Quest Journals
Journal of Research in Applied Mathematics
Volume 2~ Issue 6 (2016) pp: 19-32
ISSN(Online) : 2394-0743 ISSN (Print):2394-0735
www.questjournals.org

Research Paper

Nonstationarity of Predictor Relative Importance: Illustration and Theory

Richard Hermida[1]

Received 06 Jun, 2016; **A**ccepted 29 Jun, 2016 © The author(s) 2016. **P**ublished with open access at **www.questjournals.org**

ABSTRACT:- *Frequently in applied settings, consumers of scientific data request measures of importance for each predictor variable in multiple regression analysis. Recently, methodological advances have provided evidence that relative weights more accurately assess importance than regression weights. However, predictors may experience non-trivial amounts of variation in relative importance across various data spaces. The current study seeks to combine relative weights with the principles of local modeling to produce a method by which relative weights can be assessed across multiple local models derived by the modeler, in order to increase generalizability and decrease inferential uncertainty.*

Keywords:- *relative weights, Simpson's paradox, weighted regression, research methodology, decision-making*

I. INTRODUCTION

Applied researchers in the social sciences are often tasked by organizational clients to build models that illustrate a particular relationship between a group of predictors, and an outcome [1]. In these cases, multiple regression is an extremely common tool of choice for applied researchers in determining which predictors are associated with the outcome, and to what degree [2-4]. Within the context of traditional multiple regression, researchers generally use indices such as zero-order correlations and standardized regression weights to establish the relative contribution of individual predictors to an outcome [5-7].

However, these traditional methods have been recognized as less than ideal for estimating relative contribution of predictors to an outcome, particularly to the degree predictors in the predictor set are correlated with one another [5-6, 8-9]. To address this issue, more sophisticated techniques such as dominance analysis [6] and relative importance analysis [10] have been developed and used in applied research to gain more accurate estimations of predictor importance within a multiple regression framework [11].

Ideally, the users of relative importance analysis results will use the results in an informed decision-making process to improve an outcome of interest. However, individuals would most likely know intuitively that to some degree, the set of predictors is unlikely to have the exact same relationship to the outcome across all possible contextual factors. This idea is inherent in the idea of statistical interaction [12-13], meta-analysis [14-16], and local validation studies [17]. However, given scarcity of resources such as time and money, conducting localized studies across all possible contextual factors (and combination of factors) is usually not feasible. As a result, predictor importance is either assumed to be stationary across levels of non-studied variables, or assumed to be non-stationary across levels of non-studied variables, but simply not analyzed. This is problematic to the degree predictor importance in a global setting is different than predictor importance within various local settings, and the degree to which inferences about predictor importance in local settings are made from predictor importance statistics in global settings.

The goals of this paper are threefold. First, illustrate that consideration of predictor importance in global settings in isolation from local settings has the potential to lead to serious misinterpretations of predictor-outcome relationships. Second, illustrate local modeling procedures that can help researchers find local values for predictor-outcome relationships, as well as global values. Third, demonstrate these points with a real-life dataset and code, in order to demonstrate that these issues are not methodological minutia, and can meaningfully influence conclusions regarding predictor importance.

*Corresponding Author: Richard Hermida

Department of Psychology, George Mason University, Fairfax, Virginia, United States

II. FORMULATION OF RESEARCH PROBLEM

2.1 Relative Importance Analysis

When a researcher engages in multiple regression analysis, a key issue may be the relative importance of each individual predictor in the regression model in predicting the criterion of interest [18-20]. In order to assess importance, researchers have suggested numerous options including correlation coefficients, partial and semipartial correlations, commonality analysis, and regression weights [6, 7, 10, 21-22, 23-24].

However, to the degree that correlated predictors exist in a multiple regression equation, each of these options is flawed to a non-trivial degree [6, 13, 25-26]. Specifically, simple correlations are unable to partition variance in the criterion that is shared by multiple predictors. Similarly, partial and semipartial correlations are unable to assigned to a specific variable the variance shared by correlated predictors. Finally, standardized regression coefficients are problematic for reasons related to interpretability problems in the face of collinearity. In light of the inadequacies of these tactics in assessing importance in the presence of collinearity, two advances have emerged as being proper methodological methods for assessing importance of multiple predictors: general dominance weights and relative weights. We will discuss the later of the two now.

Relative weights, which are denoted as εj [10, 27], address issues of evaluating importance in the presence of collinearity by first deriving a set of k orthogonal weights ZXk from a newly created set of predictors which are maximally related to the set of j original predictor variables Xj, while maintain orthogonality. Second, the criterion Y is regressed onto these new orthogonal (but maximally related to the original) variables ZXk to obtain standardized regression coefficients. Third, the original variables Xj are regressed onto their orthogonal counterparts ZXk in order to obtain standardized regression coefficients. Finally, relative weights are computed by summing the products of squared standardized regression coefficients that were obtained in the previous two steps. This procedure produces a relative importance weight for each predictor in the predictor set. Moreover, as a consequence of using orthogonal transformations of the original variables, issues related to collinearity are not present in relative weights analysis. Ultimately, within the context of linear regression, the term relative importance reflects the proportionate contribution each predictor variable makes to the total predicted criterion variance taking into account a variable's contribution by itself and in combination with other predictor variables [4, 7].

There are a number of components of relative importance analysis that deserve illumination. First, the raw relative weights reflect the raw percentage of criterion variance that is uniquely attributable to the predictor in question, within the predictor set in question. Since these raw relative weights are an additive decomposition of the total model R2, the sum of the raw relative weights in a predictor set should always equal total model R2 [28]. For example, if the raw relative weights of two predictors were 0.20 and 0.15, it would follow that total model R2 was 0.35. Relatedly, these raw relative weights can be transformed into rescaled relative weights, which simply reflects the percentage of predicted variance attributable to each predictor. Using the previous example, the rescaled relative weight of the first predictor would be 57.14 (0.20/0.35). While raw and rescaled relative weights are directly linked, they are associated with different research questions and have different interpretations. Since there is little cost to reporting both raw and rescaled values, it is generally recommended to report both values for the best understanding of predictor importance.

A second set of components associated with relative importance analysis is related to confidence interval information. The confidence interval around a raw relative weight is simply a reflection of the precision of the relative weight, derived from bootstrapped samples. Each bootstrapped sample contains a raw relative weight for each predictor. Therefore, with a greater number of bootstrapped samples, the smaller will be the confidence interval surrounding the relative weight, given the alpha value in question. Assuming $\alpha = .05$, the lower bound of the confidence interval is the value (after ordering by magnitude) associated with the 2.5th percentile, while the upper bound of the confidence interval is the value associated with the 97.5th percentile, with all values between these two values representing the confidence interval of the raw relative weight.

The third and final set of components associated with relative importance analysis is related to testing the statistical significance of relative weights within a predictor set. This information is derived from examining the confidence interval of the variable in question with the relative weight of randomly generated variable [2]. Since the randomly generated variable in the sample should have a relative weight that only reflects sampling error (since the population relative weight is zero), if the confidence interval of the raw relative weight of the predictor contains zero, then the relative weight would not be statistically significant. In order to aid the clarity of our proposed procedure, readers are encouraged to examine the detailed expositions of how relative weights are formed computationally [10, 29-31].

