result, with optimal weights explaining 10.0% more criterion variance on average in the initial (i.e., derivation) sample. However, Conditions 5 and 6 (i.e., r_{xy} = .50 for two predictors and r_{xy} = .10 for two predictors) appear to inflate the average predictive power of the cross-validated optimal weights in both sampling conditions. This trend indicates that the optimal weighting technique is a more powerful regression technique as predictors have varying relationships with the criterion because optimally weighted models weigh predictors according to their predictive power. Consequentially, the optimally weighted

Table 2

Average Predictive Power of Raw and Cross-Validated Regression Models with a Sample Size of 200

	R^2_{OLS}	$R^2_{OLS,CV}$	R^2_{EW}	$R^2_{EW,CV}$
Condition 1	0.204	0.177	0.191	0.189
Condition 2	0.159	0.132	0.147	0.144
Condition 3	0.256	0.235	0.189	0.189
Condition 4	0.236	0.213	0.147	0.144
Condition 5	0.428	0.409	0.191	0.190
Condition 6	0.472	0.456	0.147	0.144
Condition 7	0.190	0.163	0.135	0.132
Condition 8	0.178	0.148	0.104	0.100
Condition 9	0.311	0.290	0.259	0.258
Condition 10	0.269	0.245	0.197	0.196

Note. R^2_{OLS} = initial R^2 with optimally weighted predictors; $R^2_{OLS,CV}$ = cross-validated R^2 with optimally weighted predictors; R^2_{EW} = initial R^2 with equally weighted predictors; $R^2_{EW,CV}$ = cross-validated R^2 with equally weighted predictors.

predictor composite explained (on average across Condition 6) 32.7% more criterion variance than equally weighted predictors. Optimal weights continued to outperform the equal weighting technique after removing Conditions 5 and 6 from consideration, but by much lower margins (5.8% with samples of 150 and 5.4% with samples of 200).

Predictive Power Upon Cross-Validation

Results demonstrate that there was less shrinkage for equally weighted composites upon cross-validation. This result reflects past research (e.g., Dawes & Corrigan, 1974). With 150 observations, the adjusted R^2 for optimally weighted regression equations averaged losses of .032 (i.e., 3.2% less criterion variance), but the adjusted R^2 for equal weights was only .004 (i.e., .4% less criterion variance). Furthermore, increasing the sample size supported past literature (e.g., Pedhazur, 1982; Trochim & Donnelly, 2008), which indicated that more observations would positively relate to predictive stability. With samples of 200 observations, the adjusted R^2 for optimal weights decreased to .024, and the average loss in predictive ability for equally weighted predictors was only .002. However, shrinkage is only one component of addressing the advantages and disadvantages of optimally and equally weighted regression techniques. Researchers and practitioners are arguably more concerned with the final cross-validated predictive ability of their regression analysis.

Optimally weighted predictors explained more criterion variance in every cross-validation sample except those in Conditions 1 and 2. The uniform predictor-criterion correlations of Conditions 1 and 2 distinguish them from the other study populations. According to the equation for partial correlation coefficients (Guion, 1998), predictors with the same validity and predictor intercorrelations will have partial correlation coefficients of similar magnitude, so the optimally weighted regression actually operates analogously to the equally weighted technique. Therefore, the results of this study provide support for Hypothesis 2. Greater variability among bivariate predictor-criterion correlations is associated with greater cross-validated R^2 values for optimally weighted (versus equally weighted) predictor composites.