A number of studies have incorporated relative importance analysis into the study of scientific relationships. For example, Snell, Tonidandel, Braddy, and Fleenor [31] used relative importance analysis to understand how political skill predicts managerial effectiveness. Scherer, Baysinger, Zolynsky, and LeBreton [32] used relative weight analysis to examine the contributions of sub-clinical psychopathy and the global traits of the Five Factor Model in the prediction of counter-productive workplace behaviors. Braddy, Gooty, Fleenor, and Yammarino [33] examined the relative importance of different leader behaviors for predicting derailment. Finally, Chao, Zhao, Kupper, and Nylander-French [34] used relative weights in investigating the factors contributing to Styrene in the reinforced-plastic industry.

Despite the use of relative importance analysis in published research, it is unclear the degree to which the results derived from relative importance analysis in published research are stationary vs. non-stationary across contextual factors. For example, it could be the case that the relative importance of a predictor changes across levels of a third variable, from a perspective of rank order (i.e.- being first, second, third, etc.…most important in a predictor set) or from a perspective of magnitude (i.e.- raw relative weights). This is due to the distinction between global and local contexts.

2.2 Global vs. Local Contexts

Traditionally within a regression format, researchers gather a set of observations and regress the outcome of interest on the predictor set of interest, with all observations contributing equally to parameter estimation and model statistics. We can consider global models as models where all aspects of the data and the distributions contribute to the estimate of fit. By contrast, local models are models where only a certain subset of the data contributes to the estimate of fit, or models where a certain subset of the data receives greater importance than other subsets in contribution to estimation of fit [35].

Simpson's paradox [36-42] illustrates that it is possible to have differing relationships amongst variables when data is aggregated to a population level; compared to when data is disaggregated on some variable at a subpopulation level (the continuous case is often referred to as Robinson's paradox, [43]). Specifically, the paradox occurs when causal inferences are drawn across different explanatory levels: from populations to subgroups, or subgroups to individuals, or from cross-sectional data to intra-individual changes over time [44]. This finding has been established by mathematicians and statisticians [39-40, 45-48], and has been found in many applied research situations including cognitive psychology [49], sports [50], and kidney stone treatment [51].

The action of the paradox is most likely derived from two major sources. The first source is a tendency for human beings to assume that causal relationships are governed by the laws of probability, along with certain causal assumptions human beings possess as a general rule. Specifically as it relates to the paradox, there is often an assumption that where there is correlation, there must be an underlying cause. Said differently, a person who falls victim to the paradox will treat proportions or rates (which are transitory) within a causal calculus [39], which implies invariant relationships. Therefore, the confusion is fairly straightforward in that there cannot be an invariant feature (i.e. - a shared cause) for a shared outcome that would lead to contradictory correlations in the subpopulations [39-40].

Therefore, the user of the model for predictive or decision purposes is left with a data analytic situation where the fit of the model in a global context (i.e. - over the entire data space) is radically different from various regions of the data space (i.e. - a particular local context). This is problematic in situations where the decision analysis in practice is likely to draw from a population that is likely to occupy the regions of the data space where fit is poor. Said differently, there is a difference between fit across the global region of data and fit across local regions, and there is difference between decision-making functions that target aggregates, and decision-making functions that target individuals or smaller, more focused groups of individuals.

This paradox has serious implications for decision-making procedures derived from statistical analysis. To the degree there is variability in stimulus-response relationships amongst local models is the degree to which consideration of local models is important relative to global models, and the degree to which considerations and recommendations using global models in isolation can lead practioners to make misleading recommendations.

Since global models are simply the aggregation of the infinite series of local models, the issues present in the aforementioned paradox would seem to be applicable to the global vs. local issues mentioned earlier. This would seem to be especially true for models where scores were clustered more strongly at some points than others. The reason for this is that the effective N differs across the continuum of the third variable, with a higher effective N being associated with focal points that are closer to the more densely populated values of the third variable, as compared to focal points that are closer to the less densely populated values of the third variable.

2.3 Local Relative Weights

Local relative weights (LRW) can be a very informative tool in observational studies, because they allow the description of variable-gradients of parameter estimates from regression models across a wide range of values based on relatively small samples. Locally-weighted averaging, used in nonparametric regression [52] and nonparametric mixed effects models [53], can be implemented in an LRW context in order to define observation weights as a function of a continuous variable and fit a series of models using differently weighted observations. Within a LRW framework, one can how information related to relative importance analyses change across the continuum of other variables not in the predictor-criterion relationship.

This process is accomplished through the use of kernel functions and kernel density estimation (KDE). To estimate the density at a given point x, the local averaging occurs within a specified neighborhood of points close to x. The closer the points are to x, the more weight they are assigned. Therefore, the density is strongest at x, with the most distant points receiving trivial or zero weights [54-55].

Using these calculated sample weights, OLS regression with multiple predictors and weighted observations can be estimated. Based on calculated sample weights for a series of focal points (see calculation steps of weights in Appendix A), LRW models can be sequentially fitted moving the weighting window along the continuous variable in question. Parameter estimates and model fit statistics for the series of models can further be plotted against the continuous variable in order to visualize their equivalence or change. Changes can then be computed for the parameter estimates and the relative importance weights. I will now provide a simple empirical example of this procedure, using applied data.

III. METHODS

3.1 Sample

The data for this study came from 3,931 individuals (55.4% female), sampled via random digit dialing, who participated in the second wave of the Midlife in the United States (MIDUS) study of adult development (MIDUS-2). The MIDUS-2 sample represents the subset of MIDUS-1 participants who were successfully recontacted and agreed to participate again, while possessing full data for all study items. At MIDUS-2, ages ranged from 30 to 84 years (M = 56.06, SD = 12.30). Participants reported their ethnicities as: 91.9% White, 3.6% African American, 0.4% Asian American, 1.5% American Indian, and 2.6% other or multiracial.

3.2 Measures

Openness to Experience. Openness to experience was measured with seven items taken from Rossi (2001) and was scored on a 1 to 4 Likert scale, with 1 = not at all, and 4 = a lot (M = 2.91, SD = 0.54). An example item is “Intelligent describes you how well”. Alpha for this scale was .77.

Conscientiousness. Conscientiousness was measured with four items taken from Rossi [56] and was scored on a 1 to 4 Likert scale, with 1 = not at all, and 4 = a lot (M = 3.46, SD = 0.45). An example item is “Responsible describes you how well”. Alpha for this scale was .76.

Extraversion. Extraversion was measured with five items taken from Rossi [56] and was scored on a 1 to 4 Likert scale, with 1 = not at all, and 4 = a lot (M = 3.10, SD = 0.57). An example item is “Outgoing describes you how well”. Alpha for this scale was .76.

Agreeableness. Agreeableness was measured with five items taken from Rossi [56] and was scored on a 1 to 4 Likert scale, with 1 = not at all, and 4 = a lot (M = 3.45, SD = 0.50). An example item is “Helpful describes you how well”. Alpha for this scale was .80.

Neuroticism. Neuroticism was measured with four items taken from Rossi [56] and was scored on a 1 to 4 Likert scale, with 1 = not at all, and 4 = a lot (M = 2.06, SD = 0.62). An example item is “Nervous describes you how well”. Alpha for this scale was .74.

Social Integration. Social integration (SI) was measured with three summed items taken from Keyes [57] with each item scored on a 1 to 7 Likert scale, with 1 = strongly disagree, and 7 = strongly agree (M = 14.73, SD = 3.99). An example item is “I feel close to other people in my community”. Alpha for this scale was .75. Age. Age was measured with a single calculated variable. The average age was 56.06 years (SD = 12.30), and ranged from 30 to 84 years of age.

3.3 Model

For this example, the predictors are the big five personality variables of openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism. The criterion variable is social integration. The “third variable” is age. Given the desire to illustrate local modeling procedures, we will conduct the procedure described in the introduction over the psychological space of the age variable. Given that age has 55 possible values for any single individual, we will conduct our analysis with 55 1 unit steps. We used a kernel function of weighting observations [58]. The following computations were carried out to define sample weights for 55 focal points between the age summed scores of 30 to 84.

The bandwidth of the kernel function was calculated by the following formula:

$$bw = (2)(N^{.2})(SDage) \qquad (3)$$

A scaled distance (z_x) was then computed by subtracting the focal age score from every observation for every focal point:

$$z_x = (Age_x - FocalAge) \,/\, (bw) \qquad (4)$$

Weights were then calculated based on the normal kernel function for every focal point:

$$K_{focalage} = (1/2\Pi) * \exp(-z_x^{\ 2}/2) \qquad (5)$$

Last, weights (W) were rescaled in order to obtain values between 0 and 1:

$$W_{focalage} = (K_{focalage} \div .399) \qquad (6)$$

First, based on these sample weights, covariance matrices amongst the predictors and outcome were generated for each of the 55 individual focal points using the WEIGHT procedure within SAS (see Appendix A). Second, using SPSS syntax (see Appendix B), raw datasets were generated for the predictors and outcomes that reflected the means and standard deviations of the variables, as well as the nature of their interrelatedness and correlation. Third, using R syntax (see Appendix C), the following information was obtained for the study model: a) raw relative weights, b) rescaled relative weights, c) confidence interval for raw relative weights, and d) confidence interval tests of statistical significance associated with Tonidandel et al. [2]. This process generated the relevant statistical information for the study model for each of the 55 individual samples of weighted observations. Reported results will include for each predictor at each focal point, a) the overall predicted variance in social integration associated with the predictor set, b) raw and rescaled relative weights, c) confidence intervals of the raw relative weights, and d) confidence interval test information for statistical significance of relative weights for each predictor.

IV. RESULTS

4.1 Global Model Relative Weights

First, relative weights were calculated for the global model. Results indicated that extraversion possessed the most relative importance in explaining variance in social integration, accounting for 37.4% of the explained variance ($R^2 = .15$). Neuroticism and agreeableness were the next most important, accounting for 30.3% and 18.5% of the explained variance, respectively. Finally, openness to experience and conscientiousness were relatively unimportant, accounting for 7.0% and 6.8% of the explained variance, respectively. While all predictors possessed statistically significant relative weights, the lower bound of the confidence interval for statistical significance was barely above zero for openness to experience and conscientiousness (0.006 and 0.005, respectively), indicating that for these two variables, that while they were statistically significant in importance, the magnitude of their importance was very modest and statistically significant in large part to a fairly robust sample size, in contrast to the predictors of extraversion, agreeableness and neuroticism, whose lower bounds ranged from 0.02 to 0.04, indicating both statistical significance and a stronger effect size.

4.2 Local Models by Predictor

Openness to experience. Openness to experience was generally of minor importance with respect to predicting variance in social integration, and ranged in rescaled relative weights from 3.8 to 9.6. The rescaled relative importance weights associated with openness peaked around age focal points of 40-50 and 60-65 as can be seen in Figure X. The standard deviation of the rescaled relative importance weights was 2.00, the lowest in the predictor set, and several times lower than extraversion, agreeableness, and neuroticism. These results illustrate that openness to experience was relatively unimportant in explaining variance in social-integration, in the presence of the other big five variables, although it was occasionally more important than conscientiousness. Pursuant to the distinction between global models and local models, the global model slightly overestimated the relative importance of openness at very early and very old ages, and slightly underestimated its relative importance towards the middle of the age distribution, as can be seen in Tables 1-2 and Fig. 1.

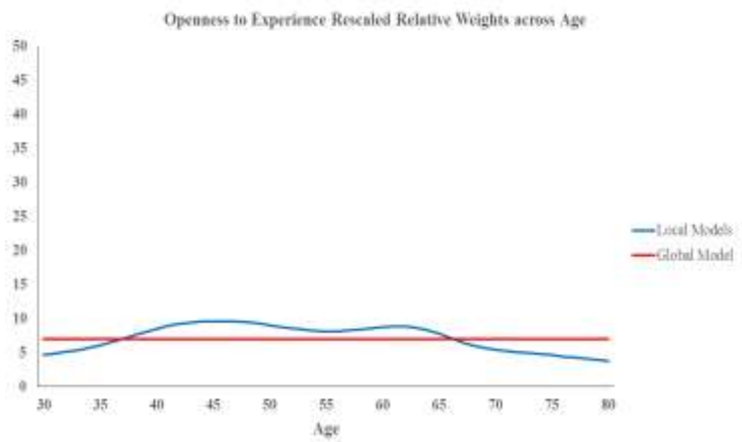

Figure 1: Openness to Experience Rescaled Relative Weights Across Age

Conscientiousness. Similarly, conscientiousness had very little importance throughout the entire range of age focal points and ranged from 0.4 to 18.0 in terms of rescaled relative importance weights. To the degree that conscientiousness experienced an upward shift in importance; the shift took place around focal points 50-60 and again at approximately ages 70-80, as can be seen in figure X. The standard deviation of the rescaled relative importance weights was 5.71, approximately the same as extraversion, agreeableness, and neuroticism. These results illustrated that conscientiousness was relatively unimportant in explaining variance in social-integration, in the presence of the other big five variables for most ages, although at select age ranges was a non-trivial contributor. Pursuant to the distinction between global models and local models, the global model noticeably overestimated the relative importance of conscientiousness at very early ages, and radically underestimated its relative importance towards the end of the age distribution, as can be seen in Tables 1-2 and Fig. 2.