The evidence for Hypothesis 2 relegates Hypothesis 1 (i.e., optimal weights will have greater cross-validated R^2 values than will equal weights when predictor intercorrelations are high) to secondary importance. Optimally weighted predictors outperformed equally weighted predictors in all but the same two conditions, regardless of predictor intercorrelation. However, certain data trends are interesting. In Conditions 1 and 2, equally weighted composites remained the superior technique regardless of predictor intercorrelations. Therefore, it appears that in the absence of variability in predictor-criterion correlations, equally weighted regression analyses may perform as well as optimally weighted regression analyses. However, in every other condition, increasing predictor intercorrelations resulted in optimal weights explaining greater criterion variance than equally weighted predictors. With predictor intercorrelations of .30, optimally weighted predictors explained 6.0% and 6.3% more criterion variance with samples of 150 and 200 observations, respectively. Increasing the predictor intercorrelation to .50 resulted in optimally weighted predictors explaining 8.9% and 9.3% more criterion variance with samples of 150 and 200 observations, respectively. In Condition 5 ($N = 150$), optimally weighted regression procedures explained 21.7% more criterion variance than equally weighted predictors, and this predictive superiority increases by 9.2% in Condition 6 (i.e., 30.9% more criterion variance). In conclusion, there is support for Hypothesis 2. In general, optimal weights will have greater cross-validated R^2 values than will equal weights when predictor intercorrelations are high.

Finally, as with Hypothesis 1, the results for Hypothesis 3 failed to surpass those for Hypothesis 2 in importance; there were not any major changes between the two sampling conditions. The weighting technique that cross-validated best with a sample of 150 observations continued to perform best with 200 observations. However, one trend was apparent; there was less shrinkage for optimal weights when sample sizes were greater. With samples of 150 observations, R^2 values decreased by 3.2% upon cross-validation for optimal weights. However, with 200 observations, this loss in predictive power was only 2.4%. By comparison, equal weights were almost unaffected by sample size (the difference in average shrinkage was only .2%). Therefore, there is some supporting evidence for Hypothesis 3; optimal weights may achieve greater cross-validated R^2 values than will equal weights when sample sizes are large.

Discussion

My study has several important implications for researchers and practitioners. My results cast doubt on the accepted wisdom (e.g., Claudy, 1972; Dorans & Drasgow, 1978; Schmidt, 1971) that equally weighted predictors should be considered the default for regression analyses. In 16 of the 20 conditions examined, the cross-validated R^2 values were greater for optimally weighted composites than for equally weighted composites. These results are most useful to those who may have otherwise ignored the potential value of optimally weighted regressions, instead preferring the advantages they associated with the equally weighted alternative. Critically, these individuals may be missing out on the incremental validity afforded by optimal weights when there is a large degree of predictor-criterion variability. However, noting this variability is just one factor to consider prior to conducting one's regression analysis.

In addition to the variability in predictor-criterion relationships, it would be wise to account for the entire bivariate correlation matrix and sample size of a dataset. The results of this study indicate that when predictor-criterion correlation variability is nominal, then equally weighted composites should be preferred. However, as predictors inevitably correlate with one another, and when predictor-criterion correlations differ by non-trivial levels, then researchers should favor optimally weighted regression equations. Not only do equally weighted procedures fail to account for various predictor validities, but this technique will also fail to address increasing communalities among the predictors, therefore resulting in subpar cross-validation. Finally, given the size of the sampling conditions in my study, organizations that select many (i.e., 150 or more) applicants at one time (e.g., colleges or military services) should be wary of defaulting to an equally weighted regression. In these applications, the precision afforded by an optimally weighted regression may provide incremental validity for predicting performance (e.g., college GPA). However, there are many other situations that my study does not account for, so there is an impetus for future research.

Directions for Future Research

I concur with previous Monte Carlo studies (e.g., Raju, 1999), which direct future research to explore other factors that affect our studies. My study addressed three critical variables for researchers and practitioners: sample size, predictor-criterion relationships, and predictor intercorrelations. However, it only addresses a small fraction of the infinite possibilities that researchers and practitioners may face. Critically, I did not even attempt to examine how varying numbers of predictors affected optimally and equally weighted regression models. Future research should address this factor. Furthermore, future research should study smaller variations in the predictor-criterion correlations to better understand when optimal weights are a more powerful regression technique, relative to equal weights.