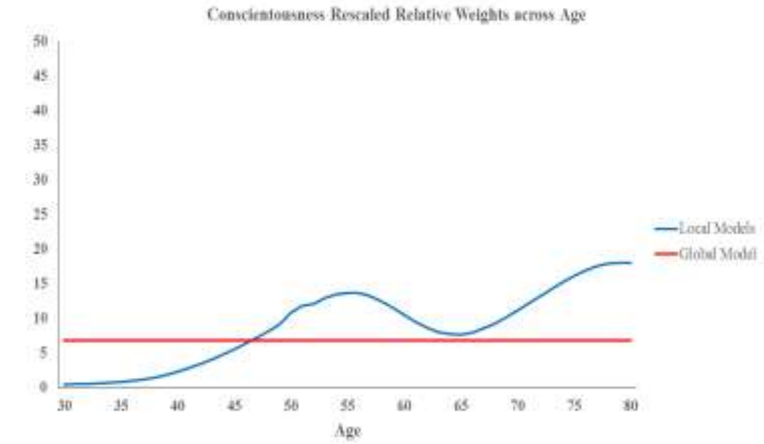

Figure 2: Conscientousness Rescaled Relative Weights Across Age

Extraversion. Extraversion was the most important variable in predicting social-integration variance, relative to the other big five variables, and possessed rescaled relative importance weights ranging from 27.4 to 50.4 with a standard deviation of 6.45. Extraversion was extremely important at lower focal points, occasionally accounting for more than 50 of the explained variance. The rescaled relative generally decreased as age increased, with minor upwards shifts around ages 60 to 70. Importantly, extraversion was the most important predictor in the predictor set for the entire set of age values.

The differences between the global model and local models showed serious differences between the two, as the global model dramatically underestimated the relative importance of extraversion at certain points, occasionally by as much as 13.0, as seen in Table 2 and Fig. 3, but also began to overestimate the rescaled relative importance of extraversion by a non-trivial degree for very high focal points (75 to 80). Ultimately, these results indicate that extraversion was important for all focal points in an absolute sense, but that the magnitude of that importance significantly varied across the spectrum of age focal points. The global model dramatically misrepresented the rescaled relative importance of extraversion at several different focal points, with the most dramatic differences involving the global model underestimating the importance of extraversion early in the age spectrum. Tables 1-2 and Fig. 3.

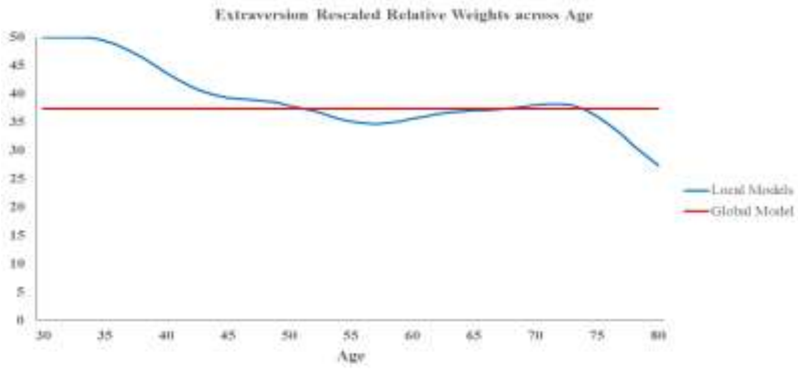

Figure 3: Extraversion Rescaled Relative Weights Across Age

Agreeableness. Agreeableness was an extremely important variable in predicting social-integration variance, relative to the other big five variables, and possessed rescaled relative importance weights ranging from 9.4 to 29.4 with a standard deviation of 6.49. Agreeableness was most important at focal points around ages 55 to 65. Agreeableness was generally the second or third most important predictor in the predictor set in predicting social-integration.

The differences between the global model and local models showed meaningful differences, as the global model underestimated the relative importance of agreeableness at certain points, occasionally by more than 10, as seen in Tables 1-2 and Fig. 4. Ultimately, these results indicate that agreeableness was moderately important for all focal points in an absolute sense, but that the magnitude of that importance somewhat varied across the spectrum of age focal points.

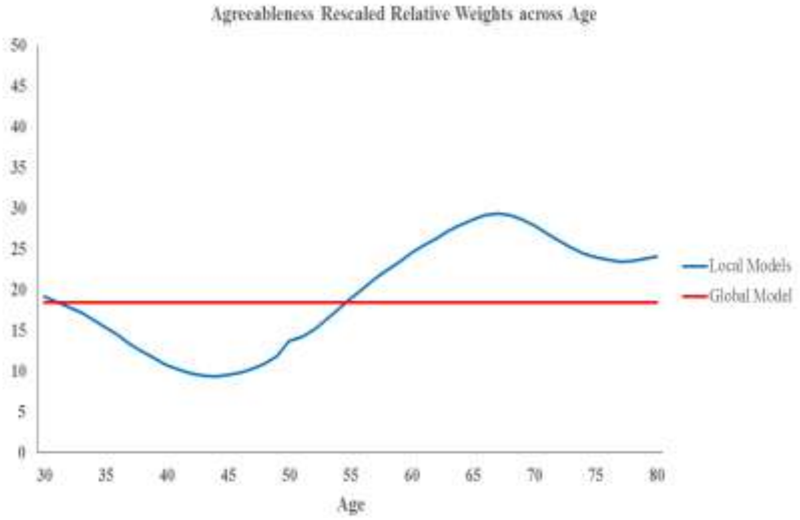

Figure 4: Agreeableness Rescaled Relative Weights Across Age

Neuroticism. Neuroticism was a somewhat important variable in predicting social-integration variance, relative to the other big five variables, and possessed rescaled relative importance weights ranging from 17.4 to 36.7 with a standard deviation of 6.04. Neuroticism was the second most important for lower to average focal points values, while being somewhat less relatively important for focal values associated with old age values.

The relative importance of neuroticism was usually overestimated according to the global model compared to the majority of local models, although in some instances underestimation did take place. This information can be seen in Tables 1-2 and Fig. 5.

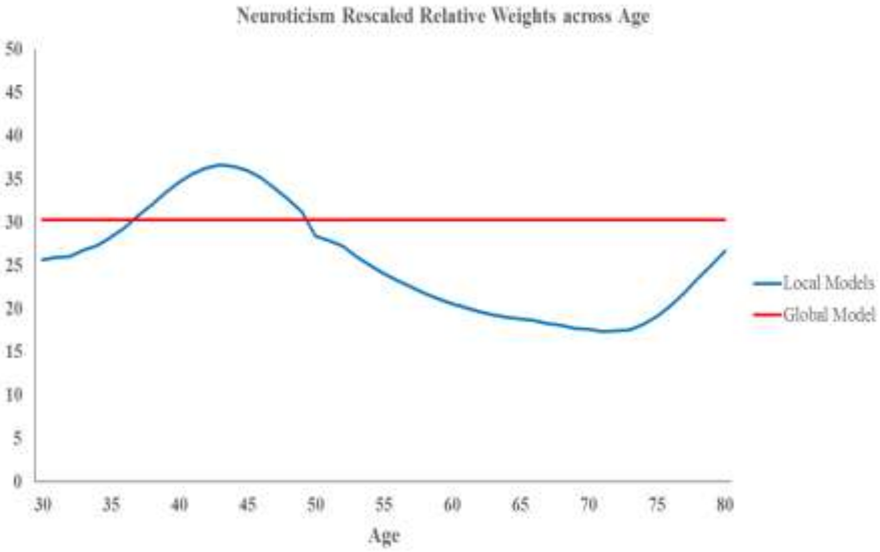

Figure 5: Neuroticism Rescaled Relative Weights Across Age

4.3 Results Summary

Taken as a whole, the global model vs. local model distinctions illustrated that at various points along the age continuum, the global model radically overestimated or underestimated the relative importance of the predictor in question, oftentimes by nearly 50%. This distinction can perhaps be seen most readily by the fact that extraversion and neuroticism were approximately the same in importance according to the global model (rescaled weights of 37 vs. 30, respectively), while at various focal points the rescaled relative weight around extraversion was more than twice that of neuroticism. Ultimately, these results in sum illustrate that the assumption of stationary predictor importance is not a reasonable assumption.

V. CONCLUSION

5.1 Limitations

While this study does provide useful information for researchers and users of regression analysis, there are several limitations as well. First, all the important limitations of relative weights analysis themselves extend to local relative weights analysis as well. Perhaps most critically, relative importance analysis is susceptible to model misspecification. That is, relative importance analysis assumes the correct model has been identified, and should not be the primary driver of predictor selection in developing a model [30], especially compared with more useful techniques [59-60]. Also of great importance is the fact that relative important weights should not be conceptualized as causal forces, and as a result require other factors such as regression weights and theory in order to suggest the best course of action [11].

Second, there are limitations specific to the procedure outlined in this study. Perhaps most important is that the procedures outlined in this study are only as useful to the degree that non-stationarity is correctly identified in the context in question. That is, in any given context, a multitude of variables could potentially cause non-stationary in a predictor set-criterion relationship in a variety of ways. One such case would be non-stationarity across two variables in competing and opposing ways. For example, if in a two predictor set the first predictor increased in relative importance as age increased, while the second predictor increased in relative importance as intelligence increased, the researcher would potentially have to discover whether age or intelligence information should be weighed more heavily in the analysis and recommendation stage of the project. Neglecting to identify all the potential variables that influence non-stationarity in a particular relationship can bias results and potentially lead researchers towards misleading inferences derived from the local relative weighting strategies.

Finally, there is room for debate over exactly which kernel function is best suited for use in weighing observations on which LRW analysis takes place. Specifically, this study used a normal kernel function for the purposes of simplicity, but other kernels exist such as the Epanechnikov and Quartic kernels [61-62]. While there are varying degrees of efficiency amongst different kernel shapes, the differences are usually trivial and not likely to meaningfully alter the type of results usually found in the social sciences and are not nearly as important as the choice of bandwidth selection [63]. Proper bandwidth selection is a frequently studied operation in a variety of contexts and contains several nuances [64-66]. For example, research has evaluated how to best minimize various error criterion through bandwidth selection [67-68]. Naturally, such issues deserved to be examined specifically within the context of LRW analysis as well.

5.2 Future Directions

One area for future research is in the direction of local modeling across the spectrum of multiple variables simultaneously. This is necessary, as a plethora of variables could theoretically cause non-stationarity in any given predictor set-criterion relationship. Expanding the current framework to finding the best ways to examine non-stationarity across multiple variables would prove extreme valuable. Relatedly, while global measures of importance are important, more narrowly focused and specific issues of importance, such as the pattern of importance [5, 69] relay different information to researchers and would benefit from application of LRW principles. Additionally, as previously mentioned, future research should examine how the specific nuances of kernel shape and bandwidth selection interact with LRW analysis. Future study can help determine and integrate best practices surrounding these issues.

Finally, moving to the more practical side of research, statistical or psychometric tools that do not interface with the tools researchers most commonly use are not likely to be implemented [70]. Therefore, another very practical direction for future research is providing researchers with the tangible and easy-to-use tools needed to carry out LRW analysis. In particular, software or programs compatible with statistical packages frequent used in applied research such as SPSS, SAS, and R would be very helpful. While a certain level of complexity is unavoidable for implementing complex statistical procedures, tools that minimize difficulty at the point of contact for researchers would be especially welcomed.

5.3 Summary

Traditionally, researchers obtain regression weights or relative importance weights from a global model and use these weights in making inferences about the data and the relationships therein. The present study suggests that in applied research, this global approach could potentially be overly broad when evaluating traditional predictor set-outcome relationships. The procedures outlined in this paper are meant to prompt researchers thinking about local modeling procedures, so as to better improve the ability of researchers findings to be properly transferred to applied settings, and for these findings to be based in a more appropriate and nuanced context.

Table 1: Raw relative weights across select values of age

Age	OPEN		CONSCI		EXTRA		AGREE		NEURO	
	LCI	UCI	LCI	UCI	LCI	UCI	LCI	UCI	LCI	UCI
30*	0.01	0.01	0.00	0.00	0.06	0.09	0.02	0.04	0.03	0.05
35*	0.01	0.01	0.00	0.00	0.07	0.09	0.02	0.03	0.04	0.06
40*	0.01	0.02	0.00	0.01	0.06	0.09	0.01	0.03	0.05	0.07
45*	0.01	0.02	0.01	0.02	0.05	0.08	0.01	0.02	0.05	0.08
50*	0.01	0.02	0.01	0.03	0.05	0.08	0.02	0.03	0.03	0.06
55*	0.01	0.02	0.02	0.03	0.05	0.07	0.02	0.04	0.03	0.05
60*	0.01	0.02	0.01	0.02	0.04	0.06	0.03	0.05	0.02	0.04
65*	0.00	0.01	0.00	0.01	0.03	0.05	0.02	0.04	0.01	0.03
70*	0.00	0.01	0.01	0.02	0.02	0.04	0.02	0.03	0.01	0.02
75*	0.00	0.01	0.01	0.02	0.02	0.04	0.01	0.03	0.01	0.03
80*	0.00	0.01	0.01	0.02	0.02	0.04	0.02	0.03	0.02	0.04
Average	0.01	0.02	0.01	0.02	0.04	0.07	0.02	0.03	0.03	0.05

Note: LCI = Lower confidence interval. UCI = Upper confidence interval.
Note: Statistically significant results are denoted by * in the LCI column.
Note: Full list of results available from the author upon request.

Table 2: Rescaled relative weights across select values of age

Age	OPEN	CONSCI	EXTRA	AGREE	NEURO
30*	4.70	0.43	49.95	19.27	25.66
35*	6.05	0.79	49.42	15.43	28.31
40*	8.46	2.26	43.78	10.82	34.68
45*	9.61	5.49	39.30	9.57	36.02
50*	9.00	10.80	37.90	13.80	28.50
55*	8.09	13.69	35.15	18.97	24.11
60*	8.76	10.47	35.56	24.56	20.64
65*	7.76	7.64	37.07	28.71	18.82
70*	5.36	11.10	37.97	27.95	17.62
75*	4.62	16.18	36.07	24.01	19.13
80*	3.76	18.04	27.39	24.12	26.69
Average (SD)	6.90 (2.00)	9.39 (5.71)	38.13 (6.47)	19.90 (6.49)	25.69 (6.04)

Note: Statistically significant results are denoted by *.
Note: Full list of results available from the author upon request.