The results of my study indicate that optimally weighted regression equations are more useful than was suggested by previous research. However, these results may not have been practical if it had not been for modern advancements in computing power. Furthermore, the results from this study would not be achievable for those who do not possess the technical skill to run Monte Carlo analyses. Therefore, my final suggestion for future research is for the design of a web-based tool that can simulate (just as my study did) any condition that a researcher or practitioner faces. I envision this product taking one of two forms. First, a database could be produced with enough datapoints to allow someone to extrapolate his or her data characteristics and determine the most appropriate regression weights. However, the second, more accurate option would be the development of a Cloud-based server that operates exactly as my study does to calculate the predictive power of optimal and equally weighted regression techniques. In either scenario, any researcher or practitioner could make the implications or policy decisions best suited to their study. Moreover, given the fact that regression analyses are not isolated to the I/O profession, this program could also become an important tool for many other professionals, promoting "science for a smarter workplace" (SIOP, n.d.).

REFERENCES

American Psychological Association. (2018). Principles for the validation and use of personnel selection procedures (5th ed.). *American Psychological Association*. https://www.apa.org/ed/accreditation/about/policies/personnel-selection-procedures.pdf

Bekker, M. H. J., Croon, M. A., & Bressers, B. (2005). Childcare involvement, job characteristics, gender and work attitudes as predictors of emotional exhaustion and sickness absence. *Work & Stress*, *19*(3), 221-237. https://doi-org.libsrv.wku.edu/10.1080/02678370500286095

Browne, M. W. (2000). Cross-validation methods. *Journal of Mathematical Psychology*, *44*, 108-132. https://doi.org/10.1006/jmps.1999.1279

Cattin, P. (1980). Estimation of the predictive power of a regression model. *Journal of Applied Psychology*, *65*(4), 407-414. https://doi.org/10.1037/0021-9010.65.4.407

Claudy, J. G. (1972). A comparison of five variable weighting procedures. *Educational and Psychological Measurement*, *32*(2), 311-322. https://doi-org.libsrv.wku.edu/10.1177/001316447203200208

Dawes, R. M., & Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin*, *81*(2), 95-106. https://doi.org/10.1037/h0037613

Dorans, N., & Drasgow, F. (1978). Alternative weighting schemes for linear prediction. *Organizational Behavior & Human Performance*, *21*(3), 316-345. https://doi-org.libsrv.wku.edu/10.1016/0030-5073(78)90057-0

Guion, R. M. (1998). *Assessment, measurement, and prediction for personnel decisions*. Lawrence Erlbaum Associates.

Herzberg, P. A. (1967). *The parameters of cross-validation* [ProQuest Information & Learning]. Dissertation Abstracts International.

McAndrew, N. S., Schiffman, R., & Leske, J. (2019). Relationships among climate of care, nursing family care and family well-being in ICUs. *Nursing Ethics*, *26*(7–8), 2494–2510. https://doi.org/10.1177/0969733019826396

Pedhazur, E. J. (1982). Elements of multiple regression analysis: Two independent variables. *Multiple regression in behavioral research: Explanation and prediction* (2nd ed.). Holt, Rinehart, Winston.

Raju, N. S., Bilgic, R., Edwards, J. E., & Fleer, P. F. (1997). Methodology review: Estimation of population validity and cross-validity, and the use of equal weights in prediction. *Applied Psychological Measurement*, *21*(4), 291–305. https://doi-org.libsrv.wku.edu/10.1177/01466216970214001

Raju, N. S., Bilgic, R., Edwards, J. E., & Fleer, P. F. (1999). Accuracy of Population Validity and Cross-Validity Estimation: An Empirical Comparison of Formula-Based, Traditional Empirical, and Equal Weights Procedures. *Applied Psychological Measurement*, *23*(2), 99–115. https://doi.org/10.1177/01466219922031220

SAS (2020). SAS university edition [computer software]. SAS Institute Inc. https://www.sas.com/en_us/software/university-edition.html

Schmidt, F. L. (1971). The relative efficiency of regression and simple unit predictor weights in applied differential psychology. *Educational and Psychological Measurement*, *31*(3), 699–714. https://doi.org/10.1177/001316447103100310

Society for Industrial and Organizational Psychology. (n.d.). *About the Society for Industrial and Organizational Psychology (SIOP)*. https://www.siop.org/About-SIOP

Trochim, W. M. K. & Donnelly, J. P. (2008). *The research methods knowledge base* (3rd ed.). Atomic Dog.