REFERENCES

[1]. Mandel, 2008, Statistical modeling and business expertise, or where is the truth? Model Assisted Statistics and Applications, 3(1): 3-20.
[2]. S. Tonidandel, and J.M. LeBreton, and J.W. Johnson, 2009, Statistical significance tests for relative weights. Psychological Methods, 14(4): 387-399.
[3]. S. Tonidandel and J.M. LeBreton, 2010, Determining the Relative Importance of Predictors in Logistic Regression: An Extension of Relative Weights Analysis. Organizational Research Methods, 13(4): 767-781.
[4]. S. Tonidandel, and J.M. LeBreton, 2015, RWA Web: A Free, Comprehensive, Web-Based, and User-Friendly Tool for Relative Weight Analyses. Journal of Business and Psychology, 30(2): 207-216.
[5]. R. Azen and D.V. Budescu, 2003, The dominance analysis approach for comparing predictors in multiple regression. Psychological Methods, 8(2): 129-148.
[6]. D.V. Budescu, 1993, Dominance analysis: A new approach to the problem of relative importance of predictors in multiple regression. Psychological Bulletin, 114(3): 542-551.
[7]. J. Johnson, and J.M. LeBreton, 2004, History and use of relative importance indices in organizational research. Organizational Research Methods, 7(3): 238-257.
[8]. W. Kruskal, and R. Majors, 1989, Concepts of relative importance in recent scientific literature. American Statistician, 43(1): 2-6.
[9]. J.H. Ward, 1962, Comments on "The paramorphic representation of clinical judgment." Psychological Bulletin, 59(1): 74-76.

[10]. J.W. Johnson, 2000, A heuristic method for estimating the relative weight of predictor variables in multiple regression. Multivariate Behavioral Research, 35(1): 1-19.
[11]. J.M. LeBreton, R.E. Ployhart, and R.T Ladd, 2004, A Monte Carlo comparison of relative importance methodologies. Organizational Research Methods, 7(3): 258–282.
[12]. J. Cohen, and P, Cohen, *Applied multiple regression/correlation analysis for the behavioral sciences* (Hillsdale, NJ: Erlbaum, 1983).
[13]. E.J. Pedhazur, *Multiple regression in behavioral research* (Fort Worth, TX: Harcourt Brace, 1997).
[14]. J.M. Cortina, 2002, Big things have small beginnings: An assortment of “minor” methodological misunderstandings. Journal of Management, 28(3): 339-362.
[15]. L.V. Hedges and I. Olkin, *Statistical methods for meta-analysis* (New York, NY: Academic Press, 1985).
[16]. J.E. Hunter and F.L. Schmidt, *Methods of meta-analysis: Correcting error and bias in research findings* (Newbury Park, CA: Sage, 1990).
[17]. D.A. Biddle, 2010, Should employers rely on local validation studies of validity generalization (VG) to support the use of employment tests in Title VII situations? Public Personnel Management, 39(4): 307-326.
[18]. P.E. Green, J.D. Carroll, and W.S. DeSarbo, 1978, A new measure of predictor variable importance in multiple regression. Journal of Marketing Research, 15(3): 356-360.
[19]. M.J.R. Healy, 1990, Measuring importance. Statistics in Medicine, 9(6): 633-637.
[20]. R.M. Johnson, 1966, The minimal transformation to orthonormality. Psychometrika, 31(1): 61-66.
[21]. R.B. Darlington, 1968, Multiple regression in psychological research and practice. Psychological Bulletin, 69(3): 161-182.
[22]. L.V. Hedges, and I. Olkin, 1981, The asymptotic distribution of communality components. Psychometrika, 46(3): 331-336.
[23]. A.M. Mood, 1971, Partitioning variance in multiple regression analysis as a tool for developing learning models. American Educational Research Journal, 8(2): 191-202.
[24]. R.G. Newton and D.J. Spurrell, 1967, Examples of the use of elements for clarifying regression analysis. Applied Statistics, 16(2): 165-172.
[25]. P.E. Green, and D.S. Tull, *Research for marketing decisions (*Englewood Cliffs, NJ: Prentice-Hall, 1975).
[26]. P.J. Hoffman, 1960, The paramorphic representation of clinical judgment. Psychological Bulletin, 57(1): 116-131.
[27]. L. Fabbris, 1980, Measures of predictor variable importance in multiple regression: An additional suggestion. Quality & Quantity, 14(6): 787–792.
[28]. J.M. LeBreton, M.B. Hargis, B. Griepentrog, F.L. Oswald, and R.E. Ployhart, 2007, A multidimensional approach for evaluating variables in organizational research and practice. Personnel Psychology, 60(2): 475-498.
[29]. J.M. LeBreton and S. Tonidandel, 2008, Multivariate relative importance: Extending relative weight analysis to multivariate criterion spaces. Journal of Applied Psychology, 93(2): 329-345.
[30]. S. Tonidandel and J.M. LeBreton, 2011, Relative Importance Analysis: A Useful Supplement to Regression Analysis. Journal of Business Psychology, 26(1): 1-9.
[31]. S.J. Snell, S. Tonidandel, P. Braddy, and J. Fleenor, 2014, The relative importance of political skill dimensions for predicting managerial effectiveness. European Journal of Work and Organizational Psychology, 23(6): 915-929.
[32]. K.T. Scherer, M. Baysinger, D. Zolynsky and J.M. LeBreton, 2013, Predicting counterproductive work behaviors with subclinical psychopathy: Beyond the Five Factor Model of personality. Personality and Individual Differences, 55(3): 300–305.
[33]. P.W. Braddy, J. Gooty, J.W. Fleenor, and F.J. Yammarino, 2013, Leader behaviors and career derailment potential: A multi-analytic method examination of rating source and self–other agreement. The Leadership Quarterly, 25(2): 373-390.
[34]. Y.C.E. Chao, Y. Zhao, L.L Kupper and L.A. Nylander-French, 2008, Quantifying the Relative Importance of Predictors in Multiple Linear Regression Analyses for Public Health Studies. Journal of Occupational and Environmental Hygiene, 5(8): 519-529.
[35]. D.J. Hand and V. Vinciotti, 2003, Local versus global models for classification problems: Fitting models where it matters. The American Statistician, 57(2): 124-130.
[36]. S.P. Bandyoapdhyay, D. Nelson, M. Greenwood, G. Brittan and J. Berwald, 2011, The logic of Simpson’s paradox. Synthese, 181(3): 185-208.
[37]. C. Blyth, 1972, On Simpson’s paradox and the sure-thing principle. Journal of the American Statistical Association, 67(338): 364–366.
[38]. R.A. Kievit, W.E. Frankenhuis, L.J. Waldorp, and D. Borsboom, 2013, Simpson’s Paradox in Psychological Science: A Practical Guide. Frontiers in Psychology, 4(1): 1-14.
[39]. J. Pearl, Simpson’s Paradox: An Anatomy, technical report, UCLA Cognitive Systems Laboratory, Los Angeles, CA, 1999.
[40]. J. Pearl, *Causality* (Cambridge: Cambridge University Press, 2000).
[41]. H. Simpson, 1951, The interpretation of interaction in contingency tables. Journal of the Royal Statistical Society, Series B, 13(2): 238–241.
[42]. G.U. Yule, 1903, Notes on the Theory of Association of Attributes in Statistics. Biometrika, 2(2): 121–134.
[43]. W.S. Robinson, 1950, Ecological correlations and the behavior of individuals. American Sociological Review, 15(3): 351-357.
[44]. R.A. Kievit, J.W. Romeijn, L.J. Waldorp, J.M.H. Wicherts, S. Scholte and D. Borsboom, 2011, Modeling Mind and Matter: Reductionism and Psychological Measurement in Cognitive Neuroscience. Psychological Inquiry, 22(2): 139-157.
[45]. A.P. Dawid, 1980, Conditional independence for statistical operations. Annals of Statistics, 8(3): 1-31.
[46]. G. Greenland, 2010, Simpson’s Paradox from adding constraints in contingency tables as an example of Bayesian noncollapsibility. The American Statistician, 64(4): 340-344.
[47]. M.A. Hernán, D. Clayton, and N. Keiding, 2011, The Simpson’s paradox unraveled. International Journal of Epidemiology, 40(3): 780-795.
[48]. Y.K. Tu and M. Gilthorpe, 2008, Simpson's paradox, Lord's paradox, and suppression effects are the same phenomenon-the reversal paradox. Emergent Themes in Epidemiology, 5(2): 1-9.
[49]. M.L. Howe, F.M. Rabinowitz and M.J. Grant, 1993, On measuring (in)dependence of cognitive processes. Psychological Review, 100(4): 737-747.
[50]. R.L. Wardrop, 1995, Simpson’s paradox and the hot hand in basketball. The American Statistician, 49(1): 24-28.
[51]. C.R. Charig, D.R Webb, S.R. Payne, and J.E.A. Wickham, 1986, Comparison of treatment of renal calculi by open surgery, percutaneous nephrolithotomy, and extracorporeal shockwave lithotripsy. British Medical Journal, 292(1): 879-882.
[52]. J. Fox, *Nonparametric simple regressions* (London, England: Sage University Papers, 2008).
[53]. H. Wu, and J.T. Zhang, *Nonparametric regression methods for longitudinal data analysis* (Hoboken, NY: Wiley-Interscience, 2006).

[54]. R.R. Wilcox, Nonparametric estimation, in D. A. Belsley and E. J. Kontoghiorghes (Eds.), *Handbook of computational econometrics*, 5 (Hoboken, NJ: Wiley, 2001) 153-182.
[55]. R.R. Wilcox, 2004, Kernel density estimators: An approach to understanding how groups differ. Understanding Statistics, 3(4): 333-348.
[56]. A.S. Rossi, *Caring and doing for others: Social responsibility in the domains of family, work, and community (*Chicago, IL: University of Chicago Press, 2001).
[57]. C. L. M. Keyes, *The social side of psychological well-being*, doctoral diss., University of Wisconsin, Madison, WI, 1995.
[58]. T, Gasser, D. Gervini and L. Molinari, Kernel estimation, shape-invariant modeling and structural analysis, in R. Hauspie, N. Cameron, and L. Molinari (Eds.), *Methods in human growth research*, 7 (Cambridge: Cambridge University Press, 2004) 179-204.
[59]. R.R. Hocking, 1976, The analysis and selection of variables in linear regression. Biometrics, 32(1): 1-49.
[60]. P.F. Thall, K.E. Russell, and R.M. Simon, 1997, Variable selection in regression via repeated data splitting. Journal of Computational and Graphical Statistics, 6(4): 416–434.
[61]. V.A. Epanechnikov, 1969, Non-Parametric estimation of a multivariate probability density. Theory of Probability & Its Applications, 14(1): 153-158.
[62]. N.I. Fisher, 1989, Smoothing a sample of circular data. Journal of Structural Geology, 11(6): 775-778.
[63]. S.T Chiu, 1996, A comparative review of bandwidth selection for kernel density estimation. Statistica Sinca, 6(1): 129-145.
[64]. N.B. Heidenreich, A. Schindler, and S. Sperlich, 2013, Bandwidth selection for kernel density estimation: a review of fully automatic selectors. AstA Advances in Statistical Analysis, 97(4): 403-443.
[65]. M. Köhler, A. Schindler, and S. Sperlich, 2014, A review and comparison of bandwidth selection methods for kernel regression. International Statistical Review, 82(2): 243-274.
[66]. D.W. Scott, 1979, On optimal and data-based histograms. Biometrika, 66(3): 605-610.
[67]. J.S. Marron, and A.B. Tsybakov, 1995, Visual Error Criteria for Qualitative Smoothing. Journal of the American Statistical Association, 90(430): 499-507.
[68]. S.J. Sheather, and M.C. Jones, 1991, A reliable data-based bandwidth selection method for Kernel Density Estimation. Journal of the Royal Statistical Society, Series B, 53(3): 683-690.
[69]. D.V. Budescu and R. Azen, 2004. Beyond global measures of relative importance: Some insights from dominance analysis. Organizational Research Methods, 7(3): 341-350.
[70]. D. Borsboom, 2006, The attack of the psychometricians. Psychometika, 71(3): 425-440.

APPENDIX A

Example SAS Syntax for Calculation of Observational Weights and Weighted Correlation Around a Focal Point

```
/*Macro: Establishes Focal Point of Focal Variable*/
%let FP3 = 65;
/*Macro: Establishes Standard Deviation of Focal Variable*/
%let SD3 = 12.29757690;
/*Macro: Establishes Sample Size*/
%let N = 3931;
/*Establish Pi*/
data MidusWorkingFile;
set   MidusWorkingFile;
pi=constant("pi");
run;
 /*Bandwidth*/
data MidusWorkingFile;
set   MidusWorkingFile;
BW = (2)*(&N ** -.2)*&SD3;
run;
/*Scaled Distance*/
data MidusWorkingFile;
set   MidusWorkingFile;
Scaled_Distance = (Age - &FP3)/(BW);
run;
/*Kernal_Function-Part 1*/
data MidusWorkingFile;
set   MidusWorkingFile;
KF1 = ((1/ (2*pi) **.5));
run;
/*Kernal_Function-Part 2*/
data MidusWorkingFile;
set   MidusWorkingFile;
KF2 = EXP((-Scaled_Distance**2)/2);
run;
/*Kernal_Function-Final*/
data MidusWorkingFile;
set   MidusWorkingFile;
```

```
KFF = (KF1 * KF2);
run;
 /*Final Weights*/
data MidusWorkingFile;
set   MidusWorkingFile;
Final_Weights = (KFF /.399);
run;
/*Organization*/
data MidusWorkingFile2;
set   MidusWorkingFile;
Final_Weights=round(Final_Weights,0.01);
run;
 data MidusWorkingFileFinal;
set   MidusWorkingFile2;
format Final_Weights f12.2;
run;
/*Weighted (by Age) Correlation*/
Proc corr Data=MidusWorkingFileFinal outp=Corr&FP3  noprob  nomiss  ;
VAR
Social_Integration
Openness Conscientiousness Extraversion Agreeableness Neuroticism;
weight Final_Weights;
RUN;
```