Wainer, H. (1976). Estimating coefficients in linear models: It don't make no nevermind. *Psychological Bulletin*, *83*(2), 213–217. https://doi.org/10.1037/0033-2909.83.2.213

APPENDIX A

Correlation Matrices for Each Study Population

Matrix 1

Moderate Predictor Validity, Moderate Predictor Intercorrelation

	1	2	3	4	5
1. Y	1.00				
2. X_1	.30	1.00			
3. X_2	.30	.30	1.00		
4. X_3	.30	.30	.30	1.00	
5. X_4	.30	.30	.30	.30	1.00

Matrix 2

Moderate Predictor Validity, High Predictor Intercorrelation

	1	2	3	4	5
1. Y	1.00				
2. X_1	.30	1.00			
3. X_2	.30	.50	1.00		
4. X_3	.30	.50	.50	1.00	
5. X_4	.30	.50	.50	.50	1.00

Matrix 3

Low/High Predictor Validity, Moderate Predictor Intercorrelation

	1	2	3	4	5
1. Y	1.00				
2. X_1	.20	1.00			
3. X_2	.20	.30	1.00		
4. X_3	.40	.30	.30	1.00	
5. X_4	.40	.30	.30	.30	1.00

Matrix 4

Low/High Predictor Validity, High Predictor Intercorrelation

	1	2	3	4	5
1. Y	1.00				
2. X_1	.20	1.00			
3. X_2	.20	.50	1.00		
4. X_3	.40	.50	.50	1.00	
5. X_4	.40	.50	.50	.50	1.00

Matrix 5

Very Low/High Predictor Validity, Moderate Predictor Intercorrelation

	1	2	3	4	5
1. Y	1.00				
2. X_1	.10	1.00			
3. X_2	.10	.30	1.00		
4. X_3	.50	.30	.30	1.00	
5. X_4	.50	.30	.30	.30	1.00

Matrix 6

Very Low/High Predictor Validity, High Predictor Intercorrelation

	1	2	3	4	5
1. Y	1.00				
2. X_1	.10	1.00			
3. X_2	.10	.50	1.00		
4. X_3	.50	.50	.50	1.00	
5. X_4	.50	.50	.50	.50	1.00

Matrix 7

3 Low/1 High Predictor Validity, Moderate Predictor Intercorrelation

	1	2	3	4	5
1. Y	1.00				
2. X_1	.20	1.00			
3. X_2	.20	.30	1.00		
4. X_3	.20	.30	.30	1.00	
5. X_4	.40	.30	.30	.30	1.00

Matrix 8

3 Low/1 High Predictor Validity, High Predictor Intercorrelation

	1	2	3	4	5
1. Y	1.00				
2. X_1	.20	1.00			
3. X_2	.20	.50	1.00		
4. X_3	.20	.50	.50	1.00	
5. X_4	.40	.50	.50	.50	1.00

Matrix 9

1 Low/3 High Predictor Validity, Moderate Predictor Intercorrelation

	1	2	3	4	5
1. Y	1.00				
2. X_1	.20	1.00			
3. X_2	.40	.30	1.00		
4. X_3	.40	.30	.30	1.00	
5. X_4	.40	.30	.30	.30	1.00

Matrix 10

1 Low/3 High Predictor Validity, High Predictor Intercorrelation

	1	2	3	4	5
1. Y	1.00				
2. X_1	.20	1.00			
3. X_2	.40	.50	1.00		
4. X_3	.40	.50	.50	1.00	
5. X_4	.40	.50	.50	.50	1.00