APPENDIX B

Example SPSS Syntax to Create Dataset for Predictors and Outcomes Based on Weighted Correlation

```
input program.
+ loop #i = 1 to 3931.
+ do repeat response=r1 to r6 .
+ compute response = normal(1) .
+ end repeat.
+ end case.
+ end loop.
+ end file.
end input program.
correlations r1 to r6 / statistics=descriptives.
* Factor procedure computes pr1 to pr5, which are standard MVN .
factor variables = r1 to r6 / print = default det
/criteria = factors(6) /save=reg (all,pr).
* use matrix to set corr matrix.
* x is a 10,000 by 5 matrix of uncorrelated standard normals .
* cor is the target covariance matrix.
* cho is the Cholesky factor of cor .
* newx is the 10,000 by 5 data matrix which has target covariance matrix .
matrix.
get x / variables=pr1 to pr6.
compute cor= {1, 0.127079284     ,0.194737955,    0.238988553,     0.226660459,     -0.212788955;
0.127079284,     1,          0.231110658      ,0.506672464,0.286711941           ,-0.179896734;
0.194737955,     0.231110658      ,1,         0.329817167,     0.322309849      ,-0.193066907;
0.238988553,     0.506672464      ,0.329817167      ,1,         0.514181278      ,-0.216299892;
0.226660459,     0.286711941,     0.322309849      ,0.514181278,     1          ,-0.124612008;
-0.212788955     ,-0.179896734     ,-0.193066907,    -0.216299892,     -0.124612008     ,1 }.
compute deter=det(cor).
print deter / title "determinant of corr matrix" / format=f10.7 .
print sval(cor) / title "singular value decomposition of corr".
print eval(cor) / title "eigenvalues of input corr".
compute condnum=mmax(sval(cor))/mmin(sval(cor)).
print condnum / title "condition number of corr matrix" / format=f10.2 .
compute cho=chol(cor).
```

```
print cho / title "cholesky factor of corr matrix" .
compute chochek=t(cho)*cho.
print chochek / title "chol factor premult by its transpose " /format=f10.2 .
compute newx=x*cho.
compute newx=newx*1 + 0.
save newx /outfile=* /variables= nr1 to nr6.
end matrix.
correlations nr1 to nr6 / statistics=descriptives.
```

APPENDIX C

Example R Syntax to Generate Relative Weights and Bootstrapped Confidence Intervals

```
rawdata<-read.csv("C:\\Users\\rhermida\\Desktop\\Professional\\Manuscripts\\Immediate
Manuscripts\\LRW\\FP44.csv", header=TRUE)
attach(rawdata)
thedata<-data.frame(Social_Integration, Openness, Conscientousness, Extraversion, Agreeableness,
Neuroticism)
Labels<-names(thedata)[2:length(thedata)]
multRegress<-function(mydata){
numVar<<-NCOL(mydata)
Variables<<- names(mydata)[2:numVar]
mydata<-cor(mydata, use="complete.obs")
RXX<-mydata[2:numVar,2:numVar]
RXY<-mydata[2:numVar,1]
RXX.eigen<-eigen(RXX)
D<-diag(RXX.eigen$val)
delta<-sqrt(D)
lambda<-RXX.eigen$vec%*%delta%*%t(RXX.eigen$vec)
lambdasq<-lambda^2
beta<-solve(lambda)%*%RXY
rsquare<<-sum(beta^2)
RawWgt<-lambdasq%*%beta^2
import<-(RawWgt/rsquare)*100
result<<-data.frame(Variables, Raw.RelWeight=RawWgt, Rescaled.RelWeight=import)
}
multBootstrap<-function(mydata, indices){
	mydata<-mydata[indices,]
	multWeights<-multRegress(mydata)
	return(multWeights$Raw.RelWeight)
}
multBootrand<-function(mydata, indices){
	mydata<-mydata[indices,]
	multRWeights<-multRegress(mydata)
	multReps<-multRWeights$Raw.RelWeight
	randWeight<-multReps[length(multReps)]
	randStat<-multReps[-(length(multReps))]-randWeight
	return(randStat)
}
#bootstrapping
install.packages("boot")
library(boot)

mybootci<-function(x){
	boot.ci(multBoot,conf=0.95, type="bca", index=x)
}

runBoot<-function(num){
	INDEX<-1:num
	test<-lapply(INDEX, FUN=mybootci)
	test2<-t(sapply(test,'[[',i=4)) #extracts confidence interval
```

```
	CIresult<<-data.frame(Variables, CI.Lower.Bound=test2[,4],CI.Upper.Bound=test2[,5])
}
myRbootci<-function(x){
	boot.ci(multRBoot,conf=0.95,type="bca",index=x)
}
runRBoot<-function(num){
	INDEX<-1:num
	test<-lapply(INDEX,FUN=myRbootci)
	test2<-t(sapply(test,'[[',i=4))
CIresult<<-data.frame(Labels,CI.Lower.Bound=test2[,4],CI.Upper.Bound=test2[,5])
}
myCbootci<-function(x){
	boot.ci(multC2Boot,conf=0.95,type="bca",index=x)
}
runCBoot<-function(num){
	INDEX<-1:num
	test<-lapply(INDEX,FUN=myCbootci)
	test2<-t(sapply(test,'[[',i=4))
CIresult<<-data.frame(Labels2,CI.Lower.Bound=test2[,4],CI.Upper.Bound=test2[,5])
}
myGbootci<-function(x){
	boot.ci(groupBoot,conf=0.95,type="bca",index=x)
}
runGBoot<-function(num){
	INDEX<-1:num
	test<-lapply(INDEX,FUN=myGbootci)
	test2<-t(sapply(test,'[[',i=4))
CIresult<<-data.frame(Labels,CI.Lower.Bound=test2[,4],CI.Upper.Bound=test2[,5])
}
multRegress(thedata)
RW.Results<-result
RSQ.Results<-rsquare
#Bootstrapped Confidence interval around the individual relative weights
#Please be patient -- This can take a few minutes to run
multBoot<-boot(thedata, multBootstrap, 5000)
multci<-boot.ci(multBoot,conf=0.95, type="bca")
runBoot(length(thedata[,2:numVar]))
CI.Results<-CIresult
#Bootstrapped Confidence interval tests of Significance
#Please be patient -- This can take a few minutes to run
randVar<-rnorm(length(thedata[,1]),0,1)
randData<-cbind(thedata,randVar)
multRBoot<-boot(randData,multBootrand, 5000)
multRci<-boot.ci(multRBoot,conf=0.95, type="bca")
runRBoot(length(randData[,2:(numVar-1)]))
CI.Significance<-CIresult
#R-squared For the Model
RSQ.Results
#The Raw and Rescaled Weights
RW.Results
#BCa Confidence Intervals around the raw weights
CI.Results
#BCa Confidence Interval Tests of significance
#If Zero is not included, Weight is Significant
CI.Significance
